EU privacy and data protection law applied to AI: unveiling the legal problems for individuals

Häuselmann, A.N.

Universiteit Leiden

EU Privacy and Data Protection Law Applied to AI: Unveiling the Legal Problems for Individuals

Andreas Nicolas Häuselmann

EU Privacy and Data Protection Law Applied to AI: Unveiling the Legal Problems for Individuals

Proefschrift

ter verkrijging van

de graad van doctor aan de Universiteit Leiden, op gezag van rector magnificus prof.dr.ir. H. Bijl, volgens besluit van het college voor promoties

te verdedigen op dinsdag 23 April 2024

klokke 15.00 uur

door

Andreas Nicolas Häuselmann

geboren te Schlieren, Zwitserland

Promotoren:	prof. dr. ir. B.H.M. Custers
	prof. dr. G.J. Zwenne
Promotiecommissie:	prof. dr. B.W. Schermer
	dr. G. Malgieri
	prof. dr. N.N. Purtova Msc LL.M. (Universiteit Utrecht)
	prof. dr. A.M. Klingenberg (Rijksuniversiteit Groningen)
	prof. dr. M. Finck LL.M. (Universität Tübingen, Duitsland)

Table of contents

Abbreviations

AC	Affective computing
AI	Artificial Intelligence
ANN	Artificial neural networks
BCI	Brain-computer interfaces
AR	Automated reasoning
DL	Deep learning
CJEU	Court of Justice of the European Union
CNN	Convolutional neural networks
CV	Computer vision
FACS	Facial Action Coding System
GDPR	General Data Protection Regulation
ECHR	European Convention on Human Rights
ECtHR	European Court of Human Rights
EDPB	European Data Protection Board
EUCFR	European Union Charter of Fundamental Rights
ePD	ePrivacy Directive
ML	Machine learning
NLP	Natural language processing
RL	Reinforcement learning
SA	Supervisory authority
TSD	Trade secrets directive

1 Introduction

This thesis aims to unveil the legal problems individuals face when EU privacy and data protection law is applied to Artificial Intelligence (AI). This chapter begins with Section 1.1, which outlines the context and social relevance of this research. Sections 1.2 and 1.3 introduce the research question and the methodologies used to answer it. Section 1.4 clarifies the scope of the thesis by explaining the corresponding limitations, and Section 1.5 outlines the structure of this thesis.

1.1 Context and social relevance

The spectacular emergence of algorithms in the last decade has paved the way for the age of AI. In this new era, self-learning algorithms, automated predictions, classifications and various forms of scoring become even more commonplace. There is a growing public concern about the role of AI in daily life.[1] Today, hiring decisions are influenced by AI-powered emotion software capable of detecting personality traits and emotional states of job applicants based on their facial expressions.[2] Companies intend to use AI to 'hack the human brain'[3] and develop the means to allow it to control devices directly with neurodata, for example, by 'typing with thoughts'.[4] In addition, virtual assistants may use 'sniffer' algorithms to identify trigger words uttered by users, indicating statements of preference (e.g. like or love), translate them into keywords, and make these keywords subsequently accessible to advertisers.[5] When analysed with AI, a speech recording can reveal a rich variety of information about an individual, which goes well beyond the individual's identity. This includes information on the individual's emotional state,[6] personality traits, sleepiness, intoxication, physical and mental health, as well as their socioeconomic status.[7]

[1] See < https://www.liberties.eu/en/stories/impact-of-artificial-intelligence-in-everyday-life/44222 > accessed 8 February 2024; Alec Tyson, Emma Kikuchi, 'Growing public concern about the role of artificial intelligence in daily life' Pew Research Center (2023) < https://www.pewresearch.org/short-reads/2023/08/28/growing-public-concern-about-the-role-of-artificial-intelligence-in-daily-life/ > accessed 8 February 2024.
[2] Patricia Nilsson, 'How AI helps recruiters track jobseeker's emotions' *The Financial Times* (New York 3 March 2018) < https://www.ft.com/content/e2e85644-05be-11e8-9650-9c0ad2d7c5b5 > accessed 8 February 2024; For instance, HumeAI which provides AI-powered tools helping recruiters to assess personality traits as well as emotional states of candidates. See < https://hume.ai/products/facial-expression-model/ > and < https://gethume.com/blog5/artificial-intelligence-for-recruiting > accessed 8 February 2024.
[3] Nick Statt, 'Kernel is Trying to Hack the Human Brain—But Neuroscience has a Long Way to Go' *The Verge* (New York 22 February 2017) <https://www.theverge.com/2017/2/22/14631122/kernel-neuroscience-bryanjohnson-human-intelligence-ai-startup> accessed 8 February 2024.
[4] John Constine, 'Facebook is building brain-computer interfaces for typing and skin-hearing' TechCrunch (San Francisco 19 April 2017) < https://techcrunch.com/2017/04/19/facebook-brain-interface/ > accessed 8 February 2024.
[5] Edara Kiran, 'Key Word Determinations From Voice Data' US Patent Number US 8798995B1 (Assignee: Amazon Technologies, Inc.) August 2014 <https://patentimages.storage.googleapis.com/bd/ed/2b/c4c67cc5a9f1ab/US8798995.pdf>, accessed 8 February 2024.
[6] Andreas Nautsch et al, 'Preserving privacy in speaker and speech characterisation' (2019) Vol 58 Computer Speech & Language 441, 444.
[7] For more detailed information and related studies, see Jacob Leon Kröger, Otto Hans/Martin Lutz, Philip Raschke, 'Privacy Implications of Voice and Speech Analysis – Information Disclosure by Inference' in Michael Friedewald et al (eds) *Privacy and Identity management. Data for Better Living: AI and Privacy* (Springer 2020) 243.

Nevertheless, real-world examples point to reasoning deficiencies in current AI systems. They generalise, but do not distinguish, and they seem to ignore context. For example, Google's AI system developed for the purpose of recognising child abuse incorrectly classified a father as a child abuser by completely neglecting the situational context of the photograph.[8] These examples inevitably raise the question whether the current EU legal framework for the fundamental rights to privacy and the protection of personal data is fit for purpose when applied to AI.

Machine learning (ML), a major discipline of AI, produces probable yet inevitably uncertain knowledge.[9] In essence, ML-based predictions constitute 'educated guesses or bets, based on large amounts of data'.[10] If such predictions are treated as *facts*, despite their probabilistic nature, this will have real impact on humans. A statistic is factual, but with error margins, and it is not 'absolute'. The output generated by AI also raises questions in terms of fairness. Reasoning deficiencies in another AI discipline called 'automated reasoning' (AR) can lead to severe adverse effects for the individual concerned. Imagine, for example, the father who has been wrongfully classified as a child abuser by Google's AI system. By means of the AI discipline affective computing (AC), machines can access the *emotional life* of individuals, information that is highly personal, intimate and private.[11] Assessing such information may allow entities to intentionally exploit the behaviour, thoughts and decision-making vulnerabilities of individuals. As emotions play an important role in the elicitation of autonomous motivated behaviour[12] and reasoning,[13] access to an individual's emotions can therefore severely affect their personal autonomy and freedoms. For this reason, emotions merit specific protection.

That new technologies have an impact on society is naturally understood. However, the effects of new technologies such as AI cannot be easily predicted until they are widely deployed. Once they have been deployed, they are difficult to change.[14] It is also commonly understood that any disruptive technology entails risks and complex policy challenges. This applies, in particular, to AI and the fundamental rights to privacy and the protection of personal data.

[8] Kashmir Hill, 'A Dad Took Photos of His Naked Toddler for the Doctor. Google Flagged Him as A Criminal' *The New York Times* (New York, 21 August 2022) < https://www.nytimes.com/2022/08/21/technology/google-surveillance-toddler-photo.html > accessed 8 February 2024.
[9] Brent Daniel Mittelstadt et al, 'The ethics of algorithms: Mapping the debate' (2016) Vol 3 Iss 2 Big Data & Society 1, 4.
[10] Teresa Scantaburlo, Andrew Charleswoth, Nello Cristianini, 'Machine Decisions and Human Consequences' in Karen Yeung and Martin Lodge (eds) *Algorithmic Regulation* (OUP 2019) 57; Lee A Bygrave, 'Machine Learning, Cognitive Sovereignty and Data Protection Rights with Respect to Automated Decisions' (2021) University of Oslo Faculty of Law Legal Studies Research Paper Series No. 202-35 at 5 <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3721118> accessed 8 February 2024.
[11] Rosalind W Picard, *Affective Computing* (MIT Press 1997) 118.
[12] Leen Vandercammen et al, 'On the Role of Specific Emotions in Autonomous and Controlled Behaviour' (2014) Vol 28 Iss 5 European Journal of Personality 413, 445.
[13] Steffen Steinert, Orsolya Friedrich, 'Wired Emotions: Ethical Issues of Affective Brain–Computer Interfaces' (2020) Vol 26 Science and Engineering Ethics 351, 352.
14 Nicolas Carr, *The Big Switch: Rewiring the World, from Edison to Google* (W. W. Norton & Company 2009) N 1 at 87.

As indicated by its title, this thesis is about EU privacy and data protection law. It investigates legislation adopted by the *European Union*.[15] The fundamental right to *privacy* according to Article 7 of the European Union Charter of Fundamental Rights (EUCFR) protects everyone's 'right to respect for his or her private and family life, home and communications'. The fundamental right to the *protection of personal data* enshrined in Article 8 EUCFR grants everyone 'the right to the protection of personal data concerning him or her' ('fundamental right to data protection'). These fundamental rights are closely linked, but not identical,[16] as they differ in terms of material and personal scope.[17] Both fundamental rights are further substantiated in EU secondary law. The most relevant legislation in EU secondary law are the General Data Protection Regulation (GDPR) and the ePrivacy Directive (ePD). Articles 7-8 EUCFR, GDPR and the ePD form the '*current legal framework*'. The focus lies on the GDPR because it is the most comprehensive and influential piece of legislation. EU secondary law on the fundamental right to privacy, next to the ePD, is scarce if not absent entirely.

AI entails many disciplines such as machine learning, natural language processing, computer vision, affective computing and automated reasoning. When the current legal framework is applied to these AI disciplines, three types of legal problems may occur: (1) legal provisions are violated, (2) legal provisions cannot be enforced and (3) legal provisions are not fit for purpose to protect the fundamental right at stake. I investigate these legal problems from the perspective of *individuals*, meaning natural persons, as they are holders of fundamental rights and thus have an interest in effective protection. The focus is on two types of provisions contained in the current legal framework: principles[18] and enforceable rights.[19] The three types of legal problem are not mutually exclusive. Likewise, a discipline of AI could lead to more than one type of legal problem.

This thesis aims to contribute to scientific knowledge on the suitability of EU privacy and data protection law concerning AI. As opposed to other studies,[20] this thesis considers different AI disciplines because there is nothing like 'the one AI'. With this thesis, I address current gaps in scholarship by

[15] Also, the human right to respect for private and family life, his home and correspondence according to Article 8 ECHR and related case law are considered, because these sources influence the interpretation of the corresponding fundamental rights enshrined in the EUCFR.
[16] Juliane Kokott, Christoph Sobotta 'The distinction between privacy and data protection in the jurisprudence of the CJEU and ECtHR' (2013) Vol 3 No 4 International Data Privacy Law 222, 223, 228; Herke Kranenborg, Commentary of Article 8 in Steve Peers et al (eds), *The EU Charter of Fundamental Rights* (Hart/Beck 2014) 229.
[17] The material scope of the fundamental right to data protection seems to be broader whereas it is more narrow in terms of personal scope as it excludes legal persons. See Juliane Kokott, Christoph Sobotta 'The distinction between privacy and data protection in the jurisprudence of the CJEU and ECtHR' (2013) Vol 3 No 4 International Data Privacy Law 222, 225; Joined Cases C–92/09 and C–93/09 *Volker und Markus Schecke and Eifert* [2010] ECR I-11063, paras 52, 53 and 87.
[18] Proportionality, Lawfulness, Fairness, Transparency, Accuracy, Purpose limitation, Data minimisation, Confidentiality, Exhaustive enumeration, Accountability.
[19] Informational privacy, bodily privacy, mental privacy, communicational privacy, right of access, right to rectification, right to erasure, right to data portability, right to object, right not to be subject to ADM.
[20] Giovanni Sartor, Francesca Lagioia, 'The impact of the General Data Protection Regulation (GDPR) on artificial intelligence' (2020) Study for the European Parliament's Panel for the Future of Science and Technology < https://www.europarl.europa.eu/RegData/etudes/STUD/2020/641530/EPRS_STU(2020)641530_EN.pdf > accessed 8 February 2024; Frederico Marengo, *Privacy and AI: Protecting Individuals' Rights in the Age of AI* (2023).

engaging with relevant strands of computer science relating to the different AI disciplines and three specific types of legal problems. Ultimately, the suitability of the current legal framework depends on the particular AI discipline deployed and the types of legal problems caused by it.

1.2 Research question

This thesis aims to answer the following research question:

To what extent do the developments in AI require a new legal framework for the fundamental rights to privacy and the protection of personal data?

The research is structured into five subquestions:

Subquestion 1: What is AI, and what disciplines exist?
Subquestion 2: What is the current legal framework?
Subquestion 3: What legal problems arise or may arise when the *principles* enshrined in the current legal framework are applied to AI?
Subquestion 4: What legal problems arise or may arise when the *enforceable rights* enshrined in the current legal framework are applied to AI?
Subquestion 5: How should the incompatibilities of the current legal framework identified in Subquestions 3 and 4 be addressed?

Subquestion 1 examines what AI is (and what it is not), the technology of AI and the terminology used. It first elaborates on existing definitions of AI and subsequently focusses on AI disciplines that are most problematic in the context of fundamental rights to privacy and data protection. These disciplines include machine learning, natural language processing, computer vision, affective computing and automated reasoning. By means of brain-computer interfaces (BCI), also known as mind-machine interfaces, machine learning facilitates the processing of mental data, meaning any data used to infer the mental states of an individual including thoughts, beliefs and the underlying psychological mechanisms. Keyword determination systems powered by approaches from the AI discipline natural language processing attempt to identify trigger words that indicate statements of preference (such as ‘like’ or ‘love’) from speech recorded by virtual assistants and translate them into keywords for advertisement purposes. Techniques from the AI discipline computer vision are used to identify individuals based on their gait because the way in which individuals walk constitutes a unique identifier. The AI discipline affective computing facilitates the processing of emotion data, information which is highly sensitive and intimate. The AI discipline of automated reasoning focusses on automated logical reasoning, probabilistic reasoning and common sense reasoning. Subquestion 1 establishes a

proper understanding of what AI is and how it works, which is required for answering Subquestions 3, 4 and 5.

Subquestion 2 introduces the current EU legal framework relating to fundamental rights to privacy and the protection of personal data. It also examines relevant secondary EU law, namely, the General Data Protection Regulation, which is the most relevant piece of secondary EU law in the field of data protection, and the ePrivacy Directive. This subquestion sets the stage for addressing Subquestions 3 and 4, namely, what legal problems arise or may arise when the current legal framework is applied to the AI disciplines introduced in Subquestion 1.

Subquestion 3 discusses what legal problems arise or may arise when the *principles* enshrined in the current legal framework are applied to AI. To identify the legal problems, three types of legal problems are identified, namely, that (1) legal provisions are violated, (2) legal provisions cannot be enforced, and (3) legal provisions are not fit for purpose to protect the fundamental right at stake. The reasoning behind the choice to focus on these three types of legal problems is as follows. Violations of fundamental rights (Type 1) must be prevented. For example, unsupervised ML processes personal data for inexplicit purposes – the processing itself determines the purpose and future use of the personal data processed. Such processing violates the purpose limitation principle, which leads to a Type 1 legal problem. Situations in which legal provisions cannot be enforced (Type 2) are also not acceptable, because they lead to negative consequences for the de facto protection of fundamental rights. For example, the unclear substantive meaning of the fairness principle reduces legal certainty and makes it less likely that it will be enforced by means of private and regulatory enforcement. Provisions that are not fit for purpose to protect the fundamental right at stake (Type 3) point to shortcomings of the current legal framework. For example, the principle of enhancing protection for special data and the legislator's approach to exhaustively enumerate such causes a Type 3 legal problem because it does not keep up with technological developments facilitated by AI. This leads to significant gaps of protection, for example, regarding the processing of new types of sensitive personal data enabled by AI like emotion data, neurodata and mental data. Insights about Type 2 and 3 legal problems are essential when considering how the incompatibilities of the current legal framework identified should be addressed, which is the aim of Subquestion 5.

Subquestion 4 discusses what legal problems arise or may arise when the *enforceable rights* enshrined in the current legal framework are applied to AI. It identifies the same three types of legal problems as discussed in Subquestion 3. Due to the wide scope of the fundamental right to privacy, Subquestion 4 focusses on four dimensions with particular relevance for AI, namely, informational, bodily, mental, and communicational privacy. First, the fundamental right to privacy provides individuals with a form

of informational self-determination,[21] which is an extremely important dimension because AI relies heavily on the processing of information. Second, physical and mental integrity, two elements falling under the term 'private life' as developed in the corresponding case law, [22] are highly relevant. Some AI disciplines, for example, AC and ML, rely on body functions and characteristics (genetic codes, biometrics, physiological information) to gain access to mental states of individuals (thoughts, feelings, emotional states). Finally, communication is an important dimension because AI is prone to interfere with the right to respect confidential communication. AI computes communications in various forms, for instance, by means of NLP and ML. Regarding the fundamental right to data protection, Subquestion 4 focusses on enforceable rights which data subjects have according to the GDPR. They implement the requirements enshrined in the fundamental right to data protection.[23] These enforceable rights are the right of access, the right to rectification, the right to erasure, the right to portability, the right to object and the right not to be subject to automated decision-making.

Subquestion 5 discusses how the incompatibilities of the current legal framework identified in Subquestions 3 and 4 could be addressed. It does so by exploring suitable legal solutions. When referring to legal solutions, I mean (i) new interpretations of existing provisions, (ii) amending existing provisions or (iii) introducing new provisions that can 'solve' the selected legal problems. The verb 'solve' in the latter sense refers to suggestions and recommendations that can contribute to actual solutions to the selected legal problems.

1.3 Methodology

Following a single methodology can limit creativity in research by imposing a standard way of investigating law.[24] For this reason, I use several methodologies. The main research question and the corresponding subquestions determine the methodology.

The *first* method is desk research,[25] consisting of the review and analysis of typical sources in legal research. These are legislation, legislative history, case law and academic literature. To answer Subquestion 1, namely, what AI is and what disciplines exist therein, I focus on the corresponding academic literature in the field. To point to real-world use cases of AI and corresponding issues, I also use a limited number of non-scientific journalistic texts. This also incorporates the most recent

[21] *Satakunnan Markkinapörssi Oy and Satamedia Oy v. Finland* App no 931/13 (ECtHR 27 June 2017) para 137.
[22] *Denisov v Ukraine* App no 76639/11 (ECtHR 25 September 2018) para 95, *S. and Marper v United Kingdom* App no 30562/04 and 30566/04 (ECtHR 4 December 2008) *Pretty v United Kingdom* App no 2346/02 (ECtHR 29 April 2002) para 63.
[23] As emphasised by the CJEU in case law relating to the GDPR's predecessor, the DPD, see Case C-131/12, *Google Spain* [2014] ECR I-317 para 69; Joined Cases C-141/12 & C-372/12, *YS, M* [2014] ECR I-2081, Opinion of AG Sharpston para 55.
[24] Reza Banakar, Max Travers, 'Introduction' in Reza Banakar, Max Travers (eds) *Theory and Method in Socio-Legal Research* (Hart Publishing 2005) x.
[25] Desk research could also be seen as a research strategy rather than a method.

developments on which no academic publications are available (yet). These complementary sources stem from renowned newspapers[26] and technology-oriented news portals.[27] The goal of desk research as a method is acquiring the necessary theoretical knowledge. Therefore, desk research is particularly used for Subquestions 1 and 2 which introduce the AI disciplines and the current legal framework, respectively. I gathered sources from multiple databases, but mainly from the e-libraries of Leiden University[28] and the Eidgenössische Technische Hochschule (ETH) Zürich.[29] The latter was mainly used for sources relating to AI, as the ETH is a renowned university in the disciplines of engineering and technology.[30] The two databases were analysed using several search terms.[31] Sources obtained from these digital repositories were selected and analysed based on relevance,[32] expertise of the author and topicality.

The *second* method investigates to what extent the law in the books *differs* from the law in action. This method is largely inspired by sociolegal studies, which investigate how law is made, interpreted, and enforced.[33] The methodology vacuum in sociolegal studies helps to maintain the field as a truly interdisciplinary one, which is open to theoretical diversity and innovation.[34] Sociolegal studies typically adopt a variety of theoretical positions and methodologies.[35] Many sociolegal studies have been designed to demonstrate the *actual operation* or *lack of efficacy*.[36] Also, emerging movements such as new legal realism develop approaches that account for how law actually *works* in *practice*.[37] Sociolegal scholar Roscoe Pound long ago noted a divergence between 'law in books' and 'law in action'.[38] As is apparent from the main research question, this thesis investigates to what extent the law in the books *differs* from the law in action when *applied* to the multidisciplinary field of AI. Applying in the latter sense means performing legal analysis concerning the legal provision in question and the AI discipline deployed. This sociolegal approach is also recognisable by the three types of legal problems I discuss in this thesis. All of them concern the law in *action*: legal provisions are violated (Type 1), cannot be enforced (Type 2) or are not fit for purpose (Type 3). To unveil the difference between

[26] https://www.theguardian.com/international; https://www.independent.co.uk/; https://www.forbes.com/.
[27] https://www.wired.com/; https://www.theverge.com/.
[28] <https://www.bibliotheek.universiteitleiden.nl/> accessed 8 February 2024.
[29] <https://library.ethz.ch/en/> accessed 8 February 2024.
[30] <https://ethz.ch/en/news-and-events/eth-news/news/2022/04/eth-in-top-10-for-16-subjects.html> accessed 8 February 2024.
[31] For instance 'AI', 'AI disciplines', 'data protection', 'privacy', 'trade secrets', 'common sense', 'emotion data', 'right to rectification', 'right of access', 'enforcement' etc.
[32] Regarding the scope and limits of this thesis, as outlined in Section 1.4.
[33] Naomi Creutzfeldt, 'Traditions of Studying the Social and the Legal' in Naomi Creutzfeld, Marc Mason, Kirsten McConnachie (eds) *Routledge Handbook of Socio-Legal Theory and Methods* (Routledge 2020) 10.
[34] Reza Banakar, Max Travers, 'Introduction' in Reza Banakar, Max Travers (eds) *Theory and Method in Socio-Legal Research* (Hart Publishing 2005) x, xi.
[35] Naomi Creutzfeldt et al, 'Socio-legal theory and methods: introduction' in Naomi Creutzfeld, Marc Mason, Kirsten McConnachie (eds) *Routledge Handbook of Socio-Legal Theory and Methods* (Routledge 2020) 3.
[36] Carrie Menkel-Meadow, 'Uses and Abuses of Socio-Legal Studies' in Naomi Creutzfeld, Marc Mason, Kirsten McConnachie (eds) *Routledge Handbook of Socio-Legal Theory and Methods* (Routledge 2020) 39.
[37] Shauhin Talesh, Elizabeth Mertz and Heinz Klug, 'Introduction to the Research Handbook on Modern Legal Realism' in Shauhin Talesh, Elizabeth Mertz and Heinz Klug (eds) *Research Handbook on Modern Legal Realism* (Edward Elgar Publishing Limited 2021) 3.
[38] Roscoe Pound, 'Law in Books and Law in Action' (1910) Vol 44 Iss 1 Amercian Law Review 12, 35.

law in the books and law in action, I apply the principles and enforceable rights enshrined in the current legal framework to the AI disciplines. I use this method to answer the main research question, as well as Subquestions 3 and 4. To be clear, I refrain from engaging in empirical research, as one might expect from research inspired by sociolegal studies.[39] Instead, I focus on the difference between law in the books and law in action by performing legal analysis, as exposed by the three types of legal problems identified within this thesis. Type 2 legal problems reveal that the law in the books cannot be enforced in practice, and Type 3 legal problems exposes the lack of efficacy when law in the books operates in practice. To exhibit this, I use text boxes that concisely name and describe each legal problem identified. In addition, I use various tables to illustrate the difference between law in the books and law in action. These tables form the basis for further analysis and serve as a tool for drawing the corresponding conclusions.

The *third* method is legal doctrinal research. Doctrinal legal research relates to the first method, that is, desk research, so it is not entirely a separate method. Traditional doctrinal exploration influences how legal academics approach legal questions.[40] This holds also true for this thesis. Doctrinal legal research is a discipline that takes normative positions and makes choices among values and interests. This is inevitable when some interpretation is preferred over alternative ones. Ultimately, the choice to favour a specific interpretation is determined by giving more weight to some interests than to competing ones.[41] I have used this method mainly in addressing Subquestion 5, when setting the scene for possible legal solutions. I have also used this method when suggesting legal solutions to the six legal problems discussed in Chapter 6. There, the focus lies on the interests of natural persons as holders of the fundamental rights to privacy and the protection of personal data.

1.4 Limitations

Since AI covers a broad range of concepts, this thesis pays particular attention to AI disciplines that are problematic in light of the fundamental rights to privacy and data protection. As already outlined in Section 1.2, these disciplines are machine learning, natural language processing, computer vision, facial recognition, affective computing and automated reasoning. Methods of AI that combine AI with Robotics, i.e. 'Embodied Artificial Intelligence', are not in the scope of this thesis. Applications of Embodied AI such as driverless vehicles, surgical robots and companions pose different questions such as liability issues or ethical issues in the context of robot-human interactions.[42] These questions,

[39] Reza Banakar, Max Travers, 'Introduction' in Reza Banakar, Max Travers (eds) *Theory and Method in Socio-Legal Research* (Hart Publishing 2005) x.
[40] Shauhin Talesh, Elizabeth Mertz and Heinz Klug, 'Introduction to the Research Handbook on Modern Legal Realism' in Shauhin Talesh, Elizabeth Mertz and Heinz Klug (eds) *Research Handbook on Modern Legal Realism* (Edward Elgar Publishing Limited 2021) 1.
[41] Mark Van Hoecke, *Methodologies of Legal Research: What Kind of Method for What Kind of Discipline?* (Hart Publishing 2011) 10.
[42] Cándido García Molyneux, Rosa Oyarzabal, 'What Is a Robot (Under EU Law)?' (2018) Vol 1 RAIL: The Journal of Robotics, AI & Law 11, 12.

however, are not in the scope of this research, as they do not primarily concern the fundamental rights to privacy and the protection of personal data.

In terms of regulatory enforcement, this thesis mainly considers binding decisions adopted by the European Data Protection Board (EDPB). The EDPB consists of representatives of national EU supervisory authorities (SAs) responsible for data protection and the European Data Protection Supervisor (EDPS). Binding decisions aim to ensure the correct and consistent application of the GDPR and come into play where SAs in EU Member States interpret provisions of the GDPR differently. This limitation is bound by the focus of this thesis on EU law and needed for practical reasons, in particular the vast number of decisions adopted by SAs. However, I refer to decisions adopted by SAs occasionally. I do so when regulatory enforcement actions adopted by a specific SA particularly relate to AI and binding decisions of the EDPB are absent. Regarding private enforcement, this thesis focusses on CJEU and ECtHR case law. Occasionally, judgements adopted by Courts on Member States level are considered, in particular if corresponding CJEU and ECtHR case law are missing.

In terms of legal problems, the scope of this research is restricted in five ways. First, and foremost, this research exclusively deals with legal problems with respect to legislation adopted by the *European Union*[43] relating to the fundamental rights to privacy and the protection of personal data. It is devoted primarily to legal problems arising from the perspective of *individuals*, that is, *natural persons*. These two notions are used interchangeably. Obviously, legal problems may also arise for other actors involved, in particular controllers that process personal data by means of AI systems. However, legal problems that arise for controllers and other actors, such as providers of electronic communication services, are not within the scope of this research.

Second, the legal problems are restricted to the *principles* and *enforceable rights* enshrined in the current legal framework and leave out other obligations.[44] The principle of confidentiality contained in the ePD is the key principle of ensuring the confidentiality of communications as protected by the fundamental right to privacy. The principles form the basis for the fundamental right to protection of personal data.[45] Similarly, strong[46] and effective data subject rights[47] constitute a prerequisite for the protection of personal data. In addition, the legislator considers principles and enforceable rights as particularly important from a normative perspective. Distinctions in terms of maximum amounts for

[43] Also, the human right to respect for private and family life, his home and correspondence according to Article 8 ECHR and related case law are considered, because these sources influence the interpretation of the corresponding fundamental rights enshrined in the EUCFR.
[44] For instance the obligations contained in Chapter IV GDPR.
[45] Cécile de Terwangne, Commentary of Article 5 in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020) 313.
[46] Case C-132/21 *Nemzeti* [2023] ECR I-2 para 42.
[47] Case C-154/21 *Österreichische Post AG* [2023] ECR I-3 para 39.

administrative fines contained in Article 83 GDPR indicate that infringements of principles and data subject rights are considered *more serious* than infringements of other provisions.[48]

Third, and as already outlined in Section 1.2, this thesis focusses on four dimensions of the fundamental right to privacy that are particularly relevant in the light of AI, namely, informational, bodily, mental and communicational privacy.

Fourth, in terms of the fundamental right to data protection, I focus on the enforceable data subject rights enshrined in the GDPR. These enforceable rights implement the requirements set out in the fundamental right to data protection.

Fifth, this thesis approaches the fundamental rights to privacy and data protection mainly from a horizontal perspective. Private parties and powerful tech companies in particular are at the forefront of the developments in AI. Therefore, the focus is on horizontal relationships between natural persons and private parties, but not on citizens and government. This becomes apparent from the examples mentioned in Chapters 4 and 5.

1.5 Structure

Each of the subquestions introduced in Section 1.2 merits its own dedicated chapter. This thesis is therefore structured as follows:

Chapter 2 answers the first subquestion, namely, what AI is and what AI disciplines exist. It starts with existing definitions of AI (Section 2.1) and then provides an overview of the AI disciplines that seem to be the most problematic in the context of the fundamental rights to privacy and the protection of personal data (Section 2.2). These disciplines include machine learning, natural language processing, computer vision, affective computing and automated reasoning.

Chapter 3 answers the second subquestion, namely, what the current legal framework is. This chapter starts by describing the current legal framework regarding the fundamental right to privacy (Section 3.1) followed by (Section 3.2), which introduces the fundamental right to data protection. Next, (Section 3.3) discusses the most relevant piece of EU secondary law in data protection, namely, the GDPR. Finally, (Section 3.4) introduces the ePrivacy Directive, whose provisions particularise the fundamental rights to privacy and the protection of personal data in the electronic communications sector.

[48] Article 29 Working Party, 'Guidelines on the application of administrative fines for the purposes of Regulation 2016/679' (WP 253, 3 October 2017) 9; European Data Protection Board, 'Guidelines on the calculation of administrative fines under the GDPR' (Guidelines 4/2022, 16 May 2022) 16.

Chapter 4 answers the third subquestion, namely, what legal problems arise or may arise when the *principles* enshrined in the current legal framework are applied to AI. To identify the legal problems that arise or may arise from applying the current legal framework to AI, in this chapter, three types of legal problems are introduced (Section 4.1). Based on this approach, legal problems are identified for each of the AI disciplines outlined in Chapter 2. It focusses on the *principles* enshrined in the current legal framework. Sections 4.2 - 4.8 deal with the principles enshrined in the GDPR, namely, the lawfulness and proportionality (Section 4.2), fairness (Section 4.3), transparency (Section 4.4), purpose limitation (Section 4.5), data minimisation (Section 4.6), accuracy (Section 4.7) and the principle to enhance protection for special categories of personal data (Section 4.8). Section 4.9 elaborates on the principle concerning the confidentiality of communications as enshrined in the ePD. Finally, Section 4.10 concludes by providing an answer to Subquestion 3, including an overview of which AI disciplines lead to which types of legal problems.

Chapter 5 answers the fourth subquestion, namely, what legal problems arise or may arise when the *enforceable rights* enshrined in the current legal framework are applied to AI. Section 5.1 introduces the approach taken to assess legal problems. Sections 5.2 to 5.5 elaborate on the fundamental right to privacy and discuss four dimensions of privacy that are derived from the elements contained in the text of the fundamental right to privacy and corresponding case law. These four dimensions are informational (Section 5.2), bodily (Section 5.3), mental (Section 5.4), and communicational privacy (Section 5.5). Sections 5.6 to 5.11 do the same for the fundamental right to data protection. I focus on the enforceable rights which data subjects have according to the GDPR because they implement the requirements enshrined in the fundamental right to data protection.[49] These enforceable rights are the right to access (Section 5.6), the right to rectification (Section 5.7), erasure (Section 5.8), portability (Section 5.9), right to object (Section 5.10) and the right not to be subject to automated decision-making (Section 5.11). Section 5.12 concludes.

Chapter 6 answers the fifth subquestion, namely, how the incompatibilities of the current legal framework identified in Chapters 4 and 5 could be addressed. Chapters 4 and 5 strongly underscore the difference between the law in books and the law in action. Chapter 6 discusses how the gaps between the law in books and the law in action can be addressed by means of legal solutions. Because it is impossible to address all legal problems identified in Chapters 4 and 5, it focusses on six specific legal problems (Sections 6.2 - 6.7). These six legal problems are chosen based on the selection criterion effectiveness, urgency and novelty.

[49] As emphasised by the CJEU in case law relating to the GDPR's predecessor, the DPD, see Case C-131/12, *Google Spain* [2014] ECR I-317 para 69; Joined Cases C-141/12 & C-372/12, *YS, M* [2014] ECR I-2081, Opinion AG Sharpston para 55.

Chapter 7 answers the main research question of this thesis, namely, to what extent the developments in AI require a new legal framework for the fundamental rights to privacy and the protection of personal data. Section 7.1 answers the main research question. Then Sections 7.2 and 7.3 provide an outlook for future research and other recommendations.

2 Artificial Intelligence (AI)

This chapter aims to answer Subquestion 1, namely, what AI is and what AI disciplines exist.[50] It starts with existing definitions of AI (Section 2.1) and then provides an overview of the AI disciplines that seem to be the most problematic ones from a privacy and data protection perspective (Section 2.2). These disciplines include machine learning (Section 2.2.1), natural language processing (Section 2.2.2), computer vision (Section 2.2.3), affective computing (Section 2.2.4) and automated reasoning (Section 2.2.5). Section 2.3 answers Subquestion 1.

2.1 Definitions of AI

There is no officially agreed definition of Artificial Intelligence (AI). AI covers a wide range of concepts and terms, making it difficult to define. Available definitions often involve ambiguous terms such as 'thinking', 'learning' and 'intelligence'. In 1968, Minsky defined AI as 'the science of making machine do things that would require intelligence if done by men'.[51] Bellman defined AI in 1978 as 'the automation of activities that we associate with human thinking, activities such as decision-making, problem solving, learning, creating, game playing, and so on'.[52] Nilsson described AI as 'activity devoted to making machines intelligent, and intelligence is that quality that enables an entity to function appropriately and with foresight in its environment.'[53] Russell and Norvig organised definitions of AI into four categories: a) thinking humanly, b) acting humanly, c) thinking rationally and d) acting rationally.[54] According to Munakata, AI involves abilities such as 'inference based on knowledge, reasoning with uncertain or incomplete information, various forms of perception and learning, and applications to problems such as control, prediction, classification, and optimization'.[55] More recent definitions are the ones adopted by the Organisation for Economic Co-operation and Development (OECD) and the National Institute of Standards and Technology (NIST). The OECD defines an AI system as a 'machine-based system that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments. Different AI systems vary in their levels of autonomy and adaptiveness after deployment'.[56] NIST defines AI as a 'branch of computer science devoted to developing data processing systems that performs functions normally associated with human intelligence, such as reasoning, learning, and self-improvement'.[57]

[50] A modified version of this chapter was published in Bart Custers, Eduard-Fosch Villaronga (eds) *Law and Artificial Intelligence* (Asser Press 2022). See Andreas Häuselmann, 'Disciplines of AI: An Overview of Approaches and Techniques' 43-70.
[51] Marvin Minsky, *Semantic Information Processing* (MIT Press 1968).
[52] Richard Bellman, *An Introduction to Artificial Intelligence: Can computers think*? (Boyd & Faser 1978) 3
[53] Nils J Nilsson, *The Quest for Artificial Intelligence: A History of Ideas and Achievements* (Cambridge University Press 2010).
[54] Stuart Russel, Peter Norvig, *Artificial Intelligence, A Modern Approach* (3rd edn, Pearson Education 2016) 2.
[55] Toshinori Munakata, *Fundamentals of the New Artificial Intelligence* (2nd edn, Springer 2008) xx.
[56] See < https://oecd.ai/en/wonk/ai-system-definition-update > accessed 8 February 2024.
[57] See < https://csrc.nist.gov/topics/technologies/artificial-intelligence > accessed 8 February 2024.

The field of AI may be divided into narrow and general AI. Narrow AI refers to systems that are able to solve a specific problem or performing a specific task. For an example on a narrow AI system, one can refer to IBM's 'Deep Blue' chess-playing computer. Deep Blue defeated the reigning world champion in chess, Garry Kasparov, in 1997.[58] This example indicates that computers can perform better than humans. However, this holds only true for a narrow domain, such as playing chess. General AI aims to build machines that generally perform on a human level and have a 'human-level' skillset. To achieve this goal, such a system must be able to mimic the functioning of the human brain in the most important aspects.[59] Unlike with narrow AI, general AI arguably has not been achieved yet despite rapid developments, for instance, ChatGPT. Although AI found its 'birth' at the Dartmouth Summer Research Project on AI in the summer of 1956 in New Hampshire,[60] there are many open challenges. According to Shi, AI research is still in its first stage since no breakthrough progress has been achieved for some key challenges such as common sense knowledge representation and uncertain reasoning.[61] Therefore, current AI systems must be considered examples of 'narrow' AI. However, computing power has become more affordable; the computers have become faster and contain larger memories. This led to the 'summer of AI' and it seems reasonable to expect major developments in the field of AI.

In his famous paper, called 'Computing Machinery and Intelligence',[62] Turing proposed the 'Imitation Game', which has later become known as the 'Turing test'.[63] Turing offered his test as a sufficient condition for the existence of AI.[64] This test involves three actors: (A) a machine, (B) a human and (C) another human called the interrogator (see Figure 1.1). In the Turing test, the human interrogator (C) stays in a room apart from the other two actors (A) and (B). The human interrogator knows the machine (A) and human (B) by labels (X) and (Y)[65] and therefore does not know which label is (A) or (B).[66] The object of the test is for the interrogator (C) to determine which of the other two actors is the human and which is the machine[67] by asking (X) and (Y) questions which they must answer.[68] In other words, the human interrogator engages in conversation with either a human or an AI natural language program which are both hidden from view. If the human interrogator cannot reliably

[58] https://www.livescience.com/59065-deep-blue-garry-kasparov-chess-match-anniversary.html, accessed 8 February 2024.
[59] Kevin Warwick, *Artificial Intelligence: The basics* (Routledge 2012) 65.
[60] Ronald R Kline, 'Cybernetics, Automata Studies, and the Dartmouth Conference on Artificial Intelligence' (2011) 4, EEE Computer Society, 5.
[61] Zhongzhi Shi, *Advanced Artificial Intelligence* (World Scientific 2011) 18.
[62] Alan Mathison Turing, 'Computing Machinery and Intelligence' (1950) Vol LIX Iss 236 Mind 433-460.
[63] Chris Bernhardt, *Turing's Vision: The Birth of Computer Science* (MIT Press 2016) 157.
[64] Stan Franklin, 'History, motivations, and core themes' in Frankish Keith and Ramsey William (eds) *The Cambridge Handbook of Artificial Intelligence* (2014) 17.
[65] Alan Mathison Turing, 'Computing Machinery and Intelligence' (1950) Vol LIX Iss 236 Mind 433-460.
[66] Chris Bernhardt, *Turing's Vision: The Birth of Computer Science* (MIT Press 2016) 157.
[67] Alan Mathison Turing, 'Computing Machinery and Intelligence' (1950) Vol LIX Iss 236 Mind 433-460.
[68] Chris Bernhardt, *Turing's Vision: The Birth of Computer Science* (MIT Press 2016) 157.

distinguish between the human and the program/machine, (artificial) intelligence is ascribed to the program.[69]

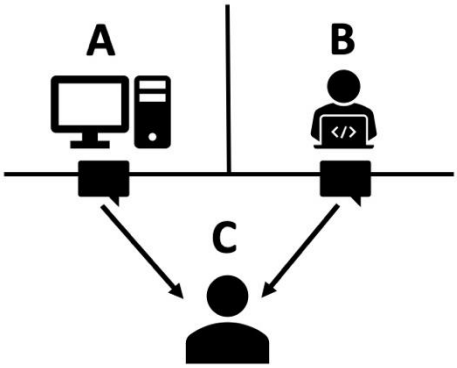

Figure 1.1 Illustration of the Turing test created by the author.

There are plenty of definitions for AI, which involve ambiguous terms such as those already mentioned. In this thesis, AI refers to adaptive machines that can autonomously execute activities and tasks that require capabilities usually associated with humans. 'Autonomously' in this sense means that the machine has the ability to make its *own* decisions and perform tasks on the designer's behalf.[70] 'Adaptive' refers to the machine's ability to learn from, and adapt to its environment in order to preserve its autonomy in dynamic environments.[71] Adaptivity is very important, since only a machine that *learns* will succeed in a vast variety of environments.[72] Learning in this context corresponds to 'adapt' the performance according to previously made experiences based on statistics and probability calculations.[73] This definition aligns well with the ones adopted by the OECD and NIST.[74]

2.2 AI disciplines

Since AI covers a broad range of concepts, this research will pay particular attention to AI disciplines which could be problematic in the light of the fundamental rights to privacy and data protection. These AI disciplines are coloured blue in Figure 1.2.[75] The remaining disciplines (white) will not be discussed in this thesis.

[69] Stan Franklin, 'History, motivations, and core themes' in Frankish Keith and Ramsey William M. (eds) *The Cambridge Handbook of Artificial Intelligence* (2014) 17, 18.
[70] Eduardo Alonso, 'Actions and agents' in Frankish Keith and Ramsey William M. (eds) *The Cambridge Handbook of Artificial Intelligence* (2014) 235, 236.
[71] Ibid 235.
[72] Stuart Russel, Peter Norvig, *Artificial Intelligence, A Modern Approach* (3rd edn, Pearson Education 2016) 39.
[73] Stefan Strauß, 'From Big Data to Deep Learning: A Leap Towards Strong AI or Intelligentia Obscura' (2018) 2 (3), Big Data and Cognitive Computing <https://www.mdpi.com/2504-2289/2/3/16> accessed 14 January 2019, 7.
[74] See < https://oecd.ai/en/wonk/ai-system-definition-update > and < https://csrc.nist.gov/topics/technologies/artificial-intelligence > respectively, accessed 8 February 2024.
[75] This figure shall not be considered as a complete overview of all AI disciplines, but serves as an illustrative overview for this thesis.

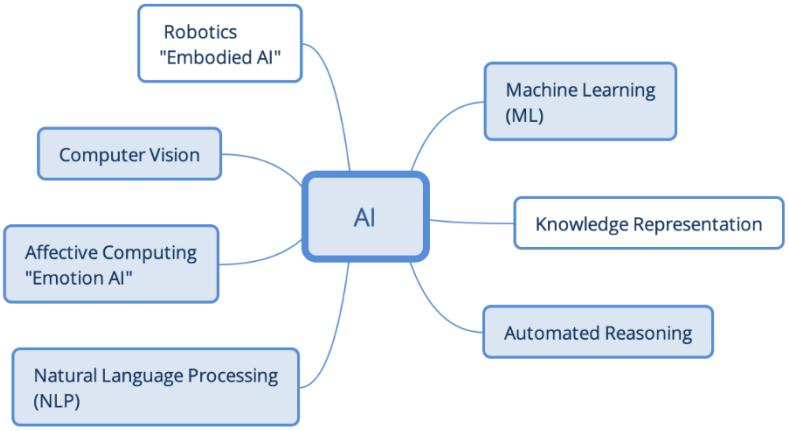

Figure 1.2 Graph created by the author outlining the AI disciplines inspired by Russel/Norvig[76] and slightly adjusted by adding the field of affective computing.

Methods of AI that combine AI with Robotics, i.e. 'Embodied Artificial Intelligence', are out of scope of this thesis. Applications of Embodied AI such as driverless vehicles, surgical robots and companions pose different questions such as liability issues or ethical issues in the context of robot-human interactions.[77] However, these questions are not in the scope of this research.

AI systems need to translate input into information or knowledge so that it can be processed to select output (action).[78] The discipline of AI research commonly referred to as knowledge representation focusses on the computers capabilities to store what it knows and hears.[79] Since research in this discipline of AI focusses on conceptual issues[80] not related to privacy and data protection, it will not be discussed here. However, the subfield of automated reasoning, which is a fundamental part of knowledge representation, will be discussed due to its implications on automated decision-making.

2.2.1 Machine learning (ML)

ML may be considered a discipline or one of the tools of AI.[81] I follow the former approach in this thesis and acknowledge that ML is often combined with other AI disciplines. Computer science has traditionally aimed to manually program computers. ML however aims to have computers program themselves based on experience.[82] In other words, the goal of ML is to adapt to new circumstances and to detect and extrapolate patterns.[83] Murphy defines ML as 'a set of methods that can

[76] Stuart Russel, Peter Norvig, *Artificial Intelligence, A Modern Approach* (3rd edn, Pearson Education 2016) 2, 3.
[77] Cándido García Molyneux, Rosa Oyarzabal, 'What Is a Robot (Under EU Law)?' (2018) Vol 1 RAIL: The Journal of Robotics, AI & Law 11, 12.
[78] Stan Franklin, 'History, motivations, and core themes' in Frankish Keith and Ramsey William M. (eds) *The Cambridge Handbook of Artificial Intelligence* (2014) 24.
[79] Stuart Russel, Peter Norvig, *Artificial Intelligence, A Modern Approach* (3rd edn, Pearson Education 2016) 2.
[80] E.g. the issue of whether or not to represent knowledge, Franklin Stan, 'History, motivations, and core themes' in Frankish Keith and Ramsey William M (eds) *The Cambridge Handbook of Artificial Intelligence* (2014) 24, 25.
[81] Vijay Kotu, Bala Deshpande, *Data Science* (2nd edn Elsevier 2019) 2.
[82] Tom M. Mitchell, 'The discipline of Machine Learning' (2006) 1 <http://www.cs.cmu.edu/~tom/pubs/MachineLearning.pdf> accessed 8 February 2024.
[83] Stuart Russel, Peter Norvig, *Artificial Intelligence, A Modern Approach* (3rd edn, Pearson Education 2016) 2.

automatically detect patterns in data, and then use the uncovered patterns to predict future data or to perform other kinds of decision making under uncertainty'.[84] ML can simply be described as the set of computational methods that use experience to improve its performance or to make accurate predictions.[85] This is achieved by using ML algorithms, algorithms that learn from experience.[86] Put simply, an algorithm is typically a numerical process that consists of a sequence of well-defined steps leading to the solution of a particular type of problem.[87] Experience refers to the data from the past available to the algorithm for analysis.[88] Learning in this context is about making computers modify or adapt their performance (actions) so that these actions become more *accurate*.[89] ML uses data-driven methods, combining fundamental concepts in computer science with approaches from statistics, probability and optimisation.[90] In fact, the probabilistic approach in ML is closely related to the field of statistics, but differs slightly in terms of its emphasis and terminology. The probabilistic approach is particularly helpful for handling ambiguous cases.[91] The main goal of ML is to generate accurate predictions for unseen data and to design efficient algorithms to produce these predictions.[92]

Before the specific *kind* of ML called deep learning (DL) will be discussed in Section 2.2.1.4, some of the most widely used ML *methods* will be elaborated on first in Sections 2.2.1.1 and 2.2.1.3. These methods are called supervised, unsupervised and reinforcement learning. In practice, the distinction between supervised and unsupervised learning is not always clear-cut. Therefore, semi-supervised learning creates a continuum between supervised and unsupervised learning: The algorithm is provided with a few labelled examples (supervised learning) but also has the task to uncover hidden patterns and structures in the data (unsupervised learning).[93] Another method deployed in ML is reinforcement learning (RL). RL is becoming increasingly relevant, in particular in natural language processing, a discipline of AI which aims to enable computers to process human language (see Section 2.2.2).

2.2.1.1 Supervised machine learning

Supervised ML aims to learn a mapping from input *x* to output *y,* given a labelled set of input-output pairs called the *training set* or training data. It can be used to make predictions on *new* input through generalisation.[94] Generalisation refers to the ability of the algorithm to categorise new examples that

[84] Kevin P Murphy, *Machine Learning: A Probabilistic Perspective* (MIT Press 2012) 1.
[85] Mehryar Mohri, Afshin Rostamizadeh, Ameet Talwalkar, *Foundations of Machine Learning* (MIT Press 2012) 1.
[86] Ian Goodfellow, Yoshua Bengio, Aaron Courville, *Deep Learning* (MIT Press 2016) 97 <www.deeplearningbook.org> accessed 8 February 2024.
[87] Yadolah Dodge, 'Algorithm' in: *The Concise Encyclopedia of Statistics* (Springer New York 2006) 1-2.
[88] Mehryar Mohri, Afshin Rostamizadeh, Ameet Talwalkar, *Foundations of Machine Learning* (MIT Press 2012) 1.
[89] Steven Marsland, *Machine Learning: An Algorithmic Perspective* (2nd edn Chapman & Hall 2015) ch 1.2.1.
[90] Mehryar Mohri, Afshin Rostamizadeh, Ameet Talwalkar, *Foundations of Machine Learning* (MIT Press 2012) 1.
[91] Kevin P Murphy, *Machine Learning: A Probabilistic Perspective* (MIT Press 2012) 1, 4.
[92] Mehryar Mohri, Afshin Rostamizadeh, Ameet Talwalkar, *Foundations of Machine Learning* (MIT Press 2012) 2.
[93] Stuart Russel, Peter Norvig, *Artificial Intelligence, A Modern Approach* (3rd edn, Pearson Education 2016) 695.
[94] Kevin P Murphy, *Machine Learning: A Probabilistic Perspective* (MIT Press 2012) 3.

differ from the ones used during the training phase.[95] In the supervised ML approach, the learning algorithm receives several examples, each *labelled* with the correct label (training data). Consider, for example, several labelled pictures with different animals (lions, horses, and cows). The goal is that the algorithm automatically recognises the correct label for the training data and *predicts* the value of unseen (unlabelled) inputs.[96] In other words, the aim is that the algorithm *generalises* accurately by producing a model that can classify input *not seen* during training.[97] The user who provides the correct labels to the algorithm is the teacher, knowing for each input the correct output. Therefore, this is called 'supervised' learning: the algorithm learns under the supervision and guidance of the teacher.[98] To measure the accuracy of the model generated by the algorithm, the teacher provides the algorithm with a set of examples that are *different* from the set of training.[99] Hence, the teacher feeds the algorithm with new pictures containing lions, horses and cows and evaluates the accuracy of the model, namely, whether the algorithm recognised the animals correctly. The algorithm learns by adjusting the relevant parameters so that the model makes the most accurate predictions on the data.[100]

There are basically two techniques used for supervised machine learning: classification and regression.[101] As indicated by its name, classification refers to situations where the predicted attribute is categorical, and regression applies to situations where the predicted attribute is numeric.[102] Classification orders data into exhaustive and exclusive groups or classes on the basis of their similarity. Consequently, all data can only be assigned to one class.[103] The example with the animal referred to the classification technique. Regression is suitable when the prediction to be made by the algorithm should be a numerical value. Regression could be described as a statistical approach that is used to identify the relationship between variables.[104] Therefore, the regression technique could be used to predict the number of people likely to click on an online advertisement based on the ad content and the user's previous surfing history. Other real-world examples using regression are predicting stock market prices given current market conditions or predicting the age of a viewer watching a given video on YouTube.[105]

[95] Christopher M Bishop, *Pattern Recognition and Machine Learning* (Springer 2006) 2.
[96] Mehryar Mohri, Afshin Rostamizadeh, Ameet Talwalkar, *Foundations of Machine Learning* (MIT Press 2012) 7.
[97] Ethem Alpaydin, *Machine Learning: The New AI* (3rd edn MIT Press 2016) 39.
[98] Toshinori Munakata, *Fundamentals of the New Artificial Intelligence* (2nd edn Springer 2008) 38.
[99] Stuart Russel, Peter Norvig, *Artificial Intelligence, A Modern Approach* (3rd edn, Pearson Education 2016) 695.
[100] Ethem Alpaydin, *Machine Learning: The New AI* (3rd edn MIT Press 2016) 39.
[101] Michele Usuelli, *R machine learning essentials* (Packt Publishing 2014) 155.
[102] Ibid 154.
[103] Toon Calders, Bart Custers, 'What is Data Mining and How Does it Work?' in Bart Custers et al. (eds) *Discrimination and Privacy in the Information Society* (Springer 2013) 32.
[104] However, note that decision tree regression would not be considered as traditional statistics.
[105] Kevin P Murphy, *Machine Learning: A Probabilistic Perspective* (MIT Press 2012) 9.

2.2.1.2 Unsupervised machine learning

Unlike supervised ML, the algorithm only receives *un*labelled training data.[106] That means that the algorithm is not told what the desired output is for each form of input and unsupervised ML does not require a human expert to manually label the data.[107] Due to the fact that there is no external comparison between actual and ideal output by the teacher, this approach is called unsupervised: There are no correct answers available.[108] Therefore, the algorithm tries to discover patterns in the input even though no explicit feedback is supplied.[109] The goal of unsupervised ML is to identify associations and patterns among a set of input data and categorise them accordingly.[110] It can be difficult to quantitatively evaluate the performance of the model, since there are no labelled examples available.[111] Two branches of techniques used for unsupervised learning are clustering and dimensionality reduction.[112]

Clustering in this context means dividing detected patterns into groups or clusters. Similar patterns are placed in the same group, while all others are put in different groups.[113] Simply put, clustering refers to the partition of unlabelled items into homogeneous regions.[114] Clusters may overlap, while classifications do not (see Section 2.2.1.1). Clustering is particularly performed to analyse very large data sets. A common example is to use clustering in the context of social network analysis, where the clustering algorithm tries to identify 'communities' within large groups of people.[115] The same applies to e-commerce, where users are clustered into groups based on their purchasing or online behaviour, which enables online shops to send customised targeted ads to each group.[116]

Dimensionality reduction aims to represent data with fewer dimensions[117] and is useful to project high-dimensional data to a lower dimensional subspace to capture the 'essence' of the data.[118] By reducing the dimensions, *hidden patterns* and *structures* in the data may be observed, and noninformative features are discarded. Dimensional representations often produce *better predictive accuracy* because they focus on the essence of the object and filter out non-essential features.[119] Dimensionality reduction is commonly used to pre-process digital images, in computer vision tasks[120] (see

[106] Mehryar Mohri, Afshin Rostamizadeh, Ameet Talwalkar, *Foundations of Machine Learning* (MIT Press 2012) 7.
[107] Kevin P Murphy, *Machine Learning: A Probabilistic Perspective* (MIT Press 2012) 9.
[108] Toshinori Munakata, *Fundamentals of the New Artificial Intelligence* (2nd edn Springer 2008) 38.
[109] Stuart Russel, Peter Norvig, *Artificial Intelligence, A Modern Approach* (3rd edn, Pearson Education 2016) 694.
[110] Hastie Trevor, Tibshirani Robert, Friedman Jerome, *The Elements of Statistical Learning* (2nd edn 2008) xi; Steven Marsland, *Machine Learning: An Algorithmic Perspective* (2nd edn Chapman & Hall 2015) ch 1.3.
[111] Mehryar Mohri, Afshin Rostamizadeh, Ameet Talwalkar, *Foundations of Machine Learning* (MIT Press 2012) 7.
[112] Michele Usuelli, *R machine learning essentials* (Packt Publishing 2014) 164.
[113] Toshinori Munakata, *Fundamentals of the New Artificial Intelligence* (2nd edn Springer 2008) 72.
[114] Mehryar Mohri, Afshin Rostamizadeh, Ameet Talwalkar, *Foundations of Machine Learning* (MIT Press 2012) 2.
[115] Mehryar Mohri, Afshin Rostamizadeh, Ameet Talwalkar, *Foundations of Machine Learning* (MIT Press 2012) 2.
[116] Kevin P Murphy, *Machine Learning: A Probabilistic Perspective* (MIT Press 2012) 11.
[117] Ethem Alpaydin, *Introduction to Machine Learning* (4th edn MIT Press 2020) 137, 138.
[118] Kevin P Murphy, *Machine Learning: A Probabilistic Perspective* (MIT Press 2012) 11.
[119] Ibid, 12.
[120] Mehryar Mohri, Afshin Rostamizadeh, Ameet Talwalkar, *Foundations of Machine Learning* (MIT Press 2012) 2.

Section 2.3) and applied in natural language processing (see Section 2.2.2), e.g., for acoustic signals.[121]

2.2.1.3 Reinforcement learning (RL)

Reinforcement learning (RL) is a distinct method in ML that differs from supervised and unsupervised ML approaches. In RL, the algorithm interacts with its environment and the method is inspired by behavioural psychology.[122] RL algorithms modify or acquire new behaviours incrementally and use *trial-and-error* experience without requiring complete knowledge or control of the environment.[123] Unlike supervised learning, RL learns with a 'critic' who does not instruct the algorithm what to do, but rather provides it with feedback in the form of a reward or punishment.[124] The reward depends on the correctness of the decision (the action by the agent).[125] In RL, the decision-maker is called the agent which interacts with everything outside the agent, called the environment. The agent and environment interact continuously: the agent selects actions, and the environment responds to these actions and presents new situations to the agent.[126] The agent has no prior knowledge of what action to take; it learns from interaction with the environment.[127] The object of the agent is to maximise its reward over a course of interactions with the environment.[128] Therefore, the agent uses the received feedback to update its knowledge so that it learns to perform actions that return the highest reward.[129]

An illustrative example is a machine (agent) that learns to play chess. The chessboard is the environment of the agent that must decide over a sequence of actions, namely, 'moves' on the chessboard (environment) to achieve a certain goal, namely, winning the game. In RL, the agent evolves and learns while analysing the consequences of its actions with the feedback received from the environment.[130] This is different from the unsupervised ML approach, where no feedback is distributed. RL also differs from supervised ML because the agent does not learn from the initially labelled training data, but from the interaction with the environment based on feedback in the form of a punishment or reward.[131] Combining it with deep learning techniques has made 'deep RL' increasingly successful in addressing challenging sequential decision-making problems such as mastering the game 'Go'[132] or

[121] Kevin P Murphy, *Machine Learning: A Probabilistic Perspective* (MIT Press 2012) 11.
[122] Ian Goodfellow, Yoshua Bengio, Aaron Courville, *Deep Learning* (MIT Press 2016) 104 <www.deeplearningbook.org> accessed 8 February 2024.
[123] Vincent François-Lavet et al, 'An Introduction to Deep Reinforcement Learning' (2018), Vol. 11, No. 3-4 Foundations and Trends in Machine Learning, 2, 15.
[124] Ethem Alpaydin, *Introduction to Machine Learning* (4th edn MIT Press 2020) 570.
[125] Andries P Engelbrecht, *Computional Intelligence – An Introduction* (2nd edn John Wiley & Sons 2007) 83.
[126] Zhongzhi Shi, *Advanced Artificial Intelligence* (World Scientific 2011) 365.
[127] Andries P. Engelbrecht, *Computational Intelligence – An Introduction* (2 edn John Wiley & Sons 2007) 83.
[128] Mehryar Mohri, Afshin Rostamizadeh, Ameet Talwalkar, *Foundations of Machine Learning* (MIT Press 2012) 8.
[129] Ethem Alpaydin, *Introduction to Machine Learning* (4th edn MIT Press 2020) 570.
[130] Zhongzhi Shi, *Advanced Artificial Intelligence* (World Scientific 2011) 362.
[131] Sumit Das et al., 'Applications of Artificial Intelligence in Machine Learning: Review and Prospect' (2015), Vol. 115, No. 9 International Journal of Computer Applications 31, 32.
[132] See <https://www.deepmind.com/research/highlighted-research/alphago> accessed 8 February 2024.

beating the world's top professionals in poker.[133] Its adaptive capabilities make RL very suitable for interactive applications. For example, deep RL is applied for dialogue systems and conversational agents, in particular for digital assistants and chatbots.[134] The most impressive and current example is ChatGPT provided by OpenAI. ChatGPT is a large language model trained to produce text. It was optimised by using reinforcement larning with human feedback.[135] Deep RL seems to possess promising potential for real-world applications such as robotics, self-driving cars, finance and smart grids.[136] Current ML applications based on the supervised method for natural language processing and speech recognition require vast amounts of labelled training data. This issue could be eliminated by applying deep RL methods.[137]

2.2.1.4 Artificial Neural Networks and deep learning

The human brain consists of a very large number of processing units called neurons.[138] These neurons have an output fibre called an axon and a terminal fibre called a synapse. The axons split up and connect to several dendrites, which are the input pathways of other neurons through the junction terminal synapse.[139] Because the neurons of the human brain are connected, it is called a neural network. Figure 1.3 shows a typical biological neuron.

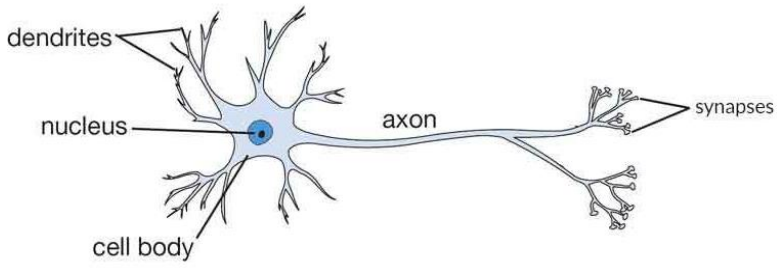

Figure 1.3 Biological neuron illustrated by Navdeep Singh.[140] Used with permission.

Although it is not entirely clear how the neural network of human brains actually works, it is considered to be the fundamental functional source of intelligence, which includes perception, learning and cognition.[141] The characteristic of a neural network is that the neurons operate in parallel and transfer

[133] See <https://www.nature.com/articles/d41586-019-02156-9> accessed 8 February 2024.
[134] Iulian Serban et al. 'A Deep Reinforcement Learning Chatbot' (2017) 1 <https://arxiv.org/pdf/1709.02349.pdf> accessed 8 February 2024.
[135] See FAQs about ChatGPT provided by OpenAI: < https://help.openai.com/en/articles/6783457-what-is-chatgpt > accessed 8 February 2024.
[136] Vincent François-Lavet et al., 'An Introduction to Deep Reinforcement Learning' (2018) Vol. 11 No. 3-4 Foundations and Trends in Machine Learning 3.
[137] Deng Li and Liu Yang, 'Epilogue: Frontiers of NLP in the Deep Learning Era' in Deng Li and Liu Yang (eds) *Deep learning in natural language processing* (Springer 2018) 316.
[138] Ethem Alpaydin, *Machine Learning: The New AI* (3rd edn MIT Press 2016) 86.
[139] Tommy Chow, Siu-Yeung Cho, *Neural Networks and Computing: Learning Algorithms and Applications* (Imperial College Press 2007) 2.
[140] Navdeep Singh Gill, 'Overview of Artificial Neural Networks and its application' <https://www.xenonstack.com/blog/artificial-neural-network-applications/> accessed 8 February 2024.
[141] Toshinori Munakata, *Fundamentals of the New Artificial Intelligence* (2nd edn, Springer 2008) 7.

information among themselves over the synapses so that the neurons are connected and influence each other.[142] The brain is believed to learn by examples and experience and to be highly capable of adapting to external changes.[143]

A single biological neuron would be too simple to make decisions like humans do. Similarly, a single artificial neuron would not be able to cope with challenging decision-making and prediction processes. Hence, to unleash the full potential of artificial neurons, they must operate in parallel and transfer information among themselves. That is why researchers such as Rumelhart and others in 1986 attempted to design artificial neural networks (ANN) with the aim to allow an arbitrarily connected neural network to develop an internal structure that is appropriate for a particular task.[144] ANNs can be simply described as an abstract model that is inspired by knowledge of the inner workings of the human brain that can be programmed on a computer. ANNs consist of artificial neurons and interconnections similar to the human brain. The network receives input, performs internal processes such as the activation of the neurons and finally yields output.[145] However, ANNs are generally not designed to be realistic models of the human brain. The neural perspective on deep learning is motivated by two main ideas: first, that the brain provides an example that intelligent behaviour is possible; and second, that it is possible to create machine learning models that shed light on the principles of the brain and human intelligence.[146] The pattern of connections between the artificial neurons is called the architecture or topology of the ANN and consists of distinct layers of neurons. The layers depend on the model used.[147] Each of the layers has a certain number of neurons which is usually determined by a specific application problem the model aims to solve. An example of a deep ANN is given in Figure 1.4.

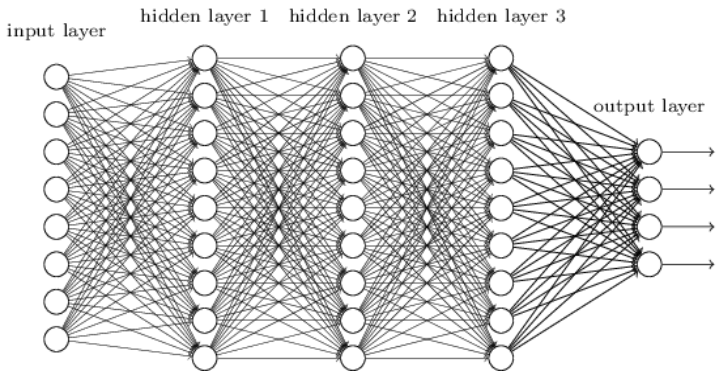

Figure 1.4 Example of a deep artificial neural network illustrated by Michael Nielsen.[148] Used with permission.

[142] Ethem Alpaydin, *Machine Learning: The New AI* (3rd edn MIT Press 2016) 86.
[143] Tommy Chow, Siu-Yeung Cho, *Neural Networks and Computing: Learning Algorithms and Applications* (Imperial College Press 2007) 2.
[144] David Rumelhart, Geoffrey Hinton, Ronald Williams 'Learning representations by backpropagating errors' (1986) Vol. 323 Nature 533.
[145] Toshinori Munakata, *Fundamentals of the New Artificial Intelligence* (2nd edn, Springer 2008) 3, 7.
[146] Ian Goodfellow, Yoshua Bengio, Aaron Courville, *Deep Learning* (MIT Press 2016) 13 <www.deeplearningbook.org> accessed 8 February 2024.
[147] Toshinori Munakata, *Fundamentals of the New Artificial Intelligence* (2nd edn, Springer 2008) 9.
[148] Michael Nielsen, 'Why are deep neural networks hard to train' in: *Neural Networks and Deep Learning* (Determination Press 2015) <http://neuralnetworksanddeeplearning.com/chap5.html> accessed 8 February 2024.

Generally, there are one input layer, one output layer and *any number* of hidden layers. Neurons of the input layer are connected to the neurons of the hidden layer through edges, and the neurons of the hidden layer(s) are connected to the output layer. A weight is associated to each edge. The input layer (see on the left side of Figure 1.4) consists of neurons that receive their input directly from the data, and its function is to merely send out input signals to the hidden layer neurons; it does not compute anything.[149] The hidden layer then applies computation methods to the inputs that depend on the model used for the neural network, transforming the received inputs to something the output layer can use. Hidden means that the values in these layers are not given in the data, but the model has the task of determining which concepts are useful for explaining the relationships in the observed data.[150] It then sends its output to the next layer, in the present case, to hidden layer 2, which sends it to hidden layer 3 and subsequently to the output layer (see the right side of Figure 1.4). Subsequently, the role of the output layer is to produce the output of the entire network. The output of ANNs can then be used to extract a prediction or a decision.

Deep learning (DL) is a particular kind of ML that represents the world as a nested hierarchy of concepts.[151] The human brain seems to execute many levels of processing with increasing levels of abstraction.[152] DL seems to resemble this by computing more abstract concepts in terms of less abstract ones.[153] Most of the models used for supervised and unsupervised ML have a simple two-layer architecture.[154] This is different with DL models, which use many different layers. Approaches in DL feed a large set of input data into the ANN that produces successive transformations of the input data, where each hidden layer combines the values in its preceding layer and learns more complicated functions of the input.[155] Then, the final transformation predicts the output.[156] The deep learning approach avoids the requirement that the human operator must specify all the knowledge which the computer requires. Deep learning solves this by enabling the computer to build complex concepts out of simpler concepts. When illustrating the approach in a graph by building the concepts on top of each other, that graph is deep, with many layers. Therefore, the approach is called deep learning (see Figure 1.4).[157] DL draws inspiration from many fields, especially from linear algebra and probabilistic statistics. Foundation models, namely models that are trained on broad data using self-supervision

[149] Toshinori Munakata, *Fundamentals of the New Artificial Intelligence* (2nd edn, Springer 2008) 10.
[150] Ian Goodfellow, Yoshua Bengio, Aaron Courville, *Deep Learning* (MIT Press 2016) 6 <www.deeplearningbook.org> accessed 8 February 2024.
[151] Ian Goodfellow, Yoshua Bengio, Aaron Courville, *Deep Learning* (MIT Press 2016) 8 <www.deeplearningbook.org> accessed 8 February 2024.
[152] Kevin P Murphy, *Machine Learning: A Probabilistic Perspective* (MIT Press 2012) 95.
[153] Ian Goodfellow, Yoshua Bengio, Aaron Courville, *Deep Learning* (MIT Press 2016) 8 <www.deeplearningbook.org> accessed 8 February 2024.
[154] Kevin P Murphy, *Machine Learning: A Probabilistic Perspective* (MIT Press 2012) 995.
[155] Ethem Alpaydin, *Machine Learning: The New AI* (3rd edn MIT Press 2016) 104.
[156] Yoav Goldberg, *Neural Network Methods in Natural Language Processing* (Morgan & Claypool Publishers 2017) 2.
[157] Ian Goodfellow, Yoshua Bengio, Aaron Courville, *Deep Learning* (MIT Press 2016) 1, 5 <www.deeplearningbook.org> accessed 8 February 2024.

and that can be adapted to a wide range of tasks, are based on deep neural networks.[158] Typical examples of foundation models include large language models (LLMs) as introduced in the AI discipline natural language processing (Section 2.2.2).

Interestingly, achievements in modern DL have been made with an astonishingly small number of neurons contained in the ANNs when compared with neural networks of the human brain. Although today's ANNs are considered quite large from a computational perspective, they are smaller than the neural networks of relatively primitive animals such as frogs. Goodfellow, Bengio and Courville, leading scholars in the field, predict that ANNs will not reach the same number of neurons as the human brain possesses before the 2050s unless new technologies enable faster scaling.[159]

However, most current DL models lack reasoning and explanatory capabilities, making them vulnerable to produce unexplainable outcomes. Despite the recent success of DL, DL methods based on ANN generally lack interpretability.[160] Foundation models and LLMs are no exception.[161] Interpretability remains a challenge due to the hierarchical and nonlinear structure of ANNs and the central concept in DL called connectionism. With deep learning models, each artificial neuron works *independently* by computing a relatively simple task, and therefore *partially* contributes to the output produced by the ANNs.[162] ANNs produce output based on the central concept in DL called *connectionism*, where the idea is that a large number of simple computational units (artificial neurons) achieve intelligent behaviour when networked together.[163] Consequently, combining the characteristic of artificial neurons to work independently with the concept of connectionism leads to a situation where thousands or hundreds of thousands of artificial neurons work in parallel in an ANN with hidden layers to jointly calculate certain output.[164] Hence, it seems neither possible to understand which artificial neuron contributed to a distinct part of the output nor to understand what happened in the intermediate (hidden) layers of the ANN.[165] In other words, it is not possible to extract any underlying rules that may be implied by the DL model.[166] This holds even true for DL algorithms using the supervised learning method, where the algorithm cannot learn without being given correct sample patterns. Therefore, even if an ANN has successfully been trained to achieve its goal, the many

[158] Rishi Bommasani et al, 'On the Opportunities and Risks of Foundation Models' (2022) Center for Research on Foundation Models Standford University 1, 3 < https://arxiv.org/pdf/2108.07258.pdf > accessed 8 February 2024.
[159] Ian Goodfellow, Yoshua Bengio, Aaron Courville, *Deep Learning* (MIT Press 2016) 21 <www.deeplearningbook.org> accessed 8 February 2024.
[160] Deng Li and Liu Yang, 'A Joint Introduction to Natural Language Processing and Deep Learning' in Deng Li and Liu Yang (eds) *Deep learning in natural language processing* (Springer 2018) 11, 12.
[161] Melanie Mitchell, David C Krakauer, 'The debate over understanding in AI's large language models' (2023) Vol 120 Iss 3 PNAS 1-5.
[162] Toshinori Munakata, *Fundamentals of the New Artificial Intelligence* (2nd edn, Springer 2008) 44.
[163] Ian Goodfellow, Yoshua Bengio, Aaron Courville, *Deep Learning* (MIT Press 2016) 16 <www.deeplearningbook.org> accessed 8 February 2024.
[164] Ethem Alpaydin, *Machine Learning: The New AI* (3rd edn MIT Press 2016) 155.
[165] Ethem Alpaydin, *Machine Learning: The New AI* (3rd edn MIT Press 2016) 155.
[166] Toshinori Munakata, *Fundamentals of the New Artificial Intelligence* (2nd edn, Springer 2008) 44.

numeric values of the weights produced by the model do not have a meaning to the supervisor.[167] Clearly, the model is parameterised by all these weights, but it remains unclear how these weights have been calculated and to what extent the various input variables contributed to the outcome. ANNs in use can be *updated dynamically* as new data are fed into the network.[168] Subsequently, this updates the weights produced by the model because they are learnt from experience. These updates contribute to further challenges regarding the interpretability of DL approaches.[169]

DL is well suited to deal with complex sensor data such as input from cameras and microphones that proved to be difficult to process when using conventional computational methods.[170] This applies in particular to cognitive tasks which include natural language processing and speech recognition or face recognition, which are discussed below.[171] Current research in DL attempts to decode speech directly from the human brain. Such approaches record the activity in the cortex to decode the characteristics of the produced speech.[172] State-of-the-art deep neural network models arguably contribute to an improved overall accuracy in speech reconstruction from neural recordings in the human auditory cortex.[173] The short-term goal of these research projects is to help individuals that are unable to communicate due to injuries or neurodegenerative disorders by creating a synthesised version of their voice that can be controlled by the activity of their brain speech centres.[174] However, the long-term goal of this could be much broader and very different. Facebook announced that it wants to 'build a non-invasive, wearable device that lets people type simply by imagining themselves talking.'[175]

2.2.2 Natural language processing (NLP)

Natural language processing (NLP), a subfield of AI, aims to give computers the ability to process human language. This interdisciplinary field comprises many concepts and methods such as speech and language processing, human language technology, natural language processing, computational

[167] Toshinori Munakata, *Fundamentals of the New Artificial Intelligence* (2nd edn, Springer 2008) 12, 25, 35.
[168] A production model has fixed weights after training. To continuously update weights is possible, but by no means necessary.
[169] Paul De Laat, 'Algorithmic Decision-Making based on Machine Learning from Big Data: Can Transparency restore Accountability' (2017) Vol. 31 Issue 4 Philosophy & Technology 14 < https://link.springer.com/article/10.1007%2Fs13347-017-0293-z> accessed 8 February 2024.
[170] Tommy Chow, Siu-Yeung Cho, *Neural Networks and Computing: Learning Algorithms and Applications* (Imperial College Press 2007), 1/2.; Mitchell Tom T., *Machine Learning* (Mc-Graw-Hill 1997) 95.
[171] Tommy Chow, Siu-Yeung Cho, *Neural Networks and Computing: Learning Algorithms and Applications* (Imperial College Press 2007) 2.
[172] David A. Moses et al, 'Real-time decoding of question-and-answer speech dialogue using human cortical activity' (2019) 10 Nature Communication <https://www.nature.com/articles/s41467-019-10994-4.pdf> accessed 8 February 2024.
[173] Minda Yang et al, 'Speech Reconstruction from Human Auditory Cortex with Deep Neural Networks' (Interspeech Conference, Dresden, September 2015) 1124 < https://dblp.org/db/conf/interspeech/interspeech2015.html > accessed 8 February 2024.
[174] Nicholas Weiler, 'Breakthrough device translates brain activity into speech' (University of California, 25 April 2019) < https://www.universityofcalifornia.edu/news/synthetic-speech-generated-brain-recordings> accessed 8 February 2024.
[175] 'Imagining a new interface: Hands-free communication without saying a word' (Tech@Facebook, 30 March 2020) <https://tech.fb.com/imagining-a-new-interface-hands-free-communication-without-saying-a-word/> accessed 8 February 2024.

linguistics, and speech recognition and synthesis.[176] NLP includes both the generation and understanding of natural language.[177] The advances in NLP have led to the development of large language models (LLMs). LLMs are advanced language models with massive parameter sizes (billions to trillions)[178] and strong learning capabilities.[179] These models can perform various NLP tasks, such as translation, text summarisation, and question-answering.[180] ChatGPT is the current prime example.

From an engineering perspective, NLP intends to develop novel practical applications to facilitate interactions between computers and human languages.[181] Current NLP systems require large amounts of labelled data.[182] Speech recognition is a typical application of NLP, and its aim is to *automatically transcribe* the sequence of spoken words. It may be defined as the process of converting a speech signal to a sequence of words by means of an algorithm implemented by a computer program.[183] In particular, speech recognition does not concern *understanding* but is simply responsible to *convert* language from spoken words to text form.[184] The observable 'physical' signal of natural language is called text in symbolic form, and its counterpart is the speech signal, that is, the continuous correspondence of spoken texts.[185] Speech recognition is based on the acoustic signal captured by a microphone as input. The classes are the words that can be uttered. A word is a sequence of phonemes that are the basic speech sounds.[186] Therefore, speech recognition converts phonemes (speech signal) into text. A specific challenge in speech recognition is that different people pronounce the same word differently due to factors related to age, gender or accent, which makes it more difficult to recognise the words.[187] Another challenge is that a common conversational utterance involves multiple queries with disfluencies such as pauses and hesitations. However, current NLP systems embedded in virtual assistants typically focus on 'unnatural' and one-sided interactions without hesitation or disfluency. For this reason, speech recognition involving conversiational speech is a challenging task.[188]

[176] Daniel Jurafsky, James H Martin, *Speech and Language Processing* (2 edn, Pearson Education Limited 2014) 1.
[177] Stan Franklin, 'History, motivations, and core themes' in Frankish Keith and Ramsey William M. (eds) *The Cambridge Handbook of Artificial Intelligence* (2014) 26.
[178] Melanie Mitchell, David C Krakauer, 'The debate over understanding in AI's large language models' Vol 120 Iss 3 PNAS 1-5; Rishi Bommasani et al, 'On the Opportunities and Risks of Foundation Models' (2022) Center for Research on Foundation Models Standford University 1, 3 < https://arxiv.org/pdf/2108.07258.pdf > accessed 8 February 2024.
[179] Yupeng Chang et al, 'A Survey on Evaluation of Large Language Models' (2023) 1, 4 < https://arxiv.org/pdf/2307.03109.pdf > accessed 8 February 2024.
[180] Yiheng Liu et al, 'Summary of ChatGPT-Related research and perspective towards the future of large language models' (2023) Vol 1 Meta-Radiology 1 – 14.
[181] Deng Li and Liu Yang, 'A Joint Introduction to Natural Language Processing and Deep Learning' in Deng Li and Liu Yang (eds) *Deep learning in natural language processing* (Springer 2018) 1.
[182] Deng Li and Liu Yang, 'Epilogue: Frontiers of NLP in the Deep Learning Era' in Deng Li and Liu Yang (eds) *Deep learning in natural language processing* (Springer 2018) 316.
[183] Abhang Priyanka, Gawali Bharti, Mehrotra Suresh, *Introduction to EEG- and speech-based emotion recognition* (Elsevier Inc 2016) 13.
[184] Gokhan Tur et al, 'Deep Learning in Conversational Language Understanding' in Deng Li and Liu Yang (eds) *Deep learning in natural language processing* (Springer 2018) 24.
[185] Ibid 24.
[186] Ethem Alpaydin, *Machine Learning: The New AI* (3rd edn MIT Press 2016) 67.
[187] Ibid.
[188] Shuo-ziin Chang et al, 'Turn-Taking Prediction for Natural Conversational Speech' (Interspeech Conference Incheon, September 2022) <https://arxiv.org/pdf/2208.13321.pdf> accessed 8 February 2024.

Speech signals cannot only reveal the intended message, but also the *identity of the speaker* because the ways in which prosodic characteristics are manifested in speech disclose important information regarding the identity of the speaker.[189] Prosody refers to the study of the intonational and rhythmic aspects of language.[190] Systems in the domain of speaker verification are capable of using the voice of an individual in order to identify an unknown person (speaker identification), verify the identity of a person (speaker verification) and classify specific characteristics like age or gender (speaker classification).[191] Text-based verification of an individual through voice analysis is technically possible with a very short text such as 'Ok Google', which takes approximately 0.6 seconds if uttered by an individual.[192] Hence, speaker identity is embedded in the speaker's voice and can be recognised using automatic speaker recognition systems, which apply DL approaches.[193]

Current research in speech recognition focusses on emotion recognition from speech signals, a major subject in human-computer interaction. This research focusses on how speech is modulated when a speaker's emotion changes from neutral to another emotional state. For example, it has been observed that speech in anger or happiness shows longer utterance duration and higher pitch and energy value with deep length. [194] Speech emotion recognition may be used for various areas, such as call centres, smart devices or self-driving cars.[195] A real-world application of affective computing (AC) that aims to derive emotional states from speech is Amazon's 'Halo' wearable, which analyses voice tones to detect user emotions.[196] The recent success in NLP and speech recognition has been powered by using the DL approach in ML, currently with supervised ML methods such as classification as described in Section 2.2.1.1. Therefore, the current bottleneck of these approaches is that they require large amounts of labelled data and lack reasoning abilities. However, it is tried to overcome this bottleneck by applying the unsupervised learning paradigm and particularly deep RL methods in NLP and speech recognition.[197] Deep learning has been successfully applied to real-world tasks in AI, in particular in

[189] Leena Mary, *Extraction of Prosody for Automatic Speaker, Language, Emotion and Speech Recognition* (2nd edn Springer 2019) 1, 8.
[190] Daniel Jurafsky, James H Martin, *Speech and Language Processing* (2 edn, Pearson Education Limited 2014) 238.
[191] Soufiane Hourri, Jamal Kharroubi, 'A deep learning approach for speaker recognition' (2020) Vol. 23 Iss. 1 International Journal of Speech and Technology 123.
[192] Gregor Heigold et al, 'End-to-End Text-Dependent Speaker Verification' (2015) < https://arxiv.org/pdf/1509.08062.pdf > accessed 8 February 2024.
[193] Leena Mary, *Extraction of Prosody for Automatic Speaker, Language, Emotion and Speech Recognition* (2nd edn Springer 2019) 7. See precedent references regarding DL approaches.
[194] Abhang Priyanka, Gawali Bharti, Mehrotra Suresh, *Introduction to EEG- and speech-based emotion recognition* (Elsevier Inc 2016) 14, 105.
[195] See services of the company audeering <https://www.audeering.com/> accessed 8 February 2024.
[196] Alex Hern, 'Amazon's Halo wristband: the fitness tracker that listens to your mood' *The Guardian* (London, 28 August 2020) <https://www.theguardian.com/technology/2020/aug/28/amazons-halo-wristband-the-fitness-tracker-that-listens-to-your-mood> accessed 8 February 2024; Austin Carr, 'Amazon's New Wearable Will Know If I'm Angry. Is That Weird?' *Bloomberg* (New York, 31 August 2020) < https://www.bloomberg.com/news/newsletters/2020-08-31/amazon-s-halo-wearable-can-read-emotions-is-that-too-weird > accessed 8 February 2024.
[197] Deng Li and Liu Yang, 'Epilogue: Frontiers of NLP in the Deep Learning Era' in Deng Li and Liu Yang (eds) *Deep learning in natural language processing* (Springer 2018) 316.

speech recognition as a part of the virtual personal assistants such as Google Assistant, Amazon Alexa, Microsoft Cortana or Apple Siri.[198]

2.2.3 Computer vision (CV)

Computer vision (CV) is a subfield of AI devoted to perceive objects, i.e. the automated understanding of visual images and comprises many fields of applications.[199] The goal of object detection is to detect all instances of objects from a known class, such as people, cars or faces in an image.[200] CV can also be described as the science and technology of machines that 'see', which refers to the ability of the machine to extract information from an image necessary to solve a task.[201] CV aims to infer properties from the observed visual data, which originate from a variety of sensors such as cameras, laser scans, etc.[202] CV algorithms reconstruct the properties of one or more images, such as such as shape, illumination and colour distributions. Researchers in computer vision develop mathematical techniques to recover the three-dimensional shape and appearance of objects in imagery. Real-world applications include optical character recognition (OCR) for automatic number plate recognitions (of vehicles), medical imaging for preoperative and intra-operative imagery, automotive safety to detect unexpected obstacles such as pedestrians on the street, surveillance to monitor intruders and fingerprint recognition for automatic access authentication.[203]

CV techniques are also currently used to identify individuals based on their gait. Biometric research implies that gait, i.e. the manner in which individuals walk, constitutes a unique identifier like a fingerprint or iris.[204] The biometrics necessary for gait identification may be captured in public places and from a distance in a rather ubiquitous manner. Methods used for identification are model-based approaches which consider the human body or its movements to acquire gait parameters (e.g., step dimensions, cadence, human skeleton, body dimensions) as well as model-free approaches that acquire gait parameters by that rely on gait dynamics and the measurement of geometric representations such as silhouettes.[205]

[198] Gokhan Tur et al, 'Deep Learning in Conversational Language Understanding' in Deng Li and Liu Yang (eds) *Deep learning in natural language processing* (Springer 2018) 23.
[199] Stuart Russel, Peter Norvig, *Artificial Intelligence, A Modern Approach* (3rd edn, Pearson Education 2016) 3, Stan Franklin, 'History, motivations, and core themes' in Frankish Keith and Ramsey William M. (eds) *The Cambridge Handbook of Artificial Intelligence* (2014) 26.
[200] Yali Amit, Pedro Felzenszwalb, 'Object Detection' in Katsushi Ikeuchi (ed) *Computer Vision – A Reference Guide* (Springer 2014) 537.
[201] Sota R. Yoshida, *Computer Vision* (Nova Science Publisher 2011) vii.
[202] Varun Jampani, 'Learning Inference Models for Computer Vision' (Dissertation, Universität Tübingen 2016) 1.
[203] Richard Szeliski, *Computer Vision: Algorithms and Applications* (Griets David, Schneider Fred Springer eds 2011) 3, 5.
[204] Ale Sokolova, Anton Konushin 'Methods of Gait Recognition in Video' (2019) Vol 45 No 4 Programming and Computer Software 213.
[205] Jure Kovač, Vitomir Štruc, Peter Peer 'Frame-based classification for cross-speed gait recognition' (2019) Vol 78 Multimedia Tools and Applications 5621, 5622.

Another real-world example is Amazon Go. Amazon Go is a checkout-free grocery store which is equipped with state-of-the-art cameras and sensors. Amazon Go is powered by computer vision, DL and sensor fusion[206] in order to track shoppers and their purchases. Sensor fusion exploits the best features of sensors (for example, cameras and small Bluetooth radio transmitters called 'beacons') installed in a given environment. It is particularly helpful in situations where the sensors themselves are not self-sufficient to achieve a certain goal, for example, comprehensive and precise tracking of shoppers.[207] In Amazon Go stores, shoppers enter by scanning an Amazon Go smartphone app and sensors track items that the shoppers take from the shelves. Once picked up, the items are automatically charged to the Amazon accounts of the shoppers when they leave the store. Where Amazon Go's inventory system cannot detect the object the user removed from the shelf, the system 'may consider past purchase history' of the user.[208]

Face recognition is one of the CV applications of particular relevance for this thesis. Section 2.2.3.1 introduces face recognition and Section 2.2.3.2 explains face recognition applications applying deep learning.

2.2.3.1 Face recognition

Face recognition refers to the technology capable of identifying or verifying the identity of subjects in images or videos based on biometric data.[209] It is one of the major biometric technologies and has become increasingly relevant due to the rapid advances in image capture devices and the availability of huge amounts of face images on the web.[210] Unlike other biometric identification methods, such as iris recognition (which requires individuals to get significantly close to a camera), face recognition can be used from a distance and in a covert manner.[211] Therefore, the range of potential applications for face recognition is wide because it can be easily deployed.[212]

[206] See https://www.amazon.com/b?ie=UTF8&node=16008589011 and Vasilios Mavroudis, Michael Veale 'Eavesdropping Whilst You're Shopping: Balancing Personalisation and Privacy in Connected Retail Spaces' (Living in the Internet of Things Conference, London, March 2018) 6 <https://ieeexplore.ieee.org/document/8379705> accessed 8 February 2024.
[207] For example, cameras offer a high level of precision, but might be too expensive to cover the whole shop. Beacons are not self-sufficient to provide tracking data for customer analysis, but can cover a wider operational range. Combined by means of sensor fusion, the sensors allow precise consumer path tracking. See Mirco Sturari et al, 'Robust and affordable retail customer profiling by vision and radio beacon sensor fusion' (2016) Vol. 81 Pattern Recognition Letters 30, 31, 40.
[208] Dilip Kumar et al. 'Detecting item interaction and movement' US Patent Number US 10268983 (Assignee: Amazon Technologies, Inc.) April 2019 at 9 <https://patentimages.storage.googleapis.com/01/0b/6e/de57009f5670ae/US20150019391A1.pdf > accessed 8 February 2024.
[209] Daniel Trigueros, Li Meng, Margaret Hartnett, 'Face recognition: From Traditional to Deep Learning Methods' (2018) 1 <https://arxiv.org/pdf/1811.00116.pdf > accessed 8 February 2024.
[210] Stan Li, Anil Jain, 'Introduction' in Li Stan, Jain Anil (eds) *Handbook of Face Recognition* (2nd edn, Springer 2011) 1.
[211] Stan Li, Anil Jain, 'Introduction' in Li Stan, Jain Anil (eds) *Handbook of Face Recognition* (2nd edn, Springer 2011) 1; Daniel Trigueros, Li Meng, Margaret Hartnett, 'Face recognition: From Traditional to Deep Learning Methods' (2018) 1 <https://arxiv.org/pdf/1811.00116.pdf > accessed 8 February 2024.
[212] Daniel Trigueros, Li Meng, Margaret Hartnett, 'Face recognition: From Traditional to Deep Learning Methods' (2018) 1 <https://arxiv.org/pdf/1811.00116.pdf > accessed 8 February 2024.

Face recognition systems operate in a face verification (authentication) and/or face identification (recognition) mode. The former involves a one-on-one match that compares a query face image of the person whose identity is claimed (e.g., for self-serviced immigration clearance using E-passports). The latter involves one-to-many matching, which compares a query face image against multiple face images in a database to associate the identity of the query face. Usually, finding the most similar face is not sufficient and a confidence threshold is specified. Therefore, only those faces whose similarity score is above the threshold are reported.[213] Face recognition systems are usually built on four building blocks:

1. Face detection, which finds the position of a face in an image
2. Face normalisation, which normalises the face geometrically and photometrically
3. Face feature extraction performed to extract salient information which is useful to distinguish faces such as reference points located at fixed locations in the face (e.g., position of eyes, nose, lips)
4. Face matching, where extracted features from the input face are matched against one or many of the enrolled faces in the database[214]

The facial features used for the third building block may be grouped into two classes of features: continuous and discrete. Continuous features are real valued numbers and are extracted using distances and angles between facial landmarks such as forehead height, eyebrow length, nose height, chin height, ears length, mouth length etc. Discrete features represent a finite number of categories, for example, the shape of the eyebrow or nose root width.[215] Figure 1.5 provides an example of such features.

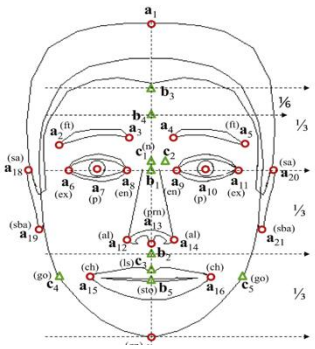

Figure 1.5 Face layout illustrated by Tome et al.[216] with examples of facial features extracted by using distances and angles between facial landmarks such as eyebrows, eyes, nose and lips. Used with permission.

[213] Stan Li, Anil Jain, 'Introduction' in Li Stan, Jain Anil (eds) *Handbook of Face Recognition* (2nd edn, Springer 2011) 3.
[214] Ibid, 4; Daniel Trigueros, Li Meng, Margaret Hartnett, 'Face recognition: From Traditional to Deep Learning Methods' (2018) 1 < https://arxiv.org/abs/1811.00116 > accessed 8 February 2024.
[215] Pedro Tome et al., 'Facial soft biometric features for forensic face recognition' (2015) Vol 257 Forensic Science International 271, 273.
[216] Ibid.

2.2.3.2 DL and face recognition

Current face recognition applications use DL methods based on convolutional neural networks (CNN) which are trained with very large datasets.[217] A CNN is a specific kind of neural network for processing data that has a known grid-like typology. For example, image data can be thought of as a 2D grid of pixels. As the name indicates, a CNN employs a mathematical operation called convolution, which is a specialised kind of linear operation.[218] Notably, the performance of a face recognition system largely depends on a variety of factors such as illumination, facial pose, expression, age span, hair and motion.[219] Whereas the building blocks of face recognition systems and the general architecture of the ANN are predetermined by the developer of the system, the ANN itself decides how to create the optimal score for determining similarity in the face matching building block mentioned in Section 2.2.3.1. Therefore, it remains often unclear how the similarity score is calculated by the ANN, even to the developer of the system.[220] Another issue is that face recognition systems perform poorly in recognising individuals of different ethnicities. For example, Hewlett Packard face recognition software could not recognise dark-coloured faces as faces.[221] A 'passport robot' in New Zealand rejected the passport picture of an Asian man because the 'subject's eyes are closed' although his eyes were open.[222]

However, face recognition systems are widely used in commercial applications and consumer products with built-in AI capabilities. Examples are cars with on-board cameras to deploy biometric identification and monitor driving behaviour[223] or connected retail spaces.[224] Furthermore, there is a trend to improve face recognition systems with the ability to monitor and analyse the emotions in real-time based on extracted biometric data and facial expressions. The gained knowledge is then used to build specific customer profiles.

[217] Daniel Trigueros, Li Meng, Margaret Hartnett, 'Face recognition: From Traditional to Deep Learning Methods' (2018) 1 <https://arxiv.org/abs/1811.00116> accessed 8 February 2024.
[218] Ian Goodfellow, Yoshua Bengio, Aaron Courville, *Deep Learning* (MIT Press 2016) 326 <www.deeplearningbook.org> accessed 8 February 2024.
[219] Stan Li, Anil Jain, 'Introduction' in Li Stan, Jain Anil (eds) *Handbook of Face Recognition* (2nd edn, Springer 2011) 3.
[220] Yana Welinder, Aeryn Palmer, 'Face Recognition, Real-Time Identification, and Beyond' in Selinger Evan, Polonetsky Jules, Tene Omer (eds) *The Cambridge Handbook of Consumer Privacy* (Cambridge University Press 2018) 104.
[221] Frederik Zuiderveen Borgesius, 'Discrimination, artificial intelligence, and algorithmic decision-making'(2019) Report for the Anti-discrimination department of the Council of Europe, 17 <https://rm.coe.int/discrimination-artificial-intelligence-and-algorithmic-decision-making/1680925d73> accessed 8 February 2024.
[222] Regan James, 'New Zealand passport robot tells applicant of Asian descent to open eyes' (2016) *Reuters* <https://www.reuters.com/article/us-newzealand-passport-error/new-zealand-passport-robot-tells-applicant-of-asian-descent-to-open-eyes-idUSKBN13W0RL> accessed 8 February 2024.
[223] See <https://visagetechnologies.com/application-fields/driver-monitoring/> accessed 8 February 2024
[224] See <https://www.einfochips.com/blog/facial-recognition-in-retail-enhance-in-store-customer-experience-and-improve-retailer-operations/> accessed 8 February 2024.

2.2.4 Affective computing (AC)

Affective computing (AC), sometimes called 'emotion AI', is computing that relates to, arises from or influences emotion.[225] AC is a scientific and engineering endeavour inspired by psychology, neuroscience, linguistics and related areas.[226] Affective states are considered to be experiential phenomena such as emotions and moods.[227] Emotions form an important part of human intelligence and daily live, be it for decision-making, social interaction, perception or learning. In other words, emotions play a pivotal role in functions considered essential to intelligence.[228] Picard, the pioneer in the field of AC, therefore, concludes that if computers are to be genuinely intelligent, they too should have emotional capabilities.[229] In this thesis, the focus lies on affect detection from facial expressions and speech, since they may be easily deployed compared to more invasive approaches that include measurement of physiological factors such as cardiac activity (heart rate) or skin conductance (sweat).

The following sections elaborate on affect detection from facial expressions (Section 2.2.4.1), speech (Section 2.2.4.2) and discuss multimodal approaches in which different methods of AC are combined to detect emotions (Section 2.2.4.3).

2.2.4.1 Facial expressions

Facial expressions are probably the most natural way humans express their emotions.[230] According to Darwin's evolutionary theory of emotions, emotion expressions help in regulating the social interaction and increase the likelihood of survival.[231] Due to the developments in technology, it is possible to detect facial information automatically in real-time, for example, with the use of a simple video camera. However, automatic detection of emotions derived from facial expressions and their interpretation is not simple and context-driven.[232] Physically, a facial expression is a change in the face due to movements of several muscles demonstrating an emotional state. An emotional state is an individual's transient reaction to specific encounters with the environment, one that occurs and disappears depending on particular conditions. For example, someone is feeling or reacting with anger at a particular time and place.[233] A facial expression is communicated by a transient flexing of facial

[225] Rosalind W Picard, 'Affective Computing' (1995) MIT Media Laboratory Perceptual Computing Section Technical Report No 321 at 1 <https://hd.media.mit.edu/tech-reports/TR-321.pdf> accessed 8 February 2024.
[226] Rafael Calvo et al, 'Introduction to Affective Computing' in Rafael Calvo et al (eds), *The Oxford Handbook of Affective Computing* (OUP 2015) 2.
[227] Steffen Steinert, Orsolya Friedrich, 'Wired Emotions: Ethical Issues of Affective Brain–Computer Interfaces' (2020) Vol 26 Science and Engineering Ethics 351, 352.
[228] Rosalind W Picard, *Affective Computing* (MIT Press 1997) 47.
[229] Ibid preface x.
[230] Rafael Calvo et al, 'Introduction to Affective Computing' in Rafael Calvo et al (eds), *The Oxford Handbook of Affective Computing* (OUP 2015) 4.
[231] Avinash Awasthi, Manas K. Mandal, 'Facial Expressions of Emotions: Research Perspectives' in Manas K. Mandal, Avinash Awasthi (eds) *Understanding Facial Expressions in Communication* (Springer 2015) 3.
[232] Catherine Marechal et al, 'Survey on AI-Based Multimodal Methods for Emotion Detection' in Joanna Kołodziej, Horacio González-Vélez (eds) *High-Performance Modelling and Simulation for Big Data Applications* (Springer 2019) 314, 315.
[233] Richard S Lazarus, *Emotion and Adaption* (OUP 1991) 46, 47.

futures such as mouth, eyes and eyebrows due to the contraction of the muscles that make up the face.[234] These muscle contractions are controlled by two different areas of the brain, one controlling voluntary movements and the other involuntary reactions.[235] Facial expressions can easily be used for emotion detection because it only requires a simple video camera to register facial information automatically and in real-time.[236] Two approaches to measuring facial expressions will be discussed here: message-based and sign-based approaches.

Based on the assumption that the face provides a direct 'readout' of emotion, the message-based approach makes inferences about emotion or the affective state by assigning facial expression and movements to 'basic emotions' according to Ekman.[237] Facial movements and the 'basic emotions' hypothesised are illustrated in Figure 1.6.

Figure 1.6 Facial movements and hypothesised 'basic' emotion categories illustrated by Barret et al.[238] Used with permission.

It should be noted that this approach is problematic since the meaning of an expression depends on the context. For example, smiles accompanied by cheek raising express enjoyment, the same smile combined with head lowering and turning to the side convey embarrassment. Additionally, facial expressions can be posed or faked.[239]

The sign-based approach measures anatomic facial signs and then uses experimental or observational methods to discover the relation between these signs and emotion.[240] In 1978, the psychologists Ekman and Friesen proposed a model for measuring facial muscle contractions involved in facial expression called 'Facial Action Coding System' (FACS).[241] FACS is now a common standard used to

[234] Alice Caplier, 'Visual Emotion Recognition: Status and Key Issues' in Catherine Pelachaud (ed) Emotion-oriented Systems (Wiley-ISTE 2012) 107, 109.
[235] Hyisung C. Hwand, David Matsumoto, 'Emotional Expression' in Catharine Abell, Joel Smith (eds) *The Expression of Emotion* (CUP 2016) 139, 140.
[236] Catherine Marechal et al, 'Survey on AI-Based Multimodal Methods for Emotion Detection' in Joanna Kolodziej, Horacio Gonzalez-Vélez (eds) *High-Performance Modelling and Simulation for Big Data Applications* (Springer 2019) 314.
[237] Jeffrey F. Cohn, Fernando De La Torre, 'Automated Face Analysis for Affective Computing' in Rafael Calvo et al (eds), *The Oxford Handbook of Affective Computing* (OUP 2015) 132, 133.
[238] Lisa Feldman Barrett et al. 'Emotional Expressions Reconsidered' (2019) Vol 20 (1) Psychological Science in the Public Interest 1, 19.
[239] Jeffrey F. Cohn, Fernando De La Torre, 'Automated Face Analysis for Affective Computing' in Rafael Calvo et al (eds), *The Oxford Handbook of Affective Computing* (OUP 2015) 132.
[240] Ibid 133.
[241] Paul Ekman, Wallace V. Friesen, 'Facial Action Coding System: A Technique for the Measurement of Facial Movements' (1978) Consulting Psychologists Press.

systematically describe and quantify visible human facial movement[242] and describes facial activity in terms of anatomically-based action units (AU).[243] FACS defines 46 AUs to describe each independent movement of the face, including head and eye movements. FACS is used to verify the physiological presence of emotion. Due to its comprehensiveness, it also allows the discovery of new patterns related to emotional states.[244] FACS-coded facial events (AUs) such as 'Inner Brow Raiser', 'Chin Raiser', 'Lip Corner Puller' are classified into emotion categories by matching facial events with emotional events coded from previous empirical studies.[245] Figure 1.7 provides some examples of AUs.

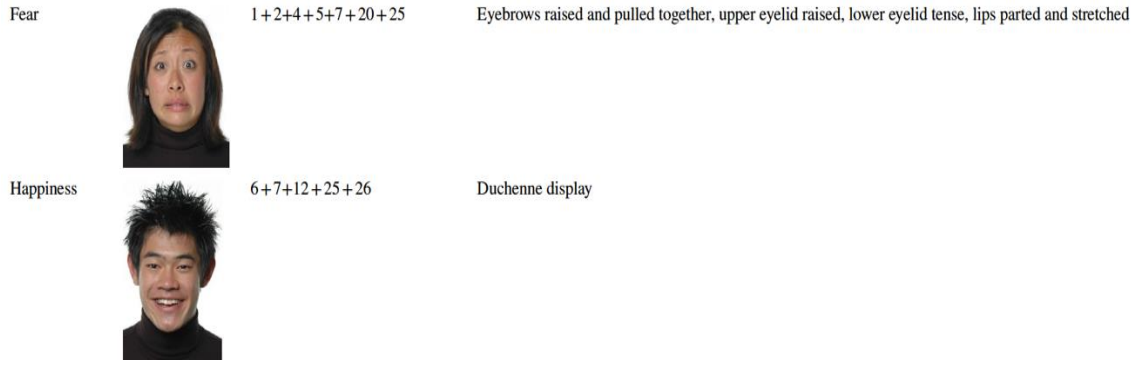

Figure 1.7 Facial expression examples for basic emotions 'fear' and 'happiness', the corresponding FACS action units and physical descriptions for each expression, illustrated by Keltner et al.[246] Used with permission.

Manual application of the FACS to videotaped behaviour is very time consuming. It takes approximately 100 hours to train a person to make judgements reliably and typically takes more than two hours to complete a one-minute video. [247]

Unsurprisingly, computer scientists started to use computer vision and graphics to automatically analyse and synthesise facial expression in automated face analysis (AFA) systems. Recently developed AFA systems claim to detect pain, frustration, emotion intensity, depression and psychological distress.[248] For example, a study aimed to predict depression, anxiety and stress levels from videos using the FACS approach built on ANN-based architecture.[249] Automated face analysis (AFA) systems seek to detect emotions using message-based and sign-based approaches. Such systems typically follow

[242] Lisa Feldman Barrett et al. 'Emotional Expressions Reconsidered' (2019) Vol 20 (1) Psychological Science in the Public Interest 1, 52.
[243] Jeffrey F. Cohn, Fernando De La Torre, 'Automated Face Analysis for Affective Computing' in Rafael Calvo et al (eds), *The Oxford Handbook of Affective Computing* (OUP 2015) 132.
[244] Marian Stewart Bartlett et al, 'Toward Automatic Recognition of Spontaneous Facial Actions' in Paul Ekman, Erika L. Rosenberg, *What the Face Reveals* (2nd edn OUP 2005) 393, 394.
[245] Erika L. Rosenberg, 'Introduction: The Study of Spontaneous Facial Expressions in Psychology' in Paul Ekman, Erika L. Rosenberg, *What the Face Reveals* (2nd edn OUP 2005) 14, 16.
[246] Dacher Keltner et al. 'Emotional Expression: Advances in Basic Emotion Theory' (2019) Vol 43 Iss 2 Journal of Nonverbal Behaviour 133, 142.
[247] Marian Stewart Bartlett et al, 'Toward Automatic Recognition of Spontaneous Facial Actions' in Paul Ekman, Erika L. Rosenberg, *What the Face Reveals* (2nd edn OUP 2005) 394.
[248] Jeffrey F. Cohn, Fernando De La Torre, 'Automated Face Analysis for Affective Computing' in Rafael Calvo et al (eds), *The Oxford Handbook of Affective Computing* (OUP 2015) 132, 133.
[249] Mihai Gavrilescu, Nicolae Vizireanu (2019) 'Predicting Depression, Anxiety, and Stress Levels from Videos Using the Facial Action Coding System' (2019) Vol 19 No 17, 1.

four steps: face detection, face registration, feature extraction and classification. In the first step, the face will be recognised using approaches from object detection in computer vision. During the face registration step, the face is rotated to an upright and frontal facing position to remove geometric differences.[250] In the feature extraction step, the algorithm extracts the main features of the face (e.g. mouth, eyebrows) and analyses movement, shape and texture composition of these regions to identify AUs.[251] After feature extraction, a machine learning (ML) component has the task to learn the relationship between the feature representation and the target facial expressions. Most of the current approaches use supervised learning[252], with a tendency to also make use of deep learning and ANN methods.[253] Fully automatic FACS coding systems use state-of-the-art ML techniques that can recognise any facial action.[254]

2.2.4.2 Speech in affective computing

Emotions of a person may be measured and quantified by observing speech signals from this person. This is exactly what speech-based emotion recognition systems aim at. Such systems are based on insight gained from research that investigates the mechanisms of emotional speech production.[255] Research in emotion recognition has shown that emotions in speech are related to prosody features such as pitch and energy.[256] Prosody refers to the study of the intonational and rhythmic aspects of language. Research has demonstrated specific associations between emotions such as fear, anger, sadness, joy and measures of pitch, voice level and speech rate.[257] Pitch is a perceptual property of a signal. The pitch of a sound is the mental sensation of fundamental frequency. In case a sound has a higher frequency, it is generally perceived as having a higher pitch.[258] The pitch of speech associated with emotions such as anger or happiness is higher than the pitch of speech associated with emotions such as sadness or disappointment.[259] In terms of speech rate, it has been shown that if the person who speaks is in an emotional state of anger or fear, the speech is usually faster. In case the person is bored or sad, then the speech is typically slower. Hence, effects of emotion tend to be present in features

[250] Michael Valstar, 'Automatic Facial Expression Analysis' in Manas K. Mandal, Avinash Awasthi (eds) *Understanding Facial Expressions in Communication* (Springer 2015) 144-150.
[251] Catherine Marechal et al, 'Survey on AI-Based Multimodal Methods for Emotion Detection' in Joanna Kołodziej, Horacio González-Vélez (eds) *High-Performance Modelling and Simulation for Big Data Applications* (Springer 2019) 315.
[252] Jeffrey F. Cohn, Fernando De La Torre, 'Automated Face Analysis for Affective Computing' in Rafael Calvo et al (eds), *The Oxford Handbook of Affective Computing* (OUP 2015) 137.
[253] Panagiotis Tzirakis et al, 'End-to-End Multimodal Emotion Recognition using Deep Neural Networks' (2015) Vol. 14 No. 8 Journal of Latex Class Files, 1.
[254] Marian Stewart Bartlett et al, 'Toward Automatic Recognition of Spontaneous Facial Actions' in Paul Ekman, Erika L. Rosenberg, *What the Face Reveals* (2nd edn OUP 2005) 395.
[255] Chi-Chun Lee et al, 'Speech in Affective Computing' in Rafael Calvo et al (eds), *The Oxford Handbook of Affective Computing* (OUP 2015) 171.
[256] Ricardo A. Calix, Leili Javadpour, Gerald M. Knapp, 'Detection of Affective States From Text and Speech For Real-Time Human-Computer Interaction' (2012) Vol 54 No 4 Human Factors and Ergonomics Society 530, 531.
[257] Christina Sobn and Murray Alpert, 'Emotion in Speech: The Acoustic Attributes of Fear, Anger, Sandess, and Joy' (1999) Vol 28 No 4 Journal of Psycholinguistic Research, 347.
[258] Daniel Jurafsky, James H Martin, *Speech and Language Processing* (2 edn, Pearson Education Limited 2014) 238.
[259] Ze-Jing Chuang, Chung-Hsien Wu, 'Multi-Modal Emotion Recognition from Speech and Text' (2004) Vol. 9 No. 2 Computional Linguistics and Chinese Language Processing, 45-62.

such as average pitch, pitch range and pitch changes, speech rate, voice quality and articulation.[260] Approaches in affective computing extract these acoustic signal features that characterise emotional speech. Machine learning algorithms map the automatically derived acoustic features described before to the desired emotion representations.[261] Research in the field aims to extract features from the voice to detect depressive people[262] or candidate stress levels during human resources interviews using ML and ANN.[263] Real-world applications of AC that aim to derive emotional states from speech are Amazon's 'Halo' wearable, which analyses voice tones to detect user emotions,[264] or Spotify's patented voice assistant,[265] which, based on commands or other utterances (e.g., 'ugh'), recognises when a user sounds sad and then offers encouragement by 'cheering' the user.[266] Methods applied to speech emotion recognition increasingly involve deep learning approaches.[267]

2.2.4.3 Multimodal approaches

Methods used in AC may be combined in multimodal approaches. For example, research in psychology aims to develop multimodal frameworks comprising audio-video fusion (facial expressions and emotions in speech) for the diagnosis, of depression to distinguish between people who suffer from depression and people who do not.[268]

Multimodal approaches may also include the detection from physiological factors such as cardiac activity (heart rate and heart rate variability). Research has shown that the variability of heart rate provides a novel marker to recognise emotions in humans.[269] Both heart rate and heart rate variability have been reported as indicators of fear, panic, anger and appreciation and are therefore used for affective computing.[270] Methods in AC can be integrated in commercial applications in order to track

[260] Rosalind W Picard, *Affective Computing* (MIT Press 1997) 179, 180.
[261] Chi-Chun Lee et al, 'Speech in Affective Computing' in Rafael Calvo et al (eds), *The Oxford Handbook of Affective Computing* (OUP 2015) 173, 177.
[262] Marius Dan Zbancioc, Silvia Monica Feraru, 'A study about the automatic recognition of the anxiety emotional state using Emo-DB' (E-Health and Bioengineering Conference, Iasi, 2015) 1.
[263] Kevin Tomba et al, 'Stress Detection Through Speech Analysis' (2018) Vol 1 ICETE 2018, 560.
[264] Alex Hern, 'Amazon's Halo wristband: the fitness tracker that listens to your mood' *The Guardian* (London, 28 August 2020) <https://www.theguardian.com/technology/2020/aug/28/amazons-halo-wristband-the-fitness-tracker-that-listens-to-your-mood> accessed 8 February 2024; Austin Carr, 'Amazon's New Wearable Will Know If I'm Angry. Is That Weird?' *Bloomberg* (New York, 31 August 2020) <https://www.bloomberg.com/news/newsletters/2020-08-31/amazon-s-halo-wearable-can-read-emotions-is-that-too-weird > accessed 8 February 2024.
[265] Daniel Bromand et al, 'Systems and Methods for Enhancing Responsiveness to Utterances Having Detectable Emotion' US Patent Number US 10566010 B2 (Assignee: Spotify AB) February 2020 11 < https://patentimages.storage.googleapis.com/2a/9d/2d/926b58a2bd956f/US10566010.pdf >, accessed 8 February 2024.
[266] Josh Mandell, 'Spotify Patents A Voice Assistant That Can Read Your Emotions' *Forbes* (New York, 12 March 2020) <https://www.forbes.com/sites/joshmandell/2020/03/12/spotify-patents-a-voice-assistant--that-can-read-your-emotions/> accessed 8 February 2024.
[267] Haytham M Fayek, Margaret Lech, Lawrence Cavedon, 'Evaluating deep learning architectures for Speech Emotion Recognition' (2017) Vol 92 Neural Networks 60.
[268] Jyoti Joshi et al, 'Multimodal assistive technologies for depression diagnosis and monitoring' (2013) Vol 7 Journal on Multimodal User Interfaces, 217.
[269] Quintana Daniel et al. 'Heart rate variability is associated with emotion recognition: Direct evidence for a relationship between the automatic nervous system and social cognition' (2012) Vol 86 No 2 International Journal of Psychophysiology 168.
[270] Jennifer Healey, 'Physiological Sensing of Emotion' in Rafael Calvo et al (eds), *The Oxford Handbook of Affective Computing* (OUP 2015) 206.

and analyse customer behaviour in retail stores, so-called behaviour inference systems. Behavioural inference systems apply deep learning (DL) and affective computing in order to monitor and analyse the shopper's behaviour based on extracted physiological factors (heart rate) and facial expressions.[271] An example for such a system comprises six modules: a speech recognition module, a biofeedback model, a facial expression and emotion recognition module, a gaze detection module, an age and gender recognition module and an identification module.[272]

2.2.5 Automated reasoning (AR)

Automated reasoning (AR) aims to develop computers that can use stored information to answer questions and draw new conclusions.[273] It may be described as the science of developing methods that intend to replace human reasoning with procedures that perform individual reasoning automatically.[274] Automated reasoning is devoted to answering questions from diverse data without human intervention and includes decision-making. As a form of reasoning, decision-making focusses on an autonomous agent trying to perform a task for a human.[275] Reasoning problems are of practical significance, they arise naturally in many applications that interact with the world, for example, reasoning about knowledge in the sciences or natural language processing. Furthermore, reasoning algorithms form the foundation for theoretical investigations into general AI (human-level AI).[276] Reasoning is the process of obtaining new knowledge from a given knowledge, where certain transformation rules are applied that depend only on knowledge and can be done exclusively in the brain without involving senses.[277] Research in automated reasoning focusses on logical reasoning, probabilistic reasoning and common sense reasoning.[278] Logical reasoning attempts to avoid any unjustified assumptions and confines itself to inferences that are infallible and beyond reasonable dispute.[279] Probabilistic reasoning deals with uncertainty about knowledge and belief. Uncertainty may be approached by applying tools from probability theory and statistics. Research in probabilistic reasoning focusses on the representation of different types of uncertainty and uncertain knowledge, reasoning with these types of knowledge, and learning them. It facilitates the development of applied systems of practical importance, such as machine vision, medical diagnosis and natural language processing. Probabilistic reasoning models are close to ML and serve as a medium between ML and AR.[280]

[271] Andrea Generosi, Silvia Ceccacci, Maura Mengoni, 'A deep learning-based system to track and analyse customer behaviour in retail store' (IEEE 8th International Conference on Consumer Electronics, Berlin 2018) 36.
[272] Ibid 37.
[273] Stuart Russel, Peter Norvig, *Artificial Intelligence, A Modern Approach* (3rd edn, Pearson Education 2016) 2.
[274] Tudor Jebelean et al, 'Automated Reasoning' in Buchberger Bruno et al (eds) *Hagenberg Research* (Springer 2009) 63.
[275] Amir Eyal, 'Reasoning and decision making' in Frankish Keith and Ramsey William M. (eds) *The Cambridge Handbook of Artificial Intelligence* (2014) 191.
[276] Ibid 191.
[277] Tudor Jebelean et al, 'Automated Reasoning' in Buchberger Bruno et al (eds) *Hagenberg Research* (Springer 2009) 63.
[278] Amir Eyal, 'Reasoning and decision making' in Frankish Keith and Ramsey William M. (eds) *The Cambridge Handbook of Artificial Intelligence* (2014) 193.
[279] John Harrison, *Handbook of Practical Logic and Automated Reasoning* (Cambridge University Press 2009) 1.
[280] Amir Eyal, 'Reasoning and decision making' in Frankish Keith and Ramsey William M. (eds) *The Cambridge Handbook of Artificial Intelligence* (2014) 201.

For a very long time, scientists and philosophers have tried to understand and formalise how humans reason and whether reasoning methods may be automatised.[281] Achieving common sense reasoning capabilities in computational systems has been one of the goals of AI since its beginning in the 1960s.[282] Common sense reasoning constitutes a central part of human behaviour and is a precondition for human intelligence. Unsurprisingly, the creation of systems that exhibit common sense reasoning is a central goal towards achieving general AI. History in AI has proven that it is more difficult to develop systems with common sense reasoning capabilities compared to systems that solve explicit reasoning problems, such as chess-playing programs or expert systems that assist in clinical diagnosis. Part of this difficulty is due to the all-encompassing aspect of common sense reasoning: It requires many different kinds of knowledge. Furthermore, most common sense knowledge is implicit and therefore difficult to explain and compute, unlike expert-knowledge which is usually explicit. Therefore, implicit common sense knowledge must be made explicit in order to develop common sense reasoning systems.[283]

Other problems that impede the development of automated common sense reasoning are the lack of a precise meaning of 'common sense reasoning', how to take into account of polysemy, ambiguity and vagueness of natural language and the difficulty in modelling the role of various forms of implicit knowledge such as context, background knowledge and tacit knowledge.[284] Therefore, common sense reasoning capabilities are still a challenge in AI applications.[285] According to Oren Etzioni, who oversees the Allen Institute for Artificial Intelligence, AI 'is devoid of common sense'.[286] Hence, to acquire common sense from massive amounts of data and implementing it in intelligent systems appears to be the next frontier in AI.[287] The lack of progress in providing general automated common sense reasoning capabilities underscores that this is a very difficult problem in the field of AI.[288] Common sense reasoning is not just the hardest problem for AI, it is also considered to be the most important problem.[289]

[281] Marco Gavanelli, Toni Mancini, 'Automated Reasoning' (2013) Vol. 7 No. 2 Intelligenza Artificiale 113.
[282] Brandon Bennet, Anthony G Cohn, 'Automated Common-sense Spatial Reasoning: Still a Hughe Challenge' in Stephen Muggleton, Nicholas Chater (eds) Human-Like Machine Intelligence (Oxford University Press 2021) 405.
[283] Ernest Davis, Leora Morgenstern, 'Introduction: Progress in formal common sense reasoning' (2004) Vol 153 Artificial Intelligence 1.
[284] Brandon Bennet, Anthony G Cohn, 'Automated Common-sense Spatial Reasoning: Still a Huge Challenge' in Stephen Muggleton, Nicholas Chater (eds) Human-Like Machine Intelligence (Oxford University Press 2021) 406.
[285] Shoham Yoav et al, 'The AI Index 2018 Annual Report' (AI Index Steering Committee Stanford University 2018) 64 <https://hai.stanford.edu/sites/default/files/2020-10/AI_Index_2018_Annual_Report.pdf > accessed 8 February 2024.
[286] Cade Metz, 'Paul Allen Wants to Teach Machines Common Sense' *The New York Times* (New York, 28 February 2018) <https://www.nytimes.com/2018/02/28/technology/paul-allen-ai-common-sense.html> accessed 8 February 2024.
[287] Niket Tandon, Aparna S. Varde, Gerard de Melo, 'Commonsense Knowledge in Machine Intelligence' (2017) Vol 46 No 4 SIGMOD Record 49.
[288] Brandon Bennet, Anthony G Cohn, 'Automated Common-sense Spatial Reasoning: Still a Hughe Challenge' in Stephen Muggleton, Nicholas Chater (eds) Human-Like Machine Intelligence (Oxford University Press 2021) 405.
[289] Gary Marcus, Ernest Davis, *Rebooting AI: Building Artificial Intelligence we can trust* (Pantheon Books 2019).

2.3 Conclusions

This chapter answered Subquestion 1, namely, what AI is and what disciplines exist therein. AI is an exciting, challenging and complex technology which accelerates at a tremendous pace. AI covers a broad range of approaches and techniques and at least five disciplines. These five disciplines are machine learning, natural language processing, computer vision, affective computing and automated reasoning.

As a major discipline of AI, *machine learning* (ML) is focussed on computers that program themselves based on experience. ML can be applied by means of several methods, ranging from supervised to unsupervised to reinforcement learning. *Deep learning* (DL) is a very powerful kind of machine learning considering that the achievements in the field have been reached with *artificial neural networks* (ANNs) comprising an astonishingly small number of neurons when compared with neural networks of the human brain. By means of *natural language processing* (NLP), machines can process human language. It includes both the generation and understanding of natural language. NLP significantly contributes to improved interactions between machines and humans. *Computer vision* (CV) facilitates the automated processing of visual images and thus enables machines to see. Face recognition, which is one of the applications of computer vision, empowers machines to identify or verify the identity of humans in images or videos based on biometric data. Because emotions form an important factor of human intelligence and daily life, *affective computing* (AC) aims to equip machines with emotional capabilities. Approaches in AC which derive emotions from facial expressions and speech may be easily deployed and widely used. Efforts in the discipline of *automated reasoning* (AR) seek to perform individual reasoning automatically.

3 The current legal framework

This chapter aims to answer the second research question, namely, what the current EU legal framework is. First, Section 3.1 of this chapter describes the current legal framework regarding the fundamental right to privacy followed by Section 3.2 which introduces the fundamental right to data protection. Next, Section 3.3 discusses the most relevant piece of EU secondary law in data protection, namely, the GDPR. Finally, the ePrivacy Directive will be introduced (Section 3.4). Section 3.5 answers Subquestion 2.

As indicated in Sections 1.1 and 1.4, and as apparent from Chapters 4 and 5, I focus on horizontal relationships and EU secondary law. This focus is also visible from the corresponding sub-sections of this chapter. The introduction of the fundamental right to data protection according to Article 8 EUCFR is brief. Nonetheless, Article 8 EUCFR is relevant because the GDPR aims to ensure a high level of protection 'of the rights guaranteed in Article 16 TFEU and *Article 8 of the Charter*'.[290] The GDPR 'implements'[291] this fundamental right and covers horizontal relationships. The principle of proportionality discussed in Section 3.2.2 is a general principle of EU law. It is relevant not only in the context of Article 8 of EUCFR but also when interpreting the GDPR.

The distinction between the fundamental right to privacy (Article 7 EUCFR) and data protection (Article 8 EUCFR) is not purely symbolic. Case law of the European Court of Justice (CJEU) shows that despite substantial overlaps, there are differences with the scope of both rights and their limitation.[292] Imagine, for example, a smart advertisement board in a supermarket powered by software that deploys computer vision and affective computing approaches to analyse the faces of customers that look at the ad board to determine their emotional states, age and sex without the possibility to identify them. Such a scenario would trigger the scope of application of the fundamental right to privacy,[293] but arguably not the right to data protection because individuals cannot be identified.[294] Nevertheless, privacy law is often used as a synonym for data protection law. Admittedly, the distinction is very semantic, similar to a debate on whether a hot dog can also be considered a sandwich.

[290] Case C-645/19 [2021] *Facebook Ireland* ECR I-483 para 45 emphasis added by the author; see also Case C-319/20, *Meta Platforms Ireland Limited* [2022] ECR I-322 para 73.
[291] Article 1 (2) GDPR which reveals the main objective of said regulation: to give meaning to this fundamental right see Hielke Hijmans, Commentary of Article 1 in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020) 56.
[292] Juliane Kokott, Christoph Sobotta 'The distinction between privacy and data protection in the jurisprudence of the CJEU and ECtHR' (2013) Vol 3 No 4 International Data Privacy Law 222.
[293] See Section 3.1.
[294] No personal data are processed; see Section 3.3.1 below.

3.1 The fundamental right to privacy

The human right to respect for private and family life as enshrined in Article 8 of the Council of Europe's European Convention for Human Rights (ECHR) and the corresponding fundamental right according to Article 7 of the EU Charter of Fundamental Rights (EUCFR) protect everyone's 'right to respect for his private and family life, his home and correspondence'.[295] Because the European Court of Justice (CJEU) held that Article 8 ECHR and Article 7 EUCFR must be interpreted *identically*,[296] I refrain from assessing the scope and meaning of these rights separately. Therefore, the following analysis applies to both rights equally. I deliberately focus on the case law of the European Court of Human Rights (ECtHR) because it is more developed than CJEU case law.[297] Within this thesis, I use the term 'private and family life' and 'privacy' interchangeably. As the attentive reader already noted, the EUCFR considers privacy to be a 'fundamental right' and the ECHR to be a 'human right'. The former is commonly used to allude to rights that are granted a special status by a certain legal order, and the latter to rights recognised in international law.[298] Because this thesis focusses on EU law, I use the term 'fundamental right'.

3.1.1 Scope

The essential object of Article 8 ECHR is to protect an individual against 'arbitrary interference by the public authorities' with its private and family life, home and correspondence.[299] This obligation is of the classic negative kind, but the ECtHR emphasised[300] that Article 8 ECHR also entails a positive obligation which requires the state to take steps to provide particular rights or to protect people against the activities of other private individuals.[301] In a Resolution, the Council of Europe stated that the right to privacy granted under Article 8 ECHR 'consists essentially in the right to live one's own life with a minimum of interference'.[302] The ECtHR cited this Resolution in its jurisprudence, including cases where non-state actors infringed the right to privacy.[303] The text in Article 8 (1) ECHR demands for respect of private and family life, home and correspondence. What the term 'respect' means is, even in the view of the ECtHR, not 'clear-cut', in particular 'where the positive obligations implicit in that concept are concerned'.[304] What seems relevant here is one of the fundamental principles in a

[295] Note that the wording of Article 7 EUCFR includes 'communications' instead of 'correspondence' as in Article 8 ECHR. However, the two terms essentially mean the same.
[296] Case C-400/10, *J. McB.* [2010] ECR I-582 para 53; Case C-450/60, *Varec SA* [2008] ECR I-91 para 48.
[297] Frederik Zuiderveen Borgesius, 'Improving Privacy Protection in the area of Behavioural Targeting' (Doctoral thesis, Universiteit van Amsterdam 2015) 99 <https://dare.uva.nl/search?identifier=c74bdba6-616c-4cd9-925e-33a5858935e5> accessed 8 February 2024.
[298] Gloria González Fuster, *The Emergence of Personal Data Protection as a Fundamental Right of the EU* (Springer 2014) 164, 166.
[299] *Niemietz v Germany* App no 13710/88 (ECtHR, 16 December 1992) para 31.
[300] *Marckx v Belgium* App no 6833/74 (ECtHR, 13 June 1979).
[301] David Harris et al, *Law of the European Convention on Human Rights* (4th edn OUP 2018) 502.
[302] Council of Europe, 'Resolution 428 Declaration on mass communication and Human Rights' (1970) para 16.
[303] *Von Hannover v Germany (No. 1)* App no 59320/00 (ECtHR 24 September 2004) para 42; *Von Hannover v Germany (No. 2)* App no 40660/08 and 60641/08 (ECtHR 07 February 2012) para 71; *Mosley v United Kingdom* App no 48009/08 (ECtHR 10 May 2011) para 56.
[304] *Mosley v United Kingdom* App no 48009/08 (ECtHR 10 May 2011) para 108.

democratic society, namely, the rule of law, which dictates the existence of measures of legal protection against arbitrary interference by public authorities with the rights protected by the ECHR.[305]

The following sections elaborate on the protected elements of the fundamental right to privacy that are specifically relevant in the context of this research. These two elements are private life (Section 3.1.1.1) and communication (Section 3.1.1.2).[306] Subsequently, the living instrument doctrine (Section 3.1.2) applied by the ECtHR will be introduced.

3.1.1.1 Private life

The notion of private life is considered to be a broad concept that includes the ability to live one's own life without arbitrary disruption or interference.[307] Thus, the most traditional aspect of the right to private life is the individual's interest in not being exposed to unwanted attention from the state or third parties.[308] In its case law, the ECtHR consistently emphasised that the concept of private life is incapable of an exhaustive definition.[309] However, the case law provides insight into the rather wide range of rights and interests covered under the notion of private life.[310] The interpretation of the term 'private life' in Article 8 is 'underpinned by the notions of personal autonomy and quality of life'.[311] Therefore, the term 'private life' is not limited to an 'inner circle' but encompasses the sphere of personal autonomy within which everyone can freely pursue the development and fulfilment of their personality and establish and develop relationships with other people and the outside world.[312] The right to respect for private life entitles the individual concerned to control the use of its image, including the right to object to the publication of a photograph and to the recording, conservation and reproduction of the image by another person.[313] An individual's image constitutes an essential attribute of personality because 'it reveals the person's unique characteristics and distinguishes the person from his or her peers.'[314] Also, secret surveillance invades an individual's private space[315] and thus interferes with the right to respect private life and correspondence.[316] A violation of the right to respect for private life may even occur when the information obtained by means of secret surveillance measures

[305] *Södermann v Sweden* App no 5786/08 (ECtHR 12 November 2013) para 75; *Tavi v Turkey* App no 11449/02 (ECtHR 9 November 2006) para 28; *Ciubotaru v Moldova* App no 27138/04 (ECtHR 27 April 2010) para 50; David Harris et al, *Law of the European Convention on Human Rights* (4th edn OUP 2018) 368.
[306] In Sections 5.2-5.5, four specific dimensions covered by these two main elements will be discussed in more detail.
[307] David Harris et al, *Law of the European Convention on Human Rights* (4th edn OUP 2018) 503, 504.
[308] Karin de Vries, 'Right to Respect for Private and Family Life' in Pieter van Dijk et al (eds) *Theory and Practice of the European Convention on Human Rights* (Intersentia 2018) 670.
[309] *Niemietz v Germany* App no 13710/88 (ECtHR, 16 December 1992) para 29; *Pretty v United Kingdom* App no 2346/02 (ECtHR 29 April 2002) para 61; *Peck v United Kingdom* App no 44647/98 (ECtHR 28 January 2003) para 57.
[310] Karin de Vries, 'Right to Respect for Private and Family Life' in Pieter van Dijk et al (eds) *Theory and Practice of the European Convention on Human Rights* (Intersentia 2018) 670.
[311] *Pretty v United Kingdom* App no 2346/02 (ECtHR 29 April 2002) paras 61and 62: Christine Goodwin v United Kingdom App no 28957/95 (ECtHR 11 July 2002) para 90.
[312] William Schabas, *The European Convention on Human Rights: A Commentary* (OUP 2015) 369.
[313] William Schabas, *The European Convention on Human Rights: A Commentary* (OUP 2015) 377.
[314] *Reklos and Davourlis v Greece* App no 1234/05 (ECtHR 15 January 2009) para 40.
[315] David Harris et al, *Law of the European Convention on Human Rights* (4th edn OUP 2018) 533.
[316] Karin de Vries, 'Right to Respect for Private and Family Life' in Pieter van Dijk et al (eds) *Theory and Practice of the European Convention on Human Rights* (Intersentia 2018) 670.

is not used subsequently.[317] Thus, it is the surveillance itself that counts as an interference with an individual's privacy.[318]

Three particular dimensions of privacy derived from the element 'private life' will be further discussed in Chapter 5, namely, when they are applied to AI. These are informational privacy (Section 5.2), bodily privacy (Section 5.3) and mental privacy (Section 5.4).

3.1.1.2 Communications

Compared to Article 8 ECHR, the wording of Article 7 EUCFR includes 'communications' instead of 'correspondence'. In essence, the two terms mean the same thing. Both mail and electronic messages fall within the scope of 'correspondence' and under 'communication'. The same applies to telephone calls and similar forms of communication[319] relying on the Internet, such as messenger apps. In other words, the right to respect correspondence protects private communications regardless of their form or content. The term 'correspondence' has been interpreted by the ECtHR in a manner that allows one to keep up with technological developments. It covers telephone, facsimile, email, Internet usage, letters and, most importantly, also other methods of communication in the future.[320] Furthermore, Article 8 of the ECHR protects both private and business-related correspondence, regardless of whether it is carried out from an office or from a private home.[321]

3.1.2 Living instrument doctrine

Article 8 requires the ECtHR to determine issues at the forefront of technology or issues that concern sensitive societal views and values. In this regard, the broad principles of Article 8 have allowed the ECtHR to continuously respond to modern legal dilemmas and human rights challenges.[322] The ECtHR has refused to define the ambit of Article 8 ECHR[323] and 'does not consider it possible or necessary to attempt an exhaustive definition of the notion of private life',[324] which allows the ECtHR to adapt the protection granted under Article 8 ECHR to new circumstances and technological and

[317] *Kopp v Switzerland* App no 23224/94 (ECtHR 25 March 1999) para 53.
[318] Karin de Vries, 'Right to Respect for Private and Family Life' in Pieter van Dijk et al (eds) *Theory and Practice of the European Convention on Human Rights* (Intersentia 2018) 670.
[319] William Schabas, *The European Convention on Human Rights: A Commentary* (OUP 2015) 400, 401.
[320] David Harris et al, *Law of the European Convention on Human Rights* (4th edn OUP 2018) 509.
[321] *Niemietz v Germany* App no 13710/88 (ECtHR, 16 December 1992) para 32; *Halford v United Kingdom* App no 20605/92 (ECtHR 25 June 1997) para 44; Karin de Vries, 'Right to Respect for Private and Family Life' in Pieter van Dijk et al (eds) *Theory and Practice of the European Convention on Human Rights* (Intersentia 2018) 671.
[322] David Harris et al, *Law of the European Convention on Human Rights* (4th edn OUP 2018) 569, 570.
[323] Frederik Zuiderveen Borgesius, 'Improving Privacy Protection in the area of Behavioural Targeting' (Doctoral thesis, Universiteit van Amsterdam 2015) 100 <https://dare.uva.nl/search?identifier=c74bdba6-616c-4cd9-925e-33a5858935e5> accessed 8 February 2024.
[324] *Niemietz v Germany* App no 13710/88 (ECtHR, 16 December 1992) para 29; *Pretty v United Kingdom* App no 2346/02 (ECtHR 29 April 2002) para 61.

societal developments.[325] This dynamic approach to interpretation has been coined the 'living instrument doctrine'.[326] According to the ECtHR, it is crucial that the ECHR is interpreted and applied in a manner which renders its rights practical and effective, not theoretical and illusory.[327] Despite the fact that the living instrument doctrine is obvious in case law,[328] in a dispute concerning covert video surveillance of an employee being suspected of theft, the ECtHR included a sort of caveat with regard to technological developments. In this case, the ECtHR declared that there was a fair balance struck between the right to respect her private life under Article 8 and the employer's interest in the protection of its property rights and the public interest in proper administration of justice.[329] However, the ECtHR stated that 'The competing interests concerned might well be given a different weight in the future, having regard to the extent to which *intrusions* into private life are made possible by new, more and more sophisticated technologies'.[330] This clearly indicates that, depending on the intrusiveness of future technology, the balancing test could have a different outcome in the future. The living instrument doctrine also affects case law adopted by the CJEU. According to the CJEU, Article 8 ECHR and Article 7 EUCFR must be interpreted identically.[331] Furthermore, the EUCFR preamble reaffirms the rights as a result, inter alia, from the ECHR and the case law of the ECtHR and the CJEU.[332] Moreover, according to Article 52 (3) EUCFR, the 'meaning and scope' of the rights contained in the EUCFR and ECHR shall be the same, provided that these rights 'correspond'. This holds true for Article 8 of the ECHR and Article 7 of the EUCFR.

3.2 The fundamental right to data protection

Article 8 EUCFR grants everyone 'the right to the protection of personal data concerning him or her'. For the sake of brevity, I term this the *fundamental right to data protection*. There is no corresponding provision on data protection in the ECHR. However, ECtHR case law under the fundamental right to privacy gave rise to a right of data protection as well. Thus, the fundamental rights to privacy and protection of personal data are closely linked but not identical.[333] The Data Protection Directive[334] has

[325] Frederik Zuiderveen Borgesius, 'Improving Privacy Protection in the area of Behavioural Targeting' (Doctoral thesis, Universiteit van Amsterdam 2015) 100 <https://dare.uva.nl/search?identifier=c74bdba6-616c-4cd9-925e-33a5858935e5> accessed 8 February 2024.
[326] Alastair Mowbray, 'The Creativity of the European Court of Human Rights' (2005) Vol 5 Iss 1 Human Rights Law Review 57-59.
[327] *Christine Goodwin v United Kingdom*, App No 28957/95 (ECtHR 11 July 2002) para 74.
[328] Adapting the protection of Article 8 ECHR to technological developments, e.g. from letters to emails etc.
[329] *Köpke v Germany*, App No 420/07 (ECtHR 05 October 2010).
[330] *Köpke v Germany*, App No 420/07 (ECtHR 05 October 2010) emphasis added.
[331] Case C-400/10, *J. McB.* [2010] ECR I-582 para 53.
[332] Giovanni Carlo Bruno, 'The Importance of the European Convention on Human Rights for the Interpretation of the Charter of Fundamental Rights of the European Union' in Giuseppe Palmisano (ed) *Making the Charter of Fundamental Rights a Living Instrument* (Brill Publishing 2014) 90.
[333] Juliane Kokott, Christoph Sobotta 'The distinction between privacy and data protection in the jurisprudence of the CJEU and ECtHR' (2013) Vol 3 No 4 International Data Privacy Law 222, 223, 228; Herke Kranenborg, Commentary of Article 8 in Steve Peers et al (eds), *The EU Charter of Fundamental Rights* (Hart/Beck 2014) 229.
[334] Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data, OJ 1995 L 281/31.

inspired Article 8 EUCFR; therefore, case law stemming from secondary EU law plays a role when interpreting Article 8 EUCFR.[335]

The following two sub-sections briefly discuss the differences between the fundamental right to privacy and data protection regarding the object and scope (Section 3.2.1) and introduce the principle of proportionality (Section 3.2.2) which plays an important role in EU data protection law.

3.2.1 Scope

The material scope of Article 8 EUCFR covers personal data, which entails all information on identified or identifiable natural persons.[336] Thus, the information protected by Article 8 EUCFR seems to be more extensive than the information covered by the right to privacy under Article 8 ECHR.[337] Additionally, the personal scope differs. The CJEU has excluded legal persons from the fundamental right to data protection,[338] whereas legal persons can rely on the fundamental right to privacy.[339] Unlike most of the other rights of the EUCFR, Article 8 contains several specifications that reflect key elements of the system of checks and balances.[340] Furthermore, Article 8 (2) EUCFR explicitly grants everyone '*right of access* to data which has been collected concerning him or her, and the *right to have it rectified*.'[341] These rights will be discussed in Section 3.3.4 below.

To what extent Article 8 EUCFR has a horizontal effect is unclear. It is argued that the provisions contained in the EUCFR do not directly create obligations for private parties because the provisions of the EUCFR are addressed solely to the institutions, bodies, offices, and agencies of the Union and to the Member States when implementing EU law.[342]As opposed to the fundamental right to privacy, there is extensive secondary EU law that regulates data protection. Secondary EU law will be discussed in Sections 3.3 and 3.4.

3.2.2 Principle of proportionality

As one of the general principles of EU law, the principle of proportionality[343] plays an important role in EU data protection law and has a decisive influence on the evaluation of whether a violation of the

[335] Herke Kranenborg, Commentary of Article 8 in Steve Peers et al (eds), *The EU Charter of Fundamental Rights* (Hart/Beck 2014) 223, 247.
[336] See Section 3.3.1.1 below for the term 'personal data'.
[337] Juliane Kokott, Christoph Sobotta 'The distinction between privacy and data protection in the jurisprudence of the CJEU and ECtHR' (2013) Vol 3 No 4 International Data Privacy Law 222, 225.
[338] Joined Cases C–92/09 and C–93/09 *Volker und Markus Schecke and Eifert* [2010] ECR I-11063, paras 52, 53 and 87.
[339] Juliane Kokott, Christoph Sobotta 'The distinction between privacy and data protection in the jurisprudence of the CJEU and ECtHR' (2013) Vol 3 No 4 International Data Privacy Law 222, 225.
[340] Herke Kranenborg, Commentary of Article 8 in Steve Peers et al (eds), *The EU Charter of Fundamental Rights* (Hart/Beck 2014) 229.
[341] Emphasis added.
[342] Juliane Kokott, Christoph Sobotta 'The distinction between privacy and data protection in the jurisprudence of the CJEU and ECtHR' (2013) Vol 3 No 4 International Data Privacy Law 222, 225.
[343] Article 5 of the consolidated version of the Treaty Establishing the European Community [2006] OJ C321E/37.

right to data protection is justified.[344] The principle of proportionality is important not only in the context of Article 8 EUCFR, but also when interpreting EU secondary law as described in Sections 3.3 and 3.4. According to case law, the principle of proportionality 'requires that measures implemented by acts of the European Union are appropriate for attaining the objective pursued and do not go beyond what is necessary to achieve it.'[345] Derogations and limitations in relation to the protection of personal data shall only apply in so far as is strictly necessary.[346]

According to EU law, the principle of proportionality has generally three components that involve the assessment of a measure's (i) suitability, (ii) necessity and (iii) proportionality stricto sensu.[347] Suitability assesses whether the measure concerned is suitable or relevant to the realisation of the goals it is aimed at meeting. Necessity raises the question whether the measure concerned is required to realise the goals it is aimed at. Proportionality stricto sensu examines non-excessiveness by determining whether the measure goes further than necessary to realise the goals it is aimed at meeting.[348] Necessity comprehends the so-called need-to-know principle and according to the CJEU, access to personal data must only be granted to authorities that have power in the specific field and not to other authorities.[349] Proportionality stricto sensu (iii) requires choosing the least onerous measure and the disadvantages caused by this measure must not be disproportionate to the aims pursued.[350] The CJEU has ruled that the Council and Commission did not comply with the principle of proportionality when requiring the publication of the names of all natural persons who were beneficiaries of agricultural funds and of the exact amounts received by those persons. It reached this conclusion because measures that would affect the fundamental right to data protection less adversely, but still would contribute to the aim pursued, had not been considered.[351]

3.3 General data protection regulation

Arguably, the most relevant and influential EU secondary data protection law is the General Data Protection Regulation (GDPR).[352] Regulations are binding legislative acts and must be applied in its entirety across the EU. With its 99 articles and 173 recitals, the GDPR must be regarded as a

[344] Charlotte Bagger Tranberg, 'Proportionality and data protection in the case law of the European Court of Justice' (2011) Vol 1 No 4 International Data Privacy Law 239-249.
[345] Joined Cases C–92/09 and C–93/09 *Schecke* [2010] ECR I-11063, para 74; *Vodafone and others* [2008] ECR I-188 para 51 and case law cited there.
[346] Case C-73/07 *Satamedia* [2008] ECR I-09831 para 56.
[347] Charlotte Bagger Tranberg, 'Proportionality and data protection in the case law of the European Court of Justice' (2011) Vol 1 No 4 International Data Privacy Law 239-249.
[348] Lee A. Bygrave, *Data Privacy Law: An International Perspective* (OUP 2014) 148.
[349] Case C-524/06 Huber v Bundesrepublik Deutschland [2008] ECR I-724, para 61.
[350] Case C-331/*88 The Queen v Ministry of Agriculture* [1990] ECR I-4023, para. 13. See also Joined Cases C-133, C-300 and C-362/93 *Crispoltoni and others / Fattoria Autonoma Tabacchi* [1994] ECR 1–4863, para 40.
[351] Joined Cases C–92/09 and C–93/09 *Schecke* [2010] ECR I-11063, para 86.
[352] *EU General Data Protection Regulation* (GDPR): Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data and repealing Directive 95/46/EC (General Data Protection Regulation), OJ 2016 L 119/1.

comprehensive piece of legislation. It sets out rules relating to the protection of natural persons regarding the processing of their personal data and rules relating to the free movement of personal data. Article 1 (2) GDPR specifically refers to the objective of the GDPR to protect fundamental rights and freedoms of natural persons and, in particular, the right to the protection of personal data according to Article 8 EUCFR. The following Sections 3.3.1-3.3.4 will shortly elaborate on the most important concepts and provisions of the GDPR in light of the context of this thesis.[353] These sections cover material and personal scope (Section 3.3.1 and 3.3.2) as the well as data protection principles (Section 3.3.3) and the rights of the data subject (Section 3.3.4).

As explained in Section 2.1, AI refers to adaptive machines that can autonomously execute activities and tasks that require capabilities usually associated with humans. Although AI has the ability to make its *own* decisions and perform tasks on the designer's behalf,[354] the GDPR does not apply to AI as such because AI does not have a legal personality. Instead, the GDPR applies to controllers and processors deploying AI systems that process personal data. Therefore, not AI itself but its deployment by companies may cause legal problems. The use of an AI system falls under the scope of the GDPR only if both the material and personal scope are triggered.

3.3.1 Material scope

In essence, the GDPR applies to the processing of personal data wholly or partly by automated means and other than by automated means when the personal data form part of a filing system or are intended to form part of such a system.[355] Thus, whether the material scope of the GDPR is triggered depends on the following key terms: personal data (Section 3.3.1.1), special categories of personal data (Section 3.3.1.2) and processing (Section 3.3.1.3).

3.3.1.1 Personal data

Personal data are defined in Article 4 (1) GDPR as a concept with four elements: i) any information ii) relating to iii) an identified or identifiable iv) natural person. The first element reflects the aim of assigning a wide scope to the concept of personal data and potentially encompasses all kinds of information.[356] The form of the information appears to be irrelevant, as the information may be available 'in written form or be contained in, *for example*, a sound or image'.[357] The second element 'relating to' is also broadly interpreted by the CJEU and is satisfied 'where the information, by reason of its

[353] I do not elaborate on the territorial scope, specific obligations of controllers and on competent supervisory authorities and possible fines.
[354] Eduardo Alonso, 'Actions and agents' in Frankish Keith and Ramsey William M. (eds) *The Cambridge Handbook of Artificial Intelligence* (2014) 235, 236.
[355] Herke Kranenborg, Commentary of Article 2 in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020) 66.
[356] Case C-434/16, *Nowak* [2017] ECR I-994 para 34.
[357] Joined Cases C-141/12 & C-372/12, *YS, M and S v Minister voor Immigratie, Integratie en Asiel* [2014] ECR I-2081, Opinion of AG Sharpston, para 45 (emphasis added).

content, purpose or effect, is linked to a particular person'.[358] What is decisive in whether information constitutes personal data depends on the third element, namely, whether the person concerned in fact is identified or identifiable. With respect to this element, a flexible approach is taken.[359] This is emphasised by the wording of Article 4 (1) and Recital 26 GDPR, in particular the references to 'singling out', 'directly or indirectly' and 'either by the controller or by another person'.[360] Regarding identification, Recital 26 states that account should be taken of 'all the means reasonably likely to be used'. According to the CJEU, this criterion would not be met if identification is prohibited by law 'or practically impossible on account of the fact that it requires a disproportionate effort in terms of time, cost, and man-power, so that the risk of identification appears in reality to be insignificant'.[361] The last element of the concept of personal data makes clear that data on corporations or other legal/juristic persons[362] as well as artificial creatures (e.g. robots) are not protected by the GDPR. Overall, personal data seems to be a broad concept.[363]

3.3.1.2 Special categories of personal data

Article 9 (1) GDPR contains an exhaustive list of special categories of personal data, namely, personal data 'revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership, and the processing of genetic data, biometric data for the purpose of uniquely identifying a natural person, data concerning health or data concerning a natural person's sex life or sexual orientation.' Whereas Article 9 GDPR solely refers to the term 'special categories of personal data' (special data), Recitals 10, 51 and 53-54 also mention the term 'sensitive data'. According to the CJEU, the rationale to ensure enhanced protection for special data is based on their particular sensitivity. Processing of special data is liable to constitute a particularly serious risk of interference with fundamental rights to privacy and data protection.[364] According to the CJEU, the rationale is to prevent significant *risks* to data subjects arising from the processing of special data, regardless of any subjective element such as the controller's *intention*.[365] Thus, there is a higher standard of protection for special data because processing of them poses a greater risk to the fundamental rights of the data

[358] Case C-434/16, *Nowak* [2017] ECR I-994 para 35.
[359] Lee A. Bygrave, Luca Tosoni, Commentary of Article 4 (1) in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020) 109.
[360] Ibid 110.
[361] Case C-582/14, *Breyer v Bundesrepublik Deutschland* [2016] ECR I-779, para 46.
[362] Lee A. Bygrave, Luca Tosoni, Commentary of Article 4 (1) in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020) 111.
[363] Case C-487/21, *F.F.* [2022] ECR I-1000 para 26; Case C-434/16, *Nowak* [2017] ECR I-994 para 34; Purtova takes the view that, in the near future, everything will be or will contain personal data due to the rapid developments in technology. Nadezhda Purtova, 'The law of everything. Broad concept of personal data and future of EU data protection law' (2018) Vol 10 Iss 1 Law, Innovation and Technology 40, 74-75 <https://www.tandfonline.com/doi/full/10.1080/17579961.2018.1452176> accessed 8 February 2024.
[364] Case C-184/20, *OT* [2022] ECR I-601, para 126; Case C-136/17, *GC and Others* [2019] ECR I-773 para 44; Recital 51 GDPR.
[365] Case C-252/21 *Meta Platforms Inc.* [2023] ECR I-537 paras 69-70; Case C-252/21, *Meta Platforms Inc.* [2022] ECR I-704, Opinion of AG Rantos para 41.

subject.[366] According to Recital 51 GDPR, this is due to the particularly sensitive nature of special data. Three of the categories of sensitive data listed in Article 9 (1) are further defined in the GDPR,[367] namely genetic data,[368] biometric data[369] and data concerning health.[370]

The definition of special categories of personal data must be interpreted broadly. The CJEU ruled that personal data which are liable to *indirectly* reveal special categories of personal data defined in Article 9 (1) GDPR are covered by the latter provision.[371] In this ruling, the CJEU followed the AG's opinion by stating that 'the verb "reveal" is consistent with the taking into account of processing not only of inherently sensitive data, but also of data revealing information of that nature *indirectly*, following an intellectual operation involving deduction or cross-referencing'.[372] Another case addresses the processing of special data in the context of websites and applications relating to Facebook users. Whether Article 9 (1) GDPR is applicable in this context depends, according to the CJEU, on the question 'whether the data collected, alone or by virtue of their association with the Facebook accounts of the users concerned, actually enable' to reveal one or more of the categories mentioned in Article 9 (1) GDPR. In certain cases, as pointed out by the CJEU, the mere act of visiting webites or the use of apps may already reveal information as referred to in Article 9 (1) GDPR.[373] Also, it is irrelevant whether a categorisation under Article 9 (1) GDPR is correct or not to fall under the scope of this provision.[374] Processing of special data is prohibited unless one of the exceptions listed in Article 9 (2) GDPR applies. These exceptions are exhaustive and must be interpreted restrictively.[375] In addition to one of the exceptions, processing of special data must always be supported by a legal basis[376] and comply with other provisions[377] of the GDPR.[378] As will be shown in Section 4.8, Article 9 GDPR is particularly relevant for the processing of arguably new types of sensitive personal data facilitated by AI (e.g., emotion data).

[366] Dara Hallinan et al, 'Neurodata and Neuroprivacy: Data Protection Outdated?' (2014) Vol 12 Iss 1 Surveillance and Society 67 <https://ojs.library.queensu.ca/index.php/surveillance-and-society/article/view/neurodata/neurodata4> accessed 8 February 2024.
[367] Ludmila Georgieva, Christopher Kuner Commentary of Article 9 in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020) 374.
[368] Article 4 (13) GDPR.
[369] Article 4 (14) GDPR.
[370] Article 4 (15) GDPR.
[371] Case C-184/20, *OT* [2022] ECR I-601, paras 117-128.
[372] Case C-184/20, *OT* [2022] ECR I-601, paras 123, emphasis added; Case C-184/20, *OT* [2022] ECR I-601, Opinion of AG Pikamäe, para 85.
[373] Case C-252/21 *Meta Platforms Inc.* [2023] ECR I-537 paras 72-73.
[374] Ibid, para 69; See also Case C-252/21, *Meta Platforms Inc.* [2022] ECR I-704, Opinion of AG Rantos paras 39 and 40.
[375] Case C-252/21 *Meta Platforms Inc.* [2023] ECR I-537 para 76.
[376] According to Article 6 GDPR; see also European Data Protection Board, 'Guidelines 3/2019 on the processing of personal data through video devices' (29 January 2020) at 17.
[377] Such as principles for processing and other rules of the GDPR; see Recital 51 GDPR.
[378] Ludmila Georgieva, Christopher Kuner Commentary of Article 9 in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020) 374, 376.

3.3.1.3 Processing

In Article 4 (2), the GDPR defines processing broadly by stating that processing refers to 'any operation or set of operations which is performed on personal data or on sets of personal data, whether or not by automated means, such as collection, recording, organisation, structuring, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, restriction, erasure or destruction'. In short, the definition of processing essentially covers any data processing operation and the use of the wording 'such as' indicates that the list entailed in the definition is not exhaustive. Processing might be further distinguished into automated and manual processing. The former refers to processing done by means of computing devices, and the latter to processing operations executed by humans without the use of computing devices.[379] It should be noted that manual processing falls only within the material scope of the GDPR if the personal data undergoing processing 'form part of a filing system or are intended to form part of a filing system'.[380]

Article 4 (6) GDPR defines a filing system as 'any structured set of personal data which is accessible according to specific criteria, whether centralised, decentralised or dispersed on a functional or geographical basis'. Due to this broad definition, any sets of data grouped together in accordance with specific criteria making such data searchable and accessible without great difficulty are likely to be covered by the definition.[381] According to the CJEU, the requirement that data must be 'structured according to specific criteria' simply demands that personal data can be easily retrieved. In the words of the CJEU, personal data do not need to be 'contained in data sheets or specific lists in another search method, in order to establish the existence of a filing system'.[382]

3.3.2 Personal scope

The GDPR distinguishes between the different actors involved in data processing. These actors are the norm addressees of the GDPR, in essence, the entities that must comply with the GDPR, namely, 'controllers' and 'processors', and the individuals that are protected by the GDPR, the 'data subjects'. The latter are not defined in the GDPR, but Recital 14 indicates that the protection afforded by the GDPR applies to natural persons, whatever their nationality or place of residence, in relation to the processing of their personal data. The definitions for the two actors having to comply with the GDPR, namely, controllers and processors, are introduced in Sections 3.3.2.1 and 3.3.2.2 respectively.

[379] Lee A. Bygrave, Luca Tosoni, Commentary of Article 4 (2) in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020) 119,120.
[380] Article 2 (1) GDPR.
[381] Luca Tosoni, Commentary of Article 4 (6) in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020) 143.
[382] Case C-25/17, *Jehova todistajat* [2018] ECR I-551 para 57-58.

3.3.2.1 Controller

Article 4 (7) GDPR defines controller as 'the natural or legal person, public authority, agency or other body which, *alone* or *jointly* with others, *determines* the *purposes* and *means* of the processing of personal data'.[383] It should be noted that the legal structure of the controller is irrelevant for being considered responsible for the legal obligations under the GDPR.[384] The concept of controller aims to primarily place responsibility for protecting personal data on the entity that actually exercises control over processing of personal data.[385] Regulatory guidance indicates that the concept of controller is functional and 'intended to allocate responsibilities where the factual influence is, and thus based on a factual rather than a formal analysis'.[386] The decisive factor for controllership is the determination of purposes and means of processing personal data. The former relates to the reason and objective of the processing (why), and the latter is to be construed broadly as how processing is exercised, encompassing both technical and organisational elements. The criterion 'determine' can broadly be described as the ability to exercise influence.[387] As the definition in Article 4 (7) GDPR indicates, controllership may be shared. Where 'several operators determine jointly the purposes and means of the processing of personal data, they participate in that processing as [joint] controllers'.[388] Whereas a wide range of joint controllership arrangements are possible, it is often difficult in practice to delineate between joint controllers, separate controllers and other actors such as processors, especially in complex data processing that involve multiple parties.[389] Joint controllership does not presuppose that both controllers involved have access to the processed data.[390]

3.3.2.2 Processor

In addition to controllers, the GDPR imposes data protection obligations on processors defined as 'natural or legal person, public authority, agency or other body which processes personal data *on behalf of the controller*'.[391] As indicated in the definition, the role of the processor is inextricably linked to that of the controller. However, the processor is an entity that is legally separate from the controller and the relationship between these two actors is one of subservience: the processor must adhere to the instructions of the controller regarding the purposes and means of the processing.[392] Any

[383] Emphasis added.
[384] Paul Voigt, Axel von dem Bussche, *The EU General Data Protection Regulation (GDPR): A Practical Guide* (Springer 2017) 18.
[385] Luca Tosoni, Commentary of Article 4 (6) in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020) 148.
[386] Art 29 Working Party, 'Opinion 1/2010 on the concepts of 'controller' and 'processor' (WP 169, 16 February 2010) at 9.
[387] Lee A. Bygrave, Luca Tosoni, Commentary of Article 4 (7) in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020) 150.
[388] Case C-40/17, *Fashion ID* [2019] ECR I-629 para 73.
[389] Lee A. Bygrave, Luca Tosoni, Commentary of Article 4 (7) in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020) 152.
[390] Case C-210/16, *Wirtschaftsakademie* [2018] ECR I-388 para 38.
[391] Article 4 (8) GDPR emphasis added.
[392] Lee A. Bygrave, Luca Tosoni, Commentary of Article 4 (7) in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020) 160.

processor that goes beyond the mandate and instructions of the controller and takes part in determining the purposes and essential means of the processing will itself become a controller.[393] A variety of actors may be deemed processors,[394] including cloud service providers[395] and other external IT service providers or payroll service providers.[396]

3.3.3 Principles

Article 5 GDPR stipulates the principles that govern any processing of personal data. These principles provide the basis for the protection of personal data and some of them are further substantiated in other provisions of the GDPR.[397] The list of principles contained in Article 5 GDPR is exhaustive. In what follows, the principles lawfulness (Section 3.3.3.1), fairness (Section 3.3.3.2), transparency (Section 3.3.3.3), purpose limitation (Section 3.3.3.4), data minimisation (Section 3.3.3.5), accuracy (Section 3.3.3.6), storage limitation (Section 3.3.3.7), confidentiality (Section 3.3.3.8) and accountability (Section 3.3.3.10) will be introduced. In addition, I discuss data protection by design and default, as defined in Article 25 GDPR (Section 3.3.3.9). Strictly speaking, this provision is not a principle in the sense of Article 5 GDPR, but is inextricably linked to the data protection principles. For this reason, I introduce it in this section. The data protection principles discussed in this section will be used in Chapter 4 for further analyses of AI systems in the context of the GDPR.

3.3.3.1 Lawfulness

Lawfulness essentially requires that data processing respects all applicable legal requirements[398] and connotes proportionality in the balancing of interests of data subjects and controllers.[399] This principle is further substantiated in Article 6 GDPR. Processing is only lawful if at least one of the lawful bases listed in the latter provision applies. These exhaustive lawful bases are (i) consent of the data subject, (ii) performance of or entering into a contract, (iii) compliance with a legal obligation, (iv) vital interests of the data subject, (v) performance of a task in the public interest and (vi) the legitimate interest pursued by the controller or third party.[400] According to regulatory guidance, there is no normative hierarchy among the lawful bases[401] and, as indicated by the wording in Article 6 (1), a specific form of processing might be based on more than one lawful basis.

[393] Paul Voigt, Axel von dem Bussche, *The EU General Data Protection Regulation (GDPR): A Practical Guide* (Springer 2017) 20.
[394] Ibid.
[395] Art 29 Working Party, 'Opinion 05/2012 on Cloud Computing' (WP 196, 1st July 2012) at 8.
[396] European Data Protection Board, 'Guidelines 07/2020 on the concepts of controller and processor in the GDPR' (2 September 2020) at 14, 26.
[397] Cécile de Terwangne, Commentary of Article 5 in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020) 313.
[398] Cécile de Terwangne, Commentary of Article 5 in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020) 314.
[399] Lee A Bygrave, *Data Privacy Law: An International Perspective* (OUP 2014) 148.
[400] Article 6 (1) lit a) to f) GDPR.
[401] Art 29 Working Party, 'Opinion 06/2014 on the notion of legitimate interests of the data controller under Article 7 of Directive 95/46/EC' (WP 217, 9 April 2014) at 10.

3.3.3.2 Fairness

Fairness requires that personal data have not been obtained or otherwise processed through unfair means, by deception or without the knowledge of the individual concerned.[402] Despite the fact that the fairness principle is a key tenet of EU data protection law and appears both in the EUCFR and GDPR, its role has thus been elusive[403] due to the lack of judicial guidance. However, both regulatory guidance[404] and regulatory enforcement at the EU level in the form binding decisions[405] adopted by the EDPB identify key elements of the fairness principle. These key elements are: autonomy of data subjects with respect to data processing, their reasonable expectations, ensuring power balance between controllers and data subjects, avoidance of deception, as well as possible adverse consequences of processing, and ensuring ethical and truthful processing.[406] In this sense, the fairness principle ensures 'that personal data shall not be processed in a way that is detrimental, discriminatory, unexpected or misleading to the data subject'.[407] Taking into account the text of Recital 39 GDPR, which stresses the link between transparency and fairness ('information to the data subjects [….] ensure fair and transparent processing'), absence of information will make processing unfair. However, fairness of processing means more than transparency[408] and has an independent meaning. This is confirmed by regulatory enforcement at the EU level. The principles of fairness, lawfulness and transparency are three *distinct* but intrinsically *linked* principles and fairness has an *independent* meaning.[409] The fairness principle focusses on proportionality in the balancing of interest of data subjects and controllers, and the latter have to take account of the reasonable expectations of data subjects when processing their personal data.[410]

[402] Cécile de Terwangne, Commentary of Article 5 in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020) 314.
[403] Damian Clifford, Jef Ausloos 'Data Protection and the Role of Fairness' (2018) Vol 37 No 1 Yearbook of European Law 130, 187, Milda Mačėnaitė, 'Protecting Children Online: Combining the Rationale and Rules of Personal Data Protection Law and Consumer Protection Law' in Mor Bakhoum et al (eds) *Personal Data in Competition, Consumer Protection and Intellectual Property Law* (Springer Nature 2018) 361.
[404] European Data Protection Board, 'Guidelines on Article 6(1)(b) GDPR' (Guidelines 2/2019, 8 October 2019), at 6; European Data Protection Board, 'Guidelines on Article 25 Data Protection by Design and Default' (Guidelines 4/2019, 20 October 2020), at 17 and 18.
[405] Article 65 GDPR.
[406] Binding Decision 3/2022 on the dispute submitted by the Irish SA on Meta Platforms Ireland Limited and its Facebook service (Art. 65 GDPR), adopted on 5 December 2022 paras 103, 219-220, 222-223, 226 478; Binding Decision 4/2022 on the dispute submitted by the Irish SA on Meta Platforms Ireland Limited and its Instagram service (Art. 65 GDPR), adopted on 5 December 2022 paras 106, 223-224, 226-227, 445; Binding Decision 5/2022 on the dispute submitted by the Irish SA on Meta Platforms Ireland Limited and its WhatsApp service (Art. 65 GDPR) adopted on 5 December 2022.
[407] European Data Protection Board, 'Guidelines on Article 25 Data Protection by Design and Default' (Guidelines 4/2019, 20 October 2020), at 17 and 18.
[408] Winston J Maxwell, 'Principle-based regulation of personal data: the case of 'fair processing' (2015) Vol 5 No 3 International Data Privacy Law 205, 208.
[409] Binding Decision 3/2022 on the dispute submitted by the Irish SA on Meta Platforms Ireland Limited and its Facebook service (Art. 65 GDPR), adopted on 5 December 2022 paras 22, 477; Binding Decision 4/2022 on the dispute submitted by the Irish SA on Meta Platforms Ireland Limited and its Instagram service (Art. 65 GDPR), adopted on 5 December 2022 para 226, 444; Binding Decision 5/2022 on the dispute submitted by the Irish SA on Meta Platforms Ireland Limited and its WhatsApp service (Art. 65 GDPR), adopted on 5 December 2022.
[410] Lee A. Bygrave, *Data Privacy Law: An International Perspective* (OUP 2014) 147, 148.

3.3.3.3 Transparency

Recital 36 GDPR specifies the principle of transparency inherent in Article 6 (1) GDPR by requiring that it must be transparent to natural persons 'that personal data concerning them are collected, used, consulted or otherwise processed.'[411] It is further substantiated in Articles 12 through 14 GDPR in the form of obligations towards the controller to provide certain information to the data subject. In view of the EDPB, these provisions are the concretisation of the transparency principle, and violations of these provisions may also amount to the violation of the transparency principle itself.[412] Article 13 GDPR applies when personal data are collected from the data subject, and Article 14 applies when personal data have not been obtained from the data subject (e.g., third party controllers, data brokers, publicly available sources).[413] Information must be easily accessible, and when informing data subjects, the controller must use clear and plain language to make the information provided easy to understand.[414] It is important to note that the GDPR obliges controllers, amongst others,[415] to inform data subjects about the purposes of the processing for which the personal data are intended and the legal basis for the processing.[416] In the case of indirect collection, controllers must also inform data subjects about the categories of personal data that are undergoing processing.[417] The description of these categories should be precise enough to allow the data subject to grasp an overall understanding of the processing in view of the fairness and transparency principle.[418]

3.3.3.4 Purpose limitation

The purpose limitation principle enshrines two requirements: (i) personal data must be collected for specified, explicit and legitimate purposes and (ii) personal data must not be further processed for incompatible purposes.[419] The principle of proportionality is embodied in the purpose limitation principle by means of the requirement that personal data should be collected for specified and legitimate purposes. Thus, the purpose limitation principle seems to be intertwined with the proportionality principle because any assessment of the proportionality relies on the identification of a processing's purpose.[420] Specification requires that purposes must be determined at the very beginning of processing,

[411] Recital 39 GDPR.
[412] EDPB, 'Binding Decision 1/2021 on the dispute arisen on the draft decision of the Irish Supervisory Authority regarding WhatsApp Ireland under Article 65 (1) lit a GDPR' (2021) paras 191, 193.
[413] Art 29 Working Party, 'Guidelines on transparency under Regulation 2016/679' (WP 260 rev.01, 11 April 2018) at 15.
[414] Recital 39 GDPR.
[415] For a full overview, see Article 13 and 14 GDPR as well as corresponding commentaries by Gabriela Zanfir-Fortuna, Commentary of Articles 13 and 14 in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020) 413 et seq.
[416] Art 13 (1) lit c and Art 14 (1) lit c GDPR.
[417] Art 14 (1) lit d GDPR.
[418] Gabriela Zanfir-Fortuna, Commentary of Article 15 in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020) 444.
[419] Case C-77/21 *Digi* [2022] ECR I-805 Opinion of AG Pikamäe para 28; Merel Elize Koning, 'The purpose and limitations of purpose limitation' (Doctoral thesis, Radboud University Nijmegen 2020) 58 < https://repository.ubn.ru.nl/bitstream/handle/2066/221665/221665.pdf?sequence=1&isAllowed=y > accessed 8 February 2024.
[420] Lee A. Bygrave, *Data Privacy Law: An International Perspective* (OUP 2014) 148. Also, regulatory guidance on legitimate interest and purpose limitation is quite similar. See Art 29 Working Party, 'Opinion 06/2014 on the notion of

namely, at the time of collection of personal data. Thus, processing of personal data for undefined or unlimited purposes is unlawful.[421] The purpose specification requirement plays a central role because all the data protection principles introduced in Section 3.3.3 are based on it.[422] The purposes must be 'explicit', that is, clearly revealed, explained or expressed towards the data subjects concerned, to ensure an unambiguous understanding of the purposes of processing.[423] Legitimacy, another component of the purpose specification principle, arguably means that personal data should only be processed for purposes 'that do not run counter to ethical and social mores that are generally deemed appropriate to govern the relationship of the controller and data subject(s).'[424]

The principle of compatible use implies that a controller may process personal data for all purposes that may be considered compatible with the initial purposes. Article 6 (4) GDPR stipulates a series of criteria to determine whether further processing for a purpose other than the one for which personal data have been initially collected is 'compatible' with this initial purpose.[425] According to the CJEU, these criteria reflect the need for a concrete, coherent and sufficiently close link between the purpose of data collection and the further processing of the data and make it possible to determine that such further processing does not detract from the legitimate expectations as to the further use of their personal data.[426] Importantly, further processing for archiving purposes in the public interest, scientific or historical research or statistical purposes is a priori not considered to be incompatible with the initial purposes provided that such processing is subject to appropriate safeguards.[427]

3.3.3.5 Data minimisation

The data minimisation principle enshrined in Article 5 (1) lit c GDPR stipulates that personal data must be 'adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed'. Its requirements that personal data must be relevant and necessary impose limits on the amount of personal data that may be processed.[428] The stipulation that personal data must be 'relevant' and 'limited' in relation to the purposes for which they are processed gives expression to the

legitimate interests of the data controller under Article 7 of Directive 95/46/EC' (WP 217, 9 April 2014) and Art 29 Working Party, 'Opinion 03/2013 on purpose limitation' (WP 203, 2 April 2013).

[421] Cécile de Terwangne, Commentary of Article 5 in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020) 315.

[422] Merel Elize Koning, 'The purpose and limitations of purpose limitation' (Doctoral thesis, Radboud University Nijmegen 2020) 102 < https://repository.ubn.ru.nl/bitstream/handle/2066/221665/221665.pdf?sequence=1&isAllowed=y > accessed 8 February 2024.

[423] Art 29 Working Party, 'Opinion 03/2013 on purpose limitation' (WP 203, 2 April 2013) at 39.

[424] Lee A. Bygrave, *Data Privacy Law: An International Perspective* (OUP 2014) 155.

[425] Cécile de Terwangne, Commentary of Article 5 in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020) 315, 316.

[426] Case C-77/21 *Digi* [2022] ECR I-805 para 36; Case C-77/21 *Digi* [2022] ECR I-805 Opinion of AG Pikamäe paras 28, 59, 60.

[427] Article 89 GDPR.

[428] Asia Biega, Michèle Finck, 'Reviving Purpose Limitation and Data Minimisation in Data-Driven Systems' (2021) Technology and Regulation 44, 56 <https://arxiv.org/abs/2101.06203> accessed 8 February 2024.

principle of proportionality.[429] The latter is a requirement arising from settled case law.[430] According to Recital 39 GDPR, personal data should only be processed if the purposes cannot reasonably be fulfilled by other means. Anything exceeding the 'minimum' amount necessary will be considered excessive and violate the data minimisation principle. If the same results can be achieved through the processing of less personal data, the exceeding part of the processing is not necessary.[431] The data minimisation principle also plays a role with respect to the storage of personal data.[432] Furthermore, what is 'necessary' refers not only to the quantity, but also to the quality of the personal data processed.[433]

3.3.3.6 Accuracy

The GDPR states that the processing of personal data must be accurate and, where necessary, kept up to date.[434] Controllers have to rectify or erase all inaccurate data and must take every reasonable step to comply with the accuracy principle.[435] The term 'reasonable' arguably implies that it is legitimate for controllers to take into account cost and resource factors when deciding on measures to rectify or delete inaccurate data.[436]

The accuracy principle intends to protect the individual concerned from being irrationally or unfairly treated based on wrong and inaccurate representations.[437] According to regulatory guidance, accurate means 'accurate as to a matter of fact'.[438] What is required to assess the accuracy of the personal data depends on the context, namely, on the purpose of the processing.[439] Thus, the accuracy principle seems to be an undefined concept in EU data protection law because questions and definitions as to exactly how accurate personal data needs to be remain unaddressed.[440]

[429] Case C-439/19 *B* [2021] ECR I-504 para 98; Lee A. Bygrave, *Data Privacy Law: An International Perspective* (OUP 2014) 148.
[430] Cases C-92/09 and C-93/09, *Schecke* [2010] ECR I-662 paras 72 and 74; Case C-58/08, *Vodafone and others* [2008] ECR I-188 para 51.
[431] Asia Biega, Michèle Finck, 'Reviving Purpose Limitation and Data Minimisation in Data-Driven Systems' (2021) Technology and Regulation 44, 56 <https://arxiv.org/abs/2101.06203> accessed 8 February 2024.
[432] Case C-77/21 *Digi* [2022] ECR I-805 para 58.
[433] Cécile de Terwangne, Commentary of Article 5 in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020) 317.
[434] Art. 5 (1) lit d GDPR.
[435] Cécile de Terwangne, Commentary of Article 5 in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020) 317.
[436] Lee A. Bygrave, *Data Privacy Law: An International Perspective* (OUP 2014) 164.
[437] Dara Hallinan, Frederik Zuiderveen Borgesius 'Opinions can be incorrect (in our opinion)! On data protection law's accuracy principle' (2020) Vol 10 No 1 IDPL 1, 9.
[438] Art 29 Working Party, 'Guidelines on the Implementation of the Court of Justice of the European Union Judgment on Google Spain and inc v. Agencia Española de Protección De Datos (AEPD) and Mario Costeja González C-131/12' (WP 225, 26 November 2014), at 15. <https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=667236> accessed 8 February 2024.
[439] Case C-434/16, *Nowak* [2017] ECR I-994 para 53.
[440] Dara Hallinan et al, 'Neurodata and Neuroprivacy: Data Protection Outdated?' (2014) Vol 12 Iss 1 Surveillance and Society 66, 67 <https://ojs.library.queensu.ca/index.php/surveillance-and-society/article/view/neurodata/neurodata4> accessed 8 February 2024.

3.3.3.7 Storage limitation

The storage limitation principle enshrined in Article 5 (1) lit d GDPR prohibits to store personal data in a form which permits identification of data subjects beyond the time necessary to achieve the purposes of processing. Storage for longer periods is permitted for or archiving purposes in the public interest, scientific or historical research purposes or statistical purposes provided that appropriate technical and organisational measures are implemented in order to safeguard the rights and freedoms of the data subjects.[441] The CJEU applied the storage limitation principle to a case where the controller 'stored personal data initially collected for other purposes in a testing and error correction database'. According to the CJEU, a controller cannot retain personal data in a database established for testing and error correction purposes for longer than what is necessary to conduct such testing and correct errors.[442]

3.3.3.8 Integrity and confidentiality

The integrity and confidentiality principle enshrined in Article 5 (1) lit f GDPR requires controllers to implement appropriate security measures to ensure that personal data are protected against unauthorised or unlawful processing and protected from accidental loss, destruction or damage.[443] Chapter IV of the GDPR further develops and substantiates this duty of security for both controllers and processors.[444] The measures taken should be commensurate with the risks involved in the processing.[445] I do not further elaborate on this principle because AI poses particular risks to information security. For instance, AI makes it easier for cybercriminals to penetrate systems without human intervention. Whereas such attacks could also compromise the protection of personal data, these attacks cause significant damage to companies whose systems were penetrated.[446] AI creates a 'cybercrime tsunami'[447] which merits dedicated research. However, such research does not fall within this thesis's scope.

3.3.3.9 Data protection by design and default

The concept of data protection by design and default enshrined in Article 25 GDPR does not appear under the principles for processing named in Article 5 of the GDPR. However, I mention it here under the principles because it is closely intertwined with them and important in the context of this thesis.[448] The concept of data protection by design and default obliges controllers to apply technical and

[441] Article 5 (1) lit e GDPR.
[442] Case C-77/21 *Digi* [2022] ECR I-805 paras 46-62.
[443] Article 5 (1) lit f GDPR.
[444] Cécile de Terwangne, Commentary of Article 5 in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020) 318.
[445] Lee A. Bygrave, *Data Privacy Law: An International Perspective* (OUP 2014) 164; see also Article 32 GDPR.
[446] Eddie Segal, 'The Impact of AI on Cybersecurity' IEEE Computer Society <https://www.computer.org/publications/tech-news/trends/the-impact-of-ai-on-cybersecurity> accessed 8 February 2024.
[447] Philip Treleaven et al, 'The Future of Cybercrime: AI and Emerging Technologies Are Creating a Cybercrime Tsunami' (2023) < https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4507244 > accessed 8 February 2024.
[448] Because this concept seems to be particularly relevant for the use of new technologies, such as AI.

organisational measures 'that are designed to implement data protection principles'.[449] It also imposes a duty on controllers to integrate necessary safeguards into the processing of personal data to ensure that processing will meet its requirements and otherwise ensure the protection of data subjects' rights.[450] The 'by design' measures are both technical and organisational and embrace not simply the design and operation of software and hardware, but also business strategies and other organisational practices. The 'by default' requirements of Article 25 (2) GDPR are mainly concerned with results that guarantee data minimisation and confidentiality.[451] It is important to note that data protection by design and default measures must be taken at both the design and processing stage.[452]

3.3.3.10 Accountability

The accountability principle in Article 5 (2) GDPR states that the controller shall be i) responsible for compliance and ii) able to demonstrate compliance with all the previous principles mentioned in Article 5 (1) GDPR.[453] It is further developed in Article 24 GDPR and requires controllers to 'implement appropriate and effective measures to ensure and be able to demonstrate' that processing of personal data occurs in accordance with the rules set out in the GDPR.[454] It follows from the accountability principle itself and from CJEU case law that the burden of proof regarding the compliance with principles enshrined in Article 5 (1) GDPR lies with the controller.[455]

3.3.4 Rights

Chapter 3 of the GDPR provides the data subject with enforceable rights. The following sections will briefly elaborate on the scope of these rights. Note that the following sections do not discuss the information obligations that controllers must comply with, although these obligations are placed in Chapter III of the GDPR termed 'rights of the data subject'.[456] Thus, transparency requirements do technically not belong to the rights of data subjects and are therefore explained in Section 3.3.3.3 dealing with the transparency principle. I have chosen not to discuss notification obligations (Article 19 GDPR) and restrictions to data subject rights (Article 23 GDPR) contained in Chapter III GDPR. These provisions do not constitute enforceable data subject rights and thus fall out of the scope of this

[449] Article 25 GDPR.
[450] Lee A. Bygrave, Commentary of Article 25 in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020) 576.
[451] Lee A. Bygrave, Commentary of Article 25 in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020) 577.
[452] Article 25, Recital 78 GDPR.
[453] Cécile de Terwangne, Commentary of Article 5 in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020) 311.
[454] Art. 24 (1), Recital 74 GDPR.
[455] Case C-175/20 *'SS' SIA* [2022] ECR I-124 paras 77, 81.
[456] I do not take the view that something as a 'right to be informed' exists under the GDPR. Rather, controllers are obliged to comply with the transparency principle, which is further substantiated in articles 12-14 GDPR. In addition, the right to restriction of processing and notification obligation regarding rectification or erasure will be left out due to the lack of direct relevance for this thesis.

thesis. Additionally, I do not specifically address the right to restriction of processing according to Article 18 GDPR, but discuss it in the context of the right to object.

In what follows, the most prominent data subject rights will be introduced. These are the right of access (Section 3.3.4.1), the right to rectification (Section 3.3.4.2), the right to erasure (Section 3.3.4.3), the right to data portability (Section 3.3.4.4), the right to object (Section 3.3.4.5) and the right not to be subject to automated decision-making (Section 3.3.4.6). These data subject rights will be further analysed in the context of AI (Chapter 5).

3.3.4.1 Right of access

The right of access according to Article 15 GDPR provides the data subject with the right to demand in-depth information on processing going beyond the general information according to Articles 13-14 GDPR, which controllers must disclose to data subjects by default.[457] Article 15 GDPR enables data subjects to receive (i) confirmation of the processing, (ii) details about the processing and (iii) access to the personal data themselves, including a copy of the personal data.[458] The first element (i) simply includes a confirmation or denial of the controller that personal data of the data subject are being processed. Details to be provided according to element (ii) overlap with the information that must be disclosed under Articles 13 and 14 GDPR when personal data are collected or received. However, the details to be provided to the data subject under the right of access must be more precise and specifically address information about the personal data related to the person making the request.[459] Such details include, where applicable, information about automated decision-making.[460]

Element (iii) of the right of access obliges the controller to 'provide a copy of personal data undergoing processing'.[461] This aims to strengthen the position of the data subject.[462] The concept of 'copy' is not defined in the GDPR and therefore must be determined in line with usual meaning in everyday language as well as in the context of Article 15 GDPR. According to current linguistic usage, the term 'copy' refers to the 'reproduction or transcription' of an original.[463] Two well-known dictionaries define the notion as 'something that has been made to be exactly like something else'[464] or 'a thing that is made to be the same as something else, especially a document or a work of art'.[465] AG Pitruzella suggests interpreting the concept of copy as the '*faithful* reproduction in intelligible form of the

[457] Paul Voigt, Axel von dem Bussche, *The EU General Data Protection Regulation (GDPR): A Practical Guide* (Springer 2017) 150.
[458] Gabriela Zanfir-Fortuna, Commentary of Article 15 in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020) 449.
[459] Ibid 463.
[460] Article 15 (1) lit h, Article 22 GDPR.
[461] Article 15 (3) GDPR.
[462] Case C-487/21, *F.F.* [2022] ECR I-1000 para 37; see also Opinion of AG Pitruzella para 69.
[463] Case C-487/21, *F.F.* [2022] ECR I-1000 para 21; see also Opinion of AG Pitruzella paras 28-30.
[464] See <https://dictionary.cambridge.org/dictionary/english/copy> accessed 8 February 2024.
[465] See <https://www.oxfordlearnersdictionaries.com/definition/english/copy_1?q=copy> accessed 8 February 2024.

personal data requested by the DS, in material and permanent form'.[466] He hesitated to clarify what is meant with 'faithful'. Dictionaries describe this notion as 'true and accurate; not changing anything'[467] and 'true or not changing any of the details, facts, style, etc. of the original'.[468] The CJEU followed AG Pitruzella's opinion. It ruled that a 'copy' refers to 'faithful reproduction or transcription' of an original. A purely general description of the data undergoing processing or a reference to categories of personal data does not correspond to that definition.[469] In addition, the right to obtain a copy includes not only personal data collected by the controller, but also information resulting from the processing of personal data, for instance, a credit score.[470] Therefore, the copy must enable the data subject to effectively exercise its right of access in full knowledge of all personal data undergoing processing, including personal data *generated* by the *controller*.[471] Article 15 (3) does not require the provision of a copy of the document but a copy of the personal data.[472] However, in some cases, controllers are required to recreate extracts from documents or even entire documents or extracts from databases containing personal that undergo processing to ensure that information is easy to understand, as required by Article 12 (1) GDPR.[473] In addition, Article 15 (3) GDPR does not provide the data subject with a right to obtain information regarding the criteria, models, rules or internal procedures (whether or not computational) used for processing the personal data.[474]

Importantly, the right of access may be restricted twofold, namely, in line with Article 23 GDPR and, more specifically, in accordance with Article 15 (4) GDPR. The latter only applies to element (iii) of the right of access. The right to obtain a copy of personal data shall not adversely affect the rights and freedoms of others,[475] 'including trade secrets or intellectual property and in particular the copyright protecting the software'.[476] The actual protection provided by the right of access must be determined contextually.[477] Rights, such as the right of access, may only be restricted when this constitutes a necessary measure to safeguard the rights and freedoms of others.[478] According to the CJEU, a balance will have to be struck in cases of conflict between right the right to obtain a full copy of personal data and rights and freedoms of others, including IP and trade secrets.[479]

[466] Case C-487/21, *F.F.* [2022] ECR I-1000, Opinion of AG Pitruzella para 70.
[467] See <https://www.oxfordlearnersdictionaries.com/definition/english/faithful?q=faithful> and < accessed 8 February 2024.
[468] See < https://dictionary.cambridge.org/dictionary/english/faithful > accessed 8 February 2024.
[469] Case C-487/21, *F.F.* [2022] ECR I-1000 para 21.
[470] Ibid, para 26.
[471] Case C-487/21, *F.F.* [2022] ECR I-1000 para 21; see also the opinion of AG Pitruzella paras 45, 70.
[472] Note however that this depends on local guidance and local case law, arguably leading to 'unharmonized' results across the EU. In a recent case in the Netherlands, the court pointed out that the GDPR does not grant a right to obtain a copy of documents, but rather a right to obtain a copy of personal data. See Rechtbank Den Haag, *C/09/572633/HA RK 19-295* ECLI:NL:RBDHA:2019:13029 para 4.5.
[473] Case C-487/21, *F.F.* [2022] ECR I-1000 para 41.
[474] Case C-487/21, *F.F.* [2022] ECR I-1000, Opinion of AG Pitruzella para 52.
[475] Article 15 (4) GDPR
[476] Recital 63 GDPR.
[477] Sandra Wachter, Brent Mittelstadt 'A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI' (2019) No 2 Columbia Business Law Review 494, 536.
[478] Case C-434/16, *Nowak* [2017] ECR I-994 para 60.
[479] Case C-487/21, *F.F.* [2022] ECR I-1000 para 44.

The right of access is closely intertwined with other data subject rights because it allows data subjects to exercise these rights as noted by the CJEU.[480] According to the CJEU, the right of access 'is necessary, inter alia, to enable the data subject to obtain, depending on the circumstances, the rectification, erasure or blocking of his data by the controller and consequently to exercise' these rights.[481] Thus, according to the CJEU, the objective of the right of access is to guarantee the protection of the right to privacy with respect to data processing, and not to ensure 'the greatest possible transparency of the decision-making process of the public authorities and to promote good administrative practices by facilitating the exercise of the right of access to documents.'[482] Also, in another case, the CJEU stressed that the right to data protection is not designed to facilitate the exercise of the right of access to documents.[483] In conclusion, the main objective of Article 15 GDPR is to allow the data subject to be aware of processing, verify the lawfulness of the latter and enforce its rights as a data subject.[484]

3.3.4.2 Right to rectification

The right to rectification according to Article 16 GDPR enables the data subject to demand the controller to rectify inaccurate personal data and to have incomplete personal data completed. Thus, in addition to the rectification of inaccurate or false data, the data subject may add missing elements in order to complete personal data by providing a supplementary statement.[485] The CJEU held that the right to rectification may also be asserted in relation to written answers submitted by the candidate in a context of a professional examination, including comments made by an examiner.[486] However, the right to rectification must be interpreted teleologically. Obviously, the right to rectification should not result in situations where a candidate for a professional examination would be allowed to correct his answers in an exam retroactively[487] or an individual to rectify the content of a legal analysis in the context of an immigration case.[488] The question of whether personal data are accurate and complete must be assessed in light of the purpose for which the data was collected.[489] Regulatory guidance states that derived or inferred data constitute (new) personal data[490] and that the right to rectification applies not only to the 'input personal data' but also to 'output data'.[491] The term rectification implicitly relies upon the notion of verification in the sense that something may demonstrably be shown to

[480] Case C-434/16, *Nowak* [2017] ECR I-994 para 57; Case C-487/21, *F.F.* [2022] ECR I-1000, Opinion of AG Pitruzella para 65.
[481] Case C-487/21, *F.F.* [2022] ECR I-1000 para 35; Joined Cases C-141/12 & C-372/12, *YS, M* [2014] ECR I-2081 para 44.
[482] Ibid paras 46-47.
[483] Case C-28/08 P, *Bavarian Lager* [2010] ECR I-6055 para 49.
[484] Case C-487/21, *F.F.* [2022] ECR I-1000 para 35; see also Opinion of AG Pitruzella para 65.
[485] Cécile de Terwangne, Commentary of Article 16 in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020) 473.
[486] Case C-434/16, *Nowak* [2017] ECR I-994 para 51.
[487] Ibid para 54.
[488] Joined Cases C-141/12 & C-372/12, *YS, M and S v Minister voor Immigratie, Integratie en Asiel* [2014] ECR I-2081, para 45.
[489] Case C-434/16, *Nowak* [2017] ECR I-994 para 53.
[490] Art 29 Working Party 'Guidelines on Automated Individual Decision-Making and Profiling for the Purposes of Regulation 2016/679', (WP251rev.01, 6 February 2018) at 8-9.
[491] Ibid at 17-18.

be inaccurate or incomplete and consequently corrected by the individual concerned.[492] Indeed, AG Sharpston takes the view that 'only information relating to *facts* about an individual can be personal data.'[493] Such facts may be expressed in different forms, for example, a person's weight may be expressed objectively in kilogrammes or in subjective terms such as 'underweight' or 'obese'.[494] Demonstration of facts might be a straightforward task when the personal data in question is verifiable (such as a name, date of birth, email address or the weight of an individual).[495] With regard to inferred data, which are defined as products of probability-based processes,[496] it is generally impossible for data subjects to prove that such data are wrong without access to the tools used to infer the data.[497]

3.3.4.3 Right to erasure

The right to erasure in Article 17 GDPR is well known as 'the right to be forgotten' and was brought to great attention of the public by the Google Spain decision of the CJEU.[498] Under the right to erasure, the data subject may demand the controller to erase his or her personal data if the personal data (i) are no longer necessary in relation to the purposes for which they are processed, (ii) have been unlawfully processed, (iii) have to be erased for compliance with a legal obligation under EU or Member State law or (iv) have been collected based on a child's consent in relation to information society services.[499] The same applies when a data subject withdraws consent or objects to the processing of personal data.[500] However, the right to erasure is not an absolute right, as indicated by the exceptions enshrined in paragraph 3 of Article 17 GDPR. These exceptions apply regardless of the ground on which the erasure is based.[501] A controller must not comply with a data subject's request for erasure to the extent that processing is necessary for (i) exercising the right to freedom of expression and information; (ii) compliance with a legal obligation of the controller that requires processing by EU or Member State law and the performance of a task carried out in the public interest; (iii) reasons of public interest in the area of public health; (iv) archiving purposes in the public interest, scientific or historical research purposes or statistical purposes; or (v) establishment, exercise or defence of legal claims.[502] The legal consequence of a successful request according to Article 17 (1) GDPR is the erasure of the personal

[492] Sandra Wachter, Brent Mittelstadt 'A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI' (2019) No 2 Columbia Business Law Review 494, 548.
[493] Joined Cases C-141/12 & C-372/12, *YS, M* [2014] ECR I-2081, Opinion of AG Sharpston para 56.
[494] Ibid para 57.
[495] Sandra Wachter, Brent Mittelstadt 'A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI' (2019) No 2 Columbia Business Law Review 494, 548.
[496] OECD Working Party on Security and Privacy in the Digital Economy JT03357584 (2014) <https://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=dsti/iccp/reg(2014)3&doclanguage=en> accessed 8 February 2024.
[497] Bart Custers, 'Profiling as inferred data. Amplifier effects and positive feedback loops' in Emre Bayamlıoğlu, Irina Baraluic, Liisa Janssens and Mireille Hildebrandt (eds), *Being Profiled: Cogitas Ergo Sum* (Amsterdam University Press 2018) 115.
[498] Case C-131/12, *Google Spain* [2014] ECR I-317.
[499] Article 17 (1) lit a, d, e, f GDPR.
[500] Ibid lit b and c.
[501] Herke Kranenborg, Commentary of Article 17 in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020) 482.
[502] Article 17 (3) GDPR. For regulatory guidance, see European Data Protection Board, 'Guidelines 05/2019 on the criteria of the Right to be Forgotten in the search engines cases under the GDPR part 1' (7 July 2020).

data.[503] The notion of 'erasure' is not defined in the GDPR, but it arguably refers to making data unusable in a way that prevents the controller, processor or any third party from processing the data by physically destroying or technically deleting the data.[504] Another legal consequence[505] is that the controller, if it has made the personal data public, is obliged to inform other controllers who are processing such data to erase any links to or copies of replications of the personal data.[506]

3.3.4.4 Right to data portability

Article 20 GDPR grants data subjects a right to indirect[507] and direct[508] data portability. Indirect data portability allows data subjects to receive their personal data and transmit them to another controller without interference from the original controller. Direct data portability enables data subjects to have their personal data transmitted directly from one controller to another.[509] As indicated in Recital 68 of the GDPR, the right to data portability 'should further strengthen the control' over personal data and is thus strongly related to the notion of control that dominated data protection reform efforts.[510] For the right to apply, three cumulative conditions have to be met: (i) the personal data have been provided directly by the data subject making the request, and processing is (ii) based on consent or a contract and (iii) carried out by automated means.[511] If one of the conditions is not met, the right cannot be invoked.[512] Condition (i) excludes personal data that is created by the controller, namely, personal data that is inferred or derived from personal data provided by the data subject.[513] Personal data like the 'online reputation' an individual develops in digital marketplaces based on customer reviews are likely excluded from the scope.[514] With regard to condition (ii), the right is limited to processing of personal data based on the lawful basis of consent[515] or performance of a contract.[516] Finally, condition (iii) excludes processing by nonautomated means.[517] When the data subject successfully invokes the right to data portability, the controller must provide the personal data in a 'structured, commonly used and machine-readable format.'[518] Recital 68 adds that the format should be interoperable and the requirement of a 'commonly used format', which is not defined in a recital or elsewhere in the GDPR,

[503] Meaning that one of the grounds in Article 17 (1) is triggered and no exception under Art. 17 (3) GDPR applies.
[504] Paul Voigt, Axel von dem Bussche, *The EU General Data Protection Regulation (GDPR): A Practical Guide* (Springer 2017) 161.
[505] Article 17 (2) GDPR.
[506] Herke Kranenborg, Commentary of Article 17 in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020) 483.
[507] Article 20 (1) GDPR.
[508] Article 20 (2) GDPR.
[509] Stephanie Elfering, *Unlocking the Right to Data Portability* (Nomos 2019) 20.
[510] Inge Graef, Martin Husovec, Nadezhda Purtova, 'Data Portability and Data Control: Lessons for an Emerging Concept in EU Law' Vol 19 No 06 German Law Journal 1359, 1365.
[511] Article 20 (1) GDPR.
[512] Stephanie Elfering, *Unlocking the Right to Data Portability* (Nomos 2019) 23.
[513] Article 29 Working Party, 'Guidelines on the right to data portability' (WP 242rev.01, 5 April 2017) at 10.
[514] Orla Lynskey, Commentary of Article 20 in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020) 503.
[515] Article 6 (1) lit a GDPR.
[516] Article 6 (1) lit b GDPR.
[517] Stephanie Elfering, *Unlocking the Right to Data Portability* (Nomos 2019) 24.
[518] Article 20 (1) GDPR.

arguably refers to a format compatible with the state of the art at the time the request is made.[519] The right to data portability is not an absolute one, as Article 20 (4) indicates that this right shall not adversely affect the rights and freedoms of others. This provision arguably also covers intellectual property rights and trade secrets, as is the case with the right of access, which is closely related to the right to data portability.[520] It has been argued that this right, next to data protection law, also has a consumer and competition law dimension[521] and that this right does not fit well with the fundamental rights nature of data protection law.[522]

3.3.4.5 Right to object

Article 21 (1) GDPR confers on the data subject the right to object to processing 'on grounds relating to his or her particular situation'. Simultaneously, it imposes a duty on the controller to cease processing unless it can demonstrate 'compelling legitimate grounds for the processing', which override the interests, rights and freedoms of the data subject or for the establishment, exercise or defence of legal claims.[523] The term 'compelling' arguably means 'overwhelming' and thus requires that the rights and interests of the data subject are overridden in a strong, significant way.[524] Thus, the compelling legitimate grounds of the controller must be so important that the purposes of processing cannot be achieved without the processing that the data subject objected to.[525] The burden of proof that the conditions in Article 21 (1) are met lies with the controller, and any rejection to comply with a data subject's objection to the processing must be explained in the correspondence with the data subject.[526] When a data subject exercises the right to object, whether successful or not, the controller must immediately restrict the processing pursuant to Article 18 (1) lit d GDPR. Where the data subject's objection to processing has merit, the controller must no longer process personal data and has the obligation to erase them[527] 'without undue delay'.[528] If the data subject objects to processing for direct marketing purposes according to Article 21 (2) GDPR, including profiling related to direct marketing, there is no need to balance interests. This provision has an absolute character and therefore it is sufficient that the data subject simply objects to such processing.[529] Other than with an objection under

[519] Stephanie Elfering, *Unlocking the Right to Data Portability* (Nomos 2019) 21.
[520] Ibid 29, 30.
[521] Inge Graef, 'Blurring Boundaries of Consumer Welfare' in Mor Bakhoum et al (eds), *Personal data in competition, consumer protection and intellectual property law* (Springer 2018) 121-151.
[522] Inge Graef, Martin Husovec, Nadezhda Purtova, 'Data Portability and Data Control: Lessons for an Emerging Concept in EU Law' Vol 19 No 06 German Law Journal 1359, 1365.
[523] Article 21 (1) GDPR.
[524] Gabriela Zanfir-Fortuna, Commentary of Article 21 in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020) 517.
[525] Paul Voigt, Axel von dem Bussche, *The EU General Data Protection Regulation (GDPR): A Practical Guide* (Springer 2017) 178.
[526] Gabriela Zanfir-Fortuna, Commentary of Article 21 in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020) 517.
[527] Gabriela Zanfir-Fortuna, Commentary of Article 21 in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020) 518.
[528] Article 17 (1) lit c GDPR.
[529] Gabriela Zanfir-Fortuna, Commentary of Article 21 in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020) 518.

Article 21 (1) GDPR, the controller does not need to erase the personal data but is simply required to cease the processing of personal data for direct marketing purposes.[530] Where personal data are processed for scientific or historical research purposes or for statistical purposes, the data subject can object to such processing according to Article 21 (6) GDPR. However, when the processing referred to in this provision is necessary for the performance of a task carried out for reasons of public interest, such public interests prevail. In this case, the controller will be obliged to prove such a necessity.[531]

3.3.4.6 Right not to be subject to automated decision making

In Article 22 (1), the GDPR grants individuals the right 'not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her'. According to the CJEU, the rationale of Article 22 GDPR is to protect data subjects effectively against 'the particular risks to their rights and freedoms associated with the automated processing of personal data'.[532] In academia, the nature of the right according to Article 22 (1) is subject to considerable disagreement.[533] The crucial question is whether Article 22 GDPR shall be interpreted as a general prohibition of automated decision-making (ADM) or if it must be interpreted as a right to be invoked, similar to a right to object.[534] In *SCHUFA*, the first case dealing with Article 22 GDPR, the CJEU interpreted this provision as a 'prohibition in principle',[535] thereby putting an end to this debate. This is in line with regulatory guidance,[536] AG Pikamäe's opinion[537] and my impressions from the oral hearing in this case.[538]

In order to apply, Article 22 (1) rests on three cumulative conditions: (i) a decision is made that is (ii) based solely on automated processing or profiling and (iii) has either legal effects or similarly significant effects.[539] Bygrave suggests that the use of the term 'including' profiling must be read as

[530] Article 21 (3) GDPR which states 'Where the data subject objects to processing for direct marketing purposes, the personal data shall no longer be processed for such purposes'.
[531] Paul Voigt, Axel von dem Bussche, *The EU General Data Protection Regulation (GDPR): A Practical Guide* (Springer 2017) 179.
[532] Case C-634/21, *SCHUFA Holding AG* [2023] ECR I-957 paras 57, 60.
[533] Lee A. Bygrave, Commentary of Article 22 in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020) 530.
[534] Arguing that Article 22 is a right to be invoked by the data subject. See Luca Tosoni, 'The right to object to automated individual decisions: resolving the ambiguity of Article 22(1) of the General Data Protection Regulation' (2021) Vol 11 Iss 2 International Data Privacy Law 145-162.
[535] Case C-634/21, *SCHUFA Holding AG* [2023] ECR I-957 paras 52, 64.
[536] Article 29 Working Party, 'Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679' (WP 251rev.01, 6 February 2018), at 9, 12, 19.
[537] Case C-634/21, *SCHUFA Holding AG* [2023] ECR I-220, Opinion AG Pikamäe para 31.
[538] Andreas Häuselmann, 'The ECJ's First Landmark Case on Automated Decision-Making – a Report from the Oral Hearing before the First Chamber' (European Law Blog, 20 February 2023) <https://europeanlawblog.eu/2023/02/20/the-ecjs-first-landmark-case-on-automated-decision-making-a-report-from-the-oral-hearing-before-the-first-chamber/> accessed 8 February 2024.
[539] Case C-634/21, *SCHUFA Holding AG* [2023] ECR I-957 para 43; Opinion AG Pikamäe paras 33 and 36; Lee A. Bygrave, Commentary of Article 22 in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020) 532.

equivalent to 'involving',[540] meaning that automated processing in the sense of Article 22 GDPR necessarily involves profiling.[541] In other words, profiling is seen as a necessary element of automated processing. However, such an interpretation is not undisputed because the term 'including' could also simply imply that automated processing *may* involve profiling. AG Pikamäe seems to support the latter interpretation. In his view, profiling is a subcategory of automated processing,[542] which implies that automated processing according to Article 22 (1) GDPR may involve profiling, but also covers other forms of automated processing. Furthermore, the regulatory guidance correctly points to the different concepts of 'ADM' and 'profiling': ADM has a different scope than profiling,[543] but may partially overlap with or result from the latter. In addition, ADM may be made with or without profiling; and profiling can take place without ADM.[544] Unfortunately, the CJEU did not address this question explicitly in *SCHUFA*, arguably because it was clear that the automated establishment of a credit score value constitutes profiling.[545] For this thesis, I interpret the reference to profiling in line with AG Pikamäe's opinion in *SCHUFA*[546] and regulatory guidance. This interpretation also matches with the rationale of Article 22 GDPR according to the CJEU, namely effective protection against risks associated with *automated processing* of personal data.[547]

The GDPR does not define the term decision contained in Article 22 (1) GDPR. However, the CJEU interprets this term broadly based on Recital 71 GDPR which also refers to 'measures'.[548] Following AG Pikamäe's opinion,[549] the CJEU ruled that a decision covers many acts which may affect individuals in several ways, including the automated establishment of a score value.[550] Bygrave suggests that a decision as required by condition (i) covers a wide range of situations and should be viewed in a fairly generic sense, provided it is formalised so that it can be distinguished from other stages that prepare, support or complement decision-making.[551] According to AG Pikamäe, the term decision implies a 'view' or 'opinion' on a particular matter from an etymological point of view. It is not necessary for the decision to have a specific form; the *effect* that the decision has on the data subject is *decisive*.[552]

[540] Lee A Bygrave, 'Minding the machine v2.0: The EU General Data Protection Regulation and Automated Decision Making' in Karen Yeung and Martin Lodge (eds) *Algorithmic Regulation* (OUP 2019) 252.
[541] Isak Mendoza, Lee A Bygrave, 'The Right Not to be Subject to Automated Decisions Based on Profiling' in Tatiana-Eleni Synodinou et al (eds) *EU Internet Law: Regulation and Enforcement* (Springer International 2017) 91.
[542] Case C-634/21, *SCHUFA Holding AG* [2023] ECR I-220, Opinion AG Pikamäe para 33.
[543] Defined in Article 4 (4) GDPR.
[544] Article 29 Working Party, 'Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679' (WP 251rev.01, 6 February 2018) at 8.
[545] Case C-634/21, *SCHUFA Holding AG* [2023] ECR I-957 para 47.
[546] Case C-634/21, *SCHUFA Holding AG* [2023] ECR I-220, Opinion AG Pikamäe para 33.
[547] The CJEU referred to automated processing and not 'only' profiling, Case C-634/21, *SCHUFA Holding AG* [2023] ECR I-957 para 57. Also, in paras 53, 62, 68 and 72 the CJEU mentions automated processing, not profiling.
[548] Case C-634/21, *SCHUFA Holding AG* [2023] ECR I-957 paras 44, 45.
[549] Case C-634/21, *SCHUFA Holding AG* [2023] ECR I-220, Opinion AG Pikamäe paras 38, 42.
[550] Case C-634/21, *SCHUFA Holding AG* [2023] ECR I-957 para 46.
[551] Lee A. Bygrave, Commentary of Article 22 in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020) 532.
[552] Case C-634/21, *SCHUFA Holding AG* [2023] ECR I-220, Opinion AG Pikamäe paras 37, 43.

The second condition (ii) means the absence of meaningful human involvement (or better: influence) in the decision process. Meaningful human involvement requires that it be carried out by a person who is competent or authorised to change a decision. Routinely applying automated decisions without any actual influence on the result (e.g., rubber-stamping automated decisions) would not be regarded as human involvement.[553] The last condition (iii) requires that the decision changes, shapes or otherwise determines an individual's rights or duties or has consequences that have a serious adverse impact.[554] Legal effects under condition (iii) means that the decision affects an individual's legal rights, such as the freedom to associate with others, vote in an election or take legal action. It also involves decisions that affect a person's legal status or rights under a contract (for example, cancellation of a contract or entitlement to social benefits).[555]

Naturally, defining what meets the threshold of 'significant effects' is more difficult. According to AG Pikamäe, these significant effects may be of economic and social nature and relate to severe consequences for freedoms and autonomy. They include adverse effects resulting from a negative score value, which significantly restricts the data subject in exercising its freedoms or even stigmatises the data subject.[556] The CJEU went a bit less far, but confirmed that the automated establishment of a probability value (credit score) meets the threshold of 'significant effects'.[557] The application of this threshold will arguably vary depending on the attributes and sensibilities of the data subject concerned.[558] Regulatory guidance indicates that the threshold may be met when the decision has the following: the potential to significantly affect the circumstances, behaviour or choices of the individual; a prolonged or permanent impact; or, in its most extreme form, the risk of leading to the exclusion or discrimination of individuals.[559]

According to Article 22 (2) GDPR, the prohibition of ADM[560] does not apply in three alternative sets of circumstances, namely, when ADM is necessary in the context of a contract, based on a statutory authority or based on consent. However, data subjects will always have the right to demand human review, to express their point of view and to contest the decision, except when ADM is based on statutory authority.[561] Whether Article 22 (3) requires controllers to provide data subjects with a right

[553] Article 29 Working Party, 'Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679' (WP 251rev.01, 6 February 2018), at 20, 21.
[554] Lee A. Bygrave, Commentary of Article 22 in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020) 534.
[555] Article 29 Working Party, 'Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679' (WP 251rev.01, 6 February 2018), at 21.
[556] Case C-634/21, *SCHUFA Holding AG* [2023] ECR I-220, Opinion AG Pikamäe paras 38, 39, 42, 43.
[557] Case C-634/21, *SCHUFA Holding AG* [2023] ECR I-957 paras 48-50.
[558] Lee A. Bygrave, Commentary of Article 22 in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020) 534.
[559] Article 29 Working Party, 'Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679' (WP 251rev.01, 6 February 2018), at 21.
[560] Case C-634/21, *SCHUFA Holding AG* [2023] ECR I-957 paras 52, 53, 64.
[561] Article 22 (3) GDPR; Lee A. Bygrave, Commentary of Article 22 in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020) 534.

of ex-post explanation (after the decision is adopted) has been subject to scholarly debate.[562] However, when considering the accountability, fairness and the transparency principles and related provisions (e.g. to provide 'meaningful information about the logic involved' in ADM) there seems to be solid ground for a right for ex-post explanation.[563] Article 22 (4) further prohibits ADM based on special categories of personal data unless such ADM is necessary for reasons of substantial public interest[564] or the data subject has provided explicit consent.[565]

3.4 ePrivacy Directive

The provisions of the ePrivacy Directive (ePD)[566] aim to 'particularise and complete'[567] the GDPR[568] in the electronic communications sector.[569] EU Directives, as opposed to EU Regulations, must be implemented in national legislation of EU Member States. This can lead to differences within the different EU Member States.[570] Whereas both the GDPR and the ePD have the object of protecting fundamental rights and freedoms,[571] the GDPR sets general rules for the processing of personal data, and the ePD regulates the fundamental right to privacy *and* data protection in the electronic communications sector.[572] Thus, in accordance with the principle lex specialis derogate legi generali,[573] provisions of the ePD that specifically regulate processing of personal data in the electronic communications sector take precedence over the general provisions of the GDPR.[574] However, this applies only where the material scope of both laws is triggered.[575] The relationship between the GDPR and ePD is

[562] Denying the existence of such a right: Sandra Wachter, Brent Mittelstadt, Luciano Floridi, 'Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation' (2017) Vol 7 Iss 2 IDPL 76-99; most scholars however disagree: Andrew Selbst, Julia Powles, 'Meaningful information and the right to explanation' (2017) Vol 7 Iss 4 IDPL 233-242; Giancludio Malgieri, Giovanni Comandé, 'Why a Right to Legibility of Automated Decision-Making Exists in the General Data Protection Regulation' (2017) Vol 7 Iss 4 IDPL 243-265; Isak Mendoza, Lee A. Bygrave, 'The Right not to be Subject to Automated Decisions based on Profiling', in Synodinou Tatiana-Eleni et al (eds), EU Internet Law: Regulation and Enforcement (Springer 2017) 77.
[563] Lee A. Bygrave, Commentary of Article 22 in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020) 538.
[564] Article 9 (2) lit g GDPR.
[565] Article 9 (2) lit a GDPR.
[566] Directive 2002/58/EC of the European Parliament and of the Council of 12 July 2002 concerning the processing of personal data and the protection of privacy in the electronic communications OJ L 201/27 further on referred to as 'ePrivacy Directive' as amended by Directive 2009/136/EC.
[567] Article 1 (2) ePrivacy Directive.
[568] Initially Data Protection Directive 95/46/EC.
[569] Tijmen H.A. Wisman, 'Privacy, Data Protection and E-Commerce' in Arno L. Lodder, Andrew D. Murray (eds) *EU regulation of e-commerce. A commentary* (Elgar 2017) 369.
[570] In this thesis, I do not elaborate on the relevant Member State laws.
[571] In the case of the GDPR, the fundamental right to the protection of personal data (Article 1 para 2 GDPR), and in the case of the ePrivacy Directive, both the fundamental right to privacy (Recital 12) and data protection (Recital 2). Note that Directive which amended the ePrivacy Directive also refers to the fundamental right to privacy and confidentiality (Recital 51) and the fundamental right to the protection of personal data (Recital 56).
[572] Christina Etteldorf, 'EDPB on the Interplay between the ePrivacy Directive and the GDPR' (2019) Iss 5 No 2 European Data Protection Law Review 224, 226.
[573] Joined Cases *T-60/06 RENV II and T-62/06 RENV II* [2016] ECR II-233 para 81.
[574] European Data Protection Board, 'Opinion 5/2019 on the interplay between the ePrivacy Directive and the GDPR, in particular regarding the competence, tasks and powers of data protection authorities' (Opinion 5/2019, 12 March 2019), at 17.
[575] Christina Etteldorf, 'EDPB on the Interplay between the ePrivacy Directive and the GDPR' (2019) Iss 5 No 2 European Data Protection Law Review 224, 226.

governed by Article 95 and Recital 173 of the GDPR. Note that the proposed ePrivacy Regulation,[576] which is supposed to repeal the ePD, is still subject to political negotiations. I will discuss this proposal to a limited extent in Sections 4.8.3, 4.9, 6.4.2, 6.4.3.

The following sections outline the material scope (Section 3.4.1) and the personal scope (Section 3.4.2) of the ePD. The last section elaborates on the provisions of the ePD that specifically regulate the use of certain types of information and the processing of personal data (Section 3.4.3).

3.4.1 Material scope

The ePD applies to the processing of personal data in connection with the provision of publicly available electronic communications services ('ECS') in public communications networks in the EU.[577] To establish what constitutes an ECS requires some effort as the ePD refers to the Framework Directive[578] which, in the context of the modernisation of the EU's telecom framework, has been repealed by the European Electronic Communications Code ('EECC').[579] The latter introduces a new definition of ECS and because Article 125 and Annex XII of the EECC specifically require that any cross reference to the repealed Framework Directive is construed to refer to the EECC, the scope of the ePD has been extended. The new definition of ECS covers Internet access services, interpersonal communications services and services consisting wholly or mainly in the conveyance of signals.[580] It also includes over-the-top (OTT) services delivering content over the Internet such as VoIP[581] solutions, messaging services and web-based email services which are functionally equivalent to the more traditional voice telephony and text message services.[582] Under the previous definition of ECS, purely Internet-based VoIP solutions were not covered and did therefore not fall under the scope of the ePD.[583] According to CJEU case law, to fall within the scope of an ECS, a service must include the *conveyance of signals.*[584] All that matters concerning the conveyance of signals is that a service provider is *responsible* vis-à-vis the end-users for *transmission* of the *signal* which ensures that they are supplied with the service to which they have subscribed.[585] In the case of web-based services, it is the *Internet Access Provider* (IAP) and the *operators* of the *various networks* of which the *open Internet* is constituted

[576] Commission, 'Proposal for a Regulation of the European Parliament and of the Council concerning the respect for private life and the protection of personal data in electronic communications and repealing Directive 2002/58/EC (Regulation on Privacy and Electronic Communications)' COM (2017) 10 final 'Proposal ePrivacy Regulation'.
[577] Article 3 (1) ePrivacy Directive.
[578] Directive 2002/21/EC of the European Parliament and of the Council of 7 March 2002 on a common regulatory framework for electronic communications networks and services OJ L 108 further on 'Framework Directive'.
[579] Directive (EU) 2018/1972 of the European Parliament establishing the European Electronic Communications Network OJ L 321/36 further on 'EECC'.
[580] Article 2 (4) EECC.
[581] VoIP solutions, for example, enable individuals to call via computer without the call being routed on to a number in the regular telephony numbering plan.
[582] Recital 15 EECC.
[583] Eleni Kosta, Jos Dumortier, 'ePrivacy Directive: Assessment of transposition, effectiveness and compatibility within the proposed Data Protections Regulaton' (2015) European Commission 36 <https://research.tilburguniversity.edu/en/publications/eprivacy-directive-assessment-of-transposition-effectiveness-and-> accessed 8 February 2024.
[584] Case C-193/18, *Google LLC* [2019] ECR I-498 para 32.
[585] Case C-475/12, *UPC* [2014] ECR I-285 para 43.

that convey the signals necessary for the functioning of web-based services.[586] It is also common understanding that providers of web-based services (e.g. Gmail) somehow participate in the conveyance of signals, for example, by means of uploading data packets to the open Internet or by splitting messages into data packets. However, this is not sufficient to be regarded as an ECS consisting 'wholly or mainly in the conveyance of signals on electronic communications networks'.[587] In essence, the ePD applies when the following cumulative conditions are met: (i) there is an ECS[588] which is (ii) offered over an electronic communications network[589] and the service and network are (iii) publicly available[590] and (iv) offered in the EU.[591] In addition, the material scope of the ePrivacy extends to the storage of information or gaining access to information already stored in the terminal equipment of a subscriber or user[592] (including cookies and other tracking technologies) and unsolicited communications (including direct marketing).[593]

3.4.2 Personal scope

As indicated by its material scope, most of the provisions of the ePD only apply to providers ECS.[594] Certain provisions of the ePD are nevertheless applicable to providers of information society services.[595] Article 5 (3) ePD as indicated by regulatory guidance applies to every entity that places on or reads information from terminal equipment including smart devices[596] and regardless of the nature of the entity.[597] This includes particularly websites operators that place cookies[598] and apps that are installed on the end-user device and access data stored on the device.[599] In addition, Article 13 ePD applies to any business, including website operators, which sends unsolicited electronic mail for direct marketing purposes.[600]

[586] Case C-193/18, *Google LLC* [2019] ECR I-498 para 36.
[587] Ibid.
[588] As defined in Article 2 (4) EECC.
[589] As defined in Article 2 (1) EECC.
[590] A service available to all members of the public on the same basis.
[591] European Data Protection Board, 'Opinion 5/2019 on the interplay between the ePrivacy Directive and the GDPR, in particular regarding the competence, tasks and powers of data protection authorities' (Opinion 5/2019, 12 March 2019), at 10.
[592] Article 5 (3) ePrivacy Directive.
[593] Article 13 ePrivacy Directive.
[594] Article 29 Working Party, 'Opinion 3/2013 on apps on smart devices' (WP 202, 27 February 2013) at 7.
[595] Eleni Kosta, Jos Dumortier, 'ePrivacy Directive: Assessment of transposition, effectiveness and compatibility within the proposed Data Protections Regulaton' (2015) European Commission 9 <https://research.tilburguniversity.edu/en/publications/eprivacy-directive-assessment-of-transposition-effectiveness-and-> accessed 8 February 2024.
[596] Including smartphones, tablets and smart TVs.
[597] Article 29 Working Party, 'Opinion 3/2013 on apps on smart devices' (WP 202, 27 February 2013), at 7.
[598] European Data Protection Board, 'Opinion 5/2019 on the interplay between the ePrivacy Directive and the GDPR, in particular regarding the competence, tasks and powers of data protection authorities' (Opinion 5/2019, 12 March 2019), at 11.
[599] Article 29 Working Party, 'Opinion 3/2013 on apps on smart devices' (WP 202, 27 February 2013), at 14.
[600] European Data Protection Board, 'Opinion 5/2019 on the interplay between the ePrivacy Directive and the GDPR, in particular regarding the competence, tasks and powers of data protection authorities' (Opinion 5/2019, 12 March 2019), at 11.

The ePD protects users who are individuals, meaning 'any natural person using a publicly available electronic communications service, for both private and business purposes, without necessarily having subscribed to this service'[601] as well as subscribers who are legal persons.[602] In this sense, the ePD complements the GDPR, as the latter does not provide protection to legal persons.

3.4.3 Specific requirements

This section elaborates on provisions of the ePD that specifically regulate the use of certain types of information and the processing of personal data. Where provisions of the ePD require consent, such consent must meet the conditions for obtaining consent according to the GDPR.[603] The CJEU already interpreted the notion of consent as required in Article 5 (3) ePD in the light of the GDPR.[604]

In what follows, I discuss provisions enshrined in the ePD with particular relevance in the context of AI. These are confidentiality of communications (Section 3.4.3.1), information stored in terminal equipment (Section 3.4.3.2) and location data (Section 3.4.3.3).

3.4.3.1 Confidentiality of communications

Article 5 (1) ePD protects the confidentiality of communications and the related traffic data.[605] It prohibits listening, tapping, storage or other types of interception and surveillance by persons other than users. Interception of communication is allowed if the user provided consent or if technical storage is necessary for the conveyance of communication.[606] Article 5 (2) ePD provides for the so-called business exception[607] and states that the protection of confidentiality shall not affect recordings for the 'purpose of providing evidence of a commercial transaction or of any other business communication'.

3.4.3.2 Information stored in terminal equipment

Article 5 (3) ePD regulates the storage of information, or gaining access to information already stored, in the terminal equipment of a subscriber or user. According to regulatory guidance, terminal equipment must be interpreted broadly, including smart devices such as smartphones, tablets, smart TVs

[601] Article 2 lit a ePrivacy Directive.
[602] Article 1 (2) ePrivacy Directive.
[603] Article 94 GDPR; European Data Protection Board, 'Guidelines 05/2020 on consent under Regulation 2016/679' (4 May 2020) at 6.
[604] Case C-673/17 *Planet 49 GmbH* [2019] ECR I-801 paras 60-65.
[605] Eleni Kosta, Jos Dumortier, 'ePrivacy Directive: Assessment of transposition, effectiveness and compatibility within the proposed Data Protections Regulaton' (2015) European Commission 10 <https://research.tilburguniversity.edu/en/publications/eprivacy-directive-assessment-of-transposition-effectiveness-and-> accessed 8 February 2024.
[606] Article 5 (1) ePrivacy Directive; Tijmen H.A. Wisman, 'Privacy, Data Protection and E-Commerce' in Arno L. Lodder, Andrew D. Murray (eds) *EU regulation of e-commerce. A commentary* (Elgar 2017) 371.
[607] Eleni Kosta, Jos Dumortier, 'ePrivacy Directive: Assessment of transposition, effectiveness and compatibility within the proposed Data Protections Regulaton' (2015) European Commission 11 <https://research.tilburguniversity.edu/en/publications/eprivacy-directive-assessment-of-transposition-effectiveness-and-> accessed 8 February 2024.

etc.[608] Article 5 (3) ePD does not require the processing of personal data.[609] Recital 24 ePD outlines that information stored on terminal equipment is 'part of the private sphere of the users requiring protection' and 'may seriously intrude upon the privacy of these users'. The legislator has amended the ePD in 2009 to make consent a requirement for storage of such information. Thus, for information to be stored in terminal equipment, whether the information constitutes personal data or not, requires the consent of the user or subscriber.[610] If placing and retrieving information through cookies or similar means is also considered to constitute processing of personal data, the GDPR applies in addition to Article 5 (3) ePD.[611] Because the provision in the ePD constitutes a lex specialis, it prevails over the GDPR and thus, consent of the user or subscriber is needed meaning that the controller cannot rely on the full range of possible lawful bases provided by Article 6 GDPR.[612] However, there are two exceptions where prior consent is not required. These are technical storage for the sole purpose of carrying out the transmission of a communication and the provision of an information society service that is explicitly requested by the user or subscriber.[613]

3.4.3.3 Location data

Article 9 ePD governs the processing of location data *other*[614] than traffic data through a public communications network or publicly available ECS.[615] This provision regulates only a fraction of location based services and thus services that are offered to members of a private network are not subject to the ePD. Therefore, Article 9 does not apply to location data transmitted through enterprise networks aimed at a private user group, or data collected and transmitted through infrared signals or GPS signals in combination with a private wireless network.[616] In addition, regulatory guidance states that 'the ePrivacy directive does not apply to the processing of location data by information society services, even when such processing is performed via a public electronic communication network'.[617] As

[608] Article 29 Working Party, 'Opinion 3/2013 on apps on smart devices' (WP 202, 27 February 2013) at 7.
[609] Case C-673/17 *Planet 49 GmbH* [2019] ECR I-801 para 69.
[610] Article 5 (3) ePrivacy Directive as amended by Directive 2009/136/EC.
[611] Article 29 Working Party, 'Opinion 2/2010 on online behavioural advertising' (WP 171, 22 June 2010) at 9.
[612] European Data Protection Board, 'Opinion 5/2019 on the interplay between the ePrivacy Directive and the GDPR, in particular regarding the competence, tasks and powers of data protection authorities' (Opinion 5/2019, 12 March 2019), at 14.
[613] Eleni Kosta, Jos Dumortier, 'ePrivacy Directive: Assessment of transposition, effectiveness and compatibility within the proposed Data Protections Regulaton' (2015) European Commission 13 <https://research.tilburguniversity.edu/en/publications/eprivacy-directive-assessment-of-transposition-effectiveness-and-> accessed 8 February 2024.
[614] Thus, location data which is also traffic data are governed by Article 6 ePrivacy Directive (see Recital 35).
[615] Tijmen H.A. Wisman, 'Privacy, Data Protection and E-Commerce' in Arno L. Lodder, Andrew D. Murray (eds) *EU regulation of e-commerce. A commentary* (Elgar 2017) 376.
[616] Eleni Kosta, Jos Dumortier, 'ePrivacy Directive: Assessment of transposition, effectiveness and compatibility within the proposed Data Protections Regulation' (2015) European Commission 14 <https://research.tilburguniversity.edu/en/publications/eprivacy-directive-assessment-of-transposition-effectiveness-and-> accessed 8 February 2024.
[617] Article 29 Working Party, 'Opinion 13/2011 on Geolocation services on smart mobile devices' (WP 185, 16 May 2011) at 9.

a result, processing location data via techniques such as Wi-Fi network proximity or IP-address databases is not covered by Article 9 ePD.[618]

Article 9 ePD allows processing of location data when they are (i) anonymous or (ii) when the user or subscriber provided consent to the extent and for the duration necessary for the provision of value-added services. Anonymisation is a controversial concept[619] considering that technology is rapidly evolving and thus facilitates better (and quicker) identifiability of individuals.[620] This seems to be particularly relevant since the exception does not restrict the processing of anonymised location data to specific purposes.[621]

3.5 Conclusions

Chapter 3 examined *Subquestion 2:* What is the current EU legal framework? The fundamental rights to privacy and the protection of personal data enshrined in the EUCFR as well as the GDPR and ePD together form the '*current legal framework*'.

The fundamental right to *privacy* according to Article 7 EUCFR protects everyone's 'right to respect for his private and family life, his home and communications'. The fundamental right to *data protection* enshrined in Article 8 EUCFR grants everyone 'the right to the protection of personal data concerning him or her'. It applies to personal data, which entails all information on identified or identifiable natural persons. The two fundamental rights[622] are closely linked, but not identical.[623] Information protected by the fundamental right to data protection seems to be more extensive as opposed to the information covered by the fundamental right to privacy.[624] In addition, the personal scope differs. Legal persons are excluded from the fundamental right to data protection[625] whereas legal persons can rely on the fundamental right to privacy.[626] Both fundamental rights are further

[618] Eleni Kosta, Jos Dumortier, 'ePrivacy Directive: Assessment of transposition, effectiveness and compatibility within the proposed Data Protections Regulation' (2015) European Commission 14 <https://research.tilburguniversity.edu/en/publications/eprivacy-directive-assessment-of-transposition-effectiveness-and-> accessed 8 February 2024.
[619] Tijmen H.A. Wisman, 'Privacy, Data Protection and E-Commerce' in Arno L. Lodder, Andrew D. Murray (eds) *EU regulation of e-commerce. A commentary* (Elgar 2017) 376.
[620] Nadezhda Purtova, 'The law of everything. Broad concept of personal data and future of EU data protection law' (2018) Vol 10 Iss 1 Law, Innovation and Technology 40, 74-75 <https://www.tandfonline.com/doi/full/10.1080/17579961.2018.1452176> accessed 8 February 2024.
[621] As opposed to exception of consent, which only allows processing for the provision of value added services.
[622] Note that the human right to respect for private and family life according to Article 8 ECHR and related ECtHR case law highly influence the interpretation of the two fundamental rights.
[623] Juliane Kokott, Christoph Sobotta 'The distinction between privacy and data protection in the jurisprudence of the CJEU and ECtHR' (2013) Vol 3 No 4 International Data Privacy Law 222, 223, 228; Herke Kranenborg, Commentary of Article 8 in Steve Peers et al (eds), *The EU Charter of Fundamental Rights* (Hart/Beck 2014) 229.
[624] Juliane Kokott, Christoph Sobotta 'The distinction between privacy and data protection in the jurisprudence of the CJEU and ECtHR' (2013) Vol 3 No 4 International Data Privacy Law 222, 225.
[625] Joined Cases C–92/09 and C–93/09 *Volker und Markus Schecke and Eifert* [2010] ECR I-11063, paras 52, 53 and 87.
[626] Juliane Kokott, Christoph Sobotta 'The distinction between privacy and data protection in the jurisprudence of the CJEU and ECtHR' (2013) Vol 3 No 4 International Data Privacy Law 222, 225.

substantiated in EU secondary law. The most relevant legislation in EU secondary law are the GDPR and the ePD.

The *GDPR* is the most comprehensive piece of legislation in data protection law and arguably also the most influential one. It contains rules relating to the protection of natural persons regarding the processing of their personal data, as well as rules aimed at facilitating the free movement of personal data. The most important provisions of the GDPR are the principles contained in Article 5 GDPR, as well as the rights of data subjects enshrined in Chapter III GDPR.

Provisions of the *ePD* aim to 'particularise and complete'[627] the GDPR[628] in the electronic communications sector.[629] Whereas both the GDPR and the ePD have the object of protecting fundamental rights and freedoms, the GDPR sets general rules for the processing of personal data. The ePD regulates the fundamental right to privacy *and* data protection in the electronic communications sector.[630] In accordance with the principle lex specialis derogate legi generali,[631] provisions of the ePD that specifically regulate processing of personal data in the electronic communications sector prevail over the general provisions of the GDPR.[632] The most important provisions of the ePD in light of AI are confidentiality of communications, information stored in terminal equipment and location data.

[627] Article 1 (2) ePrivacy Directive.
[628] Initially Data Protection Directive 95/46/EC.
[629] Tijmen H.A. Wisman, 'Privacy, Data Protection and E-Commerce' in Arno L. Lodder, Andrew D. Murray (eds) *EU regulation of e-commerce. A commentary* (Elgar 2017) 369.
[630] Christina Etteldorf, 'EDPB on the Interplay between the ePrivacy Directive and the GDPR' (2019) Iss 5 No 2 European Data Protection Law Review 224, 226.
[631] Joined Cases *T-60/06 RENV II and T-62/06 RENV II* [2016] ECR II-233 para 81.
[632] European Data Protection Board, 'Opinion 5/2019 on the interplay between the ePrivacy Directive and the GDPR, in particular regarding the competence, tasks and powers of data protection authorities' (Opinion 5/2019, 12 March 2019) at 17.

4 Legal problems: principles

This chapter aims to answer Subquestion 3, namely, what legal problems arise or may arise when the principles enshrined in the current legal framework are applied to AI. First, this chapter introduces three types of legal problems (Section 4.1). Based on this approach, legal problems are identified for each AI discipline outlined in Chapter 2 (i.e. machine learning, computer vision, natural language processing, affective computing and automated reasoning). This chapter focusses on the *principles* enshrined in the current legal framework. Sections 4.2 to 4.8 deal with the principles enshrined in the GDPR, namely, the principles of lawfulness (Section 4.2), fairness (Section 4.3), transparency (Section 4.4), purpose limitation (Section 4.5), data minimisation (Section 4.6) and accuracy (Section 4.7), as well as the principle of enhancing protection for special categories of personal data (Section 4.8).[633] Section 4.9 elaborates on the requirements with respect to the confidentiality of communication, which is regarded as a principle in a broader sense for the purpose of this thesis. Finally, Section 4.10 concludes by providing an answer to Subquestion 3, including an overview of which AI disciplines lead to which types of legal problems. Whereas AI systems may be deployed by both governmental and private actors, I focus on the latter.

4.1 Approach

When referring to legal problems, three types of legal problems are distinguished (Table 4.1).

Type	Description
1	Legal provisions are violated
2	Legal provisions cannot be enforced
3	Legal provisions are not fit for purpose to protect the fundamental right at stake

Table 4.1 Three types of legal problems.

Let me briefly explain the need to investigate these three types of legal problems in particular. Both the right to privacy and data protection are fundamental rights in the EU.[634] Violations of fundamental rights, which constitute Type 1 legal problems, must be prevented. For example, unsupervised ML approaches process personal data for inexplicit purposes – the processing itself determines the purpose and future use of the personal data. Such processing violates the purpose limitation principle, which constitutes a Type 1 legal problem. Type 2 legal problems, namely, when legal provisions cannot be enforced, are not acceptable either because they lead to negative consequences for the de facto protection of fundamental rights. For example, the unclear substantive meaning of the fairness principle reduces legal certainty and makes it less likely that this principle will be enforced by means

[633] Admittedly, this is not a traditional data protection principle. Nonetheless, it could be regarded as a principle in a broader sense, which then also aligns with the approach taken in this chapter.
[634] Article 7 and 8 CFREU.

of private litigation and by supervisory authorities, which leads to Type 2 problems. Furthermore, Type 3 legal problems, namely, legal provisions that are not fit for purpose, point to the shortcomings of the current legal framework. Legal provisions are not fit for purpose, for instance, when they fail to achieve legislative aims, are not effective or create a gap of protection. For example, the principle that special categories of personal data receive enhanced protection and the legislator's approach to exhaustively enumerate special data cause a Type 3 legal problem. This approach does not keep up with technological developments facilitated by AI. It leads to significant gaps of protection, for example, regarding the processing of new types of sensitive personal data generated by AI, such as emotion data, neurodata and mental data. Insights about this type of legal problems are essential when considering how the legal problems should be addressed, which is the aim of Subquestion 5 (see Chapter 6).

As indicated in Section 1.4, the scope of this thesis is limited to legal problems related to the fundamental rights to privacy and data protection. Thus, this chapter identifies legal problems arising primarily from the perspective of natural persons. Obviously, violations of provisions enshrined in the current legal framework (Type 1) constitute a problem for the natural persons concerned. However, legal problems related to enforcement (Type 2) are not exclusively problematic for natural persons. They also directly concern the competent supervisory authority (SA) tasked with the regulatory enforcement of the provisions enshrined in the current legal framework.[635] When the competent SA is unable to pursue regulatory enforcement, this is not only problematic for the SA itself, but also for the natural persons concerned as they have, in the case of the GDPR, a right to lodge a complaint with a SA.[636] The SA then must handle the complaint and adopt corresponding enforcement measures. Where the complaint lodged by the natural person concerns a substantively unclear provision enshrined in the current legal framework, the SA will not be able to pursue regulatory enforcement. This is problematic for both the SA and the natural person concerned. Type 3 legal problems are discussed from the perspective of *natural persons* as the primary subject of protection envisaged by fundamental rights. These types of legal problem are identified by means of the rationales and specific aims pursued by the current legal framework relevant to natural persons. Table 4.2 lists the rationales and specific objectives[637] enshrined in the current legal framework that are relevant to natural persons. The table only mentions secondary EU law because the fundamental rights to privacy and data protection enshrined in the EU Charter of Fundamental Rights (EUCFR) are less likely to cause Type 3 legal problems due to the flexibility of these rights and the living instrument doctrine adopted by the ECtHR (see also Sections 4.10 and 5.12).

[635] For instance, Supervisory Authorities that have to enforce the GDPR as described in Article 57 GDPR.
[636] Article 77 GDPR.
[637] Expressed in the form of Recitals. For an in depth discussion see Gloria González Fuster, 'Study on the essence of the fundamental rights to privacy and to protection of personal data' (2022) <https://edps.europa.eu/system/files/2023-11/study_en.pdf> accessed 8 February 2024; Gloria González Fuster, *The Emergence of Personal Data Protection as a Fundamental Right of the EU* (Springer International 2014).

GDPR
The protection of natural persons regarding their fundamental right to data protection (Recital 1)
The protection of personal data (Recital 4)
Respect for the fundamental right to privacy (Recital 4)
Processing of personal data to serve mankind (Recital 4)
Consistent and high level of protection for personal data (Recitals 6, 10)
Strong and coherent data protection framework (Recital 7)
Control for data subject over the processing of their own personal data (Recitals 7, 68)
Enhancement of legal and practical certainty for data subjects (Recital 7)
Effective protection and strengthening the rights of data subjects (Recital 11)
Same level of legally enforceable rights (Recital 13)
ePD
Full respect for the fundamental rights to privacy and data protection (Recital 2)
Guaranteeing the confidentiality of communications (Recital 3)
Protection of personal data and the privacy of the user (Recital 5)
Protection of users from risks for their personal data and privacy posed by the Internet and ECS (Recital 6)
Protection of natural persons with respect to automated storage and processing of data (Recital 7)

Table 4.2 Legislative aims pursued by EU secondary law relevant to natural persons. As indicated by Article 1 (2) GDPR, the latter's main goal is to protect the fundamental right to data protection (Article 8 EUCFR).

I do acknowledge that the rationales and specific objectives listed in Table 4.2 are to some extent arbitrary, as they solely focus on the perspective of *natural persons* as the primary subject of protection envisaged by the two fundamental rights I discuss in this thesis. However, neither the fundamental right to privacy nor the fundamental right to data protection are absolute rights. Recital 4 GDPR emphasises that the fundamental right to data protection is not an absolute right, and it must be balanced against other fundamental rights and freedoms. In its case law, also the CJEU stresses the character of this fundamental right is not absolute.[638] Interests and rights of controllers explicitly mentioned in the GDPR's recials[639] are, for instance, the freedom to conduct a business (Article 16 EUCFR), trade secrets that may be protected by the fundamental right to property (Article 17 EUCFR)[640] or intellectual property rights. Hence, Table 4.2 should not be understood as an arbitrary list. It merely contains the rationales of EU secondary law aimed at protecting natural persons in line with the focus and limitations of this thesis (see Section 1.4). Nonetheless, I do take the non-absolute nature of the fundamental right to data protection into account, which becomes particularly apparent

[638] Case C-268/21 *Norra Stockholm Bygg AB* [2023] ECR I-145 para 49; Case C-460/20, *TU* [2022] ECR I-962 para 56; Case C-136/17, *GC and Others* [2019] ECR I-773 para 57.
[639] Recitals 4, 63 GDPR.
[640] Case C-1/11 *Interseroh Scrap* [2012] ECR I-194 para 43; Case T-189/14 *Deza* [2017] para 163.

when discussing legal problems (e.g., Sections 4.2.1, 4.3.1 and 5.6.2). I also consider fundamental rights and freedoms of controllers when suggesting solutions to the legal problems identified (e.g., Sections 6.5.2 and 6.6.2).

As indicated in Sections 1.1 and 1.4, I focus on horizontal relationships. Concerning the fundamental right to data protection, I mostly elaborate on the GDPR when discussing legal problems. Article 1 (2) GDPR reveals the primary goal of this piece of EU secondary law: protecting the fundamental right to data protection according to Article 8 EUCFR. The CJEU emphasises the latter: the GDPR aims to ensure a high level of protection 'of the rights guaranteed in Article 16 TFEU and *Article 8 of the Charter*'.[641] In this sense, the GDPR 'implements'[642] this fundamental right within the realm of horizontal relationships. Whereas the primary goal of the GDPR is to guarantee the fundamental right to data protection,[643] the GDPR contains several more fine-grained objectives, as illustrated in table 4.2. For type 3 legal problems, I use these objectives to assess whether the principles contained in the GDPR are fit for purpose to protect the fundamental right to data protection[644] guaranteed by Article 8 EUCFR.

The three types of legal problems are not mutually exclusive. For example, a Type 2 legal problem may also constitute a Type 3 problem. For example, despite its role as a key tenet in EU data protection law, the substantive meaning of the fairness principle remains largely elusive, meaning it is hard to enforce (i.e. Type 2). At the same time, the fairness principle is *currently*[645] not fit for purpose (i.e. Type 3) to protect the fundamental right to data protection – a substantively unclear principle cannot ensure a high level of the protection of personal data as envisaged in EU data protection law. These three legal problems may be caused by one or more AI disciplines, as described in Chapter 2. Legal problems may be very specific to only one AI discipline or may be more general and relate to several AI disciplines. The latter applies where a provision enshrined in the current legal framework is substantively unclear (e.g., the fairness principle), which causes legal problems regardless of which discipline of AI it is applied to. Also, note that violations of the principles enshrined in the GDPR simultaneously violate the accountability principle introduced in Section 3.3.3.10. According to the

[641] Case C-645/19 [2021] *Facebook Ireland* ECR I-483 para 45 emphasis added by the author; see also Case C-319/20, *Meta Platforms Ireland Limited* [2022] ECR I-322 para 73.
[642] Article 1 (2) GDPR reveals the main objective of said regulation: to give meaning to this fundamental right. See Hielke Hijmans, Commentary of Article 1 in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020) 56.
[643] Article 1 (2) GDPR.
[644] It could be argued that the GDPR does not protect personal data but rather natural persons. It is apparent from Article 1 (1) that the GDPR protects natural persons. This also follows clearly from the concept of personal data. Protecting personal data as intended by the GDPR (Article 1 Recitals 1, 2, 4, 6, 9, 11, 89 GDPR) indispensably protects natural persons as only information relating to a natural person constitutes personal data.
[645] It is predominantly interpreted as procedural fairness. The fairness principle might be fit for purpose when substantive fairness is added to the current interpretation. See Section 6.2.

accountability principle, controllers are i) responsible for compliance and ii) must be able to demonstrate compliance with *all the principles* mentioned in Article 5 (1) GDPR.[646]

In some cases, it can be difficult to map the legal problems one-on-one with the different AI disciplines, as well as with all provisions contained in the legal framework discussed in Chapter 3. Therefore, I focus on principles enshrined in the legal framework outlined in Chapter 3 (i.e. lawfulness and proportionality, fairness, transparency, purpose limitation, data minimisation, accuracy, the principle that special categories of personal data receive enhanced protection and the principle concerning the confidentiality of communications). Principles form the basis of the fundamental right to data protection[647] and the legislator considers the infringement of principles as more *serious* than infringements of other provisions.[648] The principle of confidentiality contained in the ePD is the key principle ensuring the confidentiality of communications as protected by the fundamental right to privacy. I do not discuss the principle of integrity and confidentiality according to Article 5 (1) lit f GDPR.[649] I also skip the principle of storage limitation according to Article 5 (1) lit e GDPR because it is not particularly relevant in the context of AI.

To determine which type of legal problem arises or may arise due to the different AI disciplines, as outlined in Chapter 2, the AI disciplines are mapped with the principles contained in the current legal framework. For each principle enshrined in the current legal framework, I assess whether the principle at hand creates Type 1, 2 or 3 legal problems. When doing so, I follow the order of the AI disciplines outlined in Chapter 2.

AI refers to adaptive machines that can autonomously execute activities and tasks that require capabilities usually associated with humans. Although AI could make its *own* decisions and perform tasks on the designer's behalf,[650] AI does not have a legal personality. Thus, AI cannot itself cause the three types of legal problems discussed in this thesis. Instead, these legal problems occur when companies use AI. Hence, when concluding that AI causes legal problems, I always refer to the deployment of AI by companies. To unveil the legal problems, I rely on Chapter 2, which explains the different AI disciplines and how they work from a technological and conceptual perspective.

[646] Cécile de Terwangne, Commentary of Article 5 in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020) 311.
[647] Cécile de Terwangne, Commentary of Article 5 in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020) 313.
[648] Article 29 Working Party, 'Guidelines on the application of administrative fines for the purposes of Regulation 2016/679' (WP 253, 3 October 2017) 9; European Data Protection Board, 'Guidelines on the calculation of administrative fines under the GDPR' (Guidelines 4/2022, 16 May 2022) 16.
[649] For the reasons outlined in Section 3.3.3.8.
[650] Eduardo Alonso, 'Actions and agents' in Frankish Keith and Ramsey William M. (eds) *The Cambridge Handbook of Artificial Intelligence* (2014) 235, 236.

4.2 Lawfulness

As outlined in Section 3.3.3.1, lawfulness essentially requires that processing respects all applicable legal requirements[651] and is further substantiated in Article 6 GDPR. Processing is only lawful if at least one of the lawful bases listed in Article 6 GDPR applies, for example, consent of the data subject (lit a), performance of or entering into a contract (lit b) or the legitimate interest pursued by the controller or third party (lit f).[652] In addition, the principle of lawfulness connotes proportionality in the balancing of interests of data subjects and controllers.[653] Thus, the principle of lawfulness is closely linked to the principle of proportionality, which is one of the general principles of EU law[654] and has decisive influence on the assessment of whether a violation of a person's right to data protection is justified.[655] Thus, as already outlined in Section 3.2.2, the principle of proportionality plays an important role in EU data protection law.[656] It has generally three components which involve the assessment of a measure's (i) suitability, (ii) necessity and (iii) proportionality stricto sensu.[657] When the principle of lawfulness (and proportionality) is applied to the AI disciplines introduced in Chapter 2, Type 1 legal problems may occur.

4.2.1 Legal problems: Type 1

As explained in Section 2.1, AI refers to adaptive machines that can autonomously execute activities and tasks that require capabilities usually associated with humans. However, the GDPR does not apply to AI as such because AI does not have a legal personality. Instead, the GDPR applies to controllers and processors deploying AI systems that process personal data. Due to its autonomous and adaptive characteristics, AI has the potential to decide why and how to process personal data. With this, I do not suggest that AI systems currently can act as controllers under data protection law by determining the purposes and means as well as the legal ground for processing. Instead, I refer to the possibility that the use of AI by controllers might violate the principle of lawfulness due to the current reasoning deficiencies in the AI discipline of automated reasoning. I will illustrate this through the legal ground of legitimate interest and the deployment of unsupervised machine learning.

When the processing of personal data is based on the legal ground of the legitimate interest of the controller, the latter has to perform a Legitimate Interest Assessment (LIA).[658] This LIA requires assessing the impact of processing on the fundamental rights and freedoms of the data subject by

[651] Cécile de Terwangne, Commentary of Article 5 in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020) 314.
[652] Article 6 (1) lit a) to f) GDPR.
[653] Lee A. Bygrave, *Data Privacy Law: An International Perspective* (OUP 2014) 148.
[654] Article 5 of the consolidated version of the Treaty Establishing the European Community [2006] OJ C321E/37.
[655] Charlotte Bagger Tranberg, 'Proportionality and data protection in the case law of the European Court of Justice' (2011) Vol 1 No 4 International Data Privacy Law 239-249.
[656] Ibid.
[657] Ibid.
[658] As required by Article 6 (1) lit f GDPR

considering the nature of personal data, the way in which the information is being processed, the reasonable expectations of the data subjects, and the status of the controller and the data subject.[659] It also includes the controller's obligation to consider the proportionality of processing.[660] Before implementing an AI system, the controller needs to perform a LIA[661] and determine the input and training data to be used by the AI system. However, the AI system should be able to perform a LIA if it deploys unsupervised ML. Unsupervised ML approaches process data for *in*explicit purposes – the processing *itself* determines the purpose since its goal is to detect patterns and correlations, gain knowledge, and make accurate predictions. Also, the purpose may alter given that algorithms used in AI learn and develop over time[662] (see also Section 4.5.1). The performance of an LIA is inextricably linked to the purpose of processing because it must be assessed whether the purpose serves a legitimate interest of the controller.[663] However, in the case of unsupervised ML, the specific purpose for processing is not necessarily known in advance.

Current AI systems have been called clueless[664] to understand cause and effect and devoid of common sense.[665] The lack of progress in providing general automated common sense reasoning capabilities underscores that this is a very difficult problem in the field of AI.[666] Common sense reasoning is not just the hardest problem for AI, it is also considered to be the most important problem.[667] It seems that humans are much better then machines in this context[668] and therefore, common sense reasoning still constitutes a challenge in AI,[669] and particularly in automated reasoning (see Section 2.2.5). Apparently, there is no AI system today that has a semblance of common sense or has capabilities such as human cognition. Hence, AI systems are unable to think in a manner on par with human thinking[670] and may therefore not be capable (at least not in the near future) of appropriately weighing the

[659] Art 29 Working Party, 'Opinion 06/2014 on the notion of legitimate interests of the data controller under Article 7 of Directive 95/46/EC' (WP 217, 9 April 2014) at 36.
[660] Ibid at 33.
[661] If processing should occur based on the controller's legitimate interest.
[662] Norwegian Data Protection Authority, 'Artificial Intelligence and Privacy' (2018) 4 <https://www.datatilsynet.no/globalassets/global/english/ai-and-privacy.pdf> accessed 8 February 2024.
[663] Art 29 Working Party, 'Opinion 06/2014 on the notion of legitimate interests of the data controller under Article 7 of Directive 95/46/EC' (WP 217, 9 April 2014) at 24.
[664] Brian Bergstein, 'What AI still can't do' MIT Technology Review (Cambridge 31 January 2020) <https://www.technologyreview.com/2020/01/31/304844/ai-common-sense-reads-human-language-ai2/> accessed 8 February 2024.
[665] Cade Metz, 'Paul Allen Wants to Teach Machines Common Sense' *The New York Times* (New York, 28 February 2018) <https://www.nytimes.com/2018/02/28/technology/paul-allen-ai-common-sense.html> accessed 8 February 2024.
[666] Brandon Bennet, Anthony G Cohn, 'Automated Common-sense Spatial Reasoning: Still a Hughe Challenge' in Stephen Muggleton, Nicholas Chater (eds) Human-Like Machine Intelligence (Oxford University Press 2021) 405.
[667] Gary Marcus, Ernest Davis, *Rebooting AI: Building Artificial Intelligence we can trust* (Pantheon Books 2019).
[668] Davide Castelvecchi, 'AI pioneer: The dangers of abuse are very real' *Nature* (London, 4 April 2019) < https://www.nature.com/articles/d41586-019-00505-2> accessed 8 February 2024.
[669] Shoham Yoav et al, 'The AI Index 2018 Annual Report' (AI Index Steering Committee Stanford University 2018) 64 <https://hai.stanford.edu/sites/default/files/2020-10/AI_Index_2018_Annual_Report.pdf > accessed 8 February 2024.
[670] Lance Eliot, 'AI Ethics And The Quagmire Of Whether You Have A Legal Right To Know Of AI Inferences About You, Including Those Via AI-Based Self-Driving Cars' *Forbes* (New York, 25 May 2022) < https://www-forbes-com.cdn.ampproject.org/c/s/www.forbes.com/sites/lanceeliot/2022/05/25/ai-ethics-and-the-quagmire-of-whether-you-have-a-legal-right-to-know-of-ai-inferences-about-you-including-those-via-ai-based-self-driving-cars/amp/> accessed 8 February 2024.

fundamental rights and freedoms of the parties involved and implement the factors which must be considered according to the LIA.

This holds particularly true because the CJEU has been criticised for shortcomings in identifying the various elements that need to be balanced when assessing the proportionality of data processing based on a controller's legitimate interest and the data subject's right to data protection.[671] Indeed, the early practice of the CJEU concerning the interpretation of data protection law and the proportionality principle often left it up to the national laws, authorities and courts to carry out any concrete proportionality testing.[672] It can be said that the proportionality test is, cognitively, a difficult task due to the lack of clear elements that need to be considered within this assessment. Additionally, there seems to be a lack of concrete proportionality tests performed by the CJEU that could serve as training data for AI to learn and extract the logic of such balancing tests. As is the case with the purpose limitation and data minimisation principle,[673] computer scientists would need measurable definitions of the proportionality principle and concrete indications of how to practically and concretely implement its requirements.

In fact, the accountability principle, which is substantiated in Article 24 GDPR, requires controllers to 'implement appropriate and effective measures to ensure and be able to demonstrate' that personal data processing occurs in accordance with the rules set out in the GDPR.[674] Violations of the lawfulness principle simultaneously violate the accountability principle as introduced in Section 3.3.3.10 because controllers must be able to demonstrate compliance with *all the principles* mentioned in Article 5 (1) GDPR.[675]

The balancing problem (Type 1)

Due to the reasoning deficiencies in the AI discipline AR combined with the lack of computable requirements concerning the proportionality principle, AI systems that autonomously process personal data cannot appropriately balance the fundamental rights and freedoms and assess the proportionality of processing as required by Article 6 (1) lit f GDPR. Such processing violates both the lawfulness and proportionality principle.

[671] Audrey Guinchard, 'Taking proportioanality seriously: The use of contextual integrity for a more informed and transparent analysis in EU data protection law' (2018) Vol 24 Iss 6 European Law Journal 434, 435. For references to such criticism see footnote 5 and 6 in the latter publication.
[672] Charlotte Bagger Tranberg, 'Proportionality and data protection in the case law of the European Court of Justice' (2011) Vol 1 No 4 International Data Privacy Law 239, 242.
[673] Asia Biega, Michèle Finck, 'Reviving Purpose Limitation and Data Minimisation in Data-Driven Systems' (2021) Technology and Regulation 44, 58 <https://arxiv.org/abs/2101.06203> accessed 8 February 2024.
[674] Art. 24 (1), Recital 74 GDPR.
[675] Cécile de Terwangne, Commentary of Article 5 in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020) 311.

4.2.2 Legal problems: Type 2

When the lawfulness principle is applied to the AI disciplines introduced in Chapter 2, no specific Type 2 legal problems arise. This is mainly due to the reason that the lawfulness principle is substantively clear, as is further substantiated in Article 6 GDPR, which exhaustively enumerates six lawful bases that can be relied upon for the processing of personal data. Nevertheless, the CJEU has been criticised for shortcomings in identifying the various elements that need to be balanced when assessing the proportionality of processing.[676] These shortcomings could lead to Type 2 legal problems because substantively unclear principles are difficult to enforce. However, this problem arises regardless of whether the processing involves AI and thus does not relate specifically to AI. Therefore, I refrain from discussing this problem further.

4.2.3 Legal problems: Type 3

Similar to what I have outlined in Section 4.2.2, no specific Type 3 legal problems arises when the lawfulness is applied to AI, mainly because this principle is substantively clear from a legal point of view. It may be argued that the proportionality principle is not fit for purpose to protect the fundamental right to data protection due to the lack of clarity in terms of the various elements that need to be balanced. Likewise, it is questionable whether consent is a suitable concept to prevent the data subject from harm relating to the processing of personal data. However, these are general issues and therefore not specifically related to AI. Therefore, it will not be discussed further.

4.3 Fairness

The AI disciplines outlined in Section 2.2 create legal problems when applied to the fairness principle introduced in Section 3.3.3.2.[677] In academia, scholars seem to distinguish between two different types of fairness. According to Graef, Clifford and Valcke, *procedural fairness* in data protection law refers to formal or process-oriented requirements.[678] In the view of De Terwangne, procedural fairness considers whether or not the data involved have been obtained nor otherwise processed through unfair means, by deception or without the knowledge of the individual concerned.[679] Malgieri adds *substantive fairness* aiming to prevent adverse effects in concrete circumstances, in particular when

[676] Audrey Guinchard, 'Taking proportionality seriously: The use of contextual integrity for a more informed and transparent analysis in EU data protection law' (2018) Vol 24 Iss 6 European Law Journal 434, 435. For references to such criticism, see Footnotes 5 and 6 in the latter publication.
[677] Parts of Section 4.3 and Section 6.2 resulted in a publication see Andreas Häuselmann, Bart Custers, 'Substantive fairness in the GDPR: Fairness Elements for Article 5.1a GDPR' (2024) Vol 52 Computer Law & Security Review 105942.
[678] Inge Graef, Damien Clifford, Peggy Valcke, 'Fairness and enforcement: bridging competition, data protection, and consumer law' (2018) Vol 8 No 3 International Data Privacy Law 200, 203.
[679] Cecile de Terwangne, Commentary of Article 5 in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020) 314.

conflicting interests need to be balanced.[680] However, as pointed out in Section 3.3.3.2, the role and meaning of the fairness principle in data protection law remains elusive despite the fact that it is considered to be a key tenet of EU data protection law.[681] In addition, the CJEU has never defined the fairness principle nor the notion of fairness in data protection law.[682] Dictionaries define the term 'fairness' as 'impartial or just treatment or behaviour without favouritism'[683] or as 'the quality of treating people equally or in a way that is right or reasonable'.[684] Both regulatory guidance[685] and regulatory enforcement on EU level in the form binding decisions[686] adopted by the European Data Protection Board (EDPB)[687] identify key elements of the fairness principle. These key elements are: autonomy of data subjects with respect to data processing, their reasonable expectations, ensuring power balance between controllers and data subjects, avoidance of deception as well as possible adverse consequences of processing and ensuring ethical and truthful processing.[688] Despite the close and evident link[689] with the transparency and lawfulness principle, the fairness principle should be interpreted as having an independent meaning going[690] beyond transparency and lawfulness.[691]

Substantive fairness focusses on the adverse effects for data subjects caused by the processing of personal data and also considers the substantial circumstances and interests at stake: expectations of data subjects, effects on them, and the actual interests of the parties involved. Hence, it aims to mitigate unfair imbalances among interests of controllers and data subjects[692] which seems to be more

[680] Gianclaudio Malgieri, 'The concept of Fairness in the GDPR' (FAT* '20: Conference on Fairness, Accountability, and Transparency, Barcelona, January 2020) 2, 3 <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3517264> accessed 8 February 2024.
[681] Damian Clifford, Jef Ausloos 'Data Protection and the Role of Fairness' (2018) Vol 37 No 1 Yearbook of European Law 130, 187.
[682] Gianclaudio Malgieri, 'The concept of Fairness in the GDPR' (FAT* '20: Conference on Fairness, Accountability, and Transparency, Barcelona, January 2020) 6 <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3517264> accessed 8 February 2024.
[683] See < https://www.oxfordlearnersdictionaries.com/definition/english/fairness?q=fairness > accessed 8 February 2024.
[684] See <http://dictionary.cambridge.org/dictionary/english/fairness> accessed 8 February 2024.
[685] European Data Protection Board, 'Guidelines on Article 6(1)(b) GDPR' (Guidelines 2/2019, 8 October 2019), at 6; European Data Protection Board, 'Guidelines on Article 25 Data Protection by Design and Default' (Guidelines 4/2019, 20 October 2020), at 17 and 18.
[686] Article 65 GDPR.
[687] The EDPB consists of representatives of national EU Supervisory Authorities (SAs) responsible for data protection and the European Data Protection Supervisor (EDPS).
[688] Binding Decision 3/2022 on the dispute submitted by the Irish SA on Meta Platforms Ireland Limited and its Facebook service (Art. 65 GDPR), adopted on 5 December 2022 paras 103, 219-220, 222-223, 226 478; Binding Decision 4/2022 on the dispute submitted by the Irish SA on Meta Platforms Ireland Limited and its Instagram service (Art. 65 GDPR), adopted on 5 December 2022 paras 106, 223-224, 226-227, 445; Binding Decision 5/2022 on the dispute submitted by the Irish SA on Meta Platforms Ireland Limited and its WhatsApp service (Art. 65 GDPR), adopted on 5 December 2022.
[689] Article 5 (1) lit a GDPR mentions the three different principles together.
[690] Binding Decision 3/2022 on the dispute submitted by the Irish SA on Meta Platforms Ireland Limited and its Facebook service (Art. 65 GDPR), adopted on 5 December 2022 paras 220, 477; Binding Decision 4/2022 on the dispute submitted by the Irish SA on Meta Platforms Ireland Limited and its Instagram service (Art. 65 GDPR), adopted on 5 December 2022 paras 224, 444; Binding Decision 5/2022 on the dispute submitted by the Irish SA on Meta Platforms Ireland Limited and its WhatsApp service (Art. 65 GDPR), adopted on 5 December 2022.
[691] Winston J Maxwell, 'Principle-based regulation of personal data: the case of 'fair processing' (2015) Vol 5 No 3 International Data Privacy Law 205, 208.
[692] Gianclaudio Malgieri, 'The concept of Fairness in the GDPR' (FAT* '20: Conference on Fairness, Accountability, and Transparency, Barcelona, January 2020) 10 <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3517264> accessed 8 February 2024; Thomas Wischmeyer, 'Artificial Intelligence and Transparency: Opening the Black Box' in Thomas Wischmeyer and Timo Rademacher (eds) *Regulating Artificial Intelligence* (Springer Nature 2020) 78.

helpful when compared to procedural fairness focussing on fair ways of obtaining personal data. Substantive fairness also relates to the proportionality principle discussed in Sections 3.2.2 and 4.2.1 which requires controllers to balance the interests at hand and aims to limit the impact for the data subject caused by the processing of personal data. The CJEU uses fairness as an interpretative tool in order to balance the different interests at hand.[693] A fair balance requires specific consideration of the substantial circumstances and interests at issue.[694] The CJEU stresses the particular consideration of the data subject's interests: 'that balance may however depend, in specific cases, on the nature of the information in question and its sensitivity for the data subject's private life'.[695] Both regulatory guidance[696] and regulatory enforcement on EU level[697] point to substantive fairness by mentioning *reasonable expectations* of the data subjects, possible *adverse consequences* of processing and effects of *power imbalance* as some of the key elements of the fairness principle.

Admittedly, the following analysis of the fairness principle in Sections 4.3.1- 4.3.3 might appear quite pessimistic. This is mainly due to the *current* elusiveness surrounding this principle. I explicitly use 'current' because the fairness principle has significant potential to contribute to effective protection for individuals in the context of processing related to AI *if* interpreted substantively. Principles are open norms that allow judges to adjust the law to changing circumstances and to address contemporary problems. As open norms, principles are well suited to recalibrate data protection legislation to changing technological circumstances for achieving the goals set out by the fundamental right to data protection, including legislative goals pursued by the GDPR.[698] The fairness principle's broad scope and open texture[699] make it a suitable candidate to host normative parameters beyond transparency.[700] In Section 6.2, I discuss the fairness principle's potential to contribute to effective protection for individuals by focussing on substantive fairness.

[693] Case C-275/06 *Promusicae* [2008] ECR I-00271 paras 68, 70; Joined Cases C-92/09 and C-93/09, *Schecke* [2010] ECR I-662 para 88; Gianclaudio Malgieri, 'The concept of Fairness in the GDPR' (FAT* '20: Conference on Fairness, Accountability, and Transparency, Barcelona, January 2020) 10 <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3517264> accessed 8 February 2024.
[694] Gianclaudio Malgieri, 'The concept of Fairness in the GDPR' (FAT* '20: Conference on Fairness, Accountability, and Transparency, Barcelona, January 2020) 6 <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3517264> accessed 8 February 2024.
[695] Case C-131/12, *Google Spain* [2014] ECR I-317 para 81.
[696] European Data Protection Board, 'Guidelines on Article 6(1)(b) GDPR' (Guidelines 2/2019, 8 October 2019), at 6.
[697] Binding Decision 3/2022 on the dispute submitted by the Irish SA on Meta Platforms Ireland Limited and its Facebook service (Art. 65 GDPR), adopted on 5 December 2022 paras 219-220; Binding Decision 4/2022 on the dispute submitted by the Irish SA on Meta Platforms Ireland Limited and its Instagram service (Art. 65 GDPR), adopted on 5 December 2022 paras 223-224, 226; Binding Decision 5/2022 on the dispute submitted by the Irish SA on Meta Platforms Ireland Limited and its WhatsApp service (Art. 65 GDPR), adopted on 5 December 2022.
[698] For example a consistent and high level of protection for personal data (recitals 6 and 10), a strong and coherent data protection framework (recital 7) and effective protection (recital 11) GDPR.
[699] Lee A Bygrave, 'Minding the machine v2.0: The EU General Data Protection Regulation and Autmated Decision Making' in Karen Yeung and Martin Lodge (eds) *Algorithmic Regulation* (OUP 2019) 260.
[700] Lee A Bygrave, 'Machine Learning, Cognitive Sovereignty and Data Protection Rights with Respect to Automated Decisions' (2021) University of Oslo Faculty of Law Legal Studies Research Paper Series No. 202-35 at 22, 23 <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3721118> accessed 8 February 2024.

4.3.1 Legal problems: Type 1

AI increasingly contributes to automated decision-making (ADM). Whereas humans have been conditioned to look for causes ('why'), AI algorithms focus on correlations and probabilities ('what').[701] Current AI systems have been called to be clueless[702] to understand cause and effect and to be devoid of common sense.[703] It seems that humans are much better then machines in this context.[704] Common sense reasoning still constitutes a challenge in AI applications.[705] Apparently, there is not one AI system today which has a semblance of common sense comparable to humans. Hence, AI is unable to think in a manner on par with human thinking[706] which is underscored by the shortcomings in automated reasoning as outlined in Section 2.2.5. The lack of progress in providing general automated common sense reasoning capabilities underscores that this is a very difficult problem in the field of AI.[707] Common sense reasoning is not just the hardest problem for AI, it is also considered to be the most important problem.[708]

As outlined in Section 2.2, the term 'learning' in the context of ML does not mean 'understanding', but is about making computers modify or adapt their actions based on experience so that these actions are more accurate.[709] One of the basic skills of ML is generalisation. Generalisation, however, does not go beyond correlation and neglects reason and drawing distinctions. The AI Index acknowledges that common sense reasoning capabilities and deep natural language understanding are still a challenge in AI applications.[710] Probabilistic predictions and generalisation in the context of ML raise concerns regarding the fairness principle. It seems questionable whether ADM and automated predictions based on ML are fair for the data subjects when the algorithms *generalise* but do not *distinguish*.

[701] Viktor Mayer-Schönberger; Kenneth Cukier, *Big Data: A Revolution That Will Transform How We Live, Work and Think* (Houghton Mifflin Harcourt 2013) 14, 18.
[702] Brian Bergstein, 'What AI still can't do' MIT Technology Review (Cambridge 31 January 2020) <https://www.technologyreview.com/2020/01/31/304844/ai-common-sense-reads-human-language-ai2/> accessed 8 February 2024.
[703] Cade Metz, 'Paul Allen Wants to Teach Machines Common Sense' *The New York Times* (New York, 28 February 2018) <https://www.nytimes.com/2018/02/28/technology/paul-allen-ai-common-sense.html> accessed 8 February 2024.
[704] Davide Castelvecchi, 'AI pioneer: The dangers of abuse are very real' *Nature* (London, 4 April 2019) < https://www.nature.com/articles/d41586-019-00505-2> accessed 8 February 2024.
[705] Shoham Yoav et al, 'The AI Index 2018 Annual Report' (AI Index Steering Committee Stanford University 2018) 64 <https://hai.stanford.edu/sites/default/files/2020-10/AI_Index_2018_Annual_Report.pdf > accessed 8 February 2024.
[706] Lance Eliot, 'AI Ethics And The Quagmire Of Whether You Have A Legal Right To Know Of AI Inferences About You, Including Those Via AI-Based Self-Driving Cars' *Forbes* (New York, 25 May 2022) < https://www-forbes-com.cdn.ampproject.org/c/s/www.forbes.com/sites/lanceeliot/2022/05/25/ai-ethics-and-the-quagmire-of-whether-you-have-a-legal-right-to-know-of-ai-inferences-about-you-including-those-via-ai-based-self-driving-cars/amp/> accessed 8 February 2024.
[707] Brandon Bennet, Anthony G Cohn, 'Automated Common-sense Spatial Reasoning: Still a Hughe Challenge' in Stephen Muggleton, Nicholas Chater (eds) Human-Like Machine Intelligence (Oxford University Press 2021) 405.
[708] Gary Marcus, Ernest Davis, *Rebooting AI: Buidling Artificial Intelligence we can trust* (Pantheon Books 2019).
[709] Steven Marsland, *Machine Learning: An Algorithmic Perspective* (2nd edn Chapman & Hall 2015) ch 1.2.1.
[710] Shoham Yoav et al, 'The AI Index 2018 Annual Report' (AI Index Steering Committee Stanford University 2018) 64 <https://hai.stanford.edu/sites/default/files/2020-10/AI_Index_2018_Annual_Report.pdf > accessed 8 February 2024.

Lack of reasoning capabilities can lead to unfair decisions, and ADM based on ML can even be discriminatory. For example, the Google AI system developed to recognise child abuse wrongfully classified a father as criminal. Because his toddler had an infection on his genitals, the father took a photo, displaying himself and the infected part of the toddler's body, as advised by a nurse who said such a photo is necessary for the doctor in order to prepare for the corresponding emergency online consultation.[711] This example clearly points to the problem that ML, which is used by Google in this particular AI system, generalises but does not distinguish. ML does not understand what it classifies as 'wrong' or 'right' and neglects the context of a given picture. In this case, this wrongful classification as child abuser had severe consequences for the individual in question. The police opened an investigation and issued search warrants served on Google and his Internet service provider. Furthermore, Google disabled the account of the father, who lost all his emails, contact information and his Google Fi account, meaning he had to obtain a new phone number with another provider.[712] Thus, the wrongful and fully automated classification as a criminal (child abuser) had adverse and detrimental effects for the data subject, leaving no doubt that such processing violates the fairness principle when interpreted as 'substantive fairness' (see Section 6.2). Computational model constructions are often based on assumptions that turn out not to be true in practice.[713] ML produces probable yet inevitably uncertain knowledge and may identify significant correlations.[714] Even if strong correlations are found in datasets, this uncertain knowledge generalises by forming groups but does not distinguish between the members of this group. Data about individuals are full of correlations, but only some of these correlations meaningfully reflect the individual's actual capacity, needs or merits.[715]

This may lead to the situation that individuals are being unfairly treated, as explained in the child abuser example. In addition, it is highly doubtful whether it is fair to act upon probabilistic predictions and correlations deployed by means of ML. Actions taken based on probabilistic predictions and correlations may have real impact on human interests[716] (e.g., to receive a loan or to get a job). This holds particularly true where such predictions or correlations are essentially considered as *facts.* When individuals are treated based on simplified models or classes, concerns regarding the accuracy principle arise. It is clear that accuracy is a distinct principle, and I will discuss this separately in

[711] Kashmir Hill, 'A Dad Took Photos of His Naked Toddler for the Doctor. Google Flagged Him as A Criminal' *The New York Times* (New York, 21 August 2022) < https://www.nytimes.com/2022/08/21/technology/google-surveillance-toddler-photo.html > accessed 8 February 2024.
[712] Kashmir Hill, 'A Dad Took Photos of His Naked Toddler for the Doctor. Google Flagged Him as A Criminal' *The New York Times* (New York, 21 August 2022) < https://www.nytimes.com/2022/08/21/technology/google-surveillance-toddler-photo.html > accessed 8 February 2024.
[713] Toon Calders, Indrė Žliobaitė, 'Why Unbiased Computational Processes Can Lead to Discriminative Decision Procedures' in Bart Custers et al (eds) Discrimination and Privacy in the Information Society (Springer 2013) 45.
[714] Brent Daniel Mittelstadt et al, 'The ethics of algorithms: Mapping the debate' (2016) Vol 3 Iss 2 Big Data & Society 1, 4.
[715] Betsy A Williams, Catherine F Brooks, Yotam Shmargad, 'How Algorithms Discriminate Based on Data They Lack: Challenges, Solutions, and Policy Implications' (2018) Vol 8 Journal of Information Policy 78, 82–83.
[716] Brent Daniel Mittelstadt et al, 'The ethics of algorithms: Mapping the debate' (2016) Vol 3 Iss 2 Big Data & Society 1, 5; Solon Barocas, 'Data Mining and the Discourse on Discrimination' (2014) <https://dataethics.github.io/proceedings/DataMiningandtheDiscourseOnDiscrimination.pdf> accessed 8 February 2024.

Section 4.7. However, even if a prediction is entirely accurate from a mathematical and statistical perspective, treating individuals based on this prediction may still be unfair. Predictions generated by ML are probabilistic and relate to future conduct that has not yet happened or may never happen at all. From this perspective, applying predictions to individuals may be unfair because predictions do not reflect reality and are thus no 'facts.'

Probabilistic predictions and correlations produced by ML may thus have adverse effects on data subjects when treated as facts and, therefore, violate the fairness principle enshrined in EU data protection law. Furthermore, it seems difficult to argue that processing complies with the fairness principle when the AI system does not understand why certain patterns or correlations exist, although these patterns or correlations build the basis of ADM. With ADM generated by means of ML, the underpinning rationale of the decision is not articulated and perhaps not even known.[717] When combined with the reasoning and common sense deficiencies relating to the AI discipline of automated reasoning (see also Sections 2.2.5, 4.4.1 and 4.7.1) processing of personal data inherent to ADM seems to have substantial potential to be detrimental, discriminatory, unexpected or misleading for the data subjects concerned.

The probability problem (Type 1)

ML generates uncertain knowledge, such as predictions and correlations that are probabilistic. This may be unfair because ML mainly generalises and does not articulate the rationale of generated outputs due to the deficiencies in AR. When such outputs are essentially considered as facts, e.g. in the context of ADM, this can have adverse and detrimental effects for data subjects (e.g., when applying for a loan). This violates the fairness principle.

Face recognition systems as described in Section 2.2.3.1 and 2.2.3.2 related to computer vision might violate the principle of fairness. This is particularly due to the opacity of such systems as they may be used without any intention of or cooperation with data subjects.[718] Both the European Data Protection Board and the European Data Protection Supervisory have called for a general ban on any use of AI for automated recognition of human features such as faces in publicly accessible spaces.[719] Covert use of face recognition systems (see Section 2.2.3.1) is not only problematic in the context of law enforcement, but also when used by private actors. [720]

[717] Sue Newell, Marco Marabelli, 'The Crowd and Sensors Era: Opportunities and Challenges for Individuals, Organizations, Society, and Researchers' (ICIS, Auckland, December 2014) 11 < https://www.researchgate.net/publication/288239046_The_crowd_and_sensors_era_Opportunities_and_challenges_for_individuals_organizations_society_and_researchers > accessed 8 February 2024.
[718] Council of Europe, Consultative Committee of Convention 108, 'Guidelines on Facial Recognition' (28 January 2021) at 11 <https://rm.coe.int/guidelines-on-facial-recognition/1680a134f3> accessed 8 February 2024.
[719] European Data Protection Board and European Data Protection Supervisor, 'Joint Opinion on the Artificial Intelligence Act' (Joint Opinion 5/2021) at 32.
[720] Council of Europe, Consultative Committee of Convention 108, 'Guidelines on Facial Recognition' (28 January 2021) at 11 <https://rm.coe.int/guidelines-on-facial-recognition/1680a134f3> accessed 8 February 2024.

In the Netherlands, one supermarket chain has used facial recognition technology to prevent theft. All faces of customers who entered the supermarket were registered and consequently checked against a database containing faces of individuals who had previously been banned from entering the supermarket.[721] In Spain, a similar case occurred where a supermarket relied on a facial recognition system to identify individuals who had previously committed crimes in its stores and were banned from entering.[722] It seems questionable whether such processing is 'fair' for the data subjects concerned because it might have adverse effects on the data subjects. If detected by the system, a data subject might be confronted with the police and in any case be publicly exposed to other customers of the supermarket and very likely to be suspected of having committed a crime. Substantive fairness would require striking a fair balance between the interests at hand, namely the goal of the supermarket to prevent theft and the interests of the concerned data subjects. As outlined by the CJEU, this balance also depends on the nature of the information in question and its sensitivity for the data subject's private life.[723] The consideration of the data subjects fundamental rights to privacy and data protection would arguably outweigh the interest of the supermarket to prevent theft considering the intrusive nature of face recognition systems and the corresponding sensitivity for data subjects. In addition, applying the proportionality principle (Section 3.2.2) to this case would arguably lead to the same result.

The facial recognition problem (Type 1)

When covertly applied, face recognition systems powered by the AI discipline computer vision may violate the fairness principle due the intrusive nature of such systems and the corresponding sensitivity for the data subjects concerned, e.g., to be suspected of theft by default and/or to be publicly exposed as a criminal.

In particular, processing of personal data occurring in the context of affective computing (AC) raises the question whether such processing complies with the fairness principle. As pointed out in the probability problem, it is clear that accuracy is a distinct principle that merits dedicated analysis (Section 4.7). Nonetheless, treating individuals based on inaccurate personal data can still be unfair in the context of the fairness principle.

Generally, processing of emotion data enabled by means of AC could be misleading, specifically because the accuracy of outputs generated by AC has been questioned[724] (see also Section 4.7.1). For

[721] The Dutch Data Protection Supervisory Authority has issued a formal warning against this supermarket-chain, see < https://autoriteitpersoonsgegevens.nl/en/current/dutch-dpa-issues-formal-warning-to-supermarket-for-use-of-facial-recognition-technology > accessed 8 February 2024.
[722] Summary of Spanish SA Decision PS/00120/2021 < https://gdprhub.eu/index.php?title=AEPD_(Spain)_-_PS/00120/2021 > accessed 8 February 2024.
[723] Case C-131/12, *Google Spain* [2014] ECR I-317 para 81.
[724] Kate Crawford et al, 'AI Now Report' (2018) AI Now Institute 8 <https://ainowinstitute.org/publication/ai-now-2018-report-2> accessed 8 February 2024; Lisa Feldman Barrett et al. 'Emotional Expressions Reconsidered' (2019) Vol 20 (1) Psychological Science in the Public Interest 1; Sara Preto, 'Emotion-reading algorithms cannot predict intentions via facial

this thesis, emotion data are defined as information related to emotions of an individual ('emotion data'). Emotions refer to the six most-used emotion categories[725] in emotion research: anger, disgust, fear, happiness, sadness and surprise.[726] These six 'basic emotions'[727] are further described in Section 2.2.4.1. Processing of emotion data by AC could be both detrimental and unexpected for the individuals concerned. Imagine an employer that uses automated video assessments such as HireVue[728] to detect emotional states of applicants during these assessments. In particular, in these circumstances, processing of emotion data by means of AC might have adverse consequences for the data subject. Perhaps for precisely this reason, HireVue discontinued the use of the component of its services that analyses facial expressions of applicants.[729]

It has been argued that it should be prohibited to link recognition of emotions to the hiring of staff because it poses risks of great concern on both societal and individual levels.[730] Whereas prohibition seems to be a very restrictive measure, it is certainly valid to question the fairness of using information about the emotional states of individuals in an employment context. Considering the questionable accuracy of AC, the non-transparent manner of processing (candidates do not get to know which emotions the system detected), the sensitive nature of the personal processed (see Section 4.8.3) and the possible adverse effects for the applicant, it seems reasonable to conclude that such processing does not comply with the fairness principle. The asymmetrical power relations between employers and applicants also plays a role. When deciding to rely on AC-powered video assessments during the recruitment process to detect the applicants emotional state, the employer takes advantage of its stronger position. Substantive fairness aims to balance precisely these kind of power asymmetries and to prevent adverse effects in concrete circumstances.[731] Here, the adverse effects are obvious. Arguably inaccurate and rather sensitive personal data are processed to determine whether the applicant will receive a job offer. Undoubtedly, the latter decision has a considerable effect on the applicant.

expressions' *USC News* (Los Angeles, 4 September 2019) <https://news.usc.edu/160360/algorithms-emotions-facial-expressions-predict-intentions/> accessed 8 February 2024.

[725] These six emotions refer to research conducted by psychologists in the early seventies that developed the methodology of 'basic emotions'; see Paul Ekman, Wallace v Friesen, 'Constants across cultures in the face and emotion' (1971) Vol 17 (2) Journal of Personality and Social Psychology 124.

[726] Lisa Feldman Barrett et al. 'Emotional Expressions Reconsidered' (2019) Vol 20 (1) Psychological Science in the Public Interest 1, 52.

[727] Eiman Kanjo et al, 'Emotions in context: examining pervasive affective sensing systems, applications, and analyses' (2015) Vol 19 Personal and Ubiquitous Computing 1197, 1204 <https://link.springer.com/content/pdf/10.1007/s00779-015-0842-3.pdf> accessed 8 February 2024.

[728] Nathan Mondragon, Clemens Aicholzer, Kiki Leutner, 'The Next Generation of Assessments' (HireVue 2019) <http://hrlens.org/wp-content/uploads/2019/11/The-Next-Generation-of-Assessments-HireVue-White-Paper.pdf> accessed 8 February 2024.

[729] Will Knight, 'Job Screening Service Halts Facial Analysis of Applicants' *Wired* (New York, 12 January 2021) < https://www.wired.com/story/job-screening-service-halts-facial-analysis-applicants/ > accessed 8 February 2024. However, other providers offer similar services. HumeAI provides AI-powered tools helping recruiters to assess personality traits as well as emotional states of candidates; see < https://hume.ai/products/facial-expression-model/ > and < https://gethume.com/blog5/artificial-intelligence-for-recruiting > accessed 8 February 2024.

[730] Council of Europe, Consultative Committee of Convention 108, 'Guidelines on Facial Recognition' (28 January 2021) at 3 <https://rm.coe.int/guidelines-on-facial-recognition/1680a134f3> accessed 8 February 2024.

[731] Gianclaudio Malgieri, 'The concept of Fairness in the GDPR' (FAT* '20: Conference on Fairness, Accountability, and Transparency, Barcelona, January 2020) 2, 3 <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3517264> accessed 8 February 2024.

Video assessments powered by AC are only one of many possible examples. The use of AC might also be unfair within other domains including marketing, customer service, healthcare, insurance, retail, autonomous driving, education and gaming.[732] Thus, the use of AC in important sectors is bound to increase, as will the possibility of adverse consequences for data subjects. Although AC systems are predominantly developed in the United States, they are being sold to global marketplaces. Corresponding algorithms are hardly tweaked for racial, cultural, ethnic or gender differences.[733]

The fairness principle is prone to be violated due to the questionable accuracy of emotion data and the sensitive nature of the personal data disclosed and otherwise processed in the context of AC. As outlined in the probability problem, ML generates predictions and establishes correlations that are probabilistic and thus constitute uncertain knowledge. This means that also the output generated by means of ML can violate the fairness principle and the accuracy principle (see also Sections 4.3.1 and 4.7.1). Furthermore, such processing is likely to be detrimental to the interest of the data subject because revealing such sensitive information can very well be used to manipulate a data subject. According to research in behavioural sciences, especially psychology, emotions are powerful, pervasive and predictable drivers of human decision-making.[734]

The inaccuracy problem (Type 1)

The questionable accuracy of personal data generated by the AI disciplines AC and ML violate the fairness principle as the processing of inaccurate personal data is detrimental and misleading to the data subject.

The sensitivity problem (Type 1)

AC allows for predicting and disclosing sensitive emotion data in ways that violate the fairness principle because the subsequent use of such personal data is detrimental to the data subject, particularly in situations entailing power asymmetries, and because emotion data may be used to manipulate the data subject.

4.3.2 Legal problems: Type 2

As indicated in Section 4.3, the fairness principle has thus far managed to remain elusive despite the fact that the fairness principle is considered to be a key tenet of EU data protection law. Apart from obvious examples (such as discrimination), it largely remains unclear when processing of personal

[732] Cem Dilmegani, 'Top 24 Affective Computing (Emotion AI) Use Cases in 2023' <https://research.aimultiple.com/affective-computing-applications/> accessed 8 February 2024; Deepanshu Gahlaut, 'Top Emotion AI Companies to Watch out for in 2023' <https://deepanshugahlaut.medium.com/top-emotion-ai-companies-to-watch-out-for-in-2023-db925868fd9f> accessed 8 February 2024.
[733] Peter Mantello, Ho Manh-Tung, 'Why we need to be weary of emotional AI' (2022) AI & Society <https://link.springer.com/article/10.1007/s00146-022-01576-y> accessed 8 February 2024.
[734] Jennifer S Lerner et al, 'Emotion and Decision Making' (2015) Vol 66 Annual Review of Psychology 799, 802.

data is unfair or results in unfair consequences.[735] The fairness principle enshrined in EU data protection law lacks sufficient precision due to the absence of corresponding case law. Because the CJEU did not yet rule on the substantive meaning of the fairness principle, it is difficult to enforce the fairness principle in practice. This holds true in particular for private enforcement pursued by data subjects or actors mentioned in Article 80 GDPR that represent data subjects, such as non-profit bodies or organisations. Principle-based regulation requires controllers to make a judgement what they must do to comply and to perform risk assessments.[736] When performing such risk assessments, controllers will not only take the risks for the data subject into consideration, but also focus on interpretive risk[737] and any associated risk from enforcement action in case of non-compliance. Admittedly, the fairness principle's elusive role is not a problem caused explicitly by AI. Rather, it exists due to the lack of interpretative guidance by the CJEU. However, legal problems relating to the fairness principle are AI-specific because AI leads to many fairness issues.[738]

The unclear substantive meaning of the fairness principle reduces legal certainty and makes it less likely that it will be enforced by means of litigation in front of the courts. This is proven by means of a complete lack of case law with respect to the substantive meaning of the fairness principle on the level of the CJEU. Only one request for a preliminary ruling[739] dealt with the fairness principle, in which the CJEU ruled that 'fair processing' requires a public authority to inform the data subjects of the transfer of their personal data to another public authority that would process these data for its own purposes.[740] However, this case solely underscores the close link between the fairness and transparency principle, but does not provide any guidance with regard to the substantive meaning of the fairness principle. In case of shortcomings related to the interpretation of core provisions such as the fairness principle, compliance is a matter of risk management, and non-compliance becomes an option.[741] Controllers can assess what level of non-compliance they are prepared to risk and what the potential cost of enforcement action and reputational damage may be in case of non-compliance.[742]

It may be easy to access and read the controller's privacy notice, but it is an entirely different task to verify whether the statements made in the privacy notice are in fact honoured[743] and to what extent the fairness principle is complied with, in the case of complex AI systems in particular. Even if the

[735] Damian Clifford, Jef Ausloos 'Data Protection and the Role of Fairness' (2018) Vol 37 No 1 Yearbook of European Law 130, 187.
[736] Julia Black, 'Forms and paradoxes of principles-based regulation' (2008) Capital Markets Law Journal Vol 3 No 4 425, 454.
[737] For instance, the likelihood that the interpretation of the principle will be approved by supervisory authorities or courts.
[738] For example due to reasoning AI's deficiencies, AI enabled manipulation as discussed in Section 4.3.3.
[739] Case C-201/14 *Bara and others* [2015] ECR I-638 para 34.
[740] Tim van Canneyt et al, 'Data Protection: CJEU case law review – 1995-2020' (2021) Vol 56 Computerrecht 78, 102.
[741] Julia Black, 'Forms and paradoxes of principles-based regulation' (2008) Capital Markets Law Journal Vol 3 No 4 425, 454.
[742] Ibid.
[743] Asia Biega, Michèle Finck, 'Reviving Purpose Limitation and Data Minimisation in Data-Driven Systems' (2021) Technology and Regulation 44, 60 <https://arxiv.org/abs/2101.06203> accessed 8 February 2024.

constraints concerning technological complexity are overcome, the legal uncertainty concerning the interpretation of the principle remains. This leaves considerable discretion to the controllers on how to interpret and apply the fairness principle. Once challenged in regulatory and private enforcement, it is likely that controllers defend such interpretation rigorously and aim to reach precedents which serve their interests.[744] This is underscored by Meta's announcement to appeal both the substance and the fines of the final decisions adopted by the Irish SA based on the EDPB's binding decisions[745] which substantively interpreted the fairness principle for the first time in regulatory enforcement.

As outlined in Section 4.3.1, AI systems may process personal data in a way which is detrimental, unexpected or misleading to the data subject, ultimately resulting in unfair processing. The elusive role and meaning of the fairness principle reduces legal certainty, although the GDPR particularly aims to enhance *legal* and *practical* certainty for data subjects (Recital 7). The elusive role makes it difficult for data subjects and supervisory authorities to challenge the fairness of processing activities enabled by AI. The fact that AI and its underlying models are likely protected by trade secrets or IP laws makes this enforcement problem even bigger (see Section 5.6.2). This Type 2 legal problem occurs regardless of which AI discipline the fairness principle is being applied to because the problem is caused by the substantively unclear meaning of the fairness principle. It is therefore a general problem and relates to all AI disciplines.

The elusiveness problem (Type 2)

AI systems are likely to process personal data in a way that would typically be considered as unfair. The elusive role and meaning of the fairness principle reduces legal certainty and makes it difficult for data subjects to challenge the fairness of processing enabled by AI systems and enforce the fairness principle accordingly.

It could be argued that the meaning of the fairness principle is substantiated by means of regulatory enforcement at the EU level in the form binding decisions.[746] The EDPB identified the following key elements: autonomy of data subjects with respect to data processing, their reasonable expectations, ensuring power balance between controllers and data subjects, avoidance of deception as well as possible adverse consequences of processing and ensuring ethical and truthful processing.[747] However, the mentioning of these key elements in the EDPB's binding decisions does not establish legal certainty. These elements reflect the view of the EU's supervisory authorities (SAs). Meta has announced

[744] This does not seem to be unrealistic considering the financial resources well-known technology companies have and the legal expertise of which they can afford to make use.
[745] See < https://about.fb.com/news/2023/01/how-meta-uses-legal-bases-for-processing-ads-in-the-eu/ > accessed 8 February 2024.
[746] Article 65 GDPR.
[747] Binding Decision 3/2022 on the dispute submitted by the Irish SA on Meta Platforms Ireland Limited and its Facebook service (Art. 65 GDPR), adopted on 5 December 2022 paras 103, 219-220, 222-223, 226 478; Binding Decision 4/2022 on the dispute submitted by the Irish SA on Meta Platforms Ireland Limited and its Instagram service (Art. 65 GDPR), adopted on 5 December 2022 paras 106, 223-224, 226-227, 445; Binding Decision 5/2022 on the dispute submitted by the Irish SA on Meta Platforms Ireland Limited and its WhatsApp service (Art. 65 GDPR), adopted on 5 December 2022.

to appeal both the substance and the fines of the final decisions adopted by the Irish SA based on the EDPB's binding decisions.[748] As the controller, Meta has a right to an effective judicial remedy against the legally binding decision adopted by the Irish SA according to Article 78 (1) GDPR. As emphasised by the CJEU, the purpose of Article 78 GDPR is to examine the lawfulness of the decision adopted by a SA.[749] The Irish Court has full[750] and exclusive jurisdiction and needs to review the legality of the Irish SA's final decisions as well as the EDPB's binding decision.[751] Full jurisdiction in this context means the power to examine all questions of fact and law relevant to the dispute[752] and thus includes the question of law on how to interpret the fairness principle. It seems highly likely that the Irish Court will refer questions for a preliminary ruling to the CJEU regarding the contested decisions. In fact, it will be required to do so given the complete lack of judicial guidance regarding the interpretation of the fairness principle. The key elements of the principle of fairness mentioned by the EDPB have not yet been judicially tested. Thus, the Irish Court will arguably have doubts regarding this interpretation of the fairness principle and refer the matter to the CJEU. Hence, it may take several years until the CJEU rules on the matter. Consequently, the elusiveness of the fairness principle remains, which is notably detrimental to the GDPR's aim to enhance *legal* and *practical* certainty for data subjects (Recital 7).

4.3.3 Legal problems: Type 3

The conclusion reached in Section 4.3.2 that the substantive meaning of the GDPR's fairness principle[753] remains largely elusive and provides controllers with significant discretion on how to apply it in practice also leads to a Type 3 legal problem. Due to the lack of clarity concerning the scope and meaning of the fairness principle, the latter is *currently* not fit for purpose to protect the fundamental right to data protection for several reasons. However, as it becomes apparent from Section 6.2, I acknowledge the fairness principle's enormous potential for effective protection for individuals in an AI context if interpreted substantively.

A substantively elusive and unenforceable principle fails to achieve the GDPR's aim to establish a strong and coherent data protection framework[754] when considering that the principles provide the basis for the protection of personal data[755] in the GDPR. It also cannot ensure a consistent and high

[748] See < https://about.fb.com/news/2023/01/how-meta-uses-legal-bases-for-processing-ads-in-the-eu/ > accessed 8 February 2024.
[749] Case C-132/21 *Nemzeti* [2023] ECR I-2 para 35.
[750] Ibid para 41.
[751] Case T-709/21 *WhatsApp Ireland* [2022] ECR T-783 para 70.
[752] Case C-132/21 *Nemzeti* [2023] ECR I-2 para 41.
[753] I need to emphasise that I write about the GDPR's fairness principle. I do acknowledge that the concept of fairness is a constitutive element of the fundamental right to data protection according to Article 8 EUCFR ('fair processing').
[754] Recital 7 GDPR.
[755] Cécile de Terwangne, Commentary of Article 5 in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020) 313.

level of protection for personal data.[756] In addition, it harms the legal and practical certainty for data subjects.[757] Most importantly, the fairness principle fails to provide data subjects with *effective* protection.[758] In its case law, the CJEU has repeatedly stressed that EU data protection law aims to effectively protect the data subject's personal data against risk of misuse.[759] In this thesis, I interpret the risk of misuse broadly, referring to any unlawful use of personal data[760] to the detriment of natural persons concerned by it. A substantively elusive principle cannot prevent misuse in the form of processing of personal data that is detrimental to the data subject's interests and be perceived as unfair.

Manipulation is a typical example of personal data being processed to the detriment of the interests of the data subject. Although it is often not defined in work on the ethics of manipulation,[761] manipulation refers to hidden acts with the aim to intentionally and covertly influence a natural person by targeting and influencing this person's decision-making vulnerabilities.[762] Typically, such influence is against this person's self-interest.[763] Put simply, it perverts the way a person reaches decisions, forms preferences or adopts goals.[764] These acts are not only used to influence what the individual decides or does, but also to influence what the individual thinks or feels, i.e. the individuals thoughts.[765] Whereas manipulation is certainly not a new phenomenon, AI and particularly the disciplines ML and AC introduce new and dedicated means to manipulate decisions, behaviour and thoughts of individuals. AI powerfully enhances the range of influence that companies have in shaping behaviour and thoughts of individuals.[766] It can modify the options and choices available to individuals to manipulate their behaviour. Options or choices available to these individuals may be amended[767] to steer behaviour towards particular goals that are not for the benefit of the individuals

[756] Recitals 6, 10 GDPR; Case C-534/20, *Leistritz AG* [2022] ECR I-594 para 26; Case C-645/19 [2021] *Facebook Ireland* ECR I-483 para 45; Case C-511/18, *La Quadrature du Net* [2020] ECR I-791 para 106; Case C-131/12, *Google Spain* [2014] ECR I-317 para 66; Case C-154/21 *Österreichische Post AG* [2023] ECR I-3 para 44; Case C-132/21 *Nemzeti* [2023] ECR I-2 para 42.
[757] Recital 7 GDPR.
[758] As envisaged by Recital 11 GDPR see also Case C-645/19 [2021] *Facebook Ireland* ECR I-483 para 45; Case C-319/20, *Meta Platforms Ireland Limited* [2022] ECR I-322 para 73.
[759] Joined Cases C-293/12 and C-594/12 *Digital Rights Ireland* [2014] ECR-I 238 paras 54 and 66; Joined Cases C-203/15 and C-698/15 [2016] *Tele 2 Sverige* ECR-1 970 paras 109 and 122; Case C-362/14 *Schrems v Data Protection Commissioner* [2015] ECR I-650 para 91.
[760] See ECtHR case law to which the CJEU refers in the rulings contained in the previous footnote: *S. and Marper v United Kingdom* App no 30562/04 and 30566/04 (ECtHR 4 December 2008) para 99; *Liberty and Others v United Kingdom* App No 58243/00 (ECtHR 1 July 2008) paras 62-63; *Rotaru v Romania*, App No 28341/95 (ECtHR 4 May 2000) paras 57 to 59.
[761] Anne Barnhill, 'What is Manipulation?' in Christian Coons, Michael Weber (eds) *Manipulation* (Oxford University Press 2014) 52.
[762] Daniel Susser, Beate Roessler, Helen Nissenbaum 'Technology, autonomy, and manipulation' (2019) Vol 8 Iss 2 Internet Policy Review 1, 4.
[763] Anne Barnhill, 'What is Manipulation?' in Christian Coons, Michael Weber (eds) *Manipulation* (Oxford University Press 2014) 53.
[764] Joseph Raz, *The Morality of Freedom* (OUP 1986) 377; Cass R Sunstein, 'The Ethics of Nudging' (2015) Vol 32 Yale Journal of Regulation 413, 444.
[765] Anne Barnhill, 'What is Manipulation?' in Christian Coons, Michael Weber (eds) *Manipulation* (Oxford University Press 2014) 57.
[766] Helen Fay Nissenbaum, *Privacy in Context: Technology, Policy, and the Integrity of Social Life* (Stanford Law Books 2010) 83.
[767] Ruth Faden, Tom Beachamp, Nancy King, *A History and Theory of Informed Consent* (Oxford University Press 1986) 355.

concerned, but rather for the benefit of the company which deploys the AI system.[768] There is evidence of how intelligent artificial agents can significantly control human behaviour,[769] going clearly beyond what was previously possible. Research exploring whether it is possible for machines to learn how to influence humans indicates that by means of a computational framework based on reinforcement learning (see Section 2.2.1.3) and ANN approaches (see Section 2.2.1.4), the choices of individuals in particular decision-making tasks can be shaped toward actions or goals desired by the actor exercising influence.[770] In experiments, the machine learnt from participants' responses and identified and targeted vulnerabilities in their decision-making. These vulnerabilities were then successfully used by the machine to steer the participant's decision-making towards particular actions.[771]

It is evident that user interactions with AI-powered systems whose design has been informed by behavioural science lead to behavioural change.[772] Behavioural science concerns the study of behavioural insights and establishes a reliable understanding of behaviour and how it changes. With this information, accurate predictive models can be created.[773] That AI-powered systems, developed with insights from behavioural science, cause change of preference is less evident.[774] Preferences influence behaviour, but behaviour often predates and leads to the emergence of new preferences.[775] ML systems change not only user behaviour, but also user preferences.[776] To intentionally influence preferences of individuals severely impacts their personal autonomy.[777] The essence of autonomy is indicated by the etymology of the term: autos (self) and nomos (rule or law).[778] The ruling idea of personal autonomy is 'that people should make their own lives'[779] which means facing freely both existential and every day's choices.[780] Obviously, changing preferences of individuals influences or even violates personal autonomy as preferences no longer stem from the individuals themselves.

[768] Christopher Burr, Nello Cristianini, James Lydmann, 'An Analysis of the Interaction Between Intelligent Software Agents and Human Users' (2018) Vol 28 Minds and Machines 735, 744, 769; Christopher Burr, Nello Cristianini, 'Can machines read our mind?' (2019) Vol 29 Iss 3 Minds and Machines 461, 4464.
[769] Christopher Burr, Nello Cristianini, James Lydmann, 'An Analysis of the Interaction Between Intelligent Software Agents and Human Users' (2018) Vol 28 Minds and Machines 735, 752.
[770] Amir Dezfouli, Richard Nock, Peter Dayan, 'Adversarial vulnerabilities of human decision-making' (2020) Vol 117 Iss 46 PNAS, 29221-29228.
[771] Jon Whittle, 'AI can now learn to manipulate human behaviour' *The Conversation* (London, 18 February 2021) <https://theconversation.com/ai-can-now-learn-to-manipulate-human-behaviour-155031> accessed 8 August 2021.
[772] Matija Franklin et al, 'Recognising the importance of preference change: A call for a coordinated multidisciplinary research effort in the age of AI' (2022) <https://arxiv.org/pdf/2203.10525.pdf> 1 accessed 8 February 2024.
[773] Susan Michie, Maartje M van Stralen, Robert West, 'The behaviour change wheel: A new method for characterising and designing behaviour change interventions' (2011) Vol 6 Implementation Science 1-12 <https://implementation-science.biomedcentral.com/articles/10.1186/1748-5908-6-42> accessed 8 February 2024.
[774] Matija Franklin et al, 'Recognising the importance of preference change: A call for a coordinated multidisciplinary research effort in the age of AI' (2022) <https://arxiv.org/pdf/2203.10525.pdf> 1 accessed 8 February 2024.
[775] Dan Ariely, Michael I Norton, 'How actions create - not just reveal – preferences' (2007) Vol 12 Iss 1 Trends in Cognitive Sciences 13-16.
[776] Matija Franklin et al, 'Recognising the importance of preference change: A call for a coordinated multidisciplinary research effort in the age of AI' (2022) <https://arxiv.org/pdf/2203.10525.pdf> 1 accessed 8 February 2024.
[777] Matija Franklin et al, 'The EU's AI Act needs to address critical manipulation methods' *The OECD.AI Policy Observatory* (Paris, 21 March 2023) <https://oecd.ai/en/wonk/ai-act-manipulation-methods?utm_source=substack&utm_medium=email> accessed 8 February 2024.
[778] Gerald Dworkin, *The Theory and Practice of Autonomy* (Cambridge University Press 1988) 12, 18.
[779] Joseph Raz, *The Morality of Freedom* (Oxford University Press 1986) 369.
[780] Daniel Susser, Beate Roessler, Helen Nissenbaum 'Technology, autonomy, and manipulation' (2019) Vol 8 Iss 2 Internet Policy Review 1, 8.

Affective computing (AC) elevates the means to manipulate individuals to an even higher level. Emotions play an important role in the elicitation of autonomous motivated behaviour.[781] According to research in behavioural sciences, especially psychology, emotions constitute powerful, pervasive and predictable drivers of decision-making.[782] Emotions can have significant effects on economic transactions and play a powerful role in decision-making, reasoning[783] and everyday economic choices.[784] Because AC provides access to emotion data of individuals, it may affect people's decisions and lives in unprecedented ways. This is particularly true with regard to manipulation that operates based on facts about the subject's psychology, such as knowledge of its emotions and desires.[785] Three field experiments that reached more than 3.5 million individuals found that their behaviour can be significantly altered, measured by clicks and purchases, when provided with psychologically tailored advertisements.[786] Thus, AI and specifically the disciplines ML and AC exhibit unprecedented means to manipulate the behaviour and thoughts of individuals. Manipulations advance the manipulator's interest at the expense of the manipulated person.[787] An individual's choices, preferences and thoughts can be manipulated[788] to the detriment of individuals, which undermines or violates their personal autonomy.[789] Information regarding the emotional state of an individual might be particularly helpful to manipulate this individual because emotions play an important role in the elicitation of autonomous motivated behaviour.[790] According to research in behavioural sciences, especially psychology, emotions constitute powerful, pervasive and predictable drivers of decision-making.[791] Emotions can therefore have significant effects on economic transactions and play a powerful role in everyday economic choices.[792] Thus, manipulation enabled by AI systems harms the personal autonomy of the individuals concerned by changing their behaviour and preferences as well as by affecting their capacity for reflective choice.[793]

[781] Leen Vandercammen et al, 'On the Role of Specific Emotions in Autonomous and Controlled Behaviour' (2014) Vol 28 Iss 5 European Journal of Personality 413, 445.
[782] Jennifer S Lerner et al, 'Emotion and Decision Making' (2015) Vol 66 Annual Review of Psychology 799, 802.
[783] Steffen Steinert, Orsolya Friedrich, 'Wired Emotions: Ethical Issues of Affective Brain–Computer Interfaces' (2020) Vol 26 Science and Engineering Ethics 351, 352.
[784] Jennifer S Lerner, Deborah A Small, George Loewenstein, 'Heart Strings and Purse Strings' (2004) Vol 15 No 5 American Psychology Society 337-340.
[785] J S Blumenthal-Barby, 'A Framework for Assessing the Moral Status of Manipulation' in Christian Coons, Michael Weber (eds) *Manipulation* (Oxford University Press 2014) 123, 127.
[786] Sandra Matz et al, 'Psychological targeting as an effective approach to digital mass persuasion' (2017) Vol 114 No 48 PNAS 12714-12719.
[787] Moti Gorin, 'Towards a Theory of Interpersonal Manipulation' in Christian Coons, Michael Weber (eds) *Manipulation* (Oxford University Press 2014) 124; James Stacey Taylor, Practical Autonomy and Bioethics (Rougledge 2009) 81.
[788] Hildebrandt Mireille, Koops Bert-Jaap, 'The challenges of Ambient Law and Legal Protection in the Profiling Era' (2010) Vol. 73 (3) The Modern Law Review 428, 435.
[789] Newell Sue, Marabelli Marco, 'Strategic opportunities (and challenges) of algorithmic decision-making: A call for action on the long-term societal effects of datafication' (2015) Vol. 24 Iss. 1 The Journal of Strategic Information Systems 4.
[790] Leen Vandercammen et al, 'On the Role of Specific Emotions in Autonomous and Controlled Behaviour' (2014) Vol 28 Iss 5 European Journal of Personality 413, 445.
[791] Jennifer S Lerner et al, 'Emotion and Decision Making' (2015) Vol 66 Annual Review of Psychology 799, 802.
[792] Jennifer S Lerner, Deborah A Small, George Loewenstein, 'Heart Strings and Purse Strings' (2004) Vol 15 No 5 American Psychology Society 337-340.
[793] Matija Franklin et al, 'The EU's AI Act needs to address critical manipulation methods' *The OECD.AI Policy Observatory* (Paris, 21 March 2023) <https://oecd.ai/en/wonk/ai-act-manipulation-methods?utm_source=substack&utm_medium=email> accessed 8 February 2024.

Because the substantive meaning of the fairness principle remains elusive, it seems unclear whether processing personal data enabled by AI systems that deploy ML and AC approaches to manipulate the behaviour of individuals would, in fact, be considered as violating the fairness principle. The latter is *currently*[794] not fit for purpose to effectively protect[795] data subjects, which is detrimental to their interests and thus unfair. In its case law, the CJEU has repeatedly stressed that EU data protection law aims to effectively protect the data subject's personal data against the risk of misuse.[796] A substantively elusive and unenforceable principle leads to legal uncertainty. The principle hardly prevents misuses such as manipulations enabled by the AI disciplines AC and ML because this legal uncertainty is likely to be exploited by controllers. Thus, the elusiveness of the fairness principle fails to protect[797] data subjects from such practices effectively. It also fails to achieve other legislative aims of the GDPR, namely, to ensure a consistent and high level of protection for personal data,[798] a strong and coherent data protection framework[799] and legal and practical certainty for data subjects.[800] The substantively elusive fairness principle is also not fit for purpose to ensure that processing of personal data is designed to serve mankind[801] because it does not prevent the manipulation of data subjects and similar practices.

As the introduction (Section 4.3) indicates, this section's analysis and conclusions might appear rather negative. However, I acknowledge the fairness principle's considerable potential[802] to protect individuals from risks caused by AI effectively. I will discuss this thoroughly in Section 6.2.

Arguably, other areas of law, consumer protection law in particular, might be better equipped to prevent manipulation. Whereas this is generally a rightful observation, it should be noted that the processing of personal data enabling, for instance, the detection of emotional states of individuals is primarily governed by the GDPR. However, the fairness principle as it currently stands does not

[794] Because it is predominantly interpreted as procedural fairness and as a mere proxy for transparency see Section 6.2.
[795] Recital 11 GDPR; Case C-645/19 [2021] *Facebook Ireland* ECR I-483 para 45; Case C-319/20, *Meta Platforms Ireland Limited* [2022] ECR I-322 para 73.
[796] Joined Cases C-293/12 and C-594/12 *Digital Rights Ireland* [2014] ECR-I 238 paras 54 and 66; Joined Cases C-203/15 and C-698/15 [2016] *Tele 2 Sverige* ECR-1 970 paras 109 and 122; Case C-362/14 *Schrems v Data Protection Commissioner* [2015] ECR I-650 para 91.
[797] Case C-645/19 [2021] *Facebook Ireland* ECR I-483 para 45; Case C-319/20, *Meta Platforms Ireland Limited* [2022] ECR I-322 para 73.
[798] Recitals 6, 10 GDPR; Case C-534/20, *Leistritz AG* [2022] ECR I-594 para 26; Case C-645/19 [2021] *Facebook Ireland* ECR I-483 para 45; Case C-511/18, *La Quadrature du Net* [2020] ECR I-791 para 106; Case C-131/12, *Google Spain* [2014] ECR I-317 para 66; Case C-154/21 *Österreichische Post AG* [2023] ECR I-3 para 44; Case C-132/21 *Nemzeti* [2023] ECR I-2 para 42.
[799] Recital 7 GDPR.
[800] Recital 7 GDPR.
[801] Recital 4 GDPR.
[802] As also pointed out by Bygrave see Lee A Bygrave, 'Minding the machine v2.0: The EU General Data Protection Regulation and Autmated Decision Making' in Karen Yeung and Martin Lodge (eds) *Algorithmic Regulation* (OUP 2019) 260; Lee A Bygrave, 'Machine Learning, Cognitive Sovereignty and Data Protection Rights with Respect to Automated Decisions' (2021) University of Oslo Faculty of Law Legal Studies Research Paper Series No. 202-35 at 22, 23 <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3721118> accessed 8 February 2024.

effectively protect[803] data subjects from such processing and has its limitations due to the elusive meaning of this principle. The subsequent use of personal data generated by means of AC and ML might, in some cases, fall under the scope of consumer law, for instance, if the use of such information would be considered an unfair commercial practice. However, it is questionable whether current EU consumer law is, in fact, capable of dealing with such practices. This is indicated by the fitness check on EU consumer law launched by the European Commission in May 2022 which focusses on digital fairness.[804] Irrespective of the outcome of this fitness check, it is important that the fairness principle, as an overarching principle, ensures that personal data are not processed to the detriment of the data subjects concerned.

The manipulation problem (Type 3)

The AI disciplines AC and ML enable controllers to manipulate data subjects by intentionally and covertly exploiting their behaviour, preferences, thoughts and decision-making vulnerabilities, which can be perceived as unfair. Due to the unclear substantive meaning of the fairness principle, it remains unclear whether such processing actually violates the fairness principle. Therefore, the fairness principle is not fit for purpose to effectively protect the fundamental right to data protection and prevent misuses such as manipulations.

In addition, the unclear substantive meaning of the fairness principle also is at odds with the accountability principle which aims to strengthen the responsibility of controllers when they process personal data.[805] The accountability principle enshrined in Article 5 (2) GDPR states that the controller shall be i) responsible for compliance and ii) able to demonstrate compliance with all the principles mentioned in Article 5 (1) GDPR.[806] Shortcomings with regard to the substantive meaning of the fairness principle makes it primarily difficult for controllers to ensure compliance with it. This also affects the data subjects. Requiring controllers to demonstrate compliance with substantively unclear provisions not only fails to effectively protect[807] data subjects. It also fails to establish the responsibility and liability of controllers[808] by imposing legally enforceable obligations on controllers.[809] The accountability principle and related Article 24 GDPR demand controllers to comply with the fairness principle whose actual substantive meaning remains largely unclear. Consequently, controllers cannot, as intended, be held accountable and responsible for complying with it. This holds true regardless of which

[803] Case C-645/19 [2021] *Facebook Ireland* ECR I-483 para 45; Case C-319/20, *Meta Platforms Ireland Limited* [2022] ECR I-322 para 73; Joined Cases C-293/12 and C-594/12 *Digital Rights Ireland* [2014] ECR-I 238 paras 54 and 66; Joined Cases C-203/15 and C-698/15 [2016] *Tele 2 Svergie* ECR-1 970 paras 109 and 122; Case C-362/14 *Schrems v Data Protection Commissioner* [2015] ECR I-650 para 91.
[804] See < https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/13413-Digital-fairness-fitness-check-on-EU-consumer-law_en > accessed 8 February 2024.
[805] Recital 74 GDPR.
[806] Cécile de Terwangne, Commentary of Article 5 in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020) 311.
[807] Recital 11 GDPR; Joined Cases C-293/12 and C-594/12 *Digital Rights Ireland* [2014] ECR-I 238 paras 54 and 66; Joined Cases C-203/15 and C-698/15 [2016] *Tele 2 Svergie* ECR-1 970 paras 109 and 122; Case C-362/14 *Schrems v Data Protection Commissioner* [2015] ECR I-650 para 91.
[808] Recital 74 GDPR.
[809] Recital 13 GDPR.

AI discipline the fairness principle is being applied to. Because controllers cannot be held responsible for failing to comply with obligations that are substantively unclear, the accountability principle misses its aim.[810] This negatively affects the envisaged high level of protection[811] as well as the strong data protection framework[812] intended by the GDPR. Thus, the elusiveness of the fairness principle sabotages the accountability principle. Requiring controllers to demonstrate compliance with a substantively unclear principle is not fit for purpose to effectively protect the fundamental right to data protection. This constitutes a Type 3 legal problem. It is a reoccurring problem because the accountability principle requires compliance with all principles enlisted in Article 5 (1) GDPR. Admittedly, the sabotage problem as described here is not a problem of AI in particular, but one created by the principles contained in the GDPR.

The sabotage problem (Type 3)

Since the substantive meaning of the fairness principle remains largely unclear, it sabotages the accountability principle. Because the accountability principle demands controllers to comply with a substantively unclear principle, it is not fit for purpose to protect the fundamental right to data protection. A principle demanding compliance with substantively unclear provisions cannot hold controllers responsible, nor can it effectively protect data subjects and ensure a high level of data protection.

4.4 Transparency

As outlined in Section 3.3.3.3, the transparency principle enshrined in Article 5 (1) GDPR requires that personal data be processed in a 'transparent manner'. Recital 39 GDPR clarifies that it must be 'transparent to natural persons that personal data concerning them are collected, used, consulted or otherwise processed.'[813] Articles 13 and 14 GDPR implement the transparency principle and impose specific information duties on the controller. In view of the EDPB, these provisions are the concretisation of the transparency principle, and violations of these provisions may also amount to the violation of the transparency principle itself.[814] As will be discussed in this section, the AI disciplines introduced in Chapter 2 create legal problems concerning the transparency principle itself as well as the specific information duties imposed on controllers.

[810] To hold controllers responsible see Recital 74 GDPR.
[811] Recitals 6, 10 GDPR; Case C-534/20, *Leistritz AG* [2022] ECR I-594 para 26; Case C-645/19 [2021] *Facebook Ireland* ECR I-483 para 45; Case C-511/18, *La Quadrature du Net* [2020] ECR I-791 para 106; Case C-131/12, *Google Spain* [2014] ECR I-317 para 66; Case C-154/21 *Österreichische Post AG* [2023] ECR I-3 para 44; Case C-132/21 *Nemzeti* [2023] ECR I-2 para 42.
[812] Recital 7 GDPR.
[813] Recital 39 GDPR.
[814] EDPB, 'Binding Decision 1/2021 on the dispute arisen on the draft decision of the Irish Supervisory Authority regarding WhatsApp Ireland under Article 65 (1) lit a GDPR' (2021) paras 191, 193.

4.4.1 Legal problems: Type 1

AI systems may be rather ubiquitous[815] and all AI disciplines[816] introduced in Chapter 2 may potentially clash with the transparency principle. Research has shown that users of smart homes are unaware of the possibility that machine learning (ML) algorithms may infer highly sensitive information, including sleep patterns and home occupancy.[817] Natural language processing (NLP) embedded in virtual assistants[818] such as Alexa or Siri facilitate the interception, recording and analysis of private communications without the users being aware of it, as unveiled by the press.[819] Computer vision (CV) applications allow identification of individuals from a distance and in a covert manner by means of face detection or gait analysis, without the knowledge of the individuals concerned. Regarding affective computing (AC) applications, transparent processing would presuppose that an individual is able to see what emotion the machine recognised, a requirement that also has been propagated by the pioneer in the field of AC.[820] However, in practice, this does not seem to be the case. The automated video assessment system provided by HireVue[821] aims to detect emotional states of applicants during job assessments. Similarly, the automated border control system called IBORDERCTRL 'analyses the micro-gestures of travellers to figure out if the interviewee is lying'.[822] Both systems do not communicate the detected emotions or detected 'lies' to the individuals concerned.

In addition, the dynamic nature of AI contradicts the static nature of the transparency principle because AI systems are continuously updated and changed whereas transparency disclosure only concerns algorithms used at a given moment.[823] All the examples mentioned illustrate that applications of AI potentially violate the transparency principle because personal data are not processed in a transparent manner as required by Article 5 (1) lit a GDPR. The following example regarding ML makes

[815] Finale Doshi-Velez et al, 'Accountability of AI Under the Law: The Role of Explanation' (2017) Berkman Klein Center Working Group on Explanation and the Law Working Paper 1 <https://dash.harvard.edu/bitstream/handle/1/34372584/2017-11_aiexplainability-1.pdf> accessed 8 February 2024; Jenna Burrel, 'How the machine 'thinks': understanding opacity in machine learning algorithms' (2016) Vol 3 Iss 1 Big Data Society 1-12 <https://journals.sagepub.com/doi/pdf/10.1177/2053951715622512> accessed 8 February 2024.
[816] With the sole exception of AR, which is not problematic in this context.
[817] Zheng Serena, 'User Perceptions of Smart Home IoT Privacy' (2018) Vol. 2 Proceedings of the ACM on Human-Computer Interaction 3.
[818] See Section 4.2.6 below for an explanation of how virtual assistants work.
[819] See for example <https://www.forbes.com/sites/blakemorgan/2018/02/05/are-digital-assistants-always-listening/>, <https://www.theverge.com/2019/7/11/20690020/google-assistant-home-human-contractors-listening-recordings-vrt-nws>, < https://www.nytimes.com/2018/05/25/business/amazon-alexa-conversation-shared-echo.html> accessed 8 February 2024; see also Silvia de Conca, 'The enchanted house' Doctoral Thesis, Tilburg University 2021) <https://pure.uvt.nl/ws/portalfiles/portal/50798678/De_Conca_The_Enchanted_23_06_2021_emb_tot_23_06_2022.pdf > accessed 8 February 2024.
[820] Rosalind W Picard, *Affective Computing* (MIT Press 1997) 122.
[821] Nathan Mondragon, Clemens Aicholzer, Kiki Leutner, 'The Next Generation of Assessments' (HireVue 2019) <http://hrlens.org/wp-content/uploads/2019/11/The-Next-Generation-of-Assessments-HireVue-White-Paper.pdf> accessed 8 February 2024. HireVue halted the use of this component, but other providers offer similar services; see < https://hume.ai/products/facial-expression-model/ > and < https://gethume.com/blog5/artificial-intelligence-for-recruiting > accessed 8 February 2024.
[822] European Commission, 'Smart lie-detection system to tighten EU's busy borders' (24 October 2018) < https://ec.europa.eu/research-and-innovation/en/projects/success-stories/all/smart-lie-detection-system-tighten-eus-busy-borders > accessed 8 February 2024.
[823] Council of Europe, Committee of Convention 108, 'Guidelines on Artificial Intelligence and Data Protection' (25 January 2021) at 3 <https://rm.coe.int/guidelines-on-facial-recognition/1680a134f3> accessed 8 February 2024.

this more concrete. Litera c of Article 13 (1) and 14 (1) GDPR impose the obligation on controllers to inform data subjects about the purposes of the processing for which the personal data are intended. Where personal data are directly collected from data subjects, this information must be provided at the time when personal data are obtained, and in all other cases within one month, at the latest, after obtaining the personal data or at the time of the first communication with the data subjects. Unsupervised ML approaches process data for *un*specified and *in*explicit purposes – the processing *itself* determines the purpose of the future use of the data – since its goal is to detect patterns and correlations, gain knowledge and make accurate predictions. There is no transparency issue if the controller processes personal data within the AI system for training purposes. However, this is different when the controller intends to detect correlations, patterns, and commercially valuable insights in data by deploying unsupervised ML. In such a case, the controller will determine the *specific* purpose of processing based on the processing activity's results, i.e., after the processing. Consequently, the transparency principle as further substantiated in Articles 13 (1) and 14 (1) GDPR cannot be complied with because the purpose of processing is not known at the time of data collection or when obtained from sources other than the data subject. Indeed, regulatory guidance demands that one 'always specify the purposes of the processing at the time of collection.'[824]

Take, for example, inferred personal data defined as 'the product of probability-based processes' that are used to create predictions of behaviour deployed to categorise individuals.[825] Where the purpose of processing consists of the creation of inferred personal data as is the case with ML, regulatory guidance requires one to communicate, at the time of collection or prior to further processing, 'the *intended purpose* of creating and further processing such inferred personal data, as well as the *categories* of the inferred data processed'.[826]

ML aims to create inferred personal data by detecting patterns and correlations, gaining knowledge and making accurate predictions. Therefore, controllers will not be able to inform data subjects about the specific purposes for which personal data are further processed because this information is completely unknown at the time of data collection or prior to further processing. Controllers could certainly inform data subjects about the intended purpose of processing in rather general terms such as 'We may use your personal data for detecting patterns/correlations and make accurate predictions about you.' However, such information will ultimately not meet the level of transparency required for the purpose specification, as foreseen by regulatory guidance, which states that the phrase 'We may use your personal data to develop new services' is not sufficiently clear about the purpose of

[824] Art 29 Working Party, 'Guidelines on transparency under Regulation 2016/679' (WP 260 rev.01, 11 April 2018) at 14.
[825] OECD Working Party on Security and Privacy in the Digital Economy JT03357584 (2014) <https://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=dsti/iccp/reg(2014)3&doclanguage=en> accessed 8 February 2024.
[826] Art 29 Working Party, 'Guidelines on transparency under Regulation 2016/679' (WP 260 rev.01, 11 April 2018) at 14 emphasis added by the author.

processing.[827] Likewise, controllers are unable to inform data subjects about the categories of inferred data processed as required by regulatory guidance. The categories of inferred personal data are unknown prior to further processing. First, an AI system needs to generate the inferred personal data before the controller can inform data subjects about the categories thereof.

The opacity problem (Type 1)

Unsupervised ML approaches process data for inexplicit purposes – the processing itself determines the purpose of the future use of the data. Controllers cannot inform data subjects about the purpose of processing nor the categories of inferred personal data because this information is not known at the time of data collection or prior to further processing. This violates the transparency principle.

Transparency regarding automated decision-making (ADM) constitutes a particular issue when applied to AI. Articles 13 (2) lit f and 14 (2) lit g GDPR require controllers to provide 'meaningful information about the logic' involved in ADM. Wachter, Mittelstadt and Floridi take the view that meaningful information according to Articles 13 (2) lit f and 14 (2) lit g GDPR can logically only address system functionality, namely, information about the logic, significance, envisaged consequences and general functionality of an ADM system, but not the rationale of *specific* ADM as the latter cannot be known before the decision is made.[828] Their reasoning suggests that information according to Articles 13 (2) lit f and 14 (2) lit g GDPR can only be provided *ex-ante* because notification occurs before ADM takes place, namely, at the point when personal data are collected for processing.[829] Malgieri and Comandé argue that meaningful information about the logic involved must adhere to the standard of legibility, which requires that the information to be provided is both transparent and comprehensible, and that such information must go 'beyond the mere mathematical functionality of an algorithm' and consider contextual use, expected and actual impact, rationales and purposes.[830] More generally, there is a vivid debate in scholarship whether or not the GDPR provides a right to explanation of specific ADM.[831]

Irrespective of this debate which will be discussed in the context of the right of access (Section 5.6.2), the notion of 'meaningful information' remains elusive. It is not yet clear what 'meaningful information' precisely means in practice. This notion also appears in Article 15 (1) lit h GDPR. Custers

[827] Art 29 Working Party, 'Guidelines on transparency under Regulation 2016/679' (WP 260 rev.01, 11 April 2018) at 12.
[828] Sandra Wachter, Brent Mittelstadt, Luciano Floridi, 'Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation' (2017) Vol 7 Iss 2 IDPL 76, 78.
[829]Ibid 76, 82.
[830] Giancludio Malgieri, Giovanni Comandé, 'Why a Right to Legibility of Automated Decision-Makig Exists in the General Data Protection Regulation' (2017) Vol 7 Iss 4 IDPL 243, 245, 257, 258.
[831] Thomas Wischmeyer, 'Artificial Intelligence and Transparency: Opening the Black Box' in Thomas Wischmeyer and Timo Rademacher (eds) *Regulating Artificial Intelligence* (Springer Nature 2020) 75-101; Sandra Wachter, Brent Mittelstadt, Luciano Floridi, 'Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation' (2017) Vol 7 Iss 2 IDPL 76-99; Giancludio Malgieri, Giovanni Comandé, 'Why a Right to Legibility of Automated Decision-Makig Exists in the General Data Protection Regulation' (2017) Vol 7 Iss 4 IDPL 243-265.

and Heijne performed research related to this notion enshrined in the right of access. They suggest interpreting it as information that is useful and/or has practical value for data subjects.[832] This interpretation has a contextual component and arguably means useful and practical for data subjects to (i) become aware of processing relating to ADM, (ii) enforce their data subject rights and thus (iii) exercise control over the processing of their personal data. For this section, I interpret meaningful information as useful and/or having practical value for data subjects. This is also in line with the CJEU's focus on intelligibility with respect to Article 12 (1) GDPR, which ensures that the data subject fully understands the information sent to it.[833]

It seems clear that controllers must understand the functionality of an ADM system to be able to provide data subjects with information that is useful and/or of practical value (meaningful information). Such information can only be provided if the trained model used for the ADM system can be articulated and understood by a human.[834] Giving information about the type of input data and the expected output, explaining the variables and their weight, or shining light on the analytics architecture are various forms of transparency concerning the logic of AI algorithms.[835] However, providing such information constitutes a two-sided problem: some information might effortlessly be provided by humans, but not by AI systems and vice versa.[836] The reasons for this are as follows.

First, AI lacks common sense reasoning capabilities due to deficiencies in automated reasoning as outlined in Sections 2.2.5 and 4.2.1. Systems based on ML do not *know* why specific input should receive some label, they solely know that certain input correlate with such a label. For example, an ML model trained with a dataset in which all basketballs are orange might classify all future input that is orange as basketballs.[837] For humans, it would be common sense not to do so. Due to these reasoning deficiencies, it seems reasonable to argue that AI itself is currently not capable of displaying the logic involved in ADM systems and the rationale behind or the criteria relied on to make the automated decision. Consequently, controllers cannot provide data subjects with information that is useful or of practical value for them.

[832] Bart Custers, Anne-Sophie Heijne, 'The right of access in automated decision-making: The scope of article 15(1)(h) GDPR in theory and practice' (2022) Vol 46 Computer Law & Security Review 1, 14.
[833] Case C-487/21, *F.F.* [2022] ECR I-1000 paras 37-38; in addition, Opinion of AG Pitruzella paras 55-56.
[834] Bryce Goodman, Seth Flaxman, 'European Union regulations on algorithmic decision-making and a right to explanation' (2017) Vol 38 No 3 AI Magazine 50, 55.
[835] Council of Europe, Committee of Convention 108, 'Guidelines on Artificial Intelligence and Data Protection' (25 January 2021) at 31 <https://rm.coe.int/guidelines-on-facial-recognition/1680a134f3> accessed 8 February 2024.
[836] Finale Doshi-Velez et al, 'Accountability of AI Under the Law: The Role of Explanation' (2017) Berkman Klein Center Working Group on Explanation and the Law Working Paper 1 <https://dash.harvard.edu/bitstream/handle/1/34372584/2017-11_aiexplainability-1.pdf> accessed 8 February 2024.
[837] Zachary C Lipton, 'The Mythos of Model Interpretability' (2018) Vol 16 Iss 3 ACMQueue 3 <https://dl.acm.org/doi/pdf/10.1145/3236386.3241340?download=true> accessed 8 February 2024.

Second, the complexity of many AI systems makes it impossible to present the casual factors which have led to a decision in a manner which is understandable for data subjects.[838] In particular, the complexity of adopted ML models represents a major challenge for human cognition.[839] With some algorithms used in AI systems, it is practically impossible to retroactively connect specific input to specific output and vice versa.[840] The difficulty to establish a nexus between specific input and output and thus to derive the logic involved in ADM differs considerably between the techniques used for ML. ML algorithms deploying sparse linear models such as regression introduced in Section 2.2.1.1 tend to generate interpretable models, allowing to identify the role of each model component (e.g., weight of a feature in a linear regression model) within the whole computing process, which ultimately leads to traceability and transparency in ADM.[841] However, this is different in case of deep learning (DL) and artificial neural networks (ANN). When an ANN is used for pattern recognition in CV or NLP, an ex-post analysis of a specific ADM will likely not establish a linear causal connection which is easily comprehensible for human minds.[842] Complex processes applied in deep learning (DL) are challenging for human cognition, both in terms of explaining the logic of the algorithms and the specific ADM. Non-deterministic systems make it hard to provide detailed information about the logic involved in the processing of personal.[843] With regard to explainability seen as the identification of factors that have caused a decision,[844] ANN and DL pose perhaps the biggest challenge.[845] Most current DL models lack reasoning and explanatory capabilities, which makes them vulnerable to produce unexplainable outcomes. DL methods based on ANN generally lack interpretability.[846] It seems neither possible to understand which artificial neuron contributed to a distinct part of the output nor to understand what happened in the intermediate (hidden) layers of the ANN.[847] Therefore, humans will hardly be able to extract any underlying rules which may be used to determine the logic involved in ADM: the many numeric values of the weights produced by the model do not have a meaning to the supervisor.[848] Consequently, controllers cannot provide data subjects with meaningful information, namely, information that is useful and/or has practical value. However, the field of

[838] Thomas Wischmeyer, 'Artificial Intelligence and Transparency: Opening the Black Box' in Thomas Wischmeyer and Timo Rademacher (eds) *Regulating Artificial Intelligence* (Springer Nature 2020) 89.
[839] Zachary C Lipton, 'The Mythos of Model Interpretability' (2018) Vol 16 Iss 3 ACMQueue 18 <https://dl.acm.org/doi/pdf/10.1145/3236386.3241340?download=true> accessed 8 February 2024.
[840] Thomas Wischmeyer, 'Artificial Intelligence and Transparency: Opening the Black Box' in Thomas Wischmeyer and Timo Rademacher (eds) *Regulating Artificial Intelligence* (Springer Nature 2020) 81.
[841] Apostolos Vorras, Lilian Mitrou, 'Unboxing the Black Box of Artificial Intelligence: Algorithmic Transparency and/or a Right to Functional Explainability' in Titania-Eleni Synodinou et al (eds) *EU Inernet Law in the Digital Single Market* (Springer Nature 2021) 256.
[842] Thomas Wischmeyer, 'Artificial Intelligence and Transparency: Opening the Black Box' in Thomas Wischmeyer and Timo Rademacher (eds) *Regulating Artificial Intelligence* (Springer Nature 2020) 81.
[843] Council of Europe, Committee of Convention 108, 'Guidelines on Artificial Intelligence and Data Protection' (25 January 2021) at 3 <https://rm.coe.int/guidelines-on-facial-recognition/1680a134f3> accessed 8 February 2024.
[844] Thomas Wischmeyer, 'Artificial Intelligence and Transparency: Opening the Black Box' in Thomas Wischmeyer and Timo Rademacher (eds) *Regulating Artificial Intelligence* (Springer Nature 2020) 88.
[845] Bryce Goodman, Seth Flaxman, 'European Union regulations on algorithmic decision-making and a right to explanation' (2017) Vol 38 No 3 AI Magazine 50, 55.
[846] Deng Li and Liu Yang, 'A Joint Introduction to Natural Language Processing and Deep Learning' in Deng Li and Liu Yang (eds) *Deep learning in natural language processing* (Springer 2018) 11, 12.
[847] Ethem Alpaydin, *Machine Learning: The New AI* (3rd edn MIT Press 2016) 155.
[848] Toshinori Munakata, *Fundamentals of the New Artificial Intelligence* (2nd edn, Springer 2008) 12, 25, 35, 44.

Explainable AI ('xAI') has made significant progress in recent years. xAI aims to develop explainable techniques that empower end users to comprehend, trust, and efficiently manage AI systems.[849] Nonetheless, causal explanations, which are crucial for ADM, are still a challenge and are anticipated to be the next frontier of ML.[850]

Regulatory guidance acknowledges the challenge for humans to understand how ADM processes work in the context of ML. Nevertheless, the guidance also states that complexity is no excuse and controllers should find 'simple ways to tell the data subject about the rationale behind, or the criteria relied on reaching the decision'.[851] However, considering current deficiencies in terms of interpretability in the context of ML and ANNs and deficiencies in automated reasoning, it seems that the ideal of transparency with respect to meaningful information about the logic involved in ADM is technologically not possible (yet). Due to interpretability and reasoning deficiencies, controllers are unable to provide data subjects with meaningful information, namely, information that is useful and/or has practical value. This leads to a Type 1 legal problem, because Articles 13 (2) lit f and 14 (2) lit g GDPR are violated.

The interpretability problem (Type 1)
Due to the deficiencies in AR, AI systems cannot themselves display the logic involved in ADM systems and the rationale behind or the criteria relied on reaching the automated decision. AI systems deploying DL and ANN approaches from ML are likely to produce non-interpretable outputs. When used in the context of ADM, controllers cannot provide data subjects with meaningful information about the logic involved in ADM and thus violate the transparency principle.

4.4.2 Legal problems: Type 2

The interpretability problem (Type 2)
The interpretability problem outlined in Section 4.4.1 also leads to a Type 2 legal problem. Due to the deficiencies in terms of interpretability in the context of DL and ANN as well as deficiencies in AR, it is technologically not possible for controllers to induce meaningful information about the logic involved in ADM. Therefore, data subjects and regulators cannot enforce the transparency principle and obtain the corresponding information.

[849] Waddah Saeed, Christian Omlin, 'Explainable AI (XAI): A systematic meta-survey of current challenges and future opportunities' (2023) Vol 263 Knowledge-Based Systems 1-22.
[850] Ibid 9.
[851] Art 29 Working Party 'Guidelines on Automated Individual Decision-Making and Profiling for the Purposes of Regulation 2016/679', (WP251rev.01, 6 February 2018) at 25.

Making a methodological remark here regarding type 1 and 2 problems seems appropriate. The interpretability problem (type 1) should not automatically lead to a type 2 problem because the controller is required to cease processing after becoming aware that processing is unlawful.

Under the GDPR, a processing activity whose complexity makes it impossible for the controller to respect data protection principles should not occur.[852] This follows from the accountability principle[853] and other obligations, such as performing a Data Protection Impact Assessment (DPIA). When deploying ADM systems described in Section 4.4.1, the controller must perform a DPIA according to Article 35 GDPR. DPIAs are required if the envisaged processing is likely to result in a high risk to the rights and freedoms of data subjects. The performance of a DPIA is mandatory when the controller uses 'new technologies'[854] for processing, including AI systems. In such cases, controllers should consult the competent Supervisory Authority (SA) if the high risks cannot be mitigated.[855] If compliance with the GDPR principles is impossible, the controller should stop the processing. In addition, the competent SA may also ban such processing based on Article 58 (2) GDPR. Hence, a type 1 problem should not lead to a type 2 problem, as the processing should simply not occur. However, the possibility remains that controllers perform a cost-risk analysis and continue with such processing even after warnings or fines from SAs. The latter is not only a theoretical possibility. For instance, OpenAI continued to provide its services after bans and warnings imposed by the Italian SA.[856] Also, Clearview AI continued with its processing activities even after receiving clear signs from the EDPB concerning the lawfulness of the processing.[857] Thus, although non-compliance with data protection principles should result in ceased processing activities, this might be ignored in practice (e.g., by powerful tech companies). Ultimately, the principles are being violated and cannot be enforced simultaneously, leading to a type 2 problem.

4.4.3 Legal problems: Type 3

The GDPR contains provisions requiring controllers to inform data subjects if their personal data will be processed for a different purpose which is compatible with the one for which personal data were initially collected. Articles 13 (3) and 14 (4) GDPR specifically relate to the purpose limitation principle enshrined in Article 5 (1) lit b GDPR[858] which states that further processing for scientific research purposes or statistical purposes shall not be incompatible with the initial purpose (i.e. privileged purposes). As will be outlined in Section 4.5.3, there are reasons to argue that ML serves

[852] This also applies to the verification problem discussed in Section 4.6.2.
[853] Article 5 (2) GDPR.
[854] Article 35 (1) GDPR.
[855] Article 36 GDPR.
[856] See < https://iapp.org/news/a/garante-issues-notice-to-openai-over-alleged-gdpr-violations/ > accessed 8 February 2024.
[857] See <https://edpb.europa.eu/sites/default/files/files/file1/edpb_letter_out_2020-0052_facialrecognition.pdf> accessed 8 February 2024.
[858] Art 29 Working Party, 'Guidelines on transparency under Regulation 2016/679' (WP 260 rev.01, 11 April 2018) at 23.

research or statistical purposes, which might enable controllers to rely on these privileged purposes to generate inferred personal data. Inferred data are 'the product of probability-based processes' and are used, for instance, to create predictions of behaviour.[859]

Controllers may also further process personal data for compatible purposes other than privileged ones based on an assessment that takes into account the factors mentioned in Article 6 (4) GDPR.[860] Article 6 (4) GDPR stipulates a series of criteria to determine whether further processing for a purpose other than the one for which personal data have been initially collected is 'compatible' with this initial purpose.[861] According to the CJEU, these criteria reflect the need for a concrete, coherent and sufficiently close link between the purpose of collection and further processing of data. These criteria make it possible to determine that further processing does not detract from the data subject's legitimate expectations as to the further use of their personal data.[862] Where controllers can establish such a link, they may further process personal data in order to detect an individual's emotional state by means of AC. Provided that the purposes for further processing are compatible, either privileged[863] or otherwise compatible,[864] controllers solely need to notify data subjects in advance about these compatible purposes and with any relevant further information as referred to in paragraph *2* of Article 13 and 14 GDPR. Both provisions *do not* include information about the nature of inferred personal data or categories of personal data.

Controllers do not need to outline which personal data or categories of personal data are processed if the initial personal data are directly collected from data subjects.[865] Where the initial personal data are not directly collected from the data subject, information about the categories of personal data as received by the controller must be provided according to Article 14 (1) lit e GDPR. However, because this requirement is enshrined in paragraph 1 and *not* 2 of Article 14 to which Article 14 (4) GDPR refers, controllers do not need to inform data subjects about the actual category of the personal data inferred by means of ML or AC. In other words, controllers must indicate the categories of personal data they have received from another controller, but not the ones inferred from such data. With regard to inferred personal data, regulatory guidance on transparency requires that 'the *intended purpose* of creating and further processing such inferred personal data, as well as the *categories of the inferred*

[859] OECD Working Party on Security and Privacy in the Digital Economy JT03357584 (2014) <https://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=dsti/iccp/reg(2014)3&doclanguage=en> accessed 8 February 2024.
[860] Ibid, e.g. the link between the initial and envisaged purposes, the context of collection, the nature of the personal data (e.g., special categories) and the possible consequences for data subjects.
[861] Cécile de Terwangne, Commentary of Article 5 in Christopher Kuner, Lee A Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020) 315, 316.
[862] Case C-77/21 *Digi* [2022] ECR I-805 para 36; Case C-77/21 *Digi* [2022] ECR I-805 Opinion of AG Pikamäe paras 28, 59, 60.
[863] Research or statistical purposes in the case of ML.
[864] Article 6 (4) GDPR
[865] See Article 13 (1) GDPR and regulatory guidance which confirms that controllers must not provide individuals about the categories of personal data processed. See Art 29 Working Party, 'Guidelines on transparency under Regulation 2016/679' (WP 260 rev.01, 11 April 2018) at 36.

data processed, must always be communicated to the data subject *at the time of collection* or *prior to the further processing* for a new purpose in compliance with Article 13.3 or Article 14.4'.[866] Regulatory guidance derives this requirement *not* from the transparency principle and the related obligations contained in the GDPR, but from the fairness and purpose limitation principles.[867] By relying on the purpose limitation and fairness principle, regulatory guidance confirms, at least implicitly, the interpretation that controllers are not obliged to inform data subjects about the actual category of the inferred personal data or the detected emotional state (AC) based on Articles 13 and 14 GDPR. This is contradictory to the objectives of the transparency principle, which aim to enable data subjects to (i) become aware of processing[868] and (ii) enforce their rights.[869] It also prevents data subjects from exercising control over the processing of their personal data[870] (see also the profiling problem).

The interpretation that controllers are not obliged to inform data subjects about the actual category of the inferred personal data or the detected emotional state based on Articles 13 and 14 GDPR leads to opacity rather than transparency. Data subjects will not be informed about the inferred personal data generated by ML because there is no specific legal obligation for controllers to do so.[871] Imagine, for example, an insurance company which deploys unsupervised ML techniques to detect patterns and correlations in rather simple personal data such as sex and place of residence of their clients. The AI system detects correlations between sex and place of residence, in particular that women living in certain areas tend to live longer. Based on this correlation, the AI system automatically predicts the life expectancy of these clients and stores this information within the insurance customer relationship management system. Life expectancy constitutes inferred personal data generated by means of unsupervised ML techniques and is based on personal data directly collected from the data subjects. Therefore, the insurance company is not required under the transparency obligations enshrined in the GDPR to inform the data subjects concerning the knowledge gained, namely, the detected correlation and the predicted life expectancy. Controllers are not obliged to inform data subjects about the categories of such inferred personal data. The insurance company must inform the data subject only about the purpose of further processing and any other information mentioned in Article 13 (2) GDPR, but not about the actual personal data generated by the AI systems or at least the categories thereof. Article 13 GDPR, which is applicable in this case,[872] does not contain such a requirement, contrary to Article 14 (1) lit d GDPR.

[866] Art 29 Working Party, 'Guidelines on transparency under Regulation 2016/679' (WP 260 rev.01, 11 April 2018) in Footnote 30 at page 14, emphasis added by the author.
[867] Ibid at page 14.
[868] Recital 39 GDPR.
[869] Articles 15-22 GDPR as well as remedies contained in Articles 77-80 GDPR.
[870] Recital 7 GDPR.
[871] Provided that the privileged 'statistical purpose' or 'research purpose' apply to processing by means of ML.
[872] Because the personal data used by the system as input data was collected from the data subject.

Arguably, inferred data constitute 'new' personal data not collected from the data subject, triggering the transparency obligations contained in Article 14 GDPR. However, this view is not convincing for several reasons. First, Article 14 GDPR clearly covers situations where personal data was collected from third-party sources.[873] The latter is emphasised by the wording contained in Article 14 (2) lit f which requires controllers to disclose 'from which source the personal data originate'. Recital 61 GDPR refers to the situation where the personal data do not originate from the data subject but are 'obtained from another source.' In the example at hand, the inferred personal data originate from the data subjects but not from another source (e.g., a third-party controller). Second, this also makes sense when applying a systematic interpretation. Generating inferred personal data constitutes 'further processing' mentioned in Articles 13 (3) and 14 (4) GDPR. Article 13 (3) GDPR would be obsolete if Article 14 (4) GDPRP would govern the insurance company's further processing. Third, data subjects may enforce their right of access to obtain information about the personal data generated by the AI system. The CJEU has clarified that the scope of a copy under Article 15 (3) GDPR includes personal data *generated* by the *controller*[874] and thus inferred personal data.

The outcome will be the same when personal data are inferred by means of affective computing. Controllers are not required to inform the data subject about the specific detected emotional states or about the category of inferred personal data. Regulatory guidance that suggests otherwise, namely, that controllers need to inform data subjects about the *categories of the inferred data* processed, based on the purpose limitation and fairness principle,[875] may be easily refuted. First, Articles 13 and 14 GDPR implement the transparency principle and impose specific information duties on the controller. These provisions do not include an obligation to inform about the categories of inferred personal data, as suggested by regulatory guidance. Second, controllers already comply with the principles of transparency and purpose limitation by informing data subjects about the purpose for further processing, provided that the latter is compatible with the initial purpose. Third, as outlined in Section 4.3.2, the substantive meaning of the fairness principle is elusive.[876] This makes it easy to challenge the interpretation that controllers must inform data subjects about the categories of inferred personal data, in particular because the transparency obligations enshrined in the GDPR do not entail such a specific obligation. The conclusion that the GDPR does not require controllers to inform data subjects about inferred personal data constitutes a Type 3 legal problem. Articles 13 and 14 GDPR are not fit for purpose to protect the fundamental right to data protection. These provisions fail to achieve the objectives of the transparency principle, namely, enabling data subjects to (i) become aware of

[873] Gabriela Zanfir-Fortuna, Commentary of Article 14 in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020) 436, 445, 446.
[874] Case C-487/21, *F.F.* [2022] ECR I-1000 para 21; see also the opinion of AG Pitruzella paras 45, 70
[875] Art 29 Working Party, 'Guidelines on transparency under Regulation 2016/679' (WP 260 rev.01, 11 April 2018) in footnote 30 at page 14 emphasis added by the author.
[876] Damian Clifford, Jef Ausloos 'Data Protection and the Role of Fairness' (2018) Vol 37 No 1 Yearbook of European Law 130, 187.

processing[877] (ii) enforce their rights[878] and (iii) exercise control over the processing of their personal data[879]. Consequently, these provisions fail to effectively protect data subjects.[880] Data subjects will not be aware of the actual personal data inferred by means of ML or AC, such as the predicted life expectancy or the emotional state detected by the AI system. Therefore, data subjects cannot exercise their rights as a data subject because they are simply not aware of the inferred personal data.

The inference problem (Type 3)

ML and AC enable controllers to infer personal data such as predictions or emotion data based on personal data provided by data subjects or obtained otherwise. Transparency obligations contained in the GDPR do not require controllers to inform data subjects about personal data inferred by means of AI if the data are processed for compatible purposes. Consequently, data subjects do not become aware of such data and cannot exercise their rights. Therefore, Articles 13 and 14 GDPR are not fit for purpose to protect the fundamental right to data protection.

Controllers may predict preferences, behaviour and attitudes of data subjects using ML techniques such as regression, classification (see Section 2.2.1.1) or clustering (Section 2.2.1.2), amounting to profiling as defined in the GDPR. AC empowers controllers to predict an individual's personal state and thus to evaluate particular personal aspects related to that individual. Article 4 (4) GDPR defines profiling as 'any form of automated processing of personal data consisting of the use of personal data to evaluate certain personal aspects relating to a natural person'. In academia, profiling is referred to as 'the process of (i) inferring a set of characteristics about an individual or group of persons (i.e. the process of creating a profile) and/or (ii) treating that person or group (or other persons/groups) in light of these characteristics (i.e. the process of applying a profile)'.[881] AI, particularly ML and AC, can be used for both steps contained in the process of profiling, namely, first to infer a profile by means of unsupervised or supervised ML or predict an individual's emotional state (AC) and subsequently treat the individual accordingly. Controllers may rely on dark patterns to collect personal data required for profiling purposes.[882] Dark patterns are design practices which undermine a user's autonomy by coercing, misleading or manipulating their decision-making and behaviour.[883]

[877] Recital 39 GDPR.
[878] Articles 15-22 GDPR as well as remedies contained in Articles 77-80 GDPR.
[879] Recital 7 GDPR.
[880] Recital 11 GDPR; Case C-645/19 [2021] *Facebook Ireland* ECR I-483 para 45; Case C-319/20, *Meta Platforms Ireland Limited* [2022] ECR I-322 para 73; Joined cases C-293/12 and C-594/12 *Digital Rights Ireland* [2014] ECR-I 238 paras 54 and 66; joined cases C-203/15 and C-698/15 [2016] *Tele 2 Svergie* ECR-1 970 paras 109 and 122; Case C-362/14 *Schrems v Data Protection Commissioner* [2015] ECR I-650 para 91.
[881] Isak Mendoza, Lee A Bygrave, 'The Right Not to be Subject to Automated Decisions Based on Profiling' in Tatiana-Eleni Synodinou et al (eds) *EU Internet Law: Regulation and Enforcement* (Springer International 2017) 77.
[882] Arunesh Mathur, Jonathan Mayer, Mihir Kshirsagar, 'What Makes a Dark Pattern... Dark?' (CHI Conference on Human Factors in Computing Systems, Yokohama, May 2021) 16; Tim Kollmer, Andreas Eckhardt, 'Dark Patterns' (2022) Vol 64 Iss 6 Business & Information Systems Engineering 1.
[883] Sanju Ahuja, Jyoti Kumar, 'Conceptualizations of user autonomy within the normative evaluation of dark patterns' (2022) Vol 24 Iss 4 Ethics and Information Technology 1.

Profiling can have a highly predictive nature[884] and may generate stereotypes by assuming that certain behaviour of an individual, such as receiving good grades at a renowned university, is an indicator for a corresponding outcome, for example, securing a well-paid job.[885] Such predictive profiling may be used to predict an individual's behaviour, character, risk (e.g. score values) and to treat the individual accordingly.[886] For example, an individual's phone-charging habit is currently used as a relevant factor for determining this individual's creditworthiness. AI systems powered by ML in particular assess data points such as phone-charging habits that would commonly not be considered when determining someone's creditworthiness. Smart Finance disclosed that customers who regularly let their phone batteries drop below 12% are not considered good prospects. Another FinTech company called Lenddo states the opposite and considers hyper well-maintained smartphone batteries as a red flag because such a phone-charging habit seems to be robotic or not human enough.[887] In fact, research suggests that behaviour revealed in mobile phone usage accurately predicts the likelihood of credit repayment. By means of ML, the likelihood of repayment was predicted using behavioural features derived from mobile phone usage.[888]

The predictive nature of profiling is also emphasised by Recital 24 GDPR, which states that profiling may be used for analysing or predicting the personal preferences, behaviour and attitudes of data subjects. ML as introduced in Section 2.2.1 is the favoured way of deriving profiles[889] particularly because profiles are patterns resulting from probabilistic processing of data.[890] Apart from obvious examples such as discrimination, risks of profiling relate to the one-sided supply of information (information asymmetry) and the negative influence on the data subject's personal autonomy.[891] Profiling exacerbates the power inequality and information asymmetry between those that profile (controllers) and those that are being profiled (the data subjects).[892] It also threatens personal autonomy by surreptitiously influencing, formatting and customising individual behaviour.[893] The essence of

[884] Helena U Vrabec, 'Uncontrollable: Data Subject Rights and the Data-driven Economy' (Dissertation, Leiden University 2019) 220.
[885] Frederick F Schauer, Profiles, Probabilities, and Stereotypes (Harvard University Press 2006) 6.
[886] Helena U Vrabec, 'Uncontrollable: Data Subject Rights and the Data-driven Economy' (Dissertation, Leiden University 2019) 220; Hans Lammerant, Paul de Hert, 'Predictive profiling and its legal limits: Effectiveness gone forever' In Bart van der Sloot et al (eds) *Exploring the boundaries of big data* (2016 Amsterdam University Press/WRR) 145-173.
[887] Tanya Goodin, 'The battery life of your phone could affect your loan application' (2022) <https://tanya-goodin.com/2022/08/credit-rating-algorithmic-transparency/> accessed 8 February 2024.
[888] Daniel Björkegren, Darrell Grissen, 'Behavior Revealed in Mobile Phone Usage Predicts Credit Repayment' (2020) Vol 34 Iss 3 The World Bank Economic Review 618, 623.
[889] Lilian Edwards, Michael Veale, 'Slave to the Algorithm: Why a 'Right to Explanation' is Probably not the Remedy You are Looking for' (2017) Vol 16 Iss 1 Duke Law & Technology Review 19, 46.
[890] Helena U Vrabec, 'Uncontrollable: Data Subject Rights and the Data-driven Economy' (Dissertation, Leiden University 2019) 220.
[891] Bart Custers, 'Data Dilemmas in the Information Society' in Bart Custers et al (eds), *Discrimination and Privacy in the Information Society* (Springer 2013) 1.
[892] Mireille Hildebrandt, Bert-Jaap Koops, 'The challenges of Ambient Law and Legal Protection in the Profiling Era' (2010) Vol. 73 (3) The Modern Law Review 428, 435; Serge Gutwirth, Mireille Hildebrandt, 'Some Caveats on Profiling' in Serge Gutwirth et al (eds), *Data Protection in a Profiled World* (Springer Nature 2010) 34.
[893] Serge Gutwirth, Mireille Hildebrandt, 'Some Caveats on Profiling' in Serge Gutwirth et al (eds), *Data Protection in a Profiled World* (Springer Nature 2010) 34; Serge Gutwirth, *Privacy and the information age* (Lanham:Rowman & Littlefield Publishers 2002).

autonomy is indicated by the etymology of the term: autos (self) and nomos (rule or law).[894] Put simply, autonomy refers to a person's ability to make rational and uncoerced choices and decisions[895] or, in other words, to 'make their own lives'[896] and face freely both existential and every day's choices.[897] As noted by AG Pikamäe, profiling may reinforce existing stereotypes, increase the social divide, restrict the data subject's freedom of choice regarding certain products or services and result in the denial of services.[898] Profiling deprives data subjects not only of the means to reflect on the choices the environment makes for them, but may proactively impact the choices they make. This is called 'the autonomy trap'.[899] I now outline why the GDPR fails to address the information asymmetry concerning profiling and the subsequent negative impact on the data subject's personal autonomy.

Profiling defined in Article 4 (4) GDPR refers to any form of automated processing ('regular profiling') and also covers profiling with subsequent human involvement, as opposed to profiling used for ADM ('ADM profiling') which must be fully automated and satisfy the two other cumulative requirements of Article 22 (1) GDPR.[900] According to regulatory guidance, ADM has a different scope than regular profiling but may partially overlap with or result from profiling (see also Section 3.3.4.6). Decisions which are not solely automated according to Article 22 GDPR might also include profiling.[901] Regulatory guidance dealing with the transparency principle stresses the importance of informing data subjects about the consequences of processing, also with regard to regular profiling and not only ADM profiling which is captured by Article 22 GDPR.[902] This information duty is derived from Recital 60 GDPR stating that data subjects 'should be informed of the existence of profiling and the consequences of such profiling'. Interestingly, regulatory guidance with respect to ADM adopted *prior* to the transparency guidelines stresses that if ADM and profiling '*does not* meet the Article 22 (1) definition it is *nevertheless good practice* to provide' the information according to Article 13 (2) lit f and 14 (2) lit g.[903] These two provisions oblige controllers to inform data subjects about 'the existence of automated decision-making, *including profiling*, referred to in Article 22(1) and (4) and, *at least in those cases*, meaningful information about the logic involved, as well as the *significance and the envisaged consequences* of such processing for the data subject'.[904] Because these provisions contain the wording '*at least in those cases*', controllers are *not* legally required to inform data

[894] Gerald Dworkin, *The Theory and Practice of Autonomy* (Cambridge University Press 1988) 12, 18.
[895] Maurits Clemens Kapitein, 'Personalized Persuasion in Ambient Intelligence' (Doctoral Thesis, TU/e Eindhoven 2012) 179 < https://pure.tue.nl/ws/files/3470131/729200.pdf> accessed 8 February 2024.
[896] Joseph Raz, *The Morality of Freedom* (Oxford University Press 1986) 369.
[897] Daniel Susser, Beate Roessler, Helen Nissenbaum 'Technology, autonomy, and manipulation' (2019) Vol 8 Iss 2 Internet Policy Review 1, 8.
[898] Case C-634/21, *SCHUFA Holding AG* [2023] ECR I-220, Opinion AG Pikamäe Footnote 6 in para 19.
[899] Tal Z. Zarsky, 'Mine your own business!' (2003) 5 Yale Journal of Law and Technology 35.
[900] ADM profiling must involve a decision and has to produce legal or similarly significant effects see Case C-634/21, *SCHUFA Holding AG* [2023] ECR I-957 para 43.
[901] Article 29 Working Party, 'Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679' (WP 251rev.01, 6 February 2018) at 7 and 8.
[902] Art 29 Working Party, 'Guidelines on transparency under Regulation 2016/679' (WP 260 rev.01, 11 April 2018) at 41.
[903] Article 29 Working Party, 'Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679' (WP 251rev.01, 6 February 2018) at 25.
[904] Emphasis added by the author.

subjects about the significance and envisaged consequences of 'regular' profiling as this obligation solely applies to ADM profiling meeting the three cumulative conditions of Article 22 GDPR.[905] Even regulatory guidance confirms this reading: It is 'good practice' to provide this information also regarding regular profiling.[906]

Furthermore, the preparatory documents of the GDPR confirm this interpretation. During the GDPR negotiation process, Poland suggested to use the wording 'where applicable, information about the existence of profiling referred to in Article 4 (12a) *and/or* about automated decision-making'[907] instead the final wording of Article 13 (2) lit f and 14 (2) lit g GDPR. Therefore, it is likely that if the intent of the legislator was to oblige controllers to provide data subjects with information on the importance and implications envisaged of regular profiling, the final language of Articles 13 (2) lit f and 14 (2) lit g GDPR would contain a specific reference to the definition of profiling. The objection that controllers are in fact obliged to disclose such information regarding regular profiling based on Recital 60 GDPR is not very strong. Recitals may cast light on the interpretation to be given to a rule, but cannot in itself constitute such a rule.[908] In addition to that, recitals are legally not binding.[909] The results of regular profiling might constitute 'new' inferred personal data. However, as discussed in the inference problem, Article 14 GDPR does not apply to inferred personal data originating from data provided by the data subject. Instead, Article 14 GDPR applies where personal data have been obtained from a source other than the data subject (third party).

Thus, the transparency principle as implemented in Articles 13 (2) lit f and 14 (2) lit g GDPR does not require controllers to inform data subjects about the significance and consequences of regular profiling as defined in Article 4 (4) GDPR. These provisions fail to achieve the objectives of the transparency principle, namely, enabling data subjects to (i) become aware of processing[910] and (ii) enforce their rights.[911] The fact that data subjects will not be informed about the significance and consequences of regular profiling also sharpens the power inequality and information asymmetry

[905] The CJEU confirmed that three cumulative conditions must be met to render Article 22 GDPR applicable see Case C-634/21, *SCHUFA Holding AG* [2023] ECR I-957 para 43.
[906] Article 29 Working Party, 'Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679' (WP 251rev.01, 6 February 2018) at 25.
[907] Emphasis added by the author. Also, note that Article 4 (12a) refers to the definition of profiling as finally enshrined in Article 4 (4) GDPR. Council of the European Union, General Data Protection Regulation Interinstitutional File: 2012/0011 (COD) (2015) at 117 < https://data.consilium.europa.eu/doc/document/ST-9281-2015-INIT/en/pdf > accessed 8 February 2024.
[908] Case T-709/21, WhatsApp Ireland Ltd [2022] ECR I-783 para 71.
[909] Case C-162/97, *Nilsson* [1998] ECR I-7477, para. 54.
[910] Recital 39 GDPR.
[911] Articles 15-22 GDPR as well as remedies contained in Articles 77-80 GDPR.

between controllers and data subjects instead of mitigating them.[912] The provisions also neglect the corresponding adverse effect on the data subject's personal autonomy.[913]

Profiling can be used to predict an individual's behaviour, character, risk (e.g., score values), evaluate an individual's personal aspects (e.g., emotional states) and to treat the individual accordingly.[914] The latter may involve (i) limiting the choices available to an individual[915] or (ii) proactively pushing the individual to make a certain decision. In terms of (i), AG Pikamäe notes that profiling may restrict the data subject's freedom of choice regarding certain products or services and result in the denial of services.[916] For example, a negative score value based on profiling limits the choices available for individuals to obtain a loan or even mobile subscriptions.[917] The limited choice undermines the individual's autonomy to 'make their own lives'[918] and face freely both existential and every day's choices.[919] In terms of (ii), AI-powered profiling enables controllers to push a person towards choices it may have resisted if being aware of what is known about him or her.[920] AI entails the characteristics of a persuasive technology, which is an 'interactive computing system designed to change people's attitudes and behaviours'.[921] This holds particularly true where companies use AI to influence consumers by tailoring their products and services to their needs, interests, personality or other factors relevant for them.[922] Companies analyse any kind of customer behaviour for profiling purposes and the gained knowledge is then used to proactively change the behaviour and decisions of these customers, which is called 'actuation'.[923] Persuasion is seen as an 'attempt to change attitudes or behaviour or both' without making use of practices such as coercion or deception.[924] Behaviour also includes decisions taken by individuals.

[912] Mireille Hildebrandt, Bert-Jaap Koops, 'The challenges of Ambient Law and Legal Protection in the Profiling Era' (2010) Vol. 73 (3) The Modern Law Review 428, 435; Serge Gutwirth, Mireille Hildebrandt, 'Some Caveats on Profiling' in Serge Gutwirth et al (eds), *Data Protection in a Profiled World* (Springer Nature 2010) 34.
[913] Bart Custers, 'Data Dilemmas in the Information Society' in Bart Custers et al (eds), *Discrimination and Privacy in the Information Society* (Springer 2013) 1; Tal Z. Zarsky, 'Mine your own business!' (2003) 5 Yale Journal of Law and Technology 35.
[914] Helena U Vrabec, 'Uncontrollable: Data Subject Rights and the Data-driven Economy' (Dissertation, Leiden University 2019) 220; Hans Lammerant, Paul de Hert, 'Predictive profiling and its legal limits: Effectiveness gone forever' In Bart van der Sloot et al (eds) *Exploring the boundaries of big data* (2016 Amsterdam University Press/WRR) 145-173.
[915] Tal Z. Zarsky, 'Mine your own business!' (2003) 5 Yale Journal of Law and Technology 35
[916] Case C-634/21, *SCHUFA Holding AG* [2023] ECR I-220, Opinion AG Pikamäe Footnote 6 in para 19.
[917] Case C-203/22 *Dun & Bradstreet Austria* p 2 <https://www.ris.bka.gv.at/Dokumente/Lvwg/LVWGT_WI_20220211_VGW_101_042_791_2020_44_00/LVWGT_WI_20220211_VGW_101_042_791_2020_44_00.pdf> accessed 8 February 2024.
[918] Joseph Raz, *The Morality of Freedom* (Oxford University Press 1986) 369.
[919] Daniel Susser, Beate Roessler, Helen Nissenbaum 'Technology, autonomy, and manipulation' (2019) Vol 8 Iss 2 Internet Policy Review 1, 8.
[920] Hildebrandt Mireille, Koops Bert-Jaap, 'The challenges of Ambient Law and Legal Protection in the Profiling Era' (2010) Vol. 73 (3) The Modern Law Review 428, 436.
[921] Brian Jeffrey Fogg, *Persuasive Technology*: *Using Computers to Change What We Think and Do* (EBSCO Publishing 2003), 1.
[922] Ibid 38.
[923] Shoshana Zuboff, *The age of surveillance capitalism* (Public Affairs 2019) 204, 293.
[924] Brian Jeffrey Fogg, *Persuasive Technology*: *Using Computers to Change What We Think and Do* (EBSCO Publishing 2003) 16.

Machines may predict the emotional states of individuals which, in turn, can easily be linked with other information.[925] Advanced behavioural inference systems as discussed in Section 2.2.4.3 powered by AI (particularly AC, DL and CV) allow fine-grained tracking of shoppers' behaviour, enabling retailers to analyse the collected data for profiling purposes and place personalised offers based on nuanced insights of individuals' behaviour and profiles.[926] For example, Facebook developed CV and AC-powered systems that feed staff in a retail store with information on their customers based on customers user profiles. The information can include detected emotions of the customers and enable retailers to target them with specific products informed by their Facebook activity and detected emotional states.[927] Because emotions play an important role in the elicitation of autonomous motivated behaviour,[928] AC may be used against the interests of the person concerned, namely, to persuade or manipulate this individual.[929] Understanding emotions increases the scope to influence decision-making of individuals and makes practices such as manipulation more effective.[930] Applications of AC may affect people's decisions and lives in ways that undermine their autonomy because emotions can influence decision-making powerfully, predictably and pervasively.[931] To be autonomous presupposes that a person has the capacity of self-reflection and rationality. A person must also enjoy 'procedural independence', meaning not to be under the influence of factors that comprise her capacities for self-reflection and rationality.[932] Information about a person's emotional state has implications for procedural independence: if it becomes available, it can restrict options in ways that a person would not choose herself.[933] The capacity for emotion to influence decision-making, combined with the ability to detect emotion by means of AC, strongly impacts an individual's personal autonomy.[934]

With the help of AI, manipulation and persuasion can be automated. Research suggests that intelligent software agents can significantly influence human behaviour.[935] Automated manipulation or

[925] Holger Baumann, Sabine Dörig, 'Emotion-Oriented Systems and the Autonomy of Persons' in Paolo Petta, Catherine Pelachaud, Roddie Cowie (eds) *Emotion-Oriented Systems* (Springer 2011) 745.
[926] Vasilios Mavroudis, Michael Veale 'Eavesdropping Whilst You're Shopping: Balancing Personalisation and Privacy in Connected Retail Spaces' (Living in the Internet of Things Conference, London, March 2018)1, 2 <https://ieeexplore.ieee.org/document/8379705> accessed 8 February 2024.
[927] Katie Gibbons, 'Facebook develops facial recognition cameras that feed shop staff their customers' profile details' *The Times* (London, 01 December 2017) <https://www.thetimes.co.uk/edition/news/facebook-develops-facial-recognition-cameras-that-feed-shop-staff-their-customers-profile-details-58lx0jckt> accessed 8 February 2024.
[928] Leen Vandercammen et al, 'On the Role of Specific Emotions in Autonomous and Controlled Behaviour' (2014) Vol 28 Iss 5 European Journal of Personality 413, 445.
[929] Rosalind W Picard, *Affective Computing* (MIT Press 1997) 136.
[930] Andrew McStay, Lachlan Urquhart 'This time with feeling? Assessing EU data governance implications of out of home appraisal based emotional AI' (2019) Vol 24 No 10 First Monday <https://firstmonday.org/ojs/index.php/fm/article/view/9457/8146> accessed 8 February 2024.
[931] Jennifer S Lerner et al, 'Emotion and Decision Making' (2015) Vol 66 Annual Review of Psychology 799, 802.
[932] Holger Baumann, Sabine Dörig, 'Emotion-Oriented Systems and the Autonomy of Persons' in Paolo Petta, Catherine Pelachaud, Roddie Cowie (eds) *Emotion-Oriented Systems* (Springer 2011) 735, 736, 739.
[933] Roddy Cowie, 'Ethical Issues in Affective Computing' in Rafael Calvo et al (eds), *The Oxford Handbook of Affective Computing* (OUP 2015) 340.
[934] Damian Clifford, 'Citizen-consumers in a personalised Galaxy: Emotion influenced decision-making, a true path to the dark side?' (2017) CiTiP Working Paper 31/2017, 13 <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3037425> accessed 8 February 2024.
[935] Christopher Burr, Nello Cristianini, James Lydmann, 'An Analysis of the Interaction Between Intelligent Software Agents and Human Users' (2018) Vol 28 Minds and Machines 735, 752.

persuasion may lead to so-called 'impulse buying', where individuals make their decisions spontaneously and are dominated by emotions. Impulsive buying occurs when individuals experience an urge to buy a product, without thoughtful consideration why one needs a specific product.[936] In 2013, Amazon was granted a US patent called 'method and system for anticipatory package shipping'.[937] Based on AI-powered profiling applications that analyse a customer's historical shopping patterns by means of ML, Amazon predicts whether a customer is interested in a certain product. Then, Amazon sends that product to the customer, even if there has not been an order placed beforehand. In some situations, e.g. if the customer is 'particularly valued (e.g., *according to past ordering history, appealing demographic profile, etc.*), delivering the package to the given customer as a promotional gift may be used to build goodwill.'[938] Arguably less 'valued' customers can be provided with a discount in order to convert the potential interest in an order.[939] This is a prime example of how AI-powered profiling may be used to proactively push an individual to make a certain decision. Such profiling predicts the individual's interests and is then used to intentionally and covertly influence the person's decision-making, i.e. pushing to buy a certain product. AI-powered profiling undermines the sense of autonomy that consumers seek in their decision-making. The autonomy in choice is akin to exercising free will, and self-determination is a state of exercising one's autonomy.[940] Aggregation and analysis of data by means of profiling powerfully enhance the range of influence that marketers can have in shaping people's choices and actions.[941]

The examples in the previous paragraphs outline that AI-powered profiling may influence individuals in ways that adversely affect their autonomy and capacity to understand and author their own lives.[942] Treating individuals based on information gained from AI-powered profiling may (i) limit the choices available to an individual[943] or (ii) proactively push an individual towards a certain decision. This impacts the individual's ability to make rational and uncoerced choices and decisions.[944] It deprives data subjects not only of the means to reflect on the choices the environment makes for them, but

[936] Verhagen Tilbert, van Dolen Willemijn 'The influence of online store beliefs on consumer online impulse buying: A model and empirical application' (2011) Vol. 48 Information & Management 320.
[937] Spiegel Joel et al., 'Method and System for anticipatory Package Shipping' US Patent US 8615473B2 (Assignee: Amazon Technologies, Inc.) December 2013 <https://patentimages.storage.googleapis.com/8a/67/ff/299703230243b5/US8615473.pdf>, accessed 8 February 2024.
[938] Spiegel Joel et al., 'Method and System for anticipatory Package Shipping' US Patent US 8615473B2 (Assignee: Amazon Technologies, Inc.) December 2013 <https://patentimages.storage.googleapis.com/8a/67/ff/299703230243b5/US8615473.pdf>, accessed 8 February 2024.
[939] Ibid.
[940] André Quentin et al, 'Consumer Choice and Autonomy in the Age of Artificial Intelligence and Big Data' (2018) Vol 5 Customer Needs and Solutions 28, 29.
[941] Helen Fay Nissenbaum, *Privacy in Context: Technology, Policy, and the Integrity of Social Life* (Stanford Law Books 2010) 83.
[942] Daniel Susser, Beate Roessler, Helen Nissenbaum 'Technology, autonomy, and manipulation' (2019) Vol 8 Iss 2 Internet Policy Review 1, 13.
[943] Tal Z. Zarsky, 'Mine your own business!' (2003) 5 Yale Journal of Law and Technology 35
[944] Maurits Clemens Kapitein,'Personalized Persuasion in Ambient Intelligence' (Doctoral Thesis, TU/e Eindhoven 2012) 179 < https://pure.tue.nl/ws/files/3470131/729200.pdf> accessed 8 February 2024.

proactively impact the choices they make.[945] An individual is autonomous when its decisions and actions are its own and thus self-determined.[946] In the examples mentioned, the individual is no longer autonomous. Individuals no longer act themselves; instead, they are acted upon.[947]

By not requiring controllers to inform data subjects about the significance and consequences of regular profiling, Articles 13 (2) lit f and 14 (2) lit g GDPR fail to achieve the objectives of the transparency principle.[948] These provisions also neglect possible harms of regular profiling, in particular the sharpening of power and information asymmetries and the adverse effects on the data subject's personal autonomy. Ultimately, the concept of control is a common denominator of transparency, power symmetry and autonomy.[949] The concept of control is not defined in the GDPR, although it was one of the main reasons for the data protection reform[950] and constitutes one of the GDPR's legislative aims, namely, that 'natural persons should have control of their own personal data'.[951] The GDPR does not contain an enforceable right specifically dedicated to the concept of control. Control seems to emerge from the concept of informational self-determination. It was interpreted as individual informational control or empowerment, i.e. the ability of a natural person to control the terms under which their personal information is acquired and used.[952] Control in this sense is subsequently often presented as the hallmark of data protection law[953] and is attributed with the role of a normative anchor for personal data protection as a fundamental right.[954]

Control-related provisions in data protection law can be classified in two mechanisms: consent and data subject rights.[955] In fact, control in the context of the fundamental right to data protection, and particularly as implemented in the GDPR, grants data subjects the possibility to act,[956] i.e. to invoke their data subject rights enshrined in Articles 15-22 GDPR or enforce their rights to lodge a complaint with a SA or their right to an effective judicial remedy against the controller (Articles 77 and 79

[945] Mireille Hildebrandt, Bert-Jaap Koops, 'The challenges of Ambient Law and Legal Protection in the Profiling Era' (2010) Vol. 73 (3) The Modern Law Review 428, 435.
[946] Gerald Dworkin, *The Theory and Practice of Autonomy* (Cambridge University Press 1988) 13.
[947] See Berlin, which explains the concept of autonomy under the heading positive liberty: 'Isaiah Berlin, *Liberty* (Hendry Hardy ed Oxford University Press 1969) 185; Marijn Sax, *Between Empowerment and Manipulation* (Kluwer Law International B.V. 2021) 131.
[948] Namely enabling data subjects to (i) become aware of processing according to Recital 39 GDPR, (ii) enforce their rights according to Articles 15-22 GDPR as well as remedies contained in Articles 77-80 GDPR and (iii) exercise control over the processing of their personal data Recital 7 GDPR.
[949] Helena U Vrabec, *Data Subject Rights under the GDPR* (OUP 2021) 48.
[950] Case C-300/21 *UI* [2022] ECR I-756 Opinion AG Manuel Campos Sánchez-Bordona paras 69, 70.
[951] Recital 7 GDPR.
[952] Mary J Culnan, 'Protecting Privacy Online: Is Self-Regulation Working?' (2000) Vol 19 Iss 1 Journal of Public Policy & Marketing 20-26.
[953] Antoinette Rouvroy, Yves Poullet, 'The Right to Informational Self-determination and the Value of Self-development: Reassessing the Importance of Privacy for Democracy' in Serge Gutwirth et al (eds), *Reinventing Data Protection?* (Springer 2009) 68; Helena U Vrabec, *Data Subject Rights under the GDPR* (OUP 2021) 55.
[954] Helena U Vrabec, *Data Subject Rights under the GDPR* (OUP 2021) 54; Stefano Rodotà, 'Data Protection as a Fundamental Right' in Serge Gutwirth et al (eds), *Reinventing Data Protection?* (Springer 2009) 79; Orla Lynskey, *The Foundations of EU Data Protection Law* (OUP 2015).
[955] Christophe Lazaro, Daniel Le Métayer, 'The Control over Personal Data: True Remedy or Fairy Tale?' (2015) Vol 12 Iss 1 SCRIPT-ed 1, 16–17; Helena U Vrabec, *Data Subject Rights under the GDPR* (OUP 2021) 59.
[956] Julie E Cohen, 'Affording Fundamental Rights' (2017) Volume 4 Iss 1 Critical Analysis of Law 78, 81

GDPR). Data subjects need to invoke their rights to exercise control over the processing of their personal data. Therefore, control in the sense of the GDPR seems to be rather limited from a conceptual point of view. In a preliminary ruling, even the AG stated that 'the scope for individual action is limited' and 'confined to the exercise of those rights in specified circumstances'.[957] The AG interprets the concept of control under the GDPR as 'rights of supervision and intervention in operations carried out by others on their data, as one tool […] for the protection of those data'.[958] Also, consent, the other mechanism for data subjects to exercise control over processing, is rather limited. Consent is just one of the legal bases in the GDPR and simply empowers the data subject to accept or reject the processing of personal data suggested by a controller. It does not otherwise empower them to intervene or influence how controllers process their personal data.[959] In my view, enforceable data subject rights are the main, though limited, mechanism for data subjects to exercise control over the processing of their personal data under the GDPR.

Articles 13 and 14 GDPR fail to achieve the GDPR's objective for data subjects to be able to exercise control over the processing of their personal data.[960] Transparency is a necessary precondition for control,[961] and without being informed about the significance and possible consequences of profiling, data subjects cannot exercise control over processing by enforcing their rights (e.g., object to profiling or lodging a complaint with an SA). It could be argued that controllers need to inform data subjects about the significance and possible consequences of profiling based on Article 22 (3) GDPR. According to this provision, controllers need to 'implement suitable measures to safeguard the data subject's right and freedoms', enabling them to obtain human intervention, to express their point of view and to contest automated decision-making. However, this obligation is only triggered if profiling involves automated decision making in the sense of Article 22 GDPR. 'Regular profiling' as defined in Article 4 (4) GDPR does not trigger the obligation contained in Article 22 (3) GDPR (see also Section 5.11). This is problematic because the concept of control is a common denominator of transparency, power symmetry and autonomy.[962] With transparent data processing and effective individual control over processing of personal data, data subject's risks to autonomy generally and manipulation particularly could be reduced.[963] Therefore, these provisions are not fit for purpose to effectively protect[964] the fundamental right to data protection. The CJEU has repeatedly stressed that EU data protection law

[957] Case C-300/21 *UI* [2022] ECR I-756 Opinion AG Manuel Campos Sánchez-Bordona para 72.
[958] Ibid para 71.
[959] Ibid para 73.
[960] Recital 7 GDPR.
[961] Helena U Vrabec, *Data Subject Rights under the GDPR* (OUP 2021) 48.
[962] Helena U Vrabec, *Data Subject Rights under the GDPR* (OUP 2021) 48.
[963] Frederik Zuiderveen Borgesius, 'Improving Privacy Protection in the area of Behavioural Targeting' (Doctoral thesis, Universiteit van Amsterdam 2015) 127 <https://dare.uva.nl/search?identifier=c74bdba6-616c-4cd9-925e-33a5858935e5> accessed 8 February 2024.
[964] Case C-645/19 [2021] *Facebook Ireland* ECR I-483 para 45; Case C-319/20, *Meta Platforms Ireland Limited* [2022] ECR I-322 para 73.

aims to effectively protect the data subject's personal data against risk of misuse.[965] Because data subjects are not informed about the significance and consequences of regular profiling, they cannot exercise control[966] over such processing. Therefore, misuse of personal data with adverse effects on personal autonomy cannot be prevented. Articles 13 and 14 fail to ensure a high level of protection[967] and provide data subjects with control over the processing of their personal data.

The profiling problem (Type 3)

ML and AC facilitate profiling as defined in the GDPR. Articles 13 (2) lit f and 14 (2) lit g GDPR do not require controllers to inform data subjects about the significance and consequences of profiling not involving ADM. These provisions sharpen the information asymmetries between controllers and data subjects instead of mitigating them, which may lead to adverse effects on the data subject's personal autonomy. The transparency principle embodied in Articles 13 & 14 GDPR is not fit for purpose to effectively protect the fundamental right to data protection.

4.5 Purpose limitation

The purpose limitation principle as introduced in Section 3.3.3.4 demands data to be collected for specified, explicit and legitimate purposes. In addition, personal data shall not be further processed in a manner which is *incompatible* with those legitimate purposes.[968]

4.5.1 Legal problems: Type 1

Generally, all AI disciplines are at odds with the purpose limitation principle. This conflict has not remained unnoticed in academia.[969] Natural language processing (NLP) relies on the processing of text or speech originating from conversations in various contexts; AC relies on video footage recorded during job interviews to detect emotional states; CV uses CCTV footage initially recorded for security purposes to identify individuals based on their gait. AR is devoted to answering questions from diverse data without human intervention, including decision-making.

The tension with the purpose limitation principle particularly applies to ML, which extracts models and properties from training data and recursively derives more data. Thus, data often goes through a

[965] Joined Cases C-293/12 and C-594/12 *Digital Rights Ireland* [2014] ECR-I 238 paras 54 and 66; Joined Cases C-203/15 and C-698/15 [2016] *Tele 2 Sverige* ECR-1 970 paras 109 and 122; Case C-362/14 *Schrems v Data Protection Commissioner* [2015] ECR I-650 para 91.
[966] Recital 7 GDPR.
[967] Recitals 6, 10 GDPR; Case C-534/20, *Leistritz AG* [2022] ECR I-594 para 26; Case C-645/19 [2021] *Facebook Ireland* ECR I-483 para 45; Case C-511/18, *La Quadrature du Net* [2020] ECR I-791 para 106; Case C-131/12, *Google Spain* [2014] ECR I-317 para 66; Case C-154/21 *Österreichische Post AG* [2023] ECR I-3 para 44; Case C-132/21 *Nemzeti* [2023] ECR I-2 para 42.
[968] Cécile de Terwangne, Commentary of Article 5 in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020) 315.
[969] For an overview, see Footnote 27 in Merel Elize Koning, 'The purpose and limitations of purpose limitation' (Doctoral thesis, Radboud University Nijmegen 2020) 4 < https://repository.ubn.ru.nl/bitstream/handle/2066/221665/221665.pdf?sequence=1&isAllowed=y > accessed 8 February 2024.

series of computations arguably for varying purposes.[970] The requirement stemming from the purpose limitation principle that personal data shall be processed for predefined *explicit* purposes is difficult if not impossible to comply with in the context of ML. The explicitness requirement demands controllers to clearly reveal, explain and expose the processing purpose to ensure an unambiguous understanding of the purpose. Notably, the purpose must be explicit and determined at the time of data collection i.e. ex-ante.[971] Unsupervised ML processes data for *in*explicit purposes – the processing *itself* determines the purpose since its goal is to detect patterns and correlations, gain knowledge and make accurate predictions. This makes it impossible to comply with the explicitness requirement ex-ante, i.e. at the time of data collection.

Purpose *specification* is particularly challenging to reconcile with unsupervised ML because it is often used without very specific objectives.[972] Thus, the challenges of defining a purpose for processing and only using the corresponding personal data for that purpose are exacerbated.[973] As indicated in regulatory guidance, it may be impossible to predict what the algorithm will learn, and the purpose may alter given that algorithms used in AI learn and develop over time.[974] Unsupervised ML seems to be at odds with the very core of the purpose limitation principle because it aims to identify associations and patterns among a set of input data. This would be the case if a bank uses unsupervised ML in order to identify associations and patterns in Facebook activities of its potential customers that could be useful for the bank.[975] In general, unpredictability of outcomes in the context of ML processing is considered one of the characteristic features of ML analytics.[976] ML leads to the discovery of patterns that were unimageable previously.[977] Unsupervised ML processes data for *un*specified and *in*explicit purposes – the processing *itself* predicts the purpose of the future use of the data since its goal is to detect patterns and correlations, gain knowledge and make accurate predictions. However, processing personal data for unspecified purposes as in the case of unsupervised ML is unlawful because the

[970] Yinzhi Cao, Junfeng Yang, 'Towards Making Systems Forget with Machine Unlearning' (IEEE Symposium on Security and Privacy, San Jose, May 2015) <https://www.ieee-security.org/TC/SP2015/papers-archived/6949a463.pdf> accessed 8 February 2024.
[971] Merel Elize Koning, 'The purpose and limitations of purpose limitation' (Doctoral thesis, Radboud University Nijmegen 2020) 68, 70 < https://repository.ubn.ru.nl/bitstream/handle/2066/221665/221665.pdf?sequence=1&isAllowed=y > accessed 8 February 2024.
[972] Lee A Bygrave, 'Machine Learning, Cognitive Sovereignty and Data Protection Rights with Respect to Automated Decisions' (2021) University of Oslo Faculty of Law Legal Studies Research Paper Series No. 202-35 at 22 <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3721118> accessed 8 February 2024.
[973] Christopher Kuner et al, 'Expanding the artificial intelligence-data protection debate' (2018) Vol 8 No 4 International Data Privacy Law 289, 290.
[974] Norwegian Data Protection Authority, 'Artificial Intelligence and Privacy' (2018) 4 <https://www.datatilsynet.no/globalassets/global/english/ai-and-privacy.pdf> accessed 8 February 2024.
[975] Norwegian Data Protection Authority, 'Artificial intelligence and privacy' (2018) 17 <https://www.datatilsynet.no/globalassets/global/english/ai-and-privacy.pdf> accessed 8 February 2024.
[976] Nadezha Purtova, 'The law of everything. Broad concept of personal data and future of EU data protection law' (2018) Vol 10 No 1 Law, Innovation and Technology 40, 56.
[977] Whereas Zarsky draws this conclusion in the context of Big Data, it is also valid with regard to ML, because its aim is to detect and extrapolate patterns; Tal Z Zarsky 'Incompatible: The GDPR in the Age of Big Data' (2017) Vol 47 Iss 4 Seton Hall Law Review 996, 1006.

scope of processing is not sufficiently delineated.[978] The purpose limitation principle prohibits unspecified processing and the explicitness requirement demands controllers to ensure an unambiguous understanding of the processing purpose at the time of data collection.[979] This violates the purpose limitation principle and leads to a Type 1 legal problem.

The inexplicitness problem (Type 1)

All AI disciplines process personal data originating from various sources for a plethora of other purposes. Also, ML processes personal data for unspecific and inexplicit purposes – the processing itself determines the purpose and future use of the personal data. Such processing violates the purpose limitation principle.

AI is prone to cause function creep and secondary use. Function creep refers to situations where 'previously authorised arrangements…now being applied to purposes and targets beyond those envisaged at the time of installation.'[980] In the context of data protection law, function creep occurs when personal data initially collected for a specific purpose are subsequently used beyond what was originally understood and considered socially, ethically and legally acceptable.[981] Secondary use, i.e. using data for purposes other than the initial collection purpose, could be seen as a violation of the purpose limitation principle according to data protection law.[982] Function creep is prohibited when such secondary use goes beyond the purposes specified in advance,[983] if the purpose for further processing is not compatible with the initial purpose (see Section 4.5.2). As already outlined in Chapter 2, data needed for the development and deployment of AI are enormous. AI relies on data from different sources initially collected for different purposes.[984] In addition, ML extracts models and properties from training data and recursively derives more data. Thus, data often goes through a series of computations arguably for different purposes.[985] However, the purpose limitation principle demands data to be collected for specified, explicit and legitimate purposes and not further processed in a manner which is incompatible with those purposes.[986]

[978] Cécile de Terwangne, Commentary of Article 5 in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020) 315.
[979] Merel Elize Koning, 'The purpose and limitations of purpose limitation' (Doctoral thesis, Radboud University Nijmegen 2020) 58, 68, 70 < https://repository.ubn.ru.nl/bitstream/handle/2066/221665/221665.pdf?sequence=1&isAllowed=y > accessed 8 February 2024.
[980] Richard Fox, 'Someone to watch over us: Back to the panopticon?' (2001) Vol 1 Iss 3 Criminal Justice 251, 261.
[981] Johanne Yttri Dahl, Ann Rudinow Sætnan, 'It all happened so slowly – On controlling function creep in forensic DNA databases' (2009) Vol 37 International Journal of Law, Crime and Justice 83, 84.
[982] Frederik Zuiderveen Borgesius, 'Improving Privacy Protection in the area of Behavioural Targeting' (Doctoral thesis, Universiteit van Amsterdam 2015) 117 <https://dare.uva.nl/search?identifier=c74bdba6-616c-4cd9-925e-33a5858935e5> accessed 8 February 2024.
[983] Bart Custers, Helena Ursic, 'Big data and data reuse: a taxonomy of data reuse for balancing big data benefits and personal data protection' (2016) Vol 6 Iss 1 International Data Privacy Law 1, 6.
[984] CIPL, 'Artificial Intelligence and Data Protection in Tension' (2018) 13 <https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl_ai_first_report_-_artificial_intelligence_and_data_protection_in_te....pdf> accessed 8 February 2024.
[985] Yinzhi Cao, Junfeng Yang, 'Towards Making Systems Forget with Machine Unlearning' (IEEE Symposium on Security and Privacy, San Jose, May 2015) <https://www.ieee-security.org/TC/SP2015/papers-archived/6949a463.pdf> accessed 8 February 2024.
[986] Cécile de Terwangne, Commentary of Article 5 in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020) 315.

AI seems to pose inherent risks of function creep. Smart home technologies based on AI such as Google's 'Nest thermostat' collect data about residents' behaviour and gather data from other connected products or devices such as cars, ovens, lights, TVs, game consoles, kettles, fitness trackers, beds and Google's digital assistant.[987] Collected data can be shared with Google's patented 'Privacy-aware personalised content for the smart home' AI system, which enables secondary use of collected data by companies to draw inferences from the generated home data (e.g. when residents are at home, when they shower, when they cook, when they watch TV and when they sleep). The patent states that 'the answers to these questions may help third parties benefit consumers by providing them with interesting information, products and services as well as with providing them with targeted advertisements'.[988]

Secondary use of data is also likely to occur in the context of virtual assistants that deploy NLP and speech recognition techniques based on RL and approaches from the specific kind of ML called deep learning (DL). Amazon's US patent 'Keyword Determinations from Voice Data'[989] indicates such secondary use of data. The patent describes a system that can capture voice content when a user speaks into or near the device (e.g., Alexa), notably without activating the virtual assistant by mentioning the 'wake word' (e.g., 'hey Alexa'). Sniffer algorithms attempt to identify trigger words that indicate statements of preference (such as like or love) and translate them into keywords. The identified keywords can subsequently be transmitted to a location accessible to advertisers, who can use the keywords to select content that is likely relevant to the user.[990] Amazon has denied that it uses voice recordings for advertising and mentioned that the patent might never actually come to the market.[991] Nevertheless, incidents unveiled in the press imply that such secondary use already takes place. For example, a journalist in the US reported that she was discussing a specific kitchen gadget with her husband and some neighbours within the reach of Alexa and received ads on Amazon for the kitchen gadget the next day.[992] A marketing team within media giant Cox Media Group claims it can listen to ambient conversations of consumers through embedded microphones in smartphones, smart TVs, and

[987] Shoshana Zuboff, *The age of surveillance capitalism* (PublicAffairs 2019) 7.
[988] Zomet Asaf, Urbach Shlomo Reuben, 'Privacy-Aware Personalised Content for the Smart Home' US Patent Number US 10'453'098 (Assignee: Google LLC) October 2019 <US20160260135A1 - Privacy-aware personalized content for the smart home - Google Patents> accessed 8 February 2024.
[989] Edara Kiran, 'Key Word Determinations From Voice Data' US Patent Number US 8798995B1 (Assignee: Amazon Technologies, Inc.) August 2014 <https://patentimages.storage.googleapis.com/bd/ed/2b/c4c67cc5a9f1ab/US8798995.pdf> accessed 8 February 2024.
[990] Edara Kiran, 'Key Word Determinations From Voice Data' US Patent Number US 8798995B1 (Assignee: Amazon Technologies, Inc.) August 2014 <https://patentimages.storage.googleapis.com/bd/ed/2b/c4c67cc5a9f1ab/US8798995.pdf> accessed 8 February 2024.
[991] Griffin Andrew, 'Amazon files for Alexa patent to let it listen to people all the time and work out what they want' *The Independent* (London, 11 April 2018) <https://www.independent.co.uk/life-style/gadgets-and-tech/news/amazon-alexa-patent-listening-to-me-facebook-phone-talking-ads-a8300246.html> accessed 8 February 2024.
[992] Morgan Blake, 'Are Digital Assistants Always Listening?' *Forbes* (New York, 5 February 2018) <https://www.forbes.com/sites/blakemorgan/2018/02/05/are-digital-assistants-always-listening/#2f000e1a4eeb> accessed 8 February 2024.

other devices to gather data and use it to serve targeted ads.[993] In addition, it is not a secret that companies such as Google, Amazon, Meta and Apple maintain and improve their voice recognition devices and software by means of assessing various audio snippets recorded by such devices.[994] For example, Amazon has publicly confirmed to manually review Alexa requests to confirm that Alexa understood and responded correctly.[995]

The function creep problem (Type 1)

Particularly the AI disciplines ML and NLP significantly contribute to function creep and secondary use of personal data, which violates the purpose limitation principle.

4.5.2 Legal problems: Type 2

When the purpose limitation principle is applied to the AI disciplines introduced in Chapter 2, no specific Type 2 legal problems arise. Judicial guidance is scarce with respect to the criteria to be applied on the precision of the purpose. Research suggests that ECtHR case law makes the precision of the purpose dependent on the extent to which the data subject is affected by the processing.[996] Although the requirement to specify the purpose is a 'key element in the implementation of the European regime for the protection of personal data',[997] it does not seem to play a prominent role in CJEU case law. Cases dealing with purpose limitation do not specifically deal with the specification of purposes.[998] This is problematic when considering that the purpose specification requirement plays a central role in data protection law as all data protection principles depend on it.[999] In addition, the EU legal framework itself does not provide explicit criteria in order to determine how precisely the purposes should be specified.[1000] According to regulatory guidance, purposes which are too vague or general do not meet the criteria of being specific. For example, the guidance refers to 'elastic purposes' sometimes used by controllers such as 'future research', 'product innovation' and 'improving user experience'.[1001]

[993] Joseph Cox, 'Marketing Company Claims That It Actually Is Listening to Your Phone and Smart Speakers to Target Ads' *404 Media* (United States, 14 December 2023) <Marketing Company Claims That It Actually Is Listening to Your Phone and Smart Speakers to Target Ads (404media.co)> accessed 8 February 2024.
[994] Tine Munk, 'Does Online Privacy Exist in the GDPR Era? The Google Voice Assistant Case' in Tatiana-Eleni Synodiou et al (eds) *EU Internet Law in the Digital Single Market* (Springer Nature 2021) 480.
[995] Dorian Lynskey, 'Alexa, are you invading my privacy? the dark side of our voice assistants' *The Guardian* (London, 9 October 2019) <https://www.theguardian.com/technology/2019/oct/09/alexa-are-you-invading-my-privacy-the-dark-side-of-our-voice-assistants> accessed 8 February 2024.
[996] Merel Elize Koning, 'The purpose and limitations of purpose limitation' (Doctoral thesis, Radboud University Nijmegen 2020) 66, 162, 167 < https://repository.ubn.ru.nl/bitstream/handle/2066/221665/221665.pdf?sequence=1&isAllowed=y > accessed 8 February 2024.
[997] Case C-77/21 *Digi* [2022] ECR I-805 Opinion of AG Pikamäe para 40.
[998] Case C-77/21 *Digi* [2022] ECR I-805 para 27; Case C-175/20 *'SS' SIA* [2022] ECR I-124 para 64.
[999] Merel Elize Koning, 'The purpose and limitations of purpose limitation' (Doctoral thesis, Radboud University Nijmegen 2020) 102 < https://repository.ubn.ru.nl/bitstream/handle/2066/221665/221665.pdf?sequence=1&isAllowed=y > accessed 8 February 2024.
[1000] Maximilian von Grafenstein, *The Principle of Purpose Limitation in Data Protection Laws* (Nomos 2017) 232, 233, 244.
[1001] Art 29 Working Party, 'Opinion 03/2013 on purpose limitation' (WP 203, 2 April 2013) at 16; Art 29 Working Party, 'Opinion 02/2013 on apps on smart devices,' (WP 202, 27 February 2013) at 23.

Whereas this regulatory guidance is certainly needed and welcome, it is legally not binding[1002] and is not judicially tested. Shortcomings in terms of purpose specification may lead to Type 2 legal problems because substantively unclear principles are difficult to enforce. Nevertheless, this problem arises regardless of whether the processing involves AI and thus does not relate specifically to AI. Therefore, I refrain from discussing this problem in further detail.

4.5.3 Legal problems: Type 3

The basic idea of the purpose limitation principle is to restrict the processing of personal data. In the words of AG Pikamäe, the purpose of this principle is to 'delimit as clearly as possible the use of personal data'.[1003] However, interdisciplinary research on the application of the purpose limitation principle in personalisation and profiling systems has revealed that purpose specification hardly restricts the ways in which personal data can be processed.[1004] Where controllers do their best to define purposes with enough specificity and can demonstrate that such purposes are legitimate,[1005] any purpose is a valid purpose under the GDPR. Thus, purpose limitation does not seem to be an appropriate legal tool to ensure data processing is restricted in data-driven systems. Instead, it is a procedural criterion that at least requires controllers to consider the need and implications of processing from the beginning.[1006] This Type 3 legal problem occurs regardless of which AI discipline the purpose limitation is applied to because the principle itself is not suitable to restrict the collection and further processing of personal data. Therefore, it is a general problem and relates to all AI disciplines as introduced in Chapter 2.

> ***The restriction problem (Type 3)***
> *The purpose limitation principle does not, as intended, restrict the collection and further processing of personal data. It thus fails to achieve its aim to limit data processing and is therefore not fit for purpose to effectively protect the fundamental right to data protection.*

The purpose limitation principle enshrines two requirements: (i) personal data must be collected for specified, explicit and legitimate purposes, and (ii) personal data must not be further processed for incompatible purposes.[1007] Apart from specifically privileged purposes, any processing taking place after collection constitutes 'further processing' and must comply with the principle of compatible

[1002] Footnote 40 refers to an opinion issued by Article 29 Working Party; see Joined Cases C-141/12 & C-372/12, *YS, M* [2014] ECR I-2081, Opinion of AG Sharpston para 49.
[1003] Case C-77/21 *Digi* [2022] ECR I-805 Opinion of AG Pikamäe para 27.
[1004] Asia Biega, Michèle Finck, 'Reviving Purpose Limitation and Data Minimisation in Data-Driven Systems' (2021) Technology and Regulation 44, 49 and 55 <https://arxiv.org/abs/2101.06203> accessed 8 February 2024.
[1005] Which does not appear to be difficult.
[1006] Asia Biega, Michèle Finck, 'Reviving Purpose Limitation and Data Minimisation in Data-Driven Systems' (2021) Technology and Regulation 44, 49 and 55 <https://arxiv.org/abs/2101.06203> accessed 8 February 2024.
[1007] Case C-77/21 *Digi* [2022] ECR I-805 Opinion of AG Pikamäe para 28; Merel Elize Koning, 'The purpose and limitations of purpose limitation' (Doctoral thesis, Radboud University Nijmegen 2020) 58 < https://repository.ubn.ru.nl/bitstream/handle/2066/221665/221665.pdf?sequence=1&isAllowed=y > accessed 8 February 2024.

use.[1008] Further processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes are privileged purposes. They are a priori considered a lawful processing operation which is a compatible purpose[1009] provided that such processing is subject to appropriate safeguards.[1010]

Processing for compatible purposes does not require an additional legal basis[1011] and prevails over the interests of the data subject when she objects to such processing if it serves a public interest.[1012] Recital 159 GDPR envisages a broad interpretation of scientific research, including technological development, demonstration, fundamental and applied research and privately-funded research. Not only academic institutions but also profit-seeking companies can carry out scientific research based on this exception.[1013] Regulatory guidance requires that scientific research performed under this exception occurs in accordance with relevant sector-related methodological and ethical standards and in conformity with good practice.[1014] Whereas it is clear that publicly funded and externally published work at academic institutes fall under the research exception, [1015] this is less obvious for research performed at private companies. However, given the broad interpretation of scientific research derived from Recital 159 GDPR and relevant regulatory guidance, companies can argue that processing of personal data in the context of AI falls under the research exception.

Statistical purposes refer to the elaboration of statistical surveys or the production of statistical, aggregated results.[1016] Because ML is strongly based on statistics, it could be argued that further processing by means of ML constitutes processing for statistical purposes and is thus allowed without the need for an additional legal basis. Statistical purposes can be construed broadly, covering uses by companies for commercial gain and permitting to use this exception for big data applications and purposes.[1017] It seems that computer scientists do not come to terms whether ML is different from statistics. Some argue that ML is different from statistics, and others argue that statistics and ML are complementary.[1018] Also in the legal domain, the scope of the statistical purpose exception is not

[1008] Case C-77/21 *Digi* [2022] ECR I-805 Opinion of AG Pikamäe para 28.
[1009] Ibid, Footnote 14.
[1010] Article 89 GDPR.
[1011] Recital 50 GDPR; Waltraut Kotschy, Commentary of Article 6 in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020) 341.
[1012] Article 21 (6) GDPR.
[1013] European Data Protection Supervisor, 'A Preliminary Opinion on data protection and scientific research' (2020) 11 <https://edps.europa.eu/sites/edp/files/publication/20-01-06_opinion_research_en.pdf> accessed 8 February 2024.
[1014] Article 29 Working Party, 'Guidelines on consent under Regulation 2016/679' (WP 259rev.01, 10 April 2018) at 28.
[1015] Asia Biega, Michèle Finck, 'Reviving Purpose Limitation and Data Minimisation in Data-Driven Systems' (2021) Technology and Regulation 44, 51 <https://arxiv.org/abs/2101.06203> accessed 8 February 2024.
[1016] Recital 162; Cécile de Terwangne, Commentary of Article 5 in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020) 317.
[1017] Viktor Mayer-Schönberger, Yann Padova, 'Regime change? Enabling Big Data through Europe's new Data Protection Regulation' (2016) Vol 17 No 2 Science and Technology Law Review 315, 325-326.
[1018] Asia Biega, Michèle Finck, 'Reviving Purpose Limitation and Data Minimisation in Data-Driven Systems' (2021) Technology and Regulation 44, 52 <https://arxiv.org/abs/2101.06203> accessed 8 February 2024.

entirely clear. Some argue that it facilitates inferential analytics by means of ML approaches[1019] and the construction of ML models based on personal data.

Regulatory guidance states that the statistics exception applies in commercial settings and to 'analytical tools of websites or big data applications aimed at market research'.[1020] Personal data may be used to draw inferences and lead to a model, which then can be applied to other individuals, for example to take decisions.[1021] Whereas Recital 162 GDPR indicates that the result of processing for statistical purposes must be aggregated data, and this result must not be used to support measures or decisions regarding any particular person, its effect remains unclear. The question is what qualifies as a decision or measure in the latter sense. Both concepts require some binding effect, distinguishing them from mere recommendations.[1022] At least some forms of ML output could qualify to fall under the scope of the statistics exception, such as the prediction of customers ceasing their relationship with a company (customer churn). Whether the prediction of specific customer churn and subsequent action taken to avoid this also fall under the statistics exception is less clear[1023] since this might be considered 'a measure or decision regarding any particular person'.[1024] Targeted advertisement is another illustrative example. Displaying ads to individuals online based on their interests inferred by ML does not necessarily constitute a decision or measure regarding the individuals concerned. Arguably, such targeted ads are mere recommendations to purchase a product or subscribe to a service, lacking the binding effect of a measure or decision. In addition, the different processing stages of the ML pipeline seem to be relevant as ML produces aggregate and individual results at different processing stages.[1025] Furthermore, ML models are likely to fall under trade secrets protection and controllers could refrain from providing meaningful information (see Sections 5.6 and 5.6.2 below).

In addition, due to the opening clause contained in Article 89 GDPR, the scope of the statistical purpose exception might vary across EU Member States. Recital 162 GDPR demands the latter to 'determine statistical content, control of access, specifications for the processing of personal data for statistical purposes' within the limits of the GDPR. This opening clause and the corresponding implementation in the Member States lead to additional legal uncertainty besides the already considerable uncertainties regarding this exception.[1026]

[1019] Sandra Wachter, Brent Mittelstadt 'A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI' (2019) No 2 Columbia Business Law Review 1, 550-551; see also 549, 592.
[1020] Art 29 Working Party, 'Opinion 03/2013 on purpose limitation' (WP 203, 2 April 2013) at 29.
[1021] Sandra Wachter, Brent Mittelstadt 'A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI' (2019) No 2 Columbia Business Law Review 1, 550-551; see also 549, 592.
[1022] For the notion of a decision in the sense of Article 22 GDPR see Case C-634/21, *SCHUFA Holding AG* [2023] ECR I-957 paras 44-46; see also corresponding Opinion AG Pikamäe para 37.
[1023] Asia Biega, Michèle Finck, 'Reviving Purpose Limitation and Data Minimisation in Data-Driven Systems' (2021) Technology and Regulation 44, 52 <https://arxiv.org/abs/2101.06203> accessed 8 February 2024.
[1024] Recital 162 GDPR.
[1025] Asia Biega, Michèle Finck, 'Reviving Purpose Limitation and Data Minimisation in Data-Driven Systems' (2021) Technology and Regulation 44, 52 <https://arxiv.org/abs/2101.06203> accessed 8 February 2024.
[1026] Asia Biega, Michèle Finck, 'Reviving Purpose Limitation and Data Minimisation in Data-Driven Systems' (2021) Technology and Regulation 44, 55 <https://arxiv.org/abs/2101.06203> accessed 8 February 2024.

Arguably, both the research and statistical purposes exception for further processing undermine the GDPR's aim that data subjects should have control over their own personal data.[1027] Scholars place the purpose limitation principle in light of the concept of control, as well as informational self-determination and autonomy.[1028] The concept of control is not defined in the GDPR, although it was one of the main reasons for the data protection reform[1029] and constitutes one of the GDPR's legislative aims.[1030] As outlined in Section 4.4.3, the main mechanism for data subjects to exercise control over the processing of their personal data under the GDPR are data subject rights. However, this mechanism is rather limited. AG Campos Sánchez-Bordona correctly notes that 'the scope for individual action is limited' and 'confined to the exercise of those rights in specified circumstances'.[1031] The principle of compatible use, and the privileged purposes concerning research and statistics in particular, hinders data subjects to enforce their rights and thus to exercise control over the processing of their data. Article 17 (3) GDPR states that the right to erasure does not apply if erasure of personal data is likely to render the achievement of the objectives of processing for research or statistical purposes impossible or seriously impair these objectives. In addition, processing of personal data for scientific and statistical purposes for the performance of a task carried out for reasons of public interests prevails over the data subjects' right to object to such processing.[1032] Therefore, the concept of compatible use undermines the individual's control over the processing of personal data because it allows one to further process personal data by means of ML. This is detrimental to the aim of GDPR to provide data subjects with control over their data[1033] and ultimately leads to a problem of Type 3, that is, the concept of compatible use is not fit for purpose to protect the fundamental right to data protection.

However, It could be argued that neither the GDPR nor the EUCFR contains a 'right of control' that transforms it into an illusory objective pursued by the GDPR and the EU's data protection reform. This criticism has its merits, but the concept of compatible use still leads to a type 3 legal problem. It undermines the GDPR's objective to protect natural persons from risks related to the processing of personal data. There are considerable uncertainties regarding the interpretations of the research and statistical purposes exception.[1034] Creative controllers will utilise these considerable uncertainties surrounding the concept of compatible use. This is detrimental to the GDPR's aim to effectively protect

[1027] Recital 7 GDPR.
[1028] For an overview, see Merel Elize Koning, 'The purpose and limitations of purpose limitation' (Doctoral thesis, Radboud University Nijmegen 2020) 72 <https://repository.ubn.ru.nl/bitstream/handle/2066/221665/221665.pdf?sequence=1&isAllowed=y> accessed 8 February 2024.
[1029] Case C-300/21 *UI* [2022] ECR I-756 Opinion AG Manuel Campos Sánchez-Bordona paras 69, 70.
[1030] Recital 7 GDPR.
[1031] Case C-300/21 *UI* [2022] ECR I-756 Opinion AG Manuel Campos Sánchez-Bordona para 72.
[1032] Article 21 (6) GDPR.
[1033] Recital 7 GDPR.
[1034] Asia Biega, Michèle Finck, 'Reviving Purpose Limitation and Data Minimisation in Data-Driven Systems' (2021) Technology and Regulation 44, 55 <https://arxiv.org/abs/2101.06203>accessed 8 February 2024.

the fundamental right to data protection, which the CJEU emphasises.[1035] Neither do these uncertainties contribute to a high level of protection as envisaged by the GDPR.[1036]

The compatible use problem (Type 3)

Processing in the context of ML might fall under the concept of compatible use because it relates to the privileged statistical and/or research purposes. This undermines the data subject's control over the processing of personal data, which is detrimental to the GDPR's aim. The concept of compatible use is therefore not fit for purpose to protect the fundamental right to data protection.

4.6 Data minimisation

The data minimisation principle as introduced in Section 3.3.3.5 requires that personal data must be adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed.[1037] This wording indicates that the data minimisation principle is a manifestation of the proportionality principle as introduced in Section 3.2.3. In the CJEU's words, the data minimisation principle 'gives expression to the principle of proportionality'.[1038]

4.6.1 Legal problems: Type 1

Quite contradictory to the data minimisation principle introduced in Section 3.3.3.5, AI needs substantial amounts of data in order to operate effectively, particularly in the training phase.[1039] AI has an 'insatiable appetite' for data and contradicts the data minimisation principle.[1040] Advanced AI applications employing complex models such as deep learning (DL) and natural language processing (NLP) need to learn many parameters and require *enough* data to function properly.[1041] As outlined in Section 2.2.1, *accurate* predictions are the main goal of data processing in ML. The underlying algorithm is decisive in terms of the required amount of data. DL, a particular kind of ML, requires large-scale training data.[1042] DL applications using the supervised training method in NLP for speech

[1035] Case C-645/19 [2021] *Facebook Ireland* ECR I-483 para 45; Case C-319/20, *Meta Platforms Ireland Limited* [2022] ECR I-322 para 73; Joined Cases C-293/12 and C-594/12 *Digital Rights Ireland* [2014] ECR-I 238 paras 54 and 66; Joined Cases C-203/15 and C-698/15 [2016] *Tele 2 Svergie* ECR-1 970 paras 109 and 122; Case C-362/14 *Schrems v Data Protection Commissioner* [2015] ECR I-650 para 91.
[1036] Recitals 6, 10 GDPR; Case C-534/20, *Leistritz AG* [2022] ECR I-594 para 26; Case C-645/19 [2021] *Facebook Ireland* ECR I-483 para 45; Case C-511/18, *La Quadrature du Net* [2020] ECR I-791 para 106; Case C-131/12, *Google Spain* [2014] ECR I-317 para 66; Case C-154/21 *Österreichische Post AG* [2023] ECR I-3 para 44; Case C-132/21 *Nemzeti* [2023] ECR I-2 para 42.
[1037] Article 5 (1) lit c GDPR.
[1038] Case C-439/19 *B* [2021] ECR I-504 para 98; Case C-175/20 *'SS' SIA* [2022] ECR I-124 para 83.
[1039] CIPL, 'Artificial Intelligence and Data Protection How the GDPR Regulates AI' (2020) 13 <https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl-hunton_andrews_kurth_legal_note_-_how_gdpr_regulates_ai__12_march_2020_.pdf> accessed 8 February 2024.
[1040] Christopher Kuner et al, 'Expanding the artificial intelligence-data protection debate' (2018) Vol 8 No 4 International Data Privacy Law 289-292.
[1041] Asia Biega, Michèle Finck, 'Reviving Purpose Limitation and Data Minimisation in Data-Driven Systems' (2021) Technology and Regulation 44, 57 and 58 <https://arxiv.org/abs/2101.06203> accessed 8 February 2024.
[1042] Zhou Zhi-Hua, Feng Ji, 'Deep Forest: Towards an Alternative to Deep Neural Networks' (IJCAI Conference, Melbourne, August 2017) 1 <https://www.ijcai.org/proceedings/2017/0497.pdf > accessed 8 February 2024.

recognition require large amounts of training data with labels.[1043] Computer vision (CV) heavily relies on the processing of photographs, in particular for facial recognition and automated face analysis systems used in the AI discipline affective computing (AC). Moreover, data analytics in the context of ML does not only require vast amounts of data, but also causes more data processing and therefore creates a closed circle: with more data, more accurate models can be trained, which generates more services and users of those services, which leads to more data being processed.[1044] Ultimately, AI violates the data minimisation principle, which constitutes a Type 1 legal problem.

The data appetite problem (Type 1)

AI has an insatiable appetite for data. Contrary to the data minimisation principle, AI and particularly DL requires substantial amounts of data to function well and generate accurate output. This violates the data minimisation principle.

However, the data appetite problem does not suggest that AI and the data minimisation principle are per se incompatible. Instead, applying data minimisation to complex AI systems is difficult. This is to a significant extent due to the current incomputability of data protection principles[1045] (Section 4.7.3). It is challenging to determine which data are necessary when personal data are processed in the context of AI and thus to limit such data accordingly. The problem with data minimisation and AI lies at the core of this principle, namely, how to exactly define what should be considered necessary for processing activities based on AI applications.[1046] Recital 39 relating to the data minimisation principle simply states that personal data 'should be processed only if the purpose of the processing could not reasonably be fulfilled by other means.' It is difficult for computer scientists to determine in a given computation of individual pieces which personal data are adequate, relevant and necessary. Consequently, computer scientists apply various, often inconsistent, approaches to the data minimisation principle.[1047] This becomes most apparent in the case of unsupervised ML that processes data for *un*specified and *in*explicit purposes. With unsupervised ML, the processing *itself* determines the purpose and future use of the data since its goal is to detect patterns, correlations, gain knowledge and make accurate predictions. Thus, in the context of unsupervised ML, the purpose of processing

[1043] Deng Li and Liu Yang, 'Epilogue: Frontiers of NLP in the Deep Learning Era' in Deng Li and Liu Yang (eds) *Deep learning in natural language processing* (Springer 2018) 1.
[1044] Zhao Jianxin et al., 'Privacy-preserving Machine Learning Based Data Analytics on Edge Devices' (AIES Conference, New Orleans, January 2018) 1 <http://www.aies-conference.com/2018/contents/papers/main/AIES_2018_paper_161.pdf > accessed 8 February 2024.
[1045] Asia Biega, Michèle Finck, 'Reviving Purpose Limitation and Data Minimisation in Data-Driven Systems' (2021) Technology and Regulation 44, 58 and 60 <https://arxiv.org/abs/2101.06203> accessed 8 February 2024.
[1046] Ronald Leenes, Silvia De Conca, 'Artificial intelligence and privacy – AI enters the house through the Cloud' in Woodrow Barfield, Ugo Pagallo (eds) *Research handbook on the law of artificial intelligence* (Edward Elgar Publishing Inc. 2018) 299, See also Mireille Hildebrandt, 'Primitives of legal protection in the era of data-driven platforms' (2018) 13 <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3140594> accessed 8 February 2024.
[1047] Asia Biega, Michèle Finck, 'Reviving Purpose Limitation and Data Minimisation in Data-Driven Systems' (2021) Technology and Regulation 44, 59 <https://arxiv.org/abs/2101.06203> accessed 8 February 2024.

is not yet known and there is no supervisor who directs the machine on the purpose for processing.[1048] It cannot be determined what pieces of personal data are necessary for yet a unknown purpose.

Even if the purpose is known and defined as 'development of AI systems', limiting the use of personal data necessary to achieve this purpose seems illusory due to the insatiable appetite for data of AI (see the data appetite problem). In addition, the purpose 'development of AI systems' arguably does not meet the criteria of being 'specific'. This purpose appears to 'elastic' as controllers use phrases such as 'future research', 'product innovation', 'improving user experience', which regulators are likely to consider as too vague or general.[1049] In addition, such an elastic purpose is not suitable for proportionality decisions as required by the data minimisation principle, namely, to limit the processing of personal data to what is necessary in relation to that purpose because the purpose specification requirement is a precondition for that proportionality assessment.[1050] As a consequence, the data minimisation principle is violated. This constitutes a Type 1 legal problem.

The necessity problem (Type 1)

In the case of unsupervised ML, it is impossible to determine whether a given computation of specific pieces of personal data is necessary, and to limit the personal data processed in accordance with the proportionality principle. Such processing violates the data minimisation principle.

4.6.2 Legal problems: Type 2

Verifying whether a controller complies with the data minimisation principle is technically difficult, if not impossible. The complexity of models adopted by AI represents a major challenge for human cognition.[1051] AI equipped systems are becoming highly opaque black boxes and individuals are unable to follow the steps these machines are taking to reach whatever conclusions they reach.[1052] DL methods based on artificial neural networks (ANN) generally lack interpretability[1053] and are particularly challenging due to their hierarchical and nonlinear structure and the central concept in DL called connectionism. In connectionism, a large number of simple computational units (artificial neurons) achieve intelligent behaviour when networked together[1054] (see Section 2.2.1.4). It seems neither possible to understand which artificial neuron contributed to a distinct part of the output nor to understand

[1048] Similarly, see Christopher Kuner et al, 'Expanding the artificial intelligence-data protection debate' (2018) Vol 8 No 4 International Data Privacy Law 289, 290.
[1049] Art 29 Working Party, 'Opinion 03/2013 on purpose limitation' (WP 203, 2 April 2013) at 16; Art 29 Working Party, 'Opinion 02/2013 on apps on smart devices,' (WP 202, 27 February 2013) at 23.
[1050] Merel Elize Koning, 'The purpose and limitations of purpose limitation' (Doctoral thesis, Radboud University Nijmegen 2020) 68, 108 < https://repository.ubn.ru.nl/bitstream/handle/2066/221665/221665.pdf?sequence=1&isAllowed=y > accessed 8 February 2024.
[1051] Zachary C Lipton, 'The Mythos of Model Interpretability' (2018) Vol 16 Iss 3 ACMQueue 18 <https://dl.acm.org/doi/pdf/10.1145/3236386.3241340?download=true> accessed 8 February 2024.
[1052] Amitai Etzioni and Oren Etzioni, 'Keeping AI Legal' (2016) 19 Vand. J. Ent. & Tech. L. 133, 137.
[1053] Deng Li and Liu Yang, 'A Joint Introduction to Natural Language Processing and Deep Learning' in Deng Li and Liu Yang (eds) *Deep learning in natural language processing* (Springer 2018) 11, 12.
[1054] Ian Goodfellow, Yoshua Bengio, Aaron Courville, *Deep Learning* (MIT Press 2016) 16 <www.deeplearningbook.org> accessed 8 February 2024.

what happened in the intermediate (hidden) layers of the ANN.[1055] When an ANN is used for pattern recognition in CV or NLP, an ex-post analysis of the model used will likely not establish linear causal connections which are comprehensible for human minds.[1056] Where the model used is not interpretable, it is difficult or impossible to verify whether the processing of individual pieces of personal data are adequate, relevant and necessary for a specific purpose according to the data minimisation principle. This cannot be mediated by the AI discipline of automated reasoning. As outlined in Sections 2.2.5, 4.3.1, 4.4.1 and 4.7.1, AI systems do not have a semblance of common sense or capabilities such as human cognition and are therefore unable to think in a manner on par with human thinking.[1057] Therefore, AI systems do not deploy arguments that may be used to determine which factors, for example, personal data, are necessary or relevant for generating certain output. Therefore, the data minimisation principle cannot be enforced, whether by means of private enforcement initiated by data subjects or in the form of regulatory enforcement pursued by SAs.

The verification problem (Type 2)

The reasoning deficiencies in AR and the complexity of models adopted by AI, particularly approaches from ML (specifically DL) as well as CV and NLP, render it difficult or impossible to verify whether the processing of personal data complies with the data minimisation principle. Consequently, the data minimisation principle cannot be enforced.

4.6.3 Legal problems: Type 3

When consequently applied, the data minimisation principle might negatively affect the accuracy principle as introduced in Section 3.3.3.6. In the context of a prediction system powered by ML, deciding that a certain piece of personal data should not be used might reasonably lead to inaccurate predictions,[1058] which violates the accuracy principle. However, it could be argued that both principles are not in conflict when the purpose is defined as 'processing *all data* necessary to make accurate predictions'. The purpose specification requirement plays a central role, also regarding the data minimisation principle. In my view, this purpose is not specific enough to effectively implement the data

[1055] Ethem Alpaydin, *Machine Learning: The New AI* (3rd edn MIT Press 2016) 155.
[1056] Thomas Wischmeyer, 'Artificial Intelligence and Transparency: Opening the Black Box' in Thomas Wischmeyer and Timo Rademacher (eds) *Regulating Artificial Intelligence* (Springer Nature 2020) 81.
[1057] Lance Eliot, 'AI Ethics And The Quagmire Of Whether You Have A Legal Right To Know Of AI Inferences About You, Including Those Via AI-Based Self-Driving Cars' *Forbes* (New York, 25 May 2022) < https://www-forbes-com.cdn.ampproject.org/c/s/www.forbes.com/sites/lanceeliot/2022/05/25/ai-ethics-and-the-quagmire-of-whether-you-have-a-legal-right-to-know-of-ai-inferences-about-you-including-those-via-ai-based-self-driving-cars/amp/> accessed 8 February 2024.
[1058] Asia Biega, Michèle Finck, 'Reviving Purpose Limitation and Data Minimisation in Data-Driven Systems' (2021) Technology and Regulation 44, 57 <https://arxiv.org/abs/2101.06203> accessed 8 February 2024.

minimisation principle.[1059] An insufficiently defined purpose leads to excessive processing of personal data and violates the data minimisation principle.[1060]

Similarly, there might be trade-offs between data minimisation and the fairness principle. To ensure fairness, it might be required to process more personal data than strictly necessary according to the data minimisation principle to guard against bias and error,[1061] for example, to avoid discrimination. An empirical study suggests that the decision to not collect data on gender or other protected attributes could make it challenging or impossible to identify discrimination against those groups once the ML algorithm has been deployed.[1062] Thus, minimisation of sensitive features such as gender may diminish the ability to detect unfairness,[1063] which is detrimental to the fairness principle. To figure out means that overcome such trade-offs requires creativity and reasoning skills. However, AI currently lacks reasoning capabilities that would allow to solve the difficult task of overcoming trade-offs between data protection principles. The trade-offs between principles combined with the reasoning deficiencies of AI lead to a Type 3 legal problem. Principles leading to trade-offs are not fit for purpose to effectively[1064] protect the fundamental right to data protection, to ensure a high level of the protection of personal data[1065] and to achieve the GDPR's aim to establish a strong and coherent data protection framework.[1066] This holds particularly true when considering that principles provide the basis for the protection of personal data.[1067] This Type 3 legal problem occurs regardless of which AI discipline the data minimisation, fairness and accuracy principles are applied to because they themselves create the trade-offs between each other. Therefore, it is a general problem and relates to all AI disciplines as introduced in Chapter 2.

[1059] Merel Elize Koning, 'The purpose and limitations of purpose limitation' (Doctoral thesis, Radboud University Nijmegen 2020) 102 < https://repository.ubn.ru.nl/bitstream/handle/2066/221665/221665.pdf?sequence=1&isAllowed=y > accessed 8 February 2024.
[1060] Art 29 Working Party, 'Opinion 01/2014 on the application of necessity and proportionality concepts and data protection within the law enforcement sector' (WP 211, 27 February 2014), at 18; Art 29 Working Party, 'Opinion 05/2013 on Smart Borders' (WP 206, 6 June 2013) at 10.
[1061] Christopher Kuner et al, 'Expanding the artificial intelligence-data protection debate' (2018) Vol 8 No 4 International Data Privacy Law 289, 290.
[1062] Gemma Galdon Cavell et al, 'Auditing Algorithms: On Lessons Learned and the Risks of Data Minimization' (Proceedings of the AAAI/ACM Conference on AI, Ethics and Society, New York 2020) 266 <https://dl.acm.org/doi/pdf/10.1145/3375627.3375852> accessed 8 February 2024.
[1063] Asia Biega, Michèle Finck, 'Reviving Purpose Limitation and Data Minimisation in Data-Driven Systems' (2021) Technology and Regulation 44, 59 <https://arxiv.org/abs/2101.06203> accessed 8 February 2024; Gemma Galdon Clavell et al, 'Auditing Algorithms: On Lessons Learned and the Risks of Data Minimization', (2020) Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society (ACM) <https://dl.acm.org/doi/10.1145/3375627.3375852> accessed 8 February 2024.
[1064] Recital 11 GDPR; Joined Cases C-293/12 and C-594/12 *Digital Rights Ireland* [2014] ECR-I 238 paras 54 and 66; Joined Cases C-203/15 and C-698/15 [2016] *Tele 2 Svergie* ECR-1 970 paras 109 and 122; Case C-362/14 *Schrems v Data Protection Commissioner* [2015] ECR I-650 para 91.
[1065] Recitals 6, 10 as well as CJEU case law, such as Case C-534/20, *Leistritz AG* [2022] ECR I-594 para 26; Case C-645/19 [2021] *Facebook Ireland* ECR I-483 para 45; Case C-511/18, *La Quadrature du Net* [2020] ECR I-791 para 106; Case C-131/12, *Google Spain* [2014] ECR I-317 para 66; Case C-154/21 *Österreichische Post AG* [2023] ECR I-3 para 44; Case C-132/21 *Nemzeti* [2023] ECR I-2 para 42.
[1066] Recital 7 GDPR.
[1067] Cécile de Terwangne, Commentary of Article 5 in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020) 313.

The trade-off problem (Type 3)

When consequently applied in the context of AI, the data minimisation principle might lead to trade-offs with the accuracy and fairness principles. Due to the shortcomings in AR, AI currently lacks reasoning capabilities that may overcome these trade-offs, and may fail to adequately protect the fundamental right to data protection.

4.7 Accuracy

As outlined in Section 3.3.3.6, the GDPR states that the processing of personal data must be accurate.[1068] The accuracy principle intends to protect the individual concerned from being irrationally or unfairly treated based on wrong and inaccurate representations.[1069] According to regulatory guidance, accurate means 'accurate as to a matter of fact'.[1070] The need for personal data to mirror the reality with respect to the data subject concerned[1071] is also stressed in academia: personal data shall, at any given time, reflect reality.[1072] Case law[1073] of the CJEU indicates that the level of accuracy of personal data is determined by the purpose of the processing:[1074] the assessment whether personal data are accurate and complete depends on the *purpose* for which data were collected.[1075] Nevertheless, the precise substantive requirements of the accuracy principle remain an underexplored topic in academia, which is problematic when considering the developments in AI and its significance with regard to the right to rectification[1076] (see also Section 5.7). However, to apply the accuracy principle to the AI disciplines introduced in Section 2.2, I distinguish between two distinct types of accuracy. These are *absolute accuracy* referring to 'accurate as a matter of fact'[1077] aiming to reflect reality[1078] (e.g. date of birth) and *relative accuracy* which is more nuanced and determines accuracy based on the purpose of processing[1079] (e.g. data 'measured' by means of a percentage).

[1068] Art. 5 (1) lit d GDPR.
[1069] Dara Hallinan, Frederik Zuiderveen Borgesius, 'Opinions can be incorrect (in our opinion)! On data protection law's accuracy principle' (2020) Vol 10 No 1 IDPL 1, 9.
[1070] Art 29 Working Party, 'Guidelines on the Implementation of the Court of Justice of the European Union Judgment on Google Spain and inc v. Agencia Española de Protección De Datos (AEPD) and Mario Costeja González C-131/12' (WP 225, 26 November 2014), at 15 <https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=667236> accessed 8 February 2024.
[1071] Ibid 15; Dara Hallinan, Frederik Zuiderveen Borgesius, 'Opinions can be incorrect (in our opinion)! On data protection law's accuracy principle' (2020) Vol 10 No 1 IDPL 1, 4.
[1072] Paul Voigt, Axel von dem Bussche, *The EU General Data Protection Regulation (GDPR): A Practical Guide* (Springer 2017) 91.
[1073] Case C-434/16, *Nowak* [2017] ECR I-994 para 53.
[1074] Dara Hallinan, Frederik Zuiderveen Borgesius, 'Opinions can be incorrect (in our opinion)! On data protection law's accuracy principle' (2020) Vol 10 No 1 IDPL 1, 4.
[1075] Case C-434/16, *Nowak* [2017] ECR I-994 para 53.
[1076] Diana Dimitrova, 'The rise of the personal data quality principle: is it legal and does it have an impact on the right to rectification?' (2021) Vol 12 No 3 European Journal of Law and Technology 2.
[1077] Art 29 Working Party, 'Guidelines on the Implementation of the Court of Justice of the European Union Judgment on Google Spain and inc v. Agencia Española de Protección De Datos (AEPD) and Mario Costeja González C-131/12' (WP 225, 26 November 2014), at 15 <https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=667236> accessed 8 February 2024.
[1078] Paul Voigt, Axel von dem Bussche, *The EU General Data Protection Regulation (GDPR): A Practical Guide* (Springer 2017) 91.
[1079] Case C-434/16, *Nowak* [2017] ECR I-994 para 53.

4.7.1 Legal problems: Type 1

ML is particularly problematic in the context of the accuracy principle. Companies increasingly offer services and products with embedded ML components which aim to predict behaviour of individuals or to detect correlations, for example with regard to credit risk and health status.[1080] Such services and products involve probabilistic predictions and detected correlations.[1081] Predictions can be defined as 'the output emitted by a model of a data generating process in response to specific configurations of input'.[1082] ML is deployed to draw inferences about the behaviours, preferences and private lives of individuals, information that can subsequently be used to nudge or manipulate individuals or to take decisions on them.[1083] Put simply, inference may be described as the process whereby a conclusion is drawn without complete certainty, but with some degree of probability.[1084] Any inferential method is built on assumptions[1085] which may be correct or not. Inference enables decision-making under conditions of uncertainty.[1086] Prediction and inference are inextricably linked to each other because inference involves the systematic comparison of predictions. Both industry and academic literature focus on predictions, in particular in the AI discipline ML.[1087] The very nature of inferences, predictions and correlations increases the risk of inaccuracy[1088] because of its probabilistic nature.[1089] To be clear, the output generated by ML does not necessarily equal inaccurate data. Suppose processing aims, as a purpose, to infer a chance of something happening in the future. In that case, the probabilistic nature of such a prediction does not automatically lead to a violation of the accuracy principle. Instead, the problem in terms of accuracy emerges when predictions are treated as facts, which is context-dependent. If such predictions or correlations are essentially considered as *facts* this might lead to detrimental effects for data subjects (e.g., when applying for a job or a loan). There is experimental evidence that humans closely follow algorithmic output and cannot correctly *assess* its quality. In this online experiment, 1,263 participants received algorithmic advice and were free to choose whether to incorporate this advice in their own response. Most of the participants followed the algorithmic

[1080] Pedreschi Dino et. al., 'Open the Black Box: Data-Driven Explanation of Black Box Decision Systems' (2018) 1 <https://arxiv.org/pdf/1806.09936.pdf > accessed 8 February 2024.
[1081] Viktor Mayer-Schönberger; Kenneth Cukier, *Big Data: A Revolution That Will Transform How We Live, Work and Think* (Houghton Mifflin Harcourt 2013) 52-55.
[1082] Nathan Sanders, 'A Balanced Perspective on Prediction and Inference for Data Science in Industry' (2019) Iss 1.1 Harvard Data Science Review 1, 15 <https://assets.pubpub.org/zmmen09c/644ef4a4-5a71-43f8-9bcd-f2f6cb92ea65.pdf> accessed 8 February 2024.
[1083] Sandra Wachter, Brent Mittelstadt, 'A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI' (2019) Issue 2 Columbia Business Law Review 494, 497, 548, 549.
[1084] Michael P Cohen, 'Inference' in Paul J Lavrakas (eds) *Encyclopedia of Survey Research Methods* (Sage Publication, Inc 2008) 334.
[1085] Michael Betancourt, 'A Unified Treatment of Predictive Model Comparison' (2015) 1 <https://arxiv.org/pdf/1506.02273.pdf > accessed 8 February 2024.
[1086] Lawrence Hazelrigg, 'Inference' in Melissa Hardy, Alan Bryman (eds) *Handbook of Data Analysis* (Sage Publications 2004) 14.
[1087] Nathan Sanders, 'A Balanced Perspective on Prediction and Inference for Data Science in Industry' (2019) Iss 1.1 Harvard Data Science Review 1, 7, 21 <https://assets.pubpub.org/zmmen09c/644ef4a4-5a71-43f8-9bcd-f2f6cb92ea65.pdf> accessed 8 February 2024.
[1088] Article 29 Working Party, 'Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679' (WP 251rev.01, 6 February 2018), at 17.
[1089] Christopher Burr, Nello Cristianini, 'Can machines read our mind?' (2019) Vol 29 Iss 3 Minds and Machines 461, 483.

recommendation closely and never realised that the algorithm was purposely biased. The setup of the experiment enabled the participants to compare their prediction with the algorithm prediction, allowing them to realise that the algorithm is biased.[1090]

ML systems that aim to predict future behaviour of individuals cannot be designed with absolute accuracy due to their predictive nature and the lack of a truth as a baseline for comparison.[1091] Predictions generated by ML relate to future conduct that has *not yet happened*. Predictive accuracy will vary for each situation and this is not necessarily obvious for the ones who deploy the system or are subject to the system.[1092] What requires scrutiny is not the input data but rather the accuracy of the inferences drawn from input data,[1093] i.e. the output of the AI system. Finding correlations in data and acting on them is often considered to be good enough.[1094] Correlations based on a sufficient volume of data are increasingly seen as sufficiently credible to direct action without first establishing causality. Even if strong correlations or causal knowledge are found, this knowledge may only concern groups, whereas actions are directed towards individuals. This may lead to situations in which individuals are inaccurately described via simplified models or classes.[1095] Inferences or predictions can never be absolutely certain and are poorly verifiable or not verifiable at all (e.g. the individual is a 'high credit risk' or 'likely to buy a house in two years').[1096] Inference 'is always an invasion of the unknown, a leap from the known'.[1097] Admittedly, it might be argued that this also applies to inferences drawn by humans. However, human inferences are based on *human reasoning* and are usually not considered facts. Machine-generated inferences are more problematic because they are *not* based on human reasoning and are often treated as facts,[1098] although they are simply probabilistic and relate to future conduct that has not yet happened. Consider the following example which occurred in a case referred to the CJEU. Due to a poor credit score value allocated to a data subject, the mobile network

[1090] Jan Biermann, John Horton, Johannes Walter, 'Algorithmic Advice as a Credence Good' (2022) Centre for European Economic Research Discussion Paper No 22-071 at 2, 14 < https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4326911 > accessed 8 February 2024.
[1091] Diana Dimitrova, 'The rise of the personal data quality principle: is it legal and does it have an impact on the right to rectification?' (2021) Vol 12 No 3 European Journal of Law and Technology 21.
[1092] Mireille Hildebrandt, 'Primitives of legal protection in the era of data-driven platforms' (2018) 15 <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3140594> accessed 8 February 2024.
[1093] Omer Tene, Jules Polonetsky, 'Big Data for All: Privacy and User Control in the Age of Analytics' (2013) 11 Northwestern Journal of Technology and Intellectual Property 239, 270.
[1094] Viktor Mayer-Schönberger Viktor, Kenneth Cukier, *Big Data: A Revolution That Will Transform How We Live, Work and Think* (Houghton Mifflin Harcourt 2013) 42, 48, 49.
[1095] Brent Daniel Mittelstadt et al, 'The ethics of algorithms: Mapping the debate' (2016) Vol 3 Iss 2 Big Data & Society 1, 5; Solon Barocas, 'Data Mining and the Discourse on Discrimination' (2014) <https://dataethics.github.io/proceedings/DataMiningandtheDiscourseOnDiscrimination.pdf> accessed 8 February 2024.
[1096] Sandra Wachter, Brent Mittelstadt, 'A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI' (2019) Issue 2 Columbia Business Law Review 494, 510.
[1097] John Dewey, *The Middle Works of John Dewey, Volume 9, 1899-1924* (Carbondale Southern Illinois University Press 1980) 165.
[1098] Jan Biermann, John Horton, Johannes Walter, 'Algorithmic Advice as a Credence Good' (2022) Centre for European Economic Research Discussion Paper No 22-071 at 2, 14 < https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4326911 > accessed 8 February 2024.

operator denied to extend a mobile contract subscription with a rather low monthly fee of 10 €.[1099] This score value was used as a fact, although it was merely a prediction about future behaviour that had not yet materialised and may never do. Inferences generated by machines are highly scalable and less likely to be correct due to current reasoning deficiencies in AI (see also Sections 2.2.5, 4.3.1, 4.4.1 and 4.7.1).

Additionally, overfitting negatively affects the accuracy of predictions generated by ML and DL models. Overfitting is a common side effect of training in ML and occurs when models reach a higher accuracy on the training data than for new input data. It is inherent to training ANNs.[1100] With overfitting, the model learns how to represent well the training data, but performs poorly on new information as input.[1101] In fact, several factors determine a model's ability to generalise well, namely, the model architecture, regularisation techniques and the dataset design.[1102] Overfitting may proactively be addressed by means of lowering the number of weights an ANN has.[1103] To tune the parameters of a given model in a way that they perform well not only on training data but also on new information is a general problem in ML. Regularisation techniques are a vital tool to prevent overfitting and aim to reduce errors in predicting data that do not form part of the training set. Regularisation algorithms for ANNs may be divided into three main categories: i) data augmentation algorithms changing the input of the ANN, ii) internal algorithms changing values and inner structures of the ANN and iii) label algorithms performing their changes over the desired output.[1104] However, overfitting remains a problem despite the technical means to mitigate it. The problem of avoiding overfitting is subject to ongoing research, with regard to ANNs in particular. Overfitting mysteries in ANNs are not yet fully understood, partly due to the 'black-box' characteristics of ANNs.[1105] In any case, because overfitting occurs during the training process of an ANN, it results in high accuracy in terms of training data, but a poor prediction performance with regard to new input.[1106] Therefore, the common problem of

[1099] Case C-203/22 *Dun & Bradstreet Austria* p 2 <https://www.ris.bka.gv.at/Dokumente/Lvwg/LVWGT_WI_20220211_VGW_101_042_791_2020_44_00/LVWGT_WI_20220211_VGW_101_042_791_2020_44_00.pdf> accessed 8 February 2024.
[1100] Nicholas Carlin, 'Evaluating and Testing Unintended Memorization in Neural Networks' (*Berkeley Artificial Intelligence Research Blog*, 13 August 2019) <https://bair.berkeley.edu/blog/2019/08/13/memorization/> accessed 8 February 2024.
[1101] Claudio Filipi Gonçalves dos Santos, João Paulo Papa 'Avoiding Overfitting: A Survey on Regularization Methods for Convolutional Neural Networks' (2022) Vol 54 No Iss 10s ACM Computing Surveys 1-25 < https://dl.acm.org/doi/pdf/10.1145/3510413 > accessed 8 February 2024.
[1102] Chenlei Fang et al, 'The Overfitting Iceberg' (Machine Learning Carnegie Mellon University 31 August 2020) < https://blog.ml.cmu.edu/2020/08/31/4-overfitting/ > accessed 8 February 2024.
[1103] Joren Verspeurt, 'Applying the GDPR to AI – a practitioner's perspective on some of the main challenges' (AI Summer School Blog KU Leuven 10 January 2023) < https://www.law.kuleuven.be/ai-summer-school/AI-GDPR > accessed 8 February 2024.
[1104] Claudio Filipi Gonçalves dos Santos, João Paulo Papa 'Avoiding Overfitting: A Survey on Regularization Methods for Convolutional Neural Networks' (2022) Vol 54 No Iss 10s ACM Computing Surveys at 3, 5 < https://dl.acm.org/doi/pdf/10.1145/3510413 > accessed 8 February 2024.
[1105] Chenlei Fang et al, 'The Overfitting Iceberg' (Machine Learning Carnegie Mellon University 31 August 2020) < https://blog.ml.cmu.edu/2020/08/31/4-overfitting/ > accessed 8 February 2024.
[1106] Jianchun Chu et al, 'A novel method overcoming overfitting of artificial neural network for accurate prediction: Application on thermophysical property of natural gas'(2021) Vol 28 Case Studies in Thermal Engineering 1-13 < https://www.sciencedirect.com/science/article/pii/S2214157X21005694 > accessed 8 February 2024.

overfitting is likely to negatively affect the accuracy of predictions and is therefore detrimental to the accuracy principle.

ML produces probable, yet inevitably uncertain knowledge and may identify significant correlations. However, these correlations are rarely sufficient to posit the existence of a causal connection and to motivate action based on such uncertain knowledge[1107] (e.g., to grant or not to grant a loan). In other words, probabilistic data does not merit to be considered and treated as facts. Thus, output generated by ML can violate the accuracy principle because it is probabilistic, uncertain and likely inaccurate.[1108] This is amplified by the phenomenon called overfitting and it does not play a role whether 'absolute accuracy' or 'relative accuracy' is considered. Other aspects of ML, such as the risk of biased training data, could lead to inaccurate or wrong representations of data subjects. Output generated by biased training data typically violates the accuracy principle.[1109] Thus, ML can violate the accuracy principle, which constitutes a Type 1 legal problem. When controllers cannot prove the accuracy of the personal data processed, they simultaneously violate the accountability principle. It follows from the accountability principle itself and CJEU case law that the burden of proof concerning compliance with the principles enshrined in Article 5 (1) GDPR lies with the controller.[1110] However, in the case of output generated by means of ML, controllers are unable to prove the accuracy of the personal data processed. This violates the accountability principle.

The inaccuracy problem (Type 1)

As indicated in Section 4.3.1, ML generates output that constitutes uncertain knowledge because it is probabilistic. Overfitting amplifies this problem. Therefore, such output is likely to be inaccurate and can violate the accuracy principle. When controllers cannot prove the accuracy of such personal data, they simultaneously violate the accountability principle.

Affective computing (AC) and the underlying processing of emotion data is in direct contrast with the accuracy principle. Different studies have rebutted the idea that a person's emotional state can be accurately inferred from his facial movements[1111] as suggested by automated face analysis (AFA) systems that deploy AC approaches (see Section 2.2.4.1) to detect emotional states. Research suggests that facial movements are not diagnostic displays that reliably and specifically signal particular emotional states regardless of context, person and culture. It is not possible to confidently infer happiness

[1107] Brent Daniel Mittelstadt et al, 'The ethics of algorithms: Mapping the debate' (2016) Vol 3 Iss 2 Big Data & Society 1, 4.
[1108] Lance Eliot, 'AI Ethics And The Quagmire Of Whether You Have A Legal Right To Know Of AI Inferences About You, Including Those Via AI-Based Self-Driving Cars' *Forbes* (New York, 25 May 2022) < https://www-forbes-com.cdn.ampproject.org/c/s/www.forbes.com/sites/lanceeliot/2022/05/25/ai-ethics-and-the-quagmire-of-whether-you-have-a-legal-right-to-know-of-ai-inferences-about-you-including-those-via-ai-based-self-driving-cars/amp/> accessed 8 February 2024.
[1109] Philipp Hacker, 'Teaching fairness to artificial intelligence: Existing and novel strategies against algorithmic discrimination under EU law' (2018) Vol 55 Iss 4 Common Market Law Review 1143, 1172.
[1110] Case C-175/20 *'SS' SIA* [2022] ECR I-124 paras 77, 81.
[1111] Lisa Feldman Barrett et al. 'Emotional Expressions Reconsidered' (2019) Vol 20 (1) Psychological Science in the Public Interest 1, 46.

from a smile, anger from a scowl or sadness from a frown because these emotion categories are more variable in their facial expressions.[1112] Another study revealed that the accuracy levels of eight commercial automatic classifiers used for facial affect recognition were consistently lower when applied to spontaneous affective behaviours compared to 'posed' affective behaviours. Validation accuracy rates of the tested classifiers varied from 48% to 62%.[1113] When absolute accuracy[1114] is considered, it is obvious that such accuracy rates do not meet this level of accuracy. The same holds true about relative accuracy when AC is applied in the context of hiring procedures. The level of accuracy required for relative accuracy depends on the purpose for processing.[1115] Processing of emotion data for the purpose of recruitment[1116] by means of AC demands a particularly high level of accuracy due to the possible impact on the data subject concerned. Thus, it can be said that the accuracy for such processing essentially must reflect reality and thus ultimately achieve absolute accuracy.

In addition, other means to detect emotions, for example based on speech (see Section 2.2.4.2) and physiological data (see Section 2.2.4.3), have been called into question due to a lack of scientific consensus whether such methods can ensure accurate or even valid results.[1117] While humans can efficiently recognise emotional aspects of speech, it is still an ongoing subject of research to automate this. Research in this context has been restricted to laboratory conditions with full-bandwidth, uncompressed and noise-free audio recordings. However, recent studies indicate that speech compression, filtering, band reduction and the addition of noise reduce accuracy significantly.[1118] Despite this, speech emotion recognition (SER) is already being applied 'in the wild'. Real-world applications of AC aiming to derive emotional states from speech are Amazon's wearable 'Halo', which analyses voice tones to detect user emotions[1119] or Spotify's patented voice assistant[1120] which, based on

[1112] Lisa Feldman Barrett et al. 'Emotional Expressions Reconsidered' (2019) Vol 20 (1) Psychological Science in the Public Interest 1, 46.
[1113] Damian Dupré et al, 'A performance comparison of eight commercially available automatic classifiers for facial affect recognition' (2020) 15 (4) PLoS ONE 1, 10 <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0231968> accessed 8 February 2024.
[1114] Art 29 Working Party, 'Guidelines on the Implementation of the Court of Justice of the European Union Judgment on Google Spain and inc v. Agencia Española de Protección De Datos (AEPD) and Mario Costeja González C-131/12' (WP 225, 26 November 2014), at 15 <https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=667236> accessed 8 February 2024; Paul Voigt, Axel von dem Bussche, *The EU General Data Protection Regulation (GDPR): A Practical Guide* (Springer 2017) 91.
[1115] Case C-434/16, *Nowak* [2017] ECR I-994 para 53.
[1116] For instance, HireVue; see Nathan Mondragon, Clemens Aicholzer, Kiki Leutner, 'The Next Generation of Assessments' (HireVue 2019) <http://hrlens.org/wp-content/uploads/2019/11/The-Next-Generation-of-Assessments-HireVue-White-Paper.pdf> accessed 8 February 2024.
[1117] Kate Crawford et al, 'AI Now Report' (2019) AI Now Institute 12 <https://ainowinstitute.org/publication/ai-now-2019-report-2> accessed 8 February 2024.
[1118] Margaret Lech et al, 'Real-Time Speech Emotion Recognition Using a Pre-trained Image Classification Network: Effects of Bandwidth Reduction and Computing' (2020) Vol 2 Frontiers in Computer Science 1, 3 <https://www.frontiersin.org/articles/10.3389/fcomp.2020.00014/full> accessed 8 February 2024.
[1119] Alex Hern, 'Amazon's Halo wristband: the fitness tracker that listens to your mood' *The Guardian* (London, 28 August 2020) < https://www.theguardian.com/technology/2020/aug/28/amazons-halo-wristband-the-fitness-tracker-that-listens-to-your-mood > accessed 8 February 2024; Austin Carr, 'Amazon's New Wearable Will Know If I'm Angry. Is That Weird?' *Bloomberg* (New York, 31 August 2020) < https://www.bloomberg.com/news/newsletters/2020-08-31/amazon-s-halo-wearable-can-read-emotions-is-that-too-weird > accessed 8 February 2024.
[1120] Daniel Bromand et al, 'Systems and Methods for Enhancing Responsiveness to Utterances Having Detectable Emotion' US Patent Number US 10566010 B2 (Assignee: Spotify AB) February 2020 11 < https://patentimages.storage.googleapis.com/2a/9d/2d/926b58a2bd956f/US10566010.pdf >, accessed 8 February 2024.

commands or other utterances (e.g., 'ugh'), recognises when a user sounds sad and then offers encouragement by 'cheering' the user up.[1121] Emotions are inferred from speech recorded or streamed in daily life environments, sometimes with significantly low accuracy rates. The Hungarian supervisory authority sanctioned a bank for unlawfully processing personal data (voice recordings) based on an SER-powered AI system which promised emotion detection and measurement for customers who called the bank, resulting from voice recordings.[1122] The AI Now Institute at New York University stated AC to be based on 'debunked pseudoscience'[1123] and recommended that 'regulators should ban the use of affect recognition in important decisions that impact people's lives and access to opportunities'.[1124]

In conclusion, it is obvious that processing personal data by AC described in this section violates the accuracy principle, which constitutes a Type 1 legal problem. This holds true when absolute accuracy[1125] is considered, but arguably also in the case of relative accuracy, which is more nuanced and depends on the purpose of processing. I take the view that validation accuracy rates between 48% and 62%[1126] are not acceptable even if the purpose of processing is not particularly sensitive for the data subject concerned. Admittedly, emotions dected by humans can also be inaccurate. However, AI systems function on a much larger scale, and could therefore cause more harm. Because controllers cannot prove the accuracy of the personal data processed, they simultaneously violate the accountability principle. It follows from the accountability principle itself as well as CJEU case law that the burden of proof regarding compliance with principles enshrined in Article 5 (1) GDPR lies with the controller.[1127] However, in the case of output generated by means of AC, controllers are unable to prove the accuracy of the personal data processed. This also violates the accountability principle.

[1121] Josh Mandell, 'Spotify Patents A Voice Assistant That Can Read Your Emotions' *Forbes* (New York, 12 March 2020) < https://www.forbes.com/sites/joshmandell/2020/03/12/spotify-patents-a-voice-assistant--that-can-read-your-emotions/ > accessed 8 February 2024.
[1122] Sebastião Barros Vale, Gabriela Zanfir-Fortuna, 'Automated Decision-Making Under the GDPR: Practical Cases from Courts and Data Protection Authorities' (2022) 48 <https://fpf.org/wp-content/uploads/2022/05/FPF-ADM-Report-R2-singles.pdf> accessed 8 February 2024.
[1123] Kate Crawford et al, 'AI Now Report' (2018) AI Now Institute 8 <https://ainowinstitute.org/publication/ai-now-2018-report-2> accessed 8 February 2024.
[1124] Kate Crawford et al, 'AI Now Report' (2019) AI Now Institute 6 <https://ainowinstitute.org/publication/ai-now-2019-report-2> accessed 8 February 2024.
[1125] Art 29 Working Party, 'Guidelines on the Implementation of the Court of Justice of the European Union Judgment on Google Spain and inc v. Agencia Española de Protección De Datos (AEPD) and Mario Costeja González C-131/12' (WP 225, 26 November 2014), at 15 <https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=667236> accessed 8 February 2024; Paul Voigt, Axel von dem Bussche, *The EU General Data Protection Regulation (GDPR): A Practical Guide* (Springer 2017) 91.
[1126] Damian Dupré et al, 'A performance comparison of eight commercially available automatic classifiers for facial affect recognition' (2020) 15 (4) PLoS ONE 1, 10 <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0231968> accessed 8 February 2024.
[1127] Case C-175/20 *'SS' SIA* [2022] ECR I-124 paras 77, 81.

The rebuttal problem (Type 1)

Research rebutted that a person's emotional state can accurately be inferred from facial movements as suggested by AFA systems powered by AC. There is also a lack of scientific consensus whether other methods used in AC generate accurate results. Output generated by AC systems is likely inaccurate and violates the accuracy principle and simultaneously the accountability principle as controllers cannot prove the accuracy of such personal data.

AI currently lacks reasoning capabilities due to deficiencies in the AI discipline of automated reasoning as outlined in Sections 2.2.5, 4.3.1, 4.4.1 and 4.7.1. Common sense reasoning constitutes a central part of human behaviour and is a precondition for human intelligence. However, common sense reasoning capabilities are still a challenge in AI applications[1128] and AI has been called 'devoid of common sense'.[1129] Apparently, there is not one AI system today which has a semblance of common sense or has capabilities such as human cognition or can think in a manner on par with human thinking.[1130] The lack of progress in providing general automated common sense reasoning capabilities underscores that this is a very difficult problem in the field of AI.[1131] It is not just the hardest problem for AI, it is also considered to be the most important problem.[1132]

Due to these limited reasoning capabilities, AI systems may generate output that is potentially inaccurate and sometimes even discriminatory. Because AI systems lack reasoning capabilities and do not *know why* a specific input should receive some label, they only detect that the particular input correlates with a given label. For example, as outlined in Section 4.3.1, Google's AI system developed for recognising child abuse inaccurately classified a father as criminal[1133] which clearly points to the problem that AI generalises but does not distinguish. The system does not understand what it classifies as 'wrong' or 'right' and neglects the context. An AI system trained with a dataset in which only basketballs were orange is a harmless example. This system might classify all future inputs that are orange as basketballs,[1134] which obviously leads to inaccurate outcomes. Meanwhile, though,

[1128] Shoham Yoav et al, 'The AI Index 2018 Annual Report' (AI Index Steering Committee Stanford University 2018) 64 <https://hai.stanford.edu/sites/default/files/2020-10/AI_Index_2018_Annual_Report.pdf > accessed 8 February 2024.
[1129] Cade Metz, 'Paul Allen Wants to Teach Machines Common Sense' *The New York Times* (New York, 28 February 2018) < https://www.nytimes.com/2018/02/28/technology/paul-allen-ai-common-sense.html > accessed 8 February 2024.
[1130] Lance Eliot, 'AI Ethics And The Quagmire Of Whether You Have A Legal Right To Know Of AI Inferences About You, Including Those Via AI-Based Self-Driving Cars' *Forbes* (New York, 25 May 2022) < https://www-forbes-com.cdn.ampproject.org/c/s/www.forbes.com/sites/lanceeliot/2022/05/25/ai-ethics-and-the-quagmire-of-whether-you-have-a-legal-right-to-know-of-ai-inferences-about-you-including-those-via-ai-based-self-driving-cars/amp/> accessed 8 February 2024.
[1131] Brandon Bennet, Anthony G Cohn, 'Automated Common-sense Spatial Reasoning: Still a Hughe Challenge' in Stephen Muggleton, Nicholas Chater (eds) Human-Like Machine Intelligence (Oxford University Press 2021) 405.
[1132] Gary Marcus, Ernest Davis, *Rebooting AI: Buidling Artificial Intelligence we can trust* (Pantheon Books 2019).
[1133] Kashmir Hill, 'A Dad Took Photos of His Naked Toddler for the Doctor. Goolge Flagged Him as A Criminal' *The New York Times* (New York, 21 August 2022) < https://www.nytimes.com/2022/08/21/technology/google-surveillance-toddler-photo.html > accessed 8 February 2024.
[1134] Zachary C Lipton, 'The Mythos of Model Interpretability' (2018) Vol 16 Iss 3 ACMQueue 3 <https://dl.acm.org/doi/pdf/10.1145/3236386.3241340?download=true> accessed 8 February 2024.

Google's photo app automatically classified images of black people as gorillas.[1135] In some circumstances, neglection of context and lack of common sense can have catastrophic consequences. The case of Molly Russel is a tragic example thereof.[1136] A recommendation system showed Molly Russel, a depressed and lonely teenage girl, 20,000 images promoting depression, suicide and self-harm – including a page of images titled 'Depression content you may like'. This system was programmed to fulfil the objectives Instagram and Pinterest gave them. It is common sense that a teenage girl looking at posts about depression does not want to be made more depressed. Ultimately, Molly Russel committed suicide.[1137] In New Zealand, a man of Asian descent had his passport application rejected because the software that approves photos claimed his eyes were closed.[1138] These examples outline that AI might generate completely inaccurate output and sometimes also discriminatory and defamatory outputs. Therefore, AI reasoning deficiencies are prone to violate the accuracy principle, regardless of whether 'absolute accuracy' or 'relative accuracy' is considered. This leads to a Type 1 legal problem.

The problem of common sense (Type 1)

AI systems can generate inaccurate data due to the reasoning deficiencies in the AI discipline of automated reasoning. AI is devoid of common sense, which may lead to completely inaccurate output. Also, controllers cannot prove the accuracy of such personal data. This violates the accuracy and accountability principles.

4.7.2 Legal problems: Type 2

The accuracy principle does not outline specific levels of accuracy that personal data processed in the context of AI must reach, and there is also no one-size-fits all approach[1139] considering that the level of accuracy depends on the purpose of processing when interpreted as relative accuracy as suggested by relevant case law.[1140] In addition, regulators so far neglected the accuracy principle by not providing substantive guidance on the matter.

When looking for more specific approaches that are helpful to interpret the accuracy principle in the context of AI, it is not possible to simply refer to the concept of accuracy or information quality in

[1135] Crawford Kate, 'Artificial Intelligence's White Guy Problem' *The New York Times* (New York, 25 June 2016) < https://www.nytimes.com/2016/06/26/opinion/sunday/artificial-intelligences-white-guy-problem.html> accessed 8 February 2024.
[1136] Angus Crawford, Bethan Bell, 'Molly Russell inquest: Father makes social media plea' BBC (London, 30 September 2022) <https://www.bbc.com/news/uk-england-london-63073489> accessed 8 February 2024.
[1137] Matija Franklin et al, 'The EU's AI Act needs to address critical manipulation methods' *The OECD.AI Policy Observatory* (Paris, 21 March 2023) <https://oecd.ai/en/wonk/ai-act-manipulation-methods?utm_source=substack&utm_medium=email> accessed 8 February 2024.
[1138] Titcomb James, 'Robot passport checker reject Asian man's photo for having his eyes closed' The Telegraph (London, 7 December 2016) < https://www.telegraph.co.uk/technology/2016/12/07/robot-passport-checker-rejects-asian-mans-photo-having-eyes/> accessed 8 February 2024.
[1139] Dara Hallinan, Frederik Zuiderveen Borgesius, 'Opinions can be incorrect (in our opinion)! On data protection law's accuracy principle' (2020) Vol 10 No 1 IDPL 1, 4.
[1140] Case C-434/16, *Nowak* [2017] ECR I-994 para 53.

the field of computer science.[1141] The latter goes far beyond the principle of accuracy as enshrined in EU data protection law.[1142] Information quality in computer science is a multidimensional concept[1143] covering at least four dimensions, namely, intrinsic, contextual, representational and accessibility information quality. What exactly falls under the scope of these four dimensions seems to vary from the perspectives of academics and practitioners[1144] and further clarification and formalisation of these dimensions is required.[1145] Nevertheless, intrinsic information quality is particularly interesting in the context of the accuracy principle[1146] because accuracy is often considered an intrinsic information quality dimension.[1147] Literature discussing the intrinsic information quality dimension explicitly refers to the terms accuracy and correctness.[1148]

In computer science,[1149] definitions of accuracy vary. Accuracy has been defined as 'the closeness between a value v and a value v′, considered the correct representation of the real-life phenomenon that v aims to represent'.[1150] Another definition states that accuracy 'measures the degree of correctness of a given collection of data'.[1151] Furthermore, two distinct kinds of accuracy exist: syntactic and semantic accuracy. The former is defined as the closeness of a value v to the elements of the corresponding definition domain D and is measured by means of comparison functions.[1152] It is expressed by means of a *numeric* value called edit distance. Take, for example, the incorrect value 'computer viion' that is included in a database that describes the AI disciplines. The edit distance between 'computer viion' and the correct term 'computer vision' is equal to one because it corresponds to the

[1141] Dara Hallinan, Frederik Zuiderveen Borgesius, 'Opinions can be incorrect (in our opinion)! On data protection law's accuracy principle' (2020) Vol 10 No 1 IDPL 1, 4.
[1142] Diana Dimitrova, 'The rise of the personal data quality principle: is it legal and does it have an impact on the right to rectification?' (2021) Vol 12 No 3 European Journal of Law and Technology 9-10; Luciano Floridi, Phyllis Illari, 'Information Quality, Data and Philosophy' in Luciano Floridi, Phyllis Illari (eds) *The Philosophy of Information Quality* (Springer Nature 2014) 6.
[1143] Luciano Floridi, Phyllis Illari, 'Information Quality, Data and Philosophy' in Luciano Floridi, Phyllis Illari (eds) *The Philosophy of Information Quality* (Springer Nature 2014) 6; Leo Pipino et al, 'Developing Measurement Scales for Data Quality Dimensions' in Richard Y Wang et al (eds) *Information Quality* (Routledge 2005 1 edn) 37.
[1144] Yang W Lee et al, 'AIMQ: a methodology for information quality assessment' (2002) Vol 40 Iss 2 Information & Management 133, 134, 136; Luciano Floridi, Phyllis Illari, 'Information Quality, Data and Philosophy' in Luciano Floridi, Phyllis Illari (eds) *The Philosophy of Information Quality* (Springer Nature 2014) 6.
[1145] Carlo Batini, Matteo Palmonari, Gianluigi Viscusi, 'Opening the Closed World: A Survey of Information Quality Research in the Wild' in Luciano Floridi, Phyllis Illari (eds) *The Philosophy of Information Quality* (Springer Nature 2014) 44.
[1146] Also, contextual information quality is at least partially relevant for the accuracy principle as it often refers to the term 'completeness'. However, it also contains other less relevant aspects such as timeliness; see also Yang W Lee et al, 'AIMQ: a methodology for information quality assessment' (2002) Vol 40 Iss 2 Information & Management 133, 134, 136.
[1147] Carlo Batini, Matteo Palmonari, Gianluigi Viscusi, 'Opening the Closed World: A Survey of Information Quality Research in the Wild' in Luciano Floridi, Phyllis Illari (eds) *The Philosophy of Information Quality* (Springer Nature 2014) 60.
[1148] Yang W Lee et al, 'AIMQ: a methodology for information quality assessment' (2002) Vol 40 Iss 2 Information & Management 133, 134, 136; Carlo Batini, Monica Scannapieco, *Data Quality* (Springer 2006) 20-27; Yang W Lee et al, 'AIMQ: a methodology for information quality assessment' (2002) Vol 40 Iss 2 Information & Management 133, 134, 136; Luciano Floridi, Phyllis Illari, 'Information Quality, Data and Philosophy' in Luciano Floridi, Phyllis Illari (eds) *The Philosophy of Information Quality* (Springer Nature 2014) 7.
[1149] In the domain of Information Quality.
[1150] Carlo Batini, Monica Scannapieco, *Data Quality* (Springer 2006) 20.
[1151] Thomas C Redman, 'Measuring Data Accuracy: A Framework and Review' in Richard Y Wang et al (eds) *Information Quality* (Routledge 2005 1 edn) 24.
[1152] Carlo Batini, Monica Scannapieco, *Data Quality* (Springer 2006) 20.

insertion of the letter ‘s’ in the value ‘computer viion’. Therefore, the syntactic accuracy is 1.[1153] Semantic accuracy is more difficult to measure.[1154] Semantic accuracy coincides with the concept correctness and is measured with yes/no or correct/incorrect. For measuring the semantic accuracy of a certain value v, the true corresponding value must be known, or it must at least be possible with additional knowledge to deduce whether the value v is or is not the true value.[1155] Semantic accuracy seems to be quite similar to absolute accuracy in the legal sense as is measured with ‘correct/incorrect’. Syntactic accuracy is more nuanced and allows for development of more flexible approaches, for example, by means of defining accuracy ranges that are considered still accurate (e.g., syntactic accuracies between 1 and 10 are considered accurate enough) which could prove to be helpful regarding relative accuracy.

In addition, the interpretation of the term ‘completeness’ varies in computer science and might relate to absolute or relative accuracy in the legal sense. For example, completeness is described as ‘the extent to which data are of sufficient breadth, depth, and scope for the *task at hand*’[1156] which seems to be similar to the notion of relative accuracy in the legal sense. Another interpretation of completeness in computer science seems to be comparable to absolute accuracy in the legal sense. A data unit consisting of one or more components (such as number, file, record), is complete if each data item constituting the data unit has been assigned a value in accordance with the data definition for the data item. If the latter is not fulfilled, the data unit is incomplete.[1157]

The concepts of accuracy and completeness in computer science will not be the ultimate solution to applying the accuracy principle. With semantic accuracy, the problem is that the correct value might *not* be known, for example, in the case of predictions or inferences produced by ML which are solely probabilistic (see Section 4.7.1). Syntactic accuracy might be too imprecise because it only allows one to calculate the closeness of a value but does not indicate that a value is inaccurate or incorrect. More generally, there is no single way to measure the accuracy of the data under all circumstances. Measuring the accuracy of the data is particularly difficult due to the nature of data. Determining data accuracy must necessarily make reference to human knowledge, other data or the real world.[1158] Another issue with respect to accuracy in computer science is a phenomenon called concept drift: Even if an AI system might initially be accurate, accuracy might change over time when it is applied in

[1153] Carlo Batini, Monica Scannapieco, *Data Quality* (Springer 2006) 21.
[1154] Carlo Batini, Matteo Palmonari, Gianluigi Viscusi, ‘Opening the Closed World: A Survey of Information Quality Research in the Wild’ in Luciano Floridi, Phyllis Illari (eds) *The Philosophy of Information Quality* (Springer Nature 2014) 54.
[1155] Carlo Batini, Monica Scannapieco, *Data Quality* (Springer 2006) 20.
[1156] Richard Y Wang, Diane M Strong, ‘Beyond Accuracy: What Data Quality Means to Data Consumers’ (1996) Vol 12 No 4 Journal of Management Information Systems, 5, 32 (emphasis added).
[1157] Leo Pipino et al, ‘Developing Measurement Scales for Data Quality Dimensions’ in Richard Y Wang et al (eds) *Information Quality* (Routledge 2005 1 edn) 44.
[1158] Thomas C Redman, ‘Measuring Data Accuracy: A Framework and Review’ in Richard Y Wang et al (eds) *Information Quality* (Routledge 2005 1 edn) 23.

practice or 'real world', in particular when the behaviour of individuals that the system seeks to evaluate changes. In this case, an AI system is likely to inaccurately evaluate these individuals.[1159] Validation accuracy which tests ML models on data unseen during training to estimate how well the model is expected to perform later in real life seems to be an interesting instrument for applying the accuracy principle in practice,[1160] particularly regarding relative accuracy. Validation accuracy rates (e.g., 80, 90 or 100%) could be helpful when applying the accuracy principle in practice because the degree of accuracy to be achieved always depends on the purpose of processing.[1161]

There has been no exchange of ideas between computer science and law on the matter of information quality and accuracy.[1162] Corresponding interdisciplinary research is a relatively recent development.[1163] This is unfortunate because such interdisciplinary research could be helpful when applying the accuracy principle to AI. Nevertheless, within this section I have outlined that the concepts of information quality, accuracy, completeness and validation accuracy from research in the field of computer science might be helpful to interpret the accuracy principle in the context of AI. More interdisciplinary research is needed to develop an interpretation of the accuracy principle which is valid and practical both from a legal and computational perspective.

Consequently, when assessing the accuracy of personal data generated by means of AI, the model upon which inferred personal data are based also must be considered to ensure a comprehensive assessment. The quality of such information, i.e. the personal data generated by means of AI, is affected by the quality of the AI system used to generate it.[1164] Regulators so far neglected the accuracy principle by not providing substantive and practice-oriented guidance on the matter, which reduces legal certainty. This makes it difficult if not impossible to enforce the accuracy principle in the context of AI, both in regulatory enforcement (by SAs)[1165] and in private enforcement pursued by data subjects and their representatives. This leads to a Type 2 legal problem and is caused by the accuracy principle itself and may arise *regardless* of which AI discipline it is applied to. Nonetheless, this problem is most apparent regarding predictions, inferences and other probabilistic output generated by means of ML and AC (see also Section 4.3.1).

[1159] Diana Dimitrova, 'The rise of the personal data quality principle: is it legal and does it have an impact on the right to rectification?' (2021) Vol 12 No 3 European Journal of Law and Technology 21.
[1160] Ethem Alpaydin, *Machine Learning: The New AI* (3rd edn MIT Press 2016) 181.
[1161] Case C-434/16, *Nowak* [2017] ECR I-994 para 53.
[1162] Dara Hallinan, Frederik Zuiderveen Borgesius, 'Opinions can be incorrect (in our opinion)! On data protection law's accuracy principle' (2020) Vol 10 No 1 IDPL 1, 4.
[1163] Burkhard Schäfer, 'Information Quality and Evidence Law: A New Role for Social Media, Digital Publishing and Copyright Law?' in Luciano Floridi, Phyllis Illari (eds) *The Philosophy of Information Quality* (Springer Nature 2014) 217.
[1164] Burkhard Schäfer, 'Information Quality and Evidence Law: A New Role for Social Media, Digital Publishing and Copyright Law?' in Luciano Floridi, Phyllis Illari (eds) *The Philosophy of Information Quality* (Springer Nature 2014) 217.
[1165] See Articles 51 to 58 GDPR.

The guidance problem (Type 2)
The lack of guidance concerning the accuracy principle and the absence of interdisciplinary research in the fields of computer science and law leads to legal uncertainty and makes it difficult if not impossible to enforce in the context of AI.

4.7.3 Legal problems: Type 3

The guidance problem explained in Section 4.7.2 automatically leads to a Type 3 legal problem. The accuracy principle is not fit for purpose to effectively protect[1166] data subjects from being inaccurately represented in the form of output generated by AI. In its case law, the CJEU has repeatedly stressed that EU data protection law aims to effectively protect the data subject's personal data against risk of misuse.[1167] A principle that lacks substantive detail cannot prevent misuse in the form of inaccurate representations of data subjects. Likewise, it cannot ensure a high level of data protection.[1168] Due to the accuracy principle's lack of detail caused by absent guidance and respective interdisciplinary research, it fails to achieve the GDPR's aim to establish a strong and coherent data protection framework[1169] when considering that principles provide the basis for the protection of personal data[1170] in the GDPR.

The fairness principle as well as the accuracy principle as discussed in Sections 4.3.2 and 4.7.2 respectively have in common that they lack sufficient guidance when applied to AI. This leads to legal uncertainty and ultimately to a Type 3 legal problem. The lack of regulatory guidance and the absence of interdisciplinary research make these principles 'incomputable'. As it is the case with the purpose limitation and data minimisation principles, [1171] measurable definitions of the accuracy and fairness principles and concrete indications on how to practically implement them are needed to make them 'computable'. To replicate and apply legal reasoning, AI requires the translation of the linguistic categories used by law into mathematical functions. This is not a straightforward task, because there is an element of flexibility and contestability in natural language used to express juridical forms that cannot be completely captured by mathematical algorithms.[1172] Whereas this points more generally to

[1166] Case C-645/19 [2021] *Facebook Ireland* ECR I-483 para 45; Case C-319/20, *Meta Platforms Ireland Limited* [2022] ECR I-322 para 73.
[1167] Joined Cases C-293/12 and C-594/12 *Digital Rights Ireland* [2014] ECR-I 238 paras 54 and 66; Joined Cases C-203/15 and C-698/15 [2016] *Tele 2 Sverige* ECR-1 970 paras 109 and 122; Case C-362/14 *Schrems v Data Protection Commissioner* [2015] ECR I-650 para 91.
[1168] Recitals 6, 10 as well as CJEU case law, such as Case C-534/20, *Leistritz AG* [2022] ECR I-594 para 26; Case C-645/19 [2021] *Facebook Ireland* ECR I-483 para 45; Case C-511/18, *La Quadrature du Net* [2020] ECR I-791 para 106; Case C-131/12, *Google Spain* [2014] ECR I-317 para 66; Case C-154/21 *Österreichische Post AG* [2023] ECR I-3 para 44; Case C-132/21 *Nemzeti* [2023] ECR I-2 para 42.
[1169] Recital 7 GDPR.
[1170] Cécile de Terwangne, Commentary of Article 5 in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020) 313.
[1171] Asia Biega, Michèle Finck, 'Reviving Purpose Limitation and Data Minimisation in Data-Driven Systems' (2021) Technology and Regulation 44, 58 and 61 <https://arxiv.org/abs/2101.06203> accessed 8 February 2024.
[1172] Christopher Markou, Simon Deakin, 'Ex Machina Lex: Exploring the Limits of Legal Computability' in Simon Deakin, Christopher Markou (eds) *Is Law Computable?: Critical Perspectives on Law and Artificial Intelligence* (Hart Publishing 2020) 66.

the limits of legal computability, it holds especially true in case of principles, which are by nature less concrete and provide a great deal of flexibility when applied in practice. This is even more true where the substantive meaning of principles remains largely unclear, as is the case with the fairness and accuracy principle.

Computability of principles is an essential requirement to develop AI systems which implement data protection principles and thus comply with the concept of data protection by design and default according to Article 25 GDPR. Although the latter, as introduced in Section 3.3.3.9, does not appear under the principles for processing named in Article 5 of the GDPR, it is closely intertwined with them. It obliges controllers to put in place, both at the design *and* processing stage,[1173] technical and organisational measures 'that are designed to implement data protection principles.'[1174]

As pointed out in the elusiveness problem discussed in Section 4.3.2, little has been written what 'fair processing' really means[1175] and on the application of the fairness principle in practice.[1176] This renders the fairness principle incomputable. In addition, interdisciplinary research highlights that certain legally prohibited kinds of discrimination are too contextual, intuitive and open to judicial interpretation to be automated. Many of the available computational implementations of the fairness principle are thus not able to adequately reflect its legal requirements.[1177]

Uncertainties regarding the proper meaning of the accuracy principle render it incomputable, even when concepts of accuracy and information quality elaborated in the field of computer science are considered (see also Section 4.7.2). The incomputability of both the fairness and accuracy principles creates a Type 3 legal problem, both regarding the principles themselves as well as the concept of Data Protection by Design and Default ('DPbDD') according to Article 25 GDPR as introduced in Section 3.3.3.9. The computability of principles is an essential requirement to develop AI systems that implement data protection principles at both the design and processing stages. At first sight, the concept of DPbDD seems promising and relevant considering new technologies such as AI. However, this concept fails to deliver what it promises because it requires controllers to implement, by means of technical measures, data protection principles which are essentially *incomputable*. This is significant when considering that principles provide the basis for the protection of personal data in EU data protection law.[1178] Developers cannot implement these principles in the design phase and during the

[1173] Article 25, Recital 78 GDPR.
[1174] Article 25 GDPR.
[1175] Winston J Maxwell, 'Principle-based regulation of personal data: the case of fair processing' (2015) Vol 5 No 3 International Data Privacy Law 205.
[1176] Damian Clifford, Jef Ausloos 'Data Protection and the Role of Fairness' (2018) Vol 37 No 1 Yearbook of European Law 130, 184.
[1177] Asia Biega, Michèle Finck, 'Reviving Purpose Limitation and Data Minimisation in Data-Driven Systems' (2021) Technology and Regulation 44, 59 <https://arxiv.org/abs/2101.06203> accessed 8 February 2024.
[1178] Cécile de Terwangne, Commentary of Article 5 in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020) 313.

actual use of AI systems. Thus, the DPbDD is not fit for purpose to protect the fundamental right to data protection. This constitutes a Type 3 legal problem. Incomputable principles are not fit for purpose to achieve the GDPR's aim to establish a strong and coherent data protection framework[1179] when considering that principles provide the basis for the protection of personal data[1180] in the GDPR. This Type 3 legal problem occurs regardless of which AI discipline the fairness and accuracy principles are applied to because the incomputability of these two principles causes the legal problem. Therefore, it is a general problem and relates to all AI disciplines as introduced in Chapter 2. To be clear, I do not suggest principle-based processing in AI systems is impossible as it cannot be computed abstractly. Instead, the incomputability is caused by the need for more guidance and more interdisciplinary research. Mathematical interpretations of principles are needed to render them computable.[1181]

The incomputability problem (Type 3)

The lack of guidance concerning the fairness and accuracy principle renders them incomputable. Developers cannot encode these principles in the design phase and during the actual use of AI systems as required by the concept of data protection by design and default which obliges controllers to implement the data protection principles by technical means. Incomputable principles are not fit for purpose to protect the fundamental right to data protection.

As outlined in Section 4.7.2, the accuracy principle is difficult to enforce in practice due to the absence of specific levels of accuracy that could be considered when assessing the accuracy of personal data in the context of AI. This is particularly problematic when considering that the accuracy principle is closely intertwined with the right to rectify personal data according to Article 16 GDPR.[1182] The AI disciplines ML and AC provide new means to generate inferences, predictions and other output. In Section 4.7.1 I have outlined that such outputs can be inaccurate. The lack of guidance regarding the accuracy principle makes it difficult for data subjects to enforce their right to rectification. I discuss this problem in Section 5.7.

4.8 Enhanced protection for 'special data'

The notion of special categories of personal data is broadly interpreted by the CJEU. It ruled that personal data *indirectly* revealing special categories of personal data defined in Article 9 (1) GDPR is also covered by the latter provision.[1183] In this ruling, the CJEU followed AG Pikamäe's opinion by stating that 'the verb "reveal" is consistent with taking into account processing of inherently

[1179] Recital 7 GDPR.
[1180] Cécile de Terwangne, Commentary of Article 5 in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020) 313.
[1181] Asia Biega, Michèle Finck, 'Reviving Purpose Limitation and Data Minimisation in Data-Driven Systems' (2021) Technology and Regulation 44, 58 and 61 <https://arxiv.org/abs/2101.06203> accessed 8 February 2024.
[1182] See Recitals 6 and 10 GDPR, as well as CJEU case law, such as Case C-534/20, *Leistritz AG* [2022] ECR I-594 para 26; Case C-645/19 [2021] *Facebook Ireland* ECR I-483 para 45; Case C-511/18, *La Quadrature du Net* [2020] ECR I-791 para 106; Case C-131/12, *Google Spain* [2014] ECR I-317 para 66.
[1183] Case C-184/20, *OT* [2022] ECR I-601, paras 117-128.

sensitive data, as well as data revealing information of that nature *indirectly*, following an intellectual operation involving deduction or cross-referencing'.[1184] In the context of AI, this ruling is quite important because ML might generate personal data that indirectly reveal special categories of personal data. ML models that apply dimensionality reduction (see Section 2.2.1.2) on easily accessible digital records of behaviour, for example, Facebook likes, may reveal and predict highly sensitive personal attributes such as sexual orientation, ethnicity, religious and political views and personality traits.[1185] It is now clear that the processing of such data falls under the scope of Article 9 GDPR. However, the broad interpretation of special categories of personal data does not solve all the legal problems that might arise due to AI. This is mainly due to the principle[1186] of enhancing protection for special data and the legislator's approach to enumerate such data exhaustively. In Section 4.8.3, I outline that this approach has significant consequences considering the technological developments facilitated by AI. Both GDPR and its predecessor use the term 'special categories' of personal data, but also refer to 'sensitive personal data' in the recitals. [1187] In order to avoid confusion, I will use the term 'special data' to refer to data that *are* in fact, protected under the GDPR and 'sensitive data' to refer to data that are, in fact, *not protected* under the GDPR (although they arguably should be).

As outlined in Section 3.3.1.2, the rationale for ensuring enhanced protection for special data stems from their particular sensitive nature (Recital 51 GDPR). Processing of special data can constitute a particularly serious interference with the fundamental rights to privacy and data protection.[1188] In view of the SAs, it is needed to specifically protect special data because misuse of such data may have more severe consequences for the data subjects than misuse of 'regular' personal data.[1189] This is underscored by Recital 51 GDPR, which states that 'processing [of sensitive personal data] could create significant risks to fundamental rights and freedoms'. Nevertheless, the principle[1190] of enhancing protection for special categories of personal data is not undisputed.[1191] This will be discussed in Section 6.3.

[1184] Case C-184/20, *OT* [2022] ECR I-601, paras 123, emphasis added; Case C-184/20, *OT* [2022] ECR I-601, Opinion of AG Pikamäe, para 85.
[1185] Michal Kosinski, David Stillwell, Thore Graepel, 'Private traits and attributes are predictable from digital records of human behaviour' (2013) Vol 110 No 15 PNAS, 5802.
[1186] For the purpose of this thesis, I regard this choice as a principle so that it neatly matches the approach taken, distinguishing between principles and rights.
[1187] See Recitals 10, 51 GDPR, Recitals 34 and 70 Data Protection Directive which refer to sensitive but not 'special' categories of personal data
[1188] Case C-184/20, *OT* [2022] ECR I-601, para 126; Case C-136/17, GC and Others [2019] ECR I-773 para 44; Recital 51 GDPR.
[1189] Art 29 Working Party, 'Advice paper on special categories of data ("sensitive data")' (20 April 2011) at 4.
[1190] Admittedly, this is not a traditional data protection principle. Nonetheless, it could be regarded as a principle in a broader sense, which then also aligns with the approach taken in this chapter.
[1191] Ludmila Georgieva, Christopher Kuner Commentary of Article 9 in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020) 370; Lee A. Bygrave, *Data Privacy Law: An International Perspective* (OUP 2014) 165; Lokke Moerel, Corien Prins, 'Privacy for the Homo Digitalis: Proposal for a New Regulatory Framework for Data Protection in the Light of Big Data and the Internet of Things' (2016) p 11 and 56 <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2784123> accessed 8 February 2024.

4.8.1 Legal problems: Type 1

As will be outlined in Section 4.8.3, the problem with respect to the approach to exhaustively enumerate special data arises because AI provides unprecedented means of generating and otherwise processing new types or categories of sensitive personal data. The exhaustive list of sensitive data contained in the GDPR does not keep up with technological developments facilitated by AI. This means that the strict rules concerning the processing of sensitive data do not apply to new types of sensitive personal data facilitated by AI. Nonapplicable or nonexisting provisions cannot be violated, and therefore no specific Type 1 legal problems arise.

4.8.2 Legal problems: Type 2

As outlined in Section 4.8.1, provisions that are not applicable or do not yet exist cannot be violated. Consequently, no specific Type 2 legal problems arise.

4.8.3 Legal problems: Type 3

AI provides an unprecedented means to generate and otherwise process arguably new types or categories of sensitive personal data. This causes legal problems regarding the principle of enhancing protection for special data and the legislator's approach to define special data exhaustively. Definitions contained in the current legal framework do not keep up with technological developments facilitated by AI. I demonstrate this issue by discussing emotion data, location data, neurodata and mental data, respectively.

Emotion data

By means of AC, machines may gain access to the emotional life of individuals, information that is highly personal, intimate and private.[1192] In fact, all emotions are by definition personal[1193] and revealing them makes an individual more vulnerable.[1194] A commonly agreed definition of emotion in any of the disciplines that study this phenomenon does not exist.[1195] For the purpose of this thesis, I define emotion data as information relating to emotions of an individual. To avoid lengthy discussions on what emotions are, I simply refer to the six most-used emotion categories[1196] in emotion research:

[1192] Rosalind W Picard, *Affective Computing* (MIT Press 1997) 118.
[1193] Not meaning personal in the sense of personal data but more to the common understanding of the notion.
[1194] Aaron Ben-Ze'Ev, *The Subtlety of Emotions* (MIT Press 2000) 183.
[1195] Kevin Mulligan, Klaus R. Scherer, 'Toward a Working Definition of Emotion' (2012) Vol. 4 No. 4 Emotion Review 345-537.
[1196] These six emotions refer to research conducted by psychologists in the early seventies that developed the methodology of 'basic emotions'; see Paul Ekman, Wallace v Friesen, 'Constants across cultures in the face and emotion' (1971) Vol 17 (2) Journal of Personality and Social Psychology 124.

anger, disgust, fear, happiness, sadness and surprise.[1197] These six 'basic emotions'[1198] are further described in Section 2.2.4.1. It should be noted that emotion data constitutes a subcategory of mental data (see Figure 2.1). Emotions are felt as personal because they relate to a person's values[1199] and express what a person cares about.[1200] Because there is an inherent relationship between emotions and personhood[1201] and privacy is considered fundamental to the maintenance of human dignity and the boundary to one's personhood,[1202] information regarding emotions is sensitive and intimate.[1203] When emotion data constitute personal data because the data subject is identified or identifiable, the question arises whether such data are specifically protected as 'special data'.

Considering the special categories of personal data defined in Article 9 (1) GDPR and its corresponding recitals,[1204] emotion data itself is never protected as a special category of personal data under the GDPR, despite its sensitive and intimate nature.[1205]

Ultimately, the approach taken in AC determines whether processing of *personal data used to detect or derive* emotion data falls under the scope of Article 9 GDPR. A distinction can be made between single-modal affect recognition and multimodal affect recognition approaches in AC. [1206] Single-modal approaches are divided into text sentiment analysis, audio emotion recognition, visual emotion recognition focussing on facial expression and body gestures and physiological-based emotion recognition systems. [1207] Physiologically-based emotion recognition systems include AC systems that detect emotional states from EEG and ECG. ECG-based emotion recognition systems record the physiological changes of the human heart in order to detect the corresponding waveform transformation, which provides information for emotion recognition.[1208] For example, ECG-based emotion recognition systems can be applied when listening to music.[1209] EEG is a non-invasive method consisting in detection

[1197] Lisa Feldman Barrett et al 'Emotional Expressions Reconsidered' (2019) Vol 20 (1) Psychological Science in the Public Interest 1, 52.
[1198] Eiman Kanjo et al, 'Emotions in context: examining pervasive affective sensing systems, applications, and analyses' (2015) Vol 19 Personal and Ubiquitous Computing 1197, 1204 <https://link.springer.com/content/pdf/10.1007/s00779-015-0842-3.pdf> accessed 8 February 2024.
[1199] Heather C Lench, Zakari Koebel Capenter, 'What Do Emotions Do for Us?' in Heather C Lench (ed) *The Function of Emotions* (Springer 2018) 1, 142.
[1200] Giovanni Stanghellini, René Rosfort, *Emotions and Personhood: Exploring Fragility – Making Sense of Vulnerability* (OUP 2013) 142.
[1201] Ibid 149.
[1202] William S Brown, 'Technology, Workplace Privacy and Personhood' (1996) Vol 15 Journal of Business Ethics 1237, 1243.
[1203] Andrew McStay, 'Emotion AI, soft biometrics and the surveillance of emotional life: An unusual consensus on privacy' (2020) Vol 7 Iss 7 Big Data & Society 1, 4.
[1204] Recitals 51, 52, 53 GDPR.
[1205] Contrary to Clifford's view that argues this 'will clearly result in the processing of sensitive personal data'; see Damian Clifford, 'Citizen-consumers in a personalised Galaxy: Emotion influenced decision-making, a true path to the dark side?' (2017) CiTiP Working Paper 31/2017, 21 <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3037425> accessed 8 February 2024.
[1206] Yan Wang et al, 'A systematic review on affective computing: emotion models, databases, and recent advances' (2022) Volumes 83-84 Information Fusion 19-52.
[1207] Ibid 19, 21.
[1208] Ibid 19, 35-36.
[1209] Yu-Liang Hsu et al, 'Automatic ECG-Based Emotion Recognition in Music Listening' (2020) Vol 11 No 1 IEEE Transactions on Affective Computing 85-99.

and registration of electrical activity occurring in the brain.[1210] EEG-based emotion recognition systems directly measure changes in brain activities, which provides internal features of emotional states.[1211]

Only physiologically-based emotion recognition systems in AC involve the processing of special data as defined in the GDPR. Information processed by these systems falls under the definition of health data, which covers not only physical or mental health, but also 'any information (…) on the *physiological* or biomedical state of the data subject independent of its source'.[1212] Consider, for example, AC applications that derive emotion data from physiological data such as heart rate, blood pressure and skin conductance. Research has shown that heart rate variability provides a novel marker to recognise emotions in humans.[1213] Information relating to heart rate, blood pressure and skin conductance falls under the definition of health data and is protected as a special category of personal data according to the GDPR.[1214] Automated face analysis systems (AFA) that try to detect depression from analysing an individual's facial expressions in videos arguably process (mental) health data, even if the data subject concerned is completely healthy.[1215]

Most of the single-modal affect recognition systems pursued in AC do not amount to the processing of special data. AC systems deploying approaches such as text sentiment analysis, audio emotion recognition and visual recognition of emotion focussing on facial expressions and body gestures do *not* involve the processing of special categories of personal data.[1216] Information processed within these approaches and derived emotion data are thus not protected as special personal data under the GDPR, despite their sensitive and intimate nature.[1217] This also holds true when biometric data are used for AC to detect the emotional state of the individual concerned, for example in the context of AFA systems and emotion detection based on an individual's voice and speech.[1218] Biometric data according to Article 9 (1) GDPR is only protected as special personal data if it is used for the *purpose*

[1210] Szczepan Paszkiel, *Analysis and Classification of EEG Signals for Brain–Computer Interfaces* (Springer Nature 2020) 3.
[1211] Yan Wang et al, 'A systematic review on affective computing: emotion models, databases, and recent advances' (2022) Volumes 83-84 Information Fusion 19, 35; Jianhua Zhang et al, 'Emotion recognition using multi-modal data and machine learning techniques: A tutorial and review' (2020) Vol 59 Information Fusion 103-126.
[1212] Recital 35 GDPR (emphasis added).
[1213] Quintana Daniel et al. 'Heart rate variability is associated with emotion recognition: Direct evidence for a relationship between the automatic nervous system and social cognition' (2012) Vol 86 No 2 International Journal of Psychophysiology 168.
[1214] Article 3 (15) and 9 (1) GDPR; Recital 15 GDPR.
[1215] Marcello Ienca, Gianclaudio Malgieri, 'Mental data protection and the GDPR' (2022) Vol 9 Iss 1 Journal of Law and the Biosciences 1, 9.
[1216] Recitals 51, 52, 53 GDPR.
[1217] Contrary to Clifford's view that argues this 'will clearly result in the processing of sensitive personal data'; see Damian Clifford, 'Citizen-consumers in a personalised Galaxy: Emotion influenced decision-making, a true path to the dark side?' (2017) CiTiP Working Paper 31/2017, 21 <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3037425> accessed 8 February 2024.
[1218] Note that Article 29 WP considered voice as biometric data, Art 29 Working Party, 'Opinion 4/2007 on the concept of personal data' (WP 136, 20 June 2007) at 8.

of uniquely *identifying* an individual. This means 'processed through a specific technical means allowing the unique identification or authentication of a natural person'.[1219]

According to regulatory guidance adopted by EU supervisory authorities, biometric identification typically involves 'the process of comparing biometric data of an individual (acquired at the time of the identification) to a number of other biometric templates stored in a data database (i.e. a one-to-many matching process)'.[1220] For example, HumeAI[1221] provides AC-powered tools helping recruiters to assess personality traits and detect emotional states of job candidates disclosed during automated video assessments based on facial expressions. This system does not process biometric data in the form of facial expressions to uniquely identify the job candidate, as required by Article 9 (1) GDPR. Rather, it detects the emotional states the candidate portrays during the automated video assessment. Identification is achieved through other means beforehand: when the candidate reveals its name, the other identifiable information. The same applies to any other AC system aiming to detect emotional states from facial expressions,[1222] for instance those offered by the companies Realeyes[1223] or Tawny.[1224]

This also holds true when AC systems use biometric data in the form of speech, as discussed in Section 2.2.4.2 to detect the emotions of the individual concerned. Consider an AC system that advises a call centre agent to speak with more empathy because the customer seems to be angry according to the automated speech and voice analysis. Such a system does not process biometric data for identification purposes. Regulatory guidance generally considers voice to be biometric data[1225] as defined in Article 4 (14) GDPR, i.e. personal data 'resulting from specific technical processing relating to the physical, physiological or behavioural characteristics of a natural person, which allow or confirm the unique identification of that natural person, such as facial images or dactyloscopic data'. However, according to Article 9 (1) GDPR, biometric data *only* qualifies as a special category of personal data if it is *used* for the *purpose* of uniquely *identifying* an individual.[1226] AG Pikamäe has termed this the 'purposive approach'.[1227] This purposive approach causes the inapplicability of Article 9 GDPR when biometric data are processed for purposes other than uniquely identifying an individual. The GDPR thus links the use of biometric data *exclusively* to the purpose of identification and therefore excludes

[1219] Recital 51 GDPR, the same recital states that processing of photographs should not systematically be considered to be processing of special categories of personal data.
[1220] Art 29 Working Party, 'Opinion 3/2012 on developments in biometric technologies' (WP 193, 27 April 2012) at 5.
[1221] See < https://hume.ai/products/facial-expression-model/ > and < https://gethume.com/blog5/artificial-intelligence-for-recruiting > accessed 26 March 2023. > accessed 8 February 2024.
[1222] Provided that identification is not based on biometric data.
[1223] See <https://www.realeyesit.com/> accessed 8 February 2024.
[1224] See <https://www.tawny.ai/product> accessed 8 February 2024.
[1225] Art 29 Working Party, 'Opinion 4/2007 on the concept of personal data' (WP 136, 20 June 2007) at 8; European Data Protection Board, 'Guidelines 02/2021 on Virtual Voice Assistants' (16 May 2011) at 31.
[1226] Article 9 (1) GDPR.
[1227] Case C-184/20 [2021] *OT* ECR I-991 Opinion of AG Pikamäe para 86.

all biometric data processed for *other purposes*,[1228] such as emotion recognition purposes. Hence, emotion data are not protected as special data under the GDPR, nor as biometric data defined in Article 4 (14) GDPR. This interpretation is also in line with the regulatory enforcement pursued by the Hungarian SA. In this case, a Hungarian bank used an AI system with the aim to detect and measure emotions of customers that called the bank's customer service.[1229] In its decision, the Hungarian SA reached the conclusion that emotion data did not constitute special data according to Article 9 (1) GDPR. Voice recordings (biometric data) were not used to identify the data subject, nor did the inferences drawn by the AI system reveal data with respect to physical or mental health.[1230]

In some cases, AC systems process special personal data to *derive or detect* emotion data. This applies to physiological-based emotion recognition systems that process information like heart rate, blood pressure and skin conductance. Such information constitutes health data, which is a special category of personal data in the GDPR. Nevertheless, the highly sensitive detected emotion data itself *never* constitutes special data under the GDPR, irrespective of which affect recognition (single-modal or multimodal) approach in AC is deployed. Thus, inherently sensitive personal data are not specifically protected in EU data protection law. This leads to a significant gap in legal protection.

The EU Commission's proposed ePrivacy Regulation[1231] as well as the compromise text[1232] used for the EU's trilogue procedure label information relating to emotions as highly sensitive. This implies that emotion data might be subject to different levels of protection depending on the applicable laws. In case both the GDPR[1233] and the future ePrivacy Regulation are triggered, emotion data will be protected as sensitive data according to the ePrivacy Regulation, but not according to the GDPR.[1234] Such a situation might be confusing and disadvantageous for data subjects, but also for companies that need to comply with the GDPR and the ePrivacy Regulation. In addition, regulating emotion data by means of different levels of protection does not seem to contribute to legal certainty.

[1228] Gloria González Fuster, Michalina Nadolna Peeters, 'Person identification, human rights and ethical principles. Rethinking biometrics in the era of artificial intelligence' (2021) 2, 20, 25 <https://www.europarl.europa.eu/RegData/etudes/STUD/2021/697191/EPRS_STU(2021)697191_EN.pdf> accessed 8 February 2024.
[1229] Sebastião Barros Vale, Gabriela Zanfir-Fortuna, 'Automated Decision-Making Under the GDPR: Practical Cases from Courts and Data Protection Authorities' (2022) 48 <https://fpf.org/wp-content/uploads/2022/05/FPF-ADM-Report-R2-singles.pdf> accessed 8 February 2024.
[1230] Cesar Manso-Sayao, Summary of Hungarian SA Decision NAIH-85-3/2022 < https://gdprhub.eu/NAIH_(Hungary)_-_NAIH-85-3/2022 > accessed 8 February 2024.
[1231] Proposal for a Regulation of the European Parliament and of the Council concerning the respect for private life and the protection of personal data in electronic communications and repealing Directive 2002/58/EC (Regulation on Privacy and Electronic Communications) Recital 2 <https://data.consilium.europa.eu/doc/document/ST-12633-2019-INIT/en/pdf > accessed 8 February 2024.
[1232] Council of the EU 6087/21 recital 2 <https://data.consilium.europa.eu/doc/document/ST-6087-2021-INIT/en/pdf > accessed 8 February 2024.
[1233] Namely, where emotion data must be considered personal data because the data subject is identified or identifiable.
[1234] Provided that the proposed ePrivacy Regulation will not be amended with regard to the sensitivity of emotion data.

Emotion data are inherently sensitive due to the intrinsic relationship between emotions and personhood[1235] and therefore merit specific protection as 'special data' according to the GDPR. Furthermore, the processing of emotion data could create significant risks to fundamental rights and freedoms in the sense of Recital 51 GDPR, the personal autonomy of the data subject, in particular. As outlined in Section 4.3.3, information concerning the emotional state of an individual might be particularly helpful to manipulate this individual because emotions play an important role in the elicitation of autonomous motivated behaviour[1236] and reasoning.[1237] AC provides access to emotion data of individuals and may affect people's decisions and lives in unprecedented ways. This holds particularly true regarding manipulation that operates by relying on facts about the subject's psychology such as knowledge about its emotions and desires.[1238] Emotions can have significant effects on economic transactions and play a powerful role in everyday economic choices.[1239] This affects personal autonomy, i.e. the idea 'that people should make their own lives'[1240] when facing freely both existential and every day's choices.[1241]

The fact that emotion data do not receive specific protection under the GDPR despite its highly sensitive nature and the risks relating to the data subject's personal autonomy leads to a Type 3 legal problem. The approach to exhaustively enumerate special categories of personal data creates a protection gap with regard to the processing of new kinds of sensitive personal data facilitated by AI. Therefore, this approach is not fit for purpose to effectively[1242] protect the fundamental right to data protection as it fails to specifically protect inherently sensitive data. In its case law, the CJEU has repeatedly stressed that EU data protection law aims to effectively protect[1243] the data subject's

[1235] Giovanni Stanghellini, René Rosfort, *Emotions and Personhood: Exploring Fragility – Making Sense of Vulnerability* (OUP 2013) 149, Andrew McStay, 'Emotion AI, soft biometrics and the surveillance of emotional life: An unusual consensus on privacy' (2020) Vol 7 Iss 7 Big Data & Society 1, 4; William S Brown, 'Technology, Workplace Privacy and Personhood' (1996) Vol 15 Journal of Business Ethics 1237, 1243.
[1236] Leen Vandercammen et al, 'On the Role of Specific Emotions in Autonomous and Controlled Behaviour' (2014) Vol 28 Iss 5 European Journal of Personality 413, 445.
[1237] Steffen Steinert, Orsolya Friedrich, 'Wired Emotions: Ethical Issues of Affective Brain–Computer Interfaces' (2020) Vol 26 Science and Engineering Ethics 351, 352.
[1238] J S Blumenthal-Barby, 'A Framework for Assessing the Moral Status of Manipulation' in Christian Coons, Michael Weber (eds) *Manipulation* (Oxford University Press 2014) 123, 127.
[1239] Jennifer S Lerner, Deborah A Small, George Loewenstein, 'Heart Strings and Purse Strings' (2004) Vol 15 No 5 American Psychology Society 337-340.
[1240] Joseph Raz, *The Morality of Freedom* (Oxford University Press 1986) 369.
[1241] Daniel Susser, Beate Roessler, Helen Nissenbaum 'Technology, autonomy, and manipulation' (2019) Vol 8 Iss 2 Internet Policy Review 1, 8.
[1242] Recital 11.
[1243] Case C-645/19 [2021] *Facebook Ireland* ECR I-483 para 45; Case C-319/20, *Meta Platforms Ireland Limited* [2022] ECR I-322 para 73.

personal data against risk of misuse.[1244] It can neither ensure a high level of protection[1245] nor a strong and coherent data protection framework[1246] when considering the gap of protection it creates.

The emotion data problem (Type 3)

The AI discipline AC facilitates the processing of emotion data, information that is highly sensitive and intimate. Despite the sensitive nature, it is not protected as special data under the GDPR because the approach to enumerate special categories of personal data exhaustively cannot keep up with developments in AI. Consequently, this principle creates a significant gap of protection and is therefore not fit for purpose to protect the fundamental right to data protection.

Location data

Location data reveals where individuals live, work and shop, which bars and restaurants they visit, which political events they attend and which medical services they need,[1247] providing a very intimate insight into the private life of individuals.[1248] Therefore, location data are of sensitive nature.[1249] It is considered to be a valuable asset with a variety of commercial and public uses.[1250] As opposed to emotion data, mental data and neurodata, location data are not a 'new' type of personal data. Rather, when processed by means of AI, location data become personal data of a sensitive nature. Based on historical patterns, modelling applications that analyse user location data can predict where a user will be located at a particular time of the day. The prediction of a user's location is often based on ML,[1251] using techniques such as regression, clustering and ANNs as described in Section 2.2.1. Research has shown that the current location of a smartphone user can be predicted with an average of 90% accuracy by exploiting ML techniques to develop a hybrid AI system for location prediction with smartphone logs.[1252] ML and probabilistic reasoning techniques can infer daily activities of an individual from location data.[1253] Collecting, storing and analysing location data can have significant privacy implications and enables to infer a detailed picture of a person's routine, lifestyle and social

[1244] Joined Cases C-293/12 and C-594/12 *Digital Rights Ireland* [2014] ECR-I 238 paras 54 and 66; Joined Cases C-203/15 and C-698/15 [2016] *Tele 2 Sverige* ECR-1 970 paras 109 and 122; Case C-362/14 *Schrems v Data Protection Commissioner* [2015] ECR I-650 para 91.
[1245] See Recitals 6 and 10 GDPR, as well as CJEU case law, such as Case C-534/20, *Leistritz AG* [2022] ECR I-594 para 26; Case C-645/19 [2021] *Facebook Ireland* ECR I-483 para 45; Case C-511/18, *La Quadrature du Net* [2020] ECR I-791 para 106; Case C-131/12, *Google Spain* [2014] ECR I-317 para 66; Case C-154/21 *Österreichische Post AG* [2023] ECR I-3 para 44.
[1246] Recital 7 GDPR.
[1247] Giovanni Butarelli, 'The urgent case for a new ePrivacy law' (2018) <https://edps.europa.eu/press-publications/press-news/blog/urgent-case-new-eprivacy-law_fr> accessed 8 February 2024.
[1248] Article 29 Working Party, 'Opinion 13/2011 on Geolocation services on smart mobile devices' (WP 185, 16 May 2011) at 18.
[1249] Article 29 Working Party, 'Opinion 13/2011 on Geolocation services on smart mobile devices' (WP 185, 16 May 2011) at 13; Article 29 Working Party, 'Opinion 01/2017 on the Proposed Regulation for the ePrivacy Regulation 2002/58/EC' (WP 247, 4 April 2017) at 30.
[1250] Andrej Savin, EU Telecommunications Law (Elgar 2018) 296.
[1251] Eran Toch et al, 'Analyzing large-scale human mobility data: a survey of machine learning methods and applications (2019) Vol 58 Knowledge and Information Systems 501, 512, 513.
[1252] Sung-Bae Cho, 'Exploiting machine learning techniques for location recognition and prediction with smartphone logs' (2016) Vol 176 Neurocomputing 98-106.
[1253] Lin Liao, 'Location-Based Activity Recognition' Dissertation University of Washington 2006.

network.[1254] Location-related information such as everyday habits, daily movements, and activities can help to establish a profile of the individuals concerned. From a privacy perspective, such profiles are no *less sensitive* than the actual content of electronic communications, according to the CJEU.[1255] Key locations such as the home or workplace of a mobile user can be inferred even from pseudonymous location data.[1256] By analysing widely available location metadata in public data streams like Twitter, such key locations can be pinpointed with a high level of accuracy, making it a trivial task to identify the individual concerned.[1257]

Despite its sensitive nature, location data are not listed in the definition of special data according to Article 9 (1) GDPR. Furthermore, the ePD does not provide protection against processing sensitive location data performed by information society providers. As outlined in Section 3.4.3.3, the processing of location data is specifically regulated by Article 9 (1) ePD and requires the consent of the user or subscriber or is allowed where location data are made anonymous when processed by electronic communications services (ECS). The latter covers access services, interpersonal communications services and services consisting wholly or mainly of the conveyance of signals[1258] and over-the-top (OTT) services such as VoIP[1259] solutions, messaging services and web-based email services which are functionally equivalent to traditional voice telephony and text message services.[1260] The strict regulation of Article 9 (1) ePD however does *not* apply where location data are processed by providers of information society services, even when such processing is performed via public electronic communication networks.[1261] Information society services are defined broadly and include any 'service normally provided for remuneration, at a distance, by electronic means and at the individual request of a recipient of services'. [1262] Whereas the installation of an app on a mobile device itself requires consent according to Article 5 (3) ePD,[1263] the processing of location data itself is not regulated by the ePD in case of information society services.

[1254] Eran Toch et al, 'Analyzing large-scale human mobility data: a survey of machine learning methods and applications (2019) Vol 58 Knowledge and Information Systems 501, 517.
[1255] Joined Cases C-203/15 and C-698/15 [2016] *Tele 2 Sverige* ECR-1 970 paras 89, 99.
[1256] Julien Freudiger, Reya Shokri, Jean-Pierre Hubaux, 'Evaluating the Privacy Risk of Location-Based Services' in Danezis Georg (ed) *Financial Cryptography and Data Security* (Springer 2012) 36.
[1257] Drakonakis Kostas et al, 'Please Forget Where I Was Last Summer: The Privacy Risks of Public Location (Meta) Data' (2019) 2 <https://arxiv.org/pdf/1901.00897.pdf> accessed 8 February 2024.
[1258] Article 2 (4) EECC.
[1259] VoIP solutions, for example, enable individuals to call via computer without the call being routed on to a number in the regular telephony numbering plan
[1260] Recital 15 EECC.
[1261] Article 29 Working Party, 'Opinion 13/2011 on Geolocation services on smart mobile devices' (WP 185, 16 May 2011) at 9.
[1262] Defined as 'any service normally provided for remuneration, at a distance, by electronic means and at the individual request of a recipient of services'. See Directive (EU) 2015/1535 laying down a procedure for the provision of information in the field of technical regulations and of rules on Information Society services ('Information Society Services Directive').
[1263] Article 29 Working Party, 'Opinion 13/2011 on Geolocation services on smart mobile devices' (WP 185, 16 May 2011) at 14.

Consequently, information society service providers must comply with the GDPR for the further processing of sensitive location data gained by means of mobile devices and can legitimise such processing by a variety of lawful grounds according to Article 6 GDPR.[1264] Given that consent is one of the main legislative tools for giving individuals control over the processing of personal data[1265] – if not the 'ultimate expression of control'[1266] – data subjects seem to have few ways to exercise control over the processing of their location data (apart from exercising their rights). Controllers and particularly information society service providers may rely on a variety of legal bases other than consent. They can legally argue that there is no need to ask permission from individuals to process their location data,[1267] information that is of sensitive nature.[1268] Given that location data are not considered special data under the GDPR, controllers may deploy ML approaches to infer daily activities, behavioural patterns and predict the location of individuals in a particular time period without the need to obtain consent from the individuals concerned. Notably, also the CJEU acknowledges the sensitive nature of profiles that may derived from location-related information.[1269]

The current legal framework does not effectively[1270] protect the fundamental rights to privacy and data protection, because sensitive location data are only regulated strictly under the ePD when it is processed by ECSs,[1271] excluding a broad range of information society services. This fails to achieve the ePD's goal of protecting users from risks regarding their personal data and privacy.[1272] It also fails to fulfil the GDPR's aim to respect the fundamental right to privacy[1273] considering that location data provide a very intimate insight into the private life of individuals[1274] as it reveals where they live, work and shop, which bars and restaurants they visit, which political events they attend and which medical services they need.[1275] Thus, the approach to exhaustively enumerate special categories of personal data creates a gap of protection with regard to the processing of sensitive location data facilitated by

[1264] Note however that regulatory guidance sees informed consent as the main applicable legal ground for the processing of location data Article 29 Working Party, 'Opinion 13/2011 on Geolocation services on smart mobile devices' (WP 185, 16 May 2011) at 13.
[1265] Giovanni Butarelli, 'The urgent case for a new ePrivacy law' (2018) <https://edps.europa.eu/press-publications/press-news/blog/urgent-case-new-eprivacy-law_fr> accessed 8 February 2024.
[1266] Case C-300/21 *UI* [2022] ECR I-756 Opinion AG Manuel Campos Sánchez-Bordona para 73.
[1267] Giovanni Butarelli, 'The urgent case for a new ePrivacy law' (2018) <https://edps.europa.eu/press-publications/press-news/blog/urgent-case-new-eprivacy-law_fr> accessed 8 February 2024.
[1268] Article 29 Working Party, 'Opinion 13/2011 on Geolocation services on smart mobile devices' (WP 185, 16 May 2011) at 13, 18; Article 29 Working Party, 'Opinion 01/2017 on the Proposed Regulation for the ePrivacy Regulation 2002/58/EC' (WP 247, 4 April 2017) at 30.
[1269] Joined Cases C-203/15 and C-698/15 [2016] *Tele 2 Sverige* ECR-1 970 paras 89, 99.
[1270] Recital 11 GDPR; Joined Cases C-293/12 and C-594/12 *Digital Rights Ireland* [2014] ECR-I 238 paras 54 and 66; Joined Cases C-203/15 and C-698/15 [2016] *Tele 2 Svergie* ECR-1 970 paras 109 and 122; Case C-362/14 *Schrems v Data Protection Commissioner* [2015] ECR I-650 para 91.
[1271] Article 29 Working Party, 'Opinion 13/2011 on Geolocation services on smart mobile devices' (WP 185, 16 May 2011) at 7.
[1272] Recital 6 ePD.
[1273] Recital 4 GDPR.
[1274] Article 29 Working Party, 'Opinion 13/2011 on Geolocation services on smart mobile devices' (WP 185, 16 May 2011) at 18.
[1275] Giovanni Butarelli, 'The urgent case for a new ePrivacy law' (2018) <https://edps.europa.eu/press-publications/press-news/blog/urgent-case-new-eprivacy-law_fr> accessed 8 February 2024, see also Joined Cases C-203/15 and C-698/15 [2016] *Tele 2 Sverige* ECR-1 970 paras 89, 99.

AI. It also fails to empower data subjects for exercising control regarding the processing of their personal data.[1276] Consent is considered to be one of the main legislative tools for giving individuals control over the processing of their personal data,[1277] if not the 'ultimate expression of control'.[1278] Because information society services do not fall under the scope of the ePD, controllers may rely on a variety of legal bases other than consent for processing location data. They can argue that there is no need to ask permission from individuals to process sensitive location data.[1279] Therefore, the approach to exhaustively enumerate special data and the restricted scope of the ePD are not fit for purpose to protect the fundamental rights to data protection and privacy when considering the gap of protection they create.

Note that locational privacy, i.e. the privacy of information about someone's physical (geographic) location[1280] is protected as such under the fundamental right to privacy. The processing of location data can be regarded as an interference with an individual's fundamental right to privacy.[1281]

The location data problem (Type 3)

ML can infer daily activities of an individual from location data and the processing of such data may have significant privacy implications, allowing to draw a detailed picture about a person's routine, lifestyle and social network. Information society service providers are not obliged to obtain consent for the processing of location data according to Article 9 (1) ePD. Likewise, sensitive location data is not protected as such according to Article 9 (1) GDPR. Consequently, these provisions create significant gaps of protection and are therefore not fit for purpose to protect the fundamental rights to privacy and data protection.

Neurodata

AI is a powerful driver for neurotechnologies which interface with the brain and that sense information about or produced by the brain function and/or offer input or 'write' information into the brain to modulate function.[1282] Advancements in human neuroscience and neurotechnology facilitate unprecedented means for accessing, collecting, sharing and otherwise processing neurodata. Neurodata is any information with respect to brain functions, neural activity, brain signals and any other information relating to the human brain ('neurodata').[1283] This broad definition includes brain signals

[1276] Recital 7 GDPR.
[1277] Giovanni Butarelli, 'The urgent case for a new ePrivacy law' (2018) <https://edps.europa.eu/press-publications/press-news/blog/urgent-case-new-eprivacy-law_fr> accessed 8 February 2024.
[1278] Case C-300/21 *UI* [2022] ECR I-756 Opinion AG Manuel Campos Sánchez-Bordona para 73.
[1279] Article 29 Working Party, 'Opinion 01/2017 on the Proposed Regulation for the ePrivacy Regulation 2002/58/EC' (WP 247, 4 April 2017) at 30.
[1280] Bert-Jaap Koops et al 'A Typology of Privacy' (2017) Vol. 38 Iss. 2 University of Pennsylvania Journal of International Law 438, 500.
[1281] *Uzun v Germany United* App no 35623/05 (ECtHR 2 December 2010) paras 51-52
[1282] Karen S Rommelfanger et al, 'Mind the Gap: Lessons Learned from Neurorights' AAAS Center for Science Diplomacy (2022) < https://www.sciencediplomacy.org/article/2022/mind-gap-lessons-learned-neurorights > accessed 8 February 2024.
[1283] Marcello Ienca, Roberto Andorno, 'Towards New Human Rights in the Age of Neuroscience and Neurotechnology' (2017) Vol 13 Iss 1 Life Science, Society and Policy 1.

measured by means of electroencephalography (EEG) and neuroimaging. The latter refers to the various techniques used to create images of the structures and/or functioning of the nervous system.[1284] EEG is a non-invasive method consisting of detection and registration of electrical activity occurring in the brain. It relies on electrodes attached to the scalp that register changes of electric potential on the skin surface caused by the activity of cerebral neurons. After their amplification, they form a record, namely, an encephalogram.[1285]

Brain-computer interfaces (BCIs), also known as mind-machine interfaces, are designed to translate brain signals into computer commands. They facilitate communication between the human brain and devices.[1286] BCIs enable their users to send commands to computers by means of brain signals alone which are usually measured by means of electroencephalography (EEG).[1287] In the beginning, BCIs have largely focussed on medical assistive applications to improve the quality of life for patients, for example on applications that enable advanced communications with paralysed patients.[1288] Recently, BCIs have been developed for non-clinical applications, such as for the purpose of entertainment, mental state monitoring, virtual reality and in Internet of Things (IoT) services,[1289] device control or real-time neuromonitoring, neurosensory-based vehicle operator systems, wearables for mental well-being and virtual reality systems.[1290] Kernel intends to 'hack the human brain'[1291] and Facebook wants to develop means of controlling devices directly with neurodata.[1292]

All these BCI applications process neurodata. Data acquisition methods facilitating the collection of neurodata used for BCI applications vary and include EEG, magnetoencephalography (MEG) and functional magnetic resonance imaging (fMRI).[1293] Non-invasive BCIs, which currently are most widely used in BCI research, place sensors on the scalp to acquire EEG signals.[1294] The development

[1284] Damian Eke et al, 'Pseudonymisation of neuroimages and data protection: Increasing access to data while retaining scientific utility' (2021) Vol 1 Iss 4 Neuroimage 1-12.
[1285] Szczepan Paszkiel, *Analysis and Classification of EEG Signals for Brain–Computer Interfaces* (Springer Nature 2020) 3.
[1286] Hongchang Shan, 'Towards High Performance and Efficient Brain Computer Interface Character Speller: Convolutional Neural Network based Methods' Dissertation Universiteit Leiden 2020, 1.
[1287] Camille Jeunet, Bernard N'Kaoua, Fabien Lotte, 'Chapter 1 - Advances in user-training for mental-imagery-based BCI control: Psychological and cognitive factors and their neural correlates' in Damien Coyle (ed) *Progress in Brain-Computer Interfaces: Lab Experiments to Real-World Applications* (Elsevier 2016) 4.
[1288] Brent J. Lance et al, 'Brain-Computer Interface Technologies in the Coming Decades' (2012) Vol 100 Proceedings of the IEE 1585.
[1289] Hongchang Shan, 'Towards High Performance and Efficient Brain Computer Interface Character Speller: Convolutional Neural Network based Methods' Dissertation Universiteit Leiden 2020, 1.
[1290] Marcello Ienca, Roberto Andorno, 'Towards New Human Rights in the Age of Neuroscience and Neurotech- nology' (2017) Vol 13 Iss 1 Life Science, Society and Policy 1, 4.
[1291] Nick Statt, 'Kernel is Trying to Hack the Human Brain—But Neuroscience has a Long Way to Go' *The Verge* (New York 22 February 2017) <https://www.theverge.com/2017/2/22/14631122/kernel-neuroscience-bryanjohnson-human-intelligence-ai-startup> accessed 8 February 2024.
[1292] John Constine, 'Facebook is building brain-computer interfaces for typing and skin-hearing' TechCrunch (San Francisco 19 April 2017) < https://techcrunch.com/2017/04/19/facebook-brain-interface/ > accessed 8 February 2024.
[1293] Szczepan Paszkiel, *Analysis and Classification of EEG Signals for Brain–Computer Interfaces* (Springer Nature 2020) 1.
[1294] Hongchang Shan, 'Towards High Performance and Efficient Brain Computer Interface Character Speller: Convolutional Neural Network based Methods' Dissertation Universiteit Leiden 2020, 2.

of consumer-directed wearable devices to record brain activity based on EEGs will likely lead to the analysis of neurodata at a large scale.[1295] Before neurodata can be useful for specific purposes, it must be 'de-coded' meaning that features must be extracted and classified according to known particularities of a specific brain activity.[1296] AI proves to be very helpful for such de-coding.[1297] BCI applications use different ML techniques for the classification of EEG signals.[1298] For example, researchers have used a convolutional neural network[1299] (CNN) to decode movement-related information from EEG data.[1300] ML and particularly DL approaches modelled on ANNs will be useful for this and allow fine-grained decoding of neurodata.[1301] Classification techniques[1302] used for supervised ML[1303] introduced in Section 2.2.1.1 as well as feature extraction techniques from the AI discipline CV[1304] can adaptively decode neurodata.[1305] Because most existing EEG decoding methods separate feature extraction from classification, it has been suggested to develop deep convolutional networks from DL to decode neurodata[1306] which combine feature extraction and classification. In addition, neurodata may be used for the purpose of artificially generating speech by means of NLP. Because neurodata associated with speech can be recorded from specific articulatory motor areas in the brain, unvoiced speech can be reconstructed and realised synthetically via a speaker.[1307]

Developments of ML, CV, NLP and DL applied to BCI open the possibility to analyse neurodata. It is very likely that processing of neurodata constitutes processing of personal data, in particular due to

[1295] Philipp Kellermayr, 'Big Neurodata: On the Responsible Use of Neurodata from Clinical and Consumer-Directed Neurotechnological Devices' (2018) Vol 14 Neuroethics 83, 84 <https://link.springer.com/article/10.1007/s12152-018-9371-x> accessed 8 February 2024.
[1296] Stephen Rainey et al, 'Is the European Data Protection Regulation sufficient to deal with emerging data concerns relating to neurotechnology?' (2020) Journal of Law and the Biosciences 2 < https://academic.oup.com/jlb/advance-article/doi/10.1093/jlb/lsaa051/5864051?searchresult=1> accessed 8 February 2024.
[1297] Karen S Rommelfanger et al, 'Mind the Gap: Lessons Learned from Neurorights' AAAS Center for Sience Diplomacy (2022) < https://www.sciencediplomacy.org/article/2022/mind-gap-lessons-learned-neurorights > accessed 8 February 2024.
[1298] Mamunir Rashid et al, 'The classification of EEG Signal Using Different Machine Learning Techniques for BCI Application' in J.-H. Kim et al (Eds) *Robot Intelligence Technology and Applications* (Springer 2018) 207-221.
[1299] Type of network architecture in DL.
[1300] Philipp Kellermayr, 'Big Neurodata: On the Responsible Use of Neurodata from Clinical and Consumer-Directed Neurotechnological Devices' (2018) Vol 14 Neuroethics 83, 86 <https://link.springer.com/article/10.1007/s12152-018-9371-x> accessed 8 February 2024.
[1301] Stephen Rainey et al, 'Is the European Data Protection Regulation sufficient to deal with emerging data concerns relating to neurotechnology?' (2020) Journal of Law and the Biosciences 2, 3 < https://academic.oup.com/jlb/advance-article/doi/10.1093/jlb/lsaa051/5864051?searchresult=1> accessed 8 February 2024.
[1302] Camille Jeunet, Bernard N'Kaoua, Fabien Lotte, 'Chapter 1 - Advances in user-training for mental-imagery-based BCI control: Psychological and cognitive factors and their neural correlates' in Damien Coyle (ed) *Progress in Brain-Computer Interfaces: Lab Experiments to Real-World Applications* (Elsevier 2016) 4, 5.
[1303] Szczepan Paszkiel, *Analysis and Classification of EEG Signals for Brain–Computer Interfaces* (Springer Nature 2020) 42.
[1304] Mark Nixon, Alberto Aguado, *Feature Extraction & Image Processing for Computer Vision* (3rd edn Elsevier 2012).
[1305] Stephen Rainey et al, 'Is the European Data Protection Regulation sufficient to deal with emerging data concerns relating to neurotechnology?' (2020) Journal of Law and the Biosciences 3 < https://academic.oup.com/jlb/advance-article/doi/10.1093/jlb/lsaa051/5864051?searchresult=1> accessed 8 February 2024.
[1306] Implementing a joint space–time–frequency feature extraction scheme for EEG decoding see Dongye Zhao et al, 'Learning joint space–time–frequency features for EEG decoding on small labeled data' (2019) Vol 114 Neural Networks 67.
[1307] Stephen Rainey et al, 'Is the European Data Protection Regulation sufficient to deal with emerging data concerns relating to neurotechnology?' (2020) Journal of Law and the Biosciences 11 < https://academic.oup.com/jlb/advance-article/doi/10.1093/jlb/lsaa051/5864051?searchresult=1> accessed 8 February 2024.

the personal nature of the brain itself: brain characteristics are largely determined by genetic factors that are often unique to individuals.[1308] Additionally, certain forms of neurodata remain unique to one specific individual regardless of attempts to segregate the link between neurodata and this specific individual.[1309] Neurodata is said to provide unique insights into people[1310] and their behaviour.[1311] Neurodata are a particularly sensitive class of data due to their direct link with mental processes[1312] and the strong link to the individual's personhood.[1313] Despite this, it is clear that neurodata as such is not considered a special category of personal data according to the GDPR.[1314] However, in some cases and depending on the context, the processing of neurodata could reveal data that is protected as a special category such as genetic data,[1315] racial and ethnic origin,[1316] health data[1317] or biometric data.[1318] Apart from these very specific cases, highly sensitive neurodata do not fall under the definition of special categories of personal data. Neurodata relates to processes of the human mind, which represents a uniquely sensitive and intimate space in the individual's private sphere. Neurodata is not only sensitive because of what can be concluded from it in terms of mental states, but also in view of inferred data, such as insights into a data subject's personality, cognitive capacity and future behaviour.[1319] It may also reveal sensitive neuronal states that are associated with below average functioning something that is not health data as such. When revealed, such data may result in discrimination. For example, someone may be labelled or classified as 'stupid' simply due to the detection of uncommon neuronal states.[1320] Because of its sensitive nature[1321] and the sensitive information that can be inferred

[1308] Therefore, neurodata could be used for so called 'brain-fingerprinting'. See Kuldeep Kumar et al, 'Multi-modal brain fingerprinting: A manifold approximation based framework' (2018) Vol 183 Neuro-Image 212-226.
[1309] Dara Hallinan et al, 'Neurodata and Neuroprivacy: Data Protection Outdated?' (2014) Vol 12 Iss 1 Surveillance and Society 65 <https://ojs.library.queensu.ca/index.php/surveillance-and-society/article/view/neurodata/neurodata4> accessed 8 February 2024; See also Damian Eke et al, 'Pseudonymisation of neuroimages and data protection: Increasing access to data while retaining scientific utility' (2021) Vol 1 Iss 4 Neuroimage 1-12.
[1310] Neurodata are of highly personalised nature and allows for identification ('brain fingerprinting').
[1311] Brent J. Lance et al, 'Brain-Computer Interface Technologies in the Coming Decades' (2012) Vol 100 Proceedings of the IEE 1587.
[1312] Marcello Ienca, Roberto Andorno, 'Towards New Human Rights in the Age of Neuroscience and Neurotechnology' (2017) Vol 13 Iss 1 Life Science, Society and Policy 1, 14; Marcello Ienca, Karolina Ignatiadis, 'Artificial Intelligence in Clinical Neuroscience: Methodological and Ethical Challenges' (2020) Vol 11 Iss 2 AJOB Neuroscience 77-87; Rafael Yuste et al, 'Four ethical priorities for neurotechnolgies and AI' (2017) Vol 551 Nature 159-163.
[1313] Marcello Ienca, Roberto Andorno, 'Towards New Human Rights in the Age of Neuroscience and Neurotech- nology' (2017) Vol 13 Iss 1 Life Science, Society and Policy 1, 14.
[1314] Because the lawmaker arguably did not anticipate the use of this novel type of data, since it is not mentioned in Article 9 (1) GDPR or in corresponding recitals.
[1315] When revealing genetic features such as biomarkers.
[1316] Morphological differences between various sections of the brain in different individuals allows the identification of different ethnical groups; see Wei Liang Chee et al, 'Brain Structure in Young and Old East Asians and Westerners: Comparison of Structural Volume and Cortical Thickness' (2011) Vol 23 Iss 5 Journal of Cognitive Neuroscience <www.ncbi.nlm.nih.gov/pmc/articles/PMC3361742/> accessed 8 February 2024.
[1317] When neurological problems or brain diseases are detected.
[1318] When neurodata are used to identify an individual.
[1319] Karen S Rommelfanger et al, 'Mind the Gap: Lessons Learned from Neurorights' AAAS Center for Sience Diplomacy (2022) < https://www.sciencediplomacy.org/article/2022/mind-gap-lessons-learned-neurorights > accessed 8 February 2024; Ryan H Purcell, Karen S Rommelfanger, 'Internet-Based Brain Training Games, Citizen Scientists, and Big Data: Ethical Issues in Unprecedented Virtual Territories' (2015) Vol 86 Iss 2 Neuron 356, 357.
[1320] Jan-Hendrik Heinrichs, 'The Sensitivity of Neuroimaging Data' (2012) Vol 5 Iss 2 Neuroethics 185, 193.
[1321] Marcello Ienca, Roberto Andorno, 'Towards New Human Rights in the Age of Neuroscience and Neurotech- nology' (2017) Vol 13 Iss 1 Life Science, Society and Policy 1, 14; Marcello Ienca, Karolina Ignatiadis, 'Artificial Intelligence in Clinical Neuroscience: Methodological and Ethical Challenges' (2020) Vol 11 Iss 2 AJOB Neuroscience 77-87; Rafael Yuste et al, 'Four ethical priorities for neurotechnolgies and AI' (2017) Vol 551 Nature 159-163.

from it, neurodata should receive specific protection under the GDPR. The high level of protection of neurodata (as special data) should include neural activity occurring in the human brain which generates the neurodata.[1322] This means that neurodata would already be protected before it is 'de-coded', revealing for instance mental data (see mental data problem later in this section). This is needed to protect sensitive information that might be inferred from it. Inferences derived from neurodata can be used to influence an individual's commercial, social and political behaviour. For example, information derived from neurodata may be used to tailor content or experiences in a way that is more addictive for individuals concerned based on psychology.[1323]

Article 9 (1) GDPR does not list neurodata. Because neurodata itself does not receive specific protection under the GDPR, the approach to exhaustively enumerate special data is not fit for purpose to effectively[1324] protect the fundamental right to data protection. In its case law, the CJEU has repeatedly stressed that EU data protection law aims to effectively protect[1325] the data subject's personal data against risk of misuse.[1326] Such risk of misuse is high when considering that inferences drawn from neurodata may be used to influence an individual's commercial, social and political behaviour. Due to its direct link with mental processes[1327] and an individual's personhood,[1328] neurodata is highly sensitive and provides unique insights into an individual's behaviour. [1329] Therefore, the processing of neurodata can pose significant risks to the fundamental rights and freedoms of individuals. By virtue of its content, neurodata carries the risk of infringing the individual's fundamental right to privacy (see also the mental data problem discussed later in this section and Section 5.4) that the GDPR envisages to protect.[1330] Article 9 (1) GDPR can neither ensure a high level of protection[1331] nor a strong and coherent data protection framework[1332] when considering the gap of protection it creates.

[1322] Marcello Ienca, Roberto Andorno, 'Towards New Human Rights in the Age of Neuroscience and Neurotech- nology' (2017) Vol 13 Iss 1 Life Science, Society and Policy 14.
[1323] Karen S Rommelfanger et al, 'Mind the Gap: Lessons Learned from Neurorights' AAAS Center for Sience Diplomacy (2022) < https://www.sciencediplomacy.org/article/2022/mind-gap-lessons-learned-neurorights > accessed 8 February 2024.
[1324] Recital 11 GDPR.
[1325] Case C-645/19 [2021] *Facebook Ireland* ECR I-483 para 45; Case C-319/20, *Meta Platforms Ireland Limited* [2022] ECR I-322 para 73.
[1326] Joined Cases C-293/12 and C-594/12 *Digital Rights Ireland* [2014] ECR-I 238 paras 54 and 66; Joined Cases C-203/15 and C-698/15 [2016] *Tele 2 Sverige* ECR-1 970 paras 109 and 122; Case C-362/14 *Schrems v Data Protection Commissioner* [2015] ECR I-650 para 91.
[1327] Marcello Ienca, Roberto Andorno, 'Towards New Human Rights in the Age of Neuroscience and Neurotech- nology' (2017) Vol 13 Iss 1 Life Science, Society and Policy 1, 14; Marcello Ienca, Karolina Ignatiadis, 'Artificial Intelligence in Clinical Neuroscience: Methodological and Ethical Challenges' (2020) Vol 11 Iss 2 AJOB Neuroscience 77-87; Rafael Yuste et al, 'Four ethical priorities for neurotechnolgies and AI' (2017) Vol 551 Nature 159-163.
[1328] Marcello Ienca, Roberto Andorno, 'Towards New Human Rights in the Age of Neuroscience and Neurotech- nology' (2017) Vol 13 Iss 1 Life Science, Society and Policy 1, 14.
[1329] Brent J. Lance et al, 'Brain-Computer Interface Technologies in the Coming Decades' (2012) Vol 100 Proceedings of the IEE 1587.
[1330] Recital 4 GDPR.
[1331] See Recitals 6 and 10 GDPR, as well as CJEU case law, such as Case C-534/20, *Leistritz AG* [2022] ECR I-594 para 26; Case C-645/19 [2021] *Facebook Ireland* ECR I-483 para 45; Case C-511/18, *La Quadrature du Net* [2020] ECR I-791 para 106; Case C-131/12, *Google Spain* [2014] ECR I-317 para 66; Case C-154/21 *Österreichische Post AG* [2023] ECR I-3 para 44; Case C-132/21 *Nemzeti* [2023] ECR I-2 para 42.
[1332] Recital 7 GDPR.

The neurodata problem (Type 3)

ML, CV, NLP and DL facilitate the processing of neurodata. Neurodata provide unique insights into people and are particularly sensitive due to their direct link with mental processes and an individual's personhood. Despite this, neurodata is not protected as special data under the GDPR because the approach to enumerate special categories of personal data exhaustively cannot keep up with the developments in AI. Consequently, this approach creates a significant gap of protection and is therefore not fit for purpose to protect the fundamental right to data protection.

Mental data

Neurotechnologies powered by AI have an unprecedented ability to decode information about mental states or processes by analysing data concerning neural activity patterns and 'transcribe' mental states by modulating neural computation.[1333] When processed by AI systems, neurodata as described earlier in this section may reveal mental data, which is any information about mental states and processes of individuals ('mental data').[1334] Mental states and processes include information related to all conscious and non-conscious mental representations, events, propositional attitudes, including thoughts, beliefs, emotions, moods and underlying psychological mechanisms.[1335] Mental data constitutes information relating to the core of an individual's private sphere,[1336] including information such as thoughts, memories and intentions. The processing of neurodata by AI systems, in particular ML and DL, allows one to derive insights in an individual's mental domain[1337] and particularly insights in 'real-time' mental processes.[1338] ML and DL approaches offer powerful capabilities (e.g. to detect patterns and make predictions) to infer a variety of highly sensitive information[1339] from neurodata, including dimensions of an individual's thoughts, intentions and sometimes even information that is not known to an individual herself or beyond her control.[1340] Through the processing of neurodata by means of AI, mental data becomes accessible. This indicates a partial overlap between the two categories of

[1333] Abel Wajnerman Paz, 'Is Your Neural Data Part of Your Mind? Exploring the Conceptual Basis of Mental Privacy' (2022) Vol 32 Minds and Machines 395, 396.
[1334] Ibid; see a similar definition by Marcello Ienca, Gianclaudio Malgieri, 'Mental data protection and the GDPR' (2022) Vol 9 Iss 1 Journal of Law and the Biosciences 1, 4.
[1335] Jan-Christoph Bublitz, 'The Nascent Right to Psychological Integrity and Mental Self-Determination' in Andreas von Arnauld, Kerstin von der Decken, Mart Susi (eds) *The Cambridge Handbook of New Human Rights* (Cambridge University Press 2020) 30; Marcello Ienca, Gianclaudio Malgieri, 'Mental data protection and the GDPR' (2022) Vol 9 Iss 1 Journal of Law and the Biosciences 1, 4.
[1336] Dara Hallinan et al, 'Neurodata and Neuroprivacy: Data Protection Outdated?' (2014) Vol 12 Iss 1 Surveillance and Society 65 <https://ojs.library.queensu.ca/index.php/surveillance-and-society/article/view/neurodata/neurodata4> accessed 8 February 2024.
[1337] Marcello Ienca, Gianclaudio Malgieri, 'Mental data protection and the GDPR' (2022) Vol 9 Iss 1 Journal of Law and the Biosciences 1, 3.
[1338] Dara Hallinan et al, 'Neurodata and Neuroprivacy: Data Protection Outdated?' (2014) Vol 12 Iss 1 Surveillance and Society 65 <https://ojs.library.queensu.ca/index.php/surveillance-and-society/article/view/neurodata/neurodata4> accessed 8 February 2024.
[1339] Marcello Ienca, Roberto Andorno, 'Towards New Human Rights in the Age of Neuroscience and Neurotech- nology' (2017) Vol 13 Iss 1 Life Science, Society and Policy 1, 24.
[1340] Dara Hallinan et al, 'Neurodata and Neuroprivacy: Data Protection Outdated?' (2014) Vol 12 Iss 1 Surveillance and Society 65 <https://ojs.library.queensu.ca/index.php/surveillance-and-society/article/view/neurodata/neurodata4> accessed 8 February 2024.

data. However, not all neurodata constitute mental data and vice versa. In addition, emotion data as discussed in the first part of this section can be seen as a subcategory of mental data. The relationship between neurodata, mental data and emotion data is illustrated in Figure 2.1.

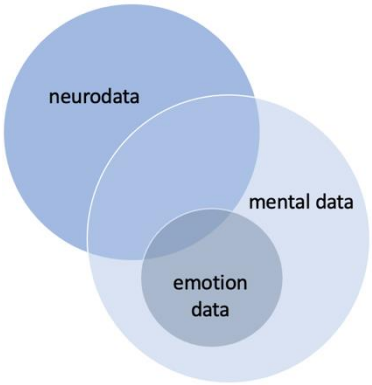

Figure 2.1
Overlaps between neurodata, mental data and emotion data.

Neurodata may be used to predict future behaviour, brain states and other aspects of an individual.[1341] When processed by AI systems, neurodata facilitates the inference of mental states of individuals. It should be noted that mental data may be generated from both neurodata and other data.[1342] Therefore, mental data and neurodata only partially overlap,[1343] as shown in Figure 2.1. For example, information regarding the emotional states of individuals might be inferred by approaches developed within the AI discipline AC as introduced in Sections 2.2.4.1 and 2.2.4.2, which do not comprise the processing of neurodata.[1344] This is illustrated in Figure 2.1 which shows that emotion data, seen as a subcategory of mental data, partially overlaps with neurodata. In addition, mental data may be inferred from digital footprints such as Facebook likes, tweets or credit card records when analysed by AI (for example, ML).[1345] Mental data form the core of an individual's private sphere[1346] and are therefore of a particularly sensitive nature. Risks associated with the processing of mental data are considerable because mental representations are the closest psychological substrate of fundamental ethical-legal notions[1347] such as personal autonomy. By using insights gained from the processing of mental data, BCI systems may influence the development of an individual's reasons by altering options to act independently, which has a negative impact to the self-determination of the individual concerned.[1348] Affective BCIs

[1341] Stephen Rainey et al, 'Is the European Data Protection Regulation sufficient to deal with emerging data concerns relating to neurotechnology?' (2020) Journal of Law and the Biosciences 3, 12, 14 < https://academic.oup.com/jlb/advance-article/doi/10.1093/jlb/lsaa051/5864051?searchresult=1> accessed 8 February 2024.
[1342] Marcello Ienca, Gianclaudio Malgieri, 'Mental data protection and the GDPR' (2022) Vol 9 Iss 1 Journal of Law and the Biosciences 1, 4.
[1343] Andrea Lavazza, 'Freedom of Thought and Mental Integrity: The Moral Requirements for Any Neural Prosthesis' (2018) Vol 12 Frontiers in Neuroscience 1-10.
[1344] The notions 'emotion data' and mental data partly overlap as the latter also covers emotion data. However, this section focusses on thoughts and other mental states.
[1345] Sandra C Matz et al, 'Privacy in the age of psychological targeting' (2020) Vol 31 Current Opinion in Psychology 116-221.
[1346] Dara Hallinan et al, 'Neurodata and Neuroprivacy: Data Protection Outdated?' (2014) Vol 12 Iss 1 Surveillance and Society 68 <https://ojs.library.queensu.ca/index.php/surveillance-and-society/article/view/neurodata/neurodata4> accessed 8 February 2024.
[1347] Marcello Ienca, Gianclaudio Malgieri, 'Mental data protection and the GDPR' (2022) Vol 9 Iss 1 Journal of Law and the Biosciences 1, 5.
[1348] Orsolya Friedrich et al, 'An Analysis of the Impact of Brain-Computer Interfaces on Autonomy' (2021) Vol 14 Iss 1 Neuroethics 17, 27.

use neurodata to extract features related to affective states such as emotions and may even stimulate and influence emotions.[1349] For example, an affective BCI system was developed to detect the current emotional state with the aim to modulate it accordingly by moving individuals from one emotional state to another.[1350] Such affective BCIs are problematic in terms of personal autonomy because they monitor, influence and directly stimulate emotional states of individuals.[1351] Influencing individuals may include manipulative forms of nudging. Nudges are 'interventions that steer people in particular directions but that also allows them to go their own way'.[1352] Nudging may be manipulative, for instance, if it is used to subvert an individual's decision-making powers.[1353]

Mental data may contain very sensitive information with respect to unexecuted behaviour such as unuttered thoughts and intended actions,[1354] information that previously was inaccessible to others. The developments in neurotechnology powered by AI can bypass the cognitive process of filtering and selectively sharing information that people typically perform to control the flow of information about them. Thus, information a person decided not to share may become available to others anyway.[1355] For example, thoughts and intentions can be disclosed by interpreting neurodata and decode it by ML and DL approaches. Researchers have achieved translating brain activity into text by means of ML and ANN approaches.[1356] Developments in neurotechnology, powered by ML and DL approaches, have unlocked the human brain to some extent.[1357] Neurodata in the form of connection patterns and activation of nerve cells are believed to constitute partial correlates of mental states an individual has at any given time.[1358] AI proves helpful to de-code such neurodata. A study has achieved to decode what the brain is neurally representing by means of CNN.[1359] However, current applications can often only decode a rather limited set of predetermined mental states from available neurodata.[1360] They are not yet able to decode mental information per se, but are sophisticated enough to establish statistically significant relations between certain patterns of neurodata and other data on the one hand,

[1349] Steffen Steinert, Orsolya Friedrich, 'Wired Emotions: Ethical Issues of Affective Brain–Computer Interfaces' (2020) Vol 26 Science and Engineering Ethics 351, 353.
[1350] Ian Daly et al, 'Affective brain–computer music interfacing' (2016) Vol 13 No 4 Journal of Neural Engineering
[1351] Steffen Steinert, Orsolya Friedrich, 'Wired Emotions: Ethical Issues of Affective Brain–Computer Interfaces' (2020) Vol 26 Science and Engineering Ethics 351, 355.
[1352] Cass R Sunstein, 'The Ethics of Nudging' (2015) Vol 32 Yale Journal of Regulation 413, 417.
[1353] Ibid, 446.
[1354] Marcello Ienca, Gianclaudio Malgieri, 'Mental data protection and the GDPR' (2022) Vol 9 Iss 1 Journal of Law and the Biosciences 1, 6.
[1355] Abel Wajnerman Paz, 'Is Mental Privacy a Component of Personal Identity?' (2021) Vol 15 Frontiers in Human Neuroscience 2.
[1356] Joseph G Makin, David A Moses, Edward F Chang, 'Machine translation of cordial activity to text with an encoder-decoder framework' (2020) Vol 23 Nature Neuroscience 575.
[1357] Marcello Ienca, Roberto Andorno, 'Towards new human rights in the age of neuroscience and neurotechnology. Life Sciences, Society and Policy' (2017) Vol 13 Life Sciences, Society and Policy 5 <https://lsspjournal.biomedcentral.com/articles/10.1186/s40504-017-0050-1> accessed 8 February 2024.
[1358] Andrea Lavazza, 'Freedom of Thought and Mental Integrity: The Moral Requirements for Any Neural Prosthesis' (2018) Vol 12 Frontiers in Neuroscience 1, 3.
[1359] Haiguang Wen et al, 'Neural Encoding and Decoding with Deep Learning for Dynamic Natural Vision' (2018) Vol 28 Iss 12 Cerebral Cortex 4136-4160.
[1360] Abel Wajnerman Paz, 'Is Your Neural Data Part of Your Mind? Exploring the Conceptual Basis of Mental Privacy' (2022) Vol 32 Minds and Machines 395, 397.

and the actual occurrence of mental states on the other hand. Information inferred from mental data and neurodata[1361] may have considerable (mental) privacy implications (see Section 5.4).

Mental data falls, as such, not under the definition of special data in the GDPR[1362] despite its highly intimate and sensitive nature. Because mental data does not receive specific protection under the GDPR, the approach to enumerate special data exhaustively is not fit for purpose to effectively[1363] protect the fundamental right to data protection. In its case law, the CJEU has repeatedly stressed that EU data protection law aims to effectively protect[1364] the data subject's personal data against risk of misuse.[1365] Such risk of misuse seems relatively high when considering that AI, sooner or later, be able to decode neurodata in a way that discloses an individual's mental states, their thoughts in particular. Thus, there is a clear conceptual and normative gap regarding the protection of mental data. It is difficult to assert that thoughts, and mental data more generally, are less sensitive than the special categories of personal data[1366] listed in the GDPR. Processing inherently sensitive mental data is prone to create significant risks to the fundamental rights and freedoms of individuals. Mental data carries the risk of infringing an individual's fundamental right to privacy (see also Section 5.4) that the GDPR also envisages to protect.[1367] Additionally, Article 9 (1) GDPR can neither ensure a high level of protection[1368] nor a strong and coherent data protection framework[1369] when considering the gap of protection it creates.

The mental data problem (Type 3)

ML and AC facilitate the processing of mental data, i.e. any data used to infer mental states of individuals including thoughts, beliefs and underlying mechanisms and processes. Mental data are inherently sensitive and form the core of an individual's private sphere. Despite this, mental data are not specifically protected under the GDPR because the approach to enumerate special categories of personal data exhaustively cannot keep up with the developments in AI. This principle creates a significant gap of protection and is not fit for purpose to protect the fundamental right to data protection.

[1361] Marcello Ienca, Gianclaudio Malgieri, 'Mental data protection and the GDPR' (2022) Vol 9 Iss 1 Journal of Law and the Biosciences 1, 6.
[1362] Stephen Rainey et al, 'Is the European Data Protection Regulation sufficient to deal with emerging data concerns relating to neurotechnology?' (2020) Journal of Law and the Biosciences 16 <https://academic.oup.com/jlb/advance-article/doi/10.1093/jlb/lsaa051/5864051?searchresult=1> accessed 8 February 2024.
[1363] Recital 11 GDPR.
[1364] Case C-645/19 [2021] *Facebook Ireland* ECR I-483 para 45; Case C-319/20, *Meta Platforms Ireland Limited* [2022] ECR I-322 para 73.
[1365] Joined Cases C-293/12 and C-594/12 *Digital Rights Ireland* [2014] ECR-I 238 paras 54 and 66; Joined Cases C-203/15 and C-698/15 [2016] *Tele 2 Sverige* ECR-1 970 paras 109 and 122; Case C-362/14 *Schrems v Data Protection Commissioner* [2015] ECR I-650 para 91.
[1366] Dara Hallinan et al, 'Neurodata and Neuroprivacy: Data Protection Outdated?' (2014) Vol 12 Iss 1 Surveillance and Society 67 <https://ojs.library.queensu.ca/index.php/surveillance-and-society/article/view/neurodata/neurodata4> accessed 8 February 2024.
[1367] Recital 4 GDPR.
[1368] See Recitals 6 and 10 GDPR, as well as CJEU case law, such as Case C-534/20, *Leistritz AG* [2022] ECR I-594 para 26; Case C-645/19 [2021] *Facebook Ireland* ECR I-483 para 45; Case C-511/18, *La Quadrature du Net* [2020] ECR I-791 para 106; Case C-131/12, *Google Spain* [2014] ECR I-317 para 66. Case C-534/20, Case C-154/21 *Österreichische Post AG* [2023] ECR I-3 para 44; Case C-132/21 *Nemzeti* [2023] ECR I-2 para 42.
[1369] Recital 7 GDPR.

A Type 3 legal problem may also be identified with regard to the protection of human minds per se, namely, the forum internum, which is denoted as a layer of the private sphere that describes the mental world of an individual.[1370] Whereas in human rights law the forum internum theoretically enjoys absolute and unconditional protection,[1371] it is doubtful whether this in fact applies in practice because the absolute, unimpugnable and fundamental nature of the forum internum seems to be undermined since individuals are not able to enforce their rights with regard to the forum internum.[1372] This will be discussed in the context of mental privacy (Section 5.4).

4.9 Confidentiality of communication

AI and people's interactions with it do not fit neatly into paradigms of communication theory that have long focussed on human–human communication.[1373] As I outline in this section, the same can be said about the legal protection concerning the confidentiality of human-machine communication. The GDPR regulates the processing of personal data, but not specifically the confidentiality of communication. This is regulated by the ePrivacy Directive ('ePD') as introduced in Section 3.4 and potentially the future ePrivacy Regulation.[1374] However, the obligation to ensure the confidentiality of communications and the general prohibition of listening, tapping, storing or other kinds of surveillance of communications and traffic data according to Article 5 (1) ePD *solely* applies to providers of publicly available electronic communication services (ECS) and providers of public electronic communication networks[1375] in the EU. Companies that provide virtual assistant services are not subject to Article 5 (1) ePD because they do not qualify as an ECS. As outlined in Section 3.4.1, an ECS covers Internet access services, interpersonal communications services and services consisting wholly or mainly in the conveyance of signals.[1376]

Clearly, virtual assistant services do not constitute Internet access services. In addition, they are not interpersonal communication services,[1377] because these services do not relate to communication

[1370] Dara Hallinan et al, 'Neurodata and Neuroprivacy: Data Protection Outdated?' (2014) Vol 12 Iss 1 Surveillance and Society 68 <https://ojs.library.queensu.ca/index.php/surveillance-and-society/article/view/neurodata/neurodata4> accessed 8 February 2024.
[1371] Article 9 ECHR.
[1372] Paul M Taylor, *Freedom of Religion UN and European Human Rights Law and Practice* (2005 Cambridge University Press) 202.
[1373] Andrea L Guzman, Seth C Lewis, 'Artificial intelligence and communication: A Human-Machine Communication agenda' (2020) Vol 22 Iss 1 New Media & Society 70-86.
[1374] Proposal for a Regulation of the European Parliament and of the Council concerning the respect for private life and the protection of personal data in electronic communications and repealing Directive 2002/58/EC (Regulation on Privacy and Electronic Communications) Recital 2 <https://data.consilium.europa.eu/doc/document/ST-12633-2019-INIT/en/pdf > accessed 8 February 2024.
[1375] Defined as 'electronic communications network used wholly or mainly for the provision of publicly available electronic communications services which support the transfer of information between network termination points'; see Article 2 (8) EECC.
[1376] Article 2 (4) EECC.
[1377] As defined in Article 2 (5) EECC: 'service normally provided for remuneration that enables *direct interpersonal and interactive exchange of information* via *electronic communications networks* between a *finite number of persons*' emphasis added.

between natural persons.[1378] Rather, they relate to communications between natural persons and a *machine*. Recital 17 EECC clarifies what interpersonal communication means: communication between *natural persons*. Communications involving legal persons fall within the definition only to a limited extent, for instance, if natural persons act on behalf of those legal persons.[1379] Therefore, human-machine communications fall outside the scope of interpersonal communication services defined in Article 2 (5) EECC. In addition, virtual assistant services do also not qualify as machine-to-machine services under the EECC. Recital 249 EECC says such services involve 'an automated transfer of data and information between devices or software-based applications with limited or no human interaction.' Virtual assistant services involve more than only limited human interaction.

A service provider is *responsible* vis-à-vis the end-users for *transmission* of the *signal* which ensures that users are supplied with the service to which they have subscribed.[1380] Clearly, providers of virtual assistant services are not responsible for the transmission of the signal. Rather, the Internet Access Providers (IAPs) and the *operators* of the *various networks* of which the open web is constituted are responsible for this.[1381]

Services facilitating human-machine communications do not qualify as an ECS which is problematic with regard to confidentiality. As will be described in Section 4.9.3, this applies particularly to the confidentiality of human-machine communications enabled by the AI disciplines NLP and AC when embedded in virtual assistants and smart devices connected to the Internet of Things ('IoT'). The IoT is the cyber-physical ecosystem of interconnected physical and potentially virtual sensors and actuators.[1382] It consists of devices such as smartphones, wearables and even toothbrushes which are connected together.[1383] The growing use of virtual assistants and smart home devices causes serious concerns about the confidentiality of communication and how related data are processed and controlled.[1384] For example, Amazon's virtual assistant Alexa is bound to be embedded in toilets, e-bikes, beds, cars and other everyday objects.[1385]

To be clear, providers of human-machine communication services need to adhere to the GDPR when processing personal data. Whereas both the GDPR and the ePD aim to protect fundamental rights and

[1378] Article 2 (5) EECC and Recital 17.
[1379] It seems unclear what the phrase 'or are at least involved on one side of the communication' contained in Recital 15 EECC precisely means.
[1380] Case C-475/12, *UPC* [2014] ECR I-285 para 43.
[1381] Case C-193/18, *Google LLC* [2019] ECR I-498 para 36.
[1382] European Union Agency for Network and Information Security, 'Good Practices for Security of Internet of Things' (2018) 45 <https://www.enisa.europa.eu/publications/good-practices-for-security-of-iot/@@download/fullReport> accessed 8 February 2024.
[1383] Matt Burgess, 'What is the Internet of Things? WIRED explains' *Wired* (New York, 16 February 2018) <https://www.wired.co.uk/article/internet-of-things-what-is-explained-iot> accessed 8 February 2024.
[1384] Tine Munk, 'Does Online Privacy Exist in the GDPR Era? The Google Voice Assistant Case' in Tatiana-Eleni Synodiou et al (eds) *EU Internet Law in the Digital Single Market* (Springer Nature 2021) 489.
[1385] Amrita Khalid, 'Alexa was everywhere at CES' *Quartz* (New York, 10 January 2020) <https://qz.com/1783414/amazons-alexa-was-everywhere-at-ces-2020/> accessed 8 February 2024.

freedoms,[1386] the GDPR sets general rules for processing personal data, and the ePD regulates the fundamental right to privacy *and* data protection in the electronic communications sector.[1387] Thus, merely because providers of human-machine communication services fall outside the scope of the ePD does not lead to a complete lacuna in legal protection. However, the provisions of the GDPR are less strict than Article 5 (1) ePD, which requires consent for the surveillance of interpersonal communications. Arguably, and as outlined in Section 4.9.3, human-machine communications deserve the same level of confidentiality as interpersonal communications.

4.9.1 Legal problems: Type 1

As will be outlined in Section 4.9.3, problems arise with respect to the confidentiality of human-machine and interpersonal communication in human-machine communication services, such as virtual assistant services, which do not qualify as ECS. Such services are therefore excluded from the scope of Article 5 (1) ePD, which prohibits surveillance of communications and related traffic data without consent of the user. Provisions that are not applicable to the company processing data cannot be violated. Therefore, no specific Type 1 legal problems arise.

4.9.2 Legal problems: Type 2

As outlined in Section 4.9.1, provisions enshrined in the current legal framework that are not applicable to providers of human-machine communication services, cannot be violated. Consequently, no specific Type 2 legal problems arise because provisions that are not applicable to a certain processing cannot be violated, and thus they also do not need to be enforced.

4.9.3 Legal problems: Type 3

Because of the restricted material scope of the ePD, the prohibition of listening, tapping, storage or other kinds of interception or surveillance of communications without consent of the individual concerned does not apply to human-machine and interpersonal communications occurring in the context of virtual assistants and smart home technologies powered by NLP and ML. Omission to subject such services to the material scope of the ePD creates a loophole for the providers of the services. A loophole exists where a failure to include something in the law allows someone to do something generally considered illegal.[1388] This occurs here due to the omission of not including virtual assistant and smart home services in the scope of the ePD. Due to this omission, providers of such services are not subject

[1386] In the case of the GDPR, the fundamental right to the protection of personal data, and in the case of the ePrivacy Directive, both the fundamental right to privacy (Recital 12) and data protection (Recital 2). Note that the Directive which amended the ePrivacy Directive also refers to the fundamental right to privacy and confidentiality (Recital 51) and the fundamental right to the protection of personal data (Recital 56).
[1387] Christina Etteldorf, 'EDPB on the Interplay between the ePrivacy Directive and the GDPR' (2019) Iss 5 No 2 European Data Protection Law Review 224, 226.
[1388] See <https://dictionary.cambridge.org/dictionary/english/loophole> accessed 8 February 2024.

to the confidentiality obligation enshrined in the ePD. They may intercept human-machine and interpersonal communications without needing to seek consent for intercepting such communications.[1389] This is particularly problematic when considering the extensive use of virtual assistants, smart home applications and similar services. Today, people routinely communicate with virtual assistants such as Amazon Alexa, Siri or Google Assistant.

In essence, virtual assistants are *software applications* equipped with the capabilities to interpret human speech as a question or instruction, perform tasks and respond using synthesised voices.[1390] Virtual assistants are made of several components designed to resolve specific challenges, for example, understanding and producing speech.[1391] They employ sophisticated NLP capabilities enabling users to interact with them conversationally. Put simply, virtual assistants work as follows. The virtual assistant permanently analyses every sound in its environment to recognise its 'wake word', which activates the recording of the user's request.[1392] A request is sent to the virtual assistant's service platform (thus *not* kept on the device) where speech is converted into text by means of speech recognition powered by NLP which translates the text into machine-readable instructions.[1393] Because virtual assistants permanently listen to detect the wake word which activates recording, virtual assistants are referred to as 'always-on' microphone-enabled devices.[1394] Accidental recordings are common in virtual assistant services and occur where virtual assistants activate, transmit and/or record audio from their environment when the wake word is *not* spoken.[1395] Such recordings are caused by accidental triggers, namely, sounds that wrongfully trigger virtual assistants, and occur within the whole range of virtual assistants available on the market, including Amazon Alexa, Google Assistant and Siri.[1396] Activating the wake word by accidental triggers is problematic because it leads to the recording (and upload to the cloud) of potentially sensitive audio data.[1397]

NLP provides powerful means to analyse voice and speech data obtained by means of virtual assistants (VA), in particular when combined with classification techniques adopted in the AI discipline

[1389] Giovanni Butarelli, 'The urgent case for a new ePrivacy law' (2018) <https://edps.europa.eu/press-publications/press-news/blog/urgent-case-new-eprivacy-law_fr> accessed 8 February 2024.
[1390] See, for an overview Roberto Pieraccini, *AI Assistants* (MIT Press 2021).
[1391] Roberto Pieraccini, *AI Assistants* (MIT Press 2021) 7.
[1392] Lea Schönherr et al, 'Unacceptable, where is my privacy? Exploring Accidental Triggers of Smart Speakers' (2020) at 1 < https://arxiv.org/pdf/2008.00508.pdf > accessed 8 February 2024.
[1393] Tom Bolton et al, 'On the Security and Privacy Challenges of Virtual Assistants' (2021) Vol 21 Iss 7 Sensors 1-3.
[1394] Yousra Javed, Shashank Sethi, Akshay Jadoun, 'Alexa's Voice Recording Behavior: A Survey of User Understanding and Awareness' (ARES '19, Canterbury 26-29 August 2019) 3 <https://dl.acm.org/doi/10.1145/3339252.3340330> accessed 8 February 2024.
[1395] Daniel J Dubois et al, 'When Speakers Are All Ears: Characterizing Misactivations of IoT Smart Speakers' (2020) Iss 4 Proceedings on Privacy Enhancing Technologies 255-276.
[1396] Daniel J Dubois et al, 'When Speakers Are All Ears: Characterizing Misactivations of IoT Smart Speakers' (2020) Iss 4 Proceedings on Privacy Enhancing Technologies 255-276; Lea Schönherr et al, 'Unacceptable, where is my privacy? Exploring Accidental Triggers of Smart Speakers' (2020) < https://arxiv.org/pdf/2008.00508.pdf > accessed 8 February 2024.
[1397] Lea Schönherr et al, 'Unacceptable, where is my privacy? Exploring Accidental Triggers of Smart Speakers' (2020) at 2 < https://arxiv.org/pdf/2008.00508.pdf > accessed 8 February 2024.

ML (see Section 2.2.1.1). With NLP and ML, rather sensitive information can be derived from human speech and other acoustic elements in recorded audio. In addition to the linguistic content of speech, a speaker's voice characteristics and manner of expression may contain a rich array of personal information, including clues about the speaker's biometric identity, personality, physical traits, geographical origin, level of intoxication/sleepiness, age, gender, health condition and even an individual's socioeconomic status.[1398]

As outlined in Section 2.2.4.2 regarding the AI discipline AC, speech-based emotion recognition systems measure and quantify emotions of a person by observing speech signals.[1399] Research has demonstrated specific associations between emotions such as fear, anger, sadness, joy and features of speech such as pitch, voice level and speech rate.[1400] Human-machine communication intercepted by means of virtual assistants can therefore also be used to detect the emotional state of the user. Amazon's patented technology enabling Alexa to detect the user's emotional state derived from the user's voice underscores this claim.[1401] Another real-world application is Amazon's wearable 'Halo', which analyses voice tones to detect user emotions.[1402] Information concerning the emotional state of an individual might be particularly helpful to manipulate this individual because emotions play an important role in the elicitation of autonomous motivated behaviour.[1403] According to research in behavioural sciences, especially psychology, emotions constitute powerful, pervasive and predictable drivers of decision-making.[1404] Emotions can therefore have significant effects on economic transactions and play a powerful role in everyday economic choices.[1405]

Companies such as Apple, Amazon, Google and the like offer virtual assistants and intercept, analyse and otherwise process human-machine communication for a plethora of purposes and infer sensitive information by means of ML, NLP and AC, without falling under the scope of the ePD. It should be noted that this lacuna in the current legal framework does not solely apply to human-machine

[1398] Jacob Leon Kröger, Otto Hans-Martin Lutz, Philip Raschke, 'Privacy Implications of Voice and Speech Analysis – Information Disclosure by Inference' in Michael Friedewald et al (eds) *Privacy and Identity management. Data for Better Living: AI and Privacy* (Springer 2020) 242.
[1399] Chi-Chun Lee et al, 'Speech in Affective Computing' in Rafael Calvo et al (eds), *The Oxford Handbook of Affective Computing* (OUP 2015) 171.
[1400] Christina Sobn and Murray Alpert, 'Emotion in Speech: The Acoustic Attributes of Fear, Anger, Sandess, and Joy' (1999) Vol 28 No 4 Journal of Psycholinguistic Research, 347.
[1401] Huafeng Jin, Shuo Wang 'Voice-Based Determination of Physical and Emotional Characteristics of Users' US Patent Number US 10096319 B1 (Assignee: Amazon Technologies, Inc.) October 2018 <https://patentimages.storage.googleapis.com/f6/a2/36/d99e36720ad953/US10096319.pdf> accessed 8 February 2024.
[1402] Alex Hern, 'Amazon's Halo wristband: the fitness tracker that listens to your mood' *The Guardian* (London, 28 August 2020) < https://www.theguardian.com/technology/2020/aug/28/amazons-halo-wristband-the-fitness-tracker-that-listens-to-your-mood > accessed 8 February 2024; Austin Carr, 'Amazon's New Wearable Will Know If I'm Angry. Is That Weird?' *Bloomberg* (New York, 31 August 2020) < https://www.bloomberg.com/news/newsletters/2020-08-31/amazon-s-halo-wearable-can-read-emotions-is-that-too-weird > accessed 8 February 2024.
[1403] Leen Vandercammen et al, 'On the Role of Specific Emotions in Autonomous and Controlled Behaviour' (2014) Vol 28 Iss 5 European Journal of Personality 413, 445.
[1404] Jennifer S Lerner et al, 'Emotion and Decision Making' (2015) Vol 66 Annual Review of Psychology 799, 802.
[1405] Jennifer S Lerner, Deborah A Small, George Loewenstein, 'Heart Strings and Purse Strings' (2004) Vol 15 No 5 American Psychology Society 337-340.

communication but also to *interpersonal communications.* All major players in the virtual assistant market (Amazon, Google, Microsoft and Apple) revealed that audio recordings made by their virtual assistants were listened to by either employees or subcontractors to categorise utterances, improve the quality of wake word detection and the performance of speech transaction.[1406] For example, in 2019 Google Assistant recordings were leaked to the Belgian news site VRT NWS. The corresponding report published by the news site outlined that Google employees systematically listened to audio files recorded by Google Home smart speakers and the Google Assistant smartphone app.[1407] Unavoidably, these audio recordings also include interpersonal communications. In the case of Google, the audio snippets contained a wide range of highly sensitive recordings, including private conversations about health status, domestic violence, sexual relationships and drug deals.[1408] In addition, a former Apple employee revealed that he had listened to hundreds of Siri recordings every day, including unintended recordings, for the purpose of quality control.[1409] These recordings concerned sensitive interpersonal communications such as discussions between doctors and patients, business deals, seemingly criminal acts and sexual encounters.[1410] Press coverage points to similar practices at Amazon.[1411]

Of course, providers of human-machine communication services must comply with the provisions enshrined in the GDPR when processing personal data in this context. However, the provisions of the GDPR are less strict than Article 5 (1) ePD, which requires consent for the surveillance of interpersonal communications. According to the GDPR, consent is only one of six legal bases. As outlined in Section 4.4.2, consent is one of the main legislative tools for giving individuals control over the processing of their personal data,[1412] if not the 'ultimate expression of control'.[1413] By excluding human-machine communication services from its scope, the ePD fails to meet its legislative aims to guarantee the confidentiality of communications,[1414] to protect natural persons with respect to the automated storage and processing of data[1415] and ultimately to protect personal data and the privacy of

[1406] CNIL, 'Exploring the ethical, technical and legal issues of voice assistants' (2020) 40 <https://www.cnil.fr/sites/default/files/atoms/files/cnil_white-paper-on_the_record.pdf> accessed 8 February 2024.
[1407] Tim Verheyden et al, 'Hey Google, are you listening?' *VRTB* (Brussels 10 July 2019) <https://www.vrt.be/vrtnws/en/2019/07/10/google-employees-are-eavesdropping-even-in-flemish-living-rooms/ > accessed 8 February 2024.
[1408] Tine Munk, 'Does Online Privacy Exist in the GDPR Era? The Google Voice Assistant Case' in Tatiana-Eleni Synodiou et al (eds) *EU Internet Law in the Digital Single Market* (Springer Nature 2021) 497.
[1409] Alex Hern, 'Apple contractors regularly hear confidential details on Siri recordings' *The Guardian* (London, 26 July 2019) < https://www.theguardian.com/technology/2019/jul/26/apple-contractors-regularly-hear-confidential-details-on-siri-recordings > accessed 8 February 2024.
[1410] Ibid.
[1411] Alex Hern, 'Amazon staff listen to customers' Alexa recordings, report says' *The Guardian* (London, 11 April 2019) <https://www.theguardian.com/technology/2019/apr/11/amazon-staff-listen-to-customers-alexa-recordings-report-says> accessed 8 February 2024.
[1412] Giovanni Butarelli, 'The urgent case for a new ePrivacy law' (2018) <https://edps.europa.eu/press-publications/press-news/blog/urgent-case-new-eprivacy-law_fr> accessed 8 February 2024.
[1413] Case C-300/21 *UI* [2022] ECR I-756 Opinion AG Manuel Campos Sánchez-Bordona para 73.
[1414] Recital 3 ePD.
[1415] Recital 7 ePD.

users.[1416] Human-machine communications deserve the same level of confidentiality as applicable to interpersonal communications under of Article 5 (1) ePD when considering the sensitivity of information captured by human-machine communications and the sensitive information that can be derived from it. Due to this gap in legal protection, Article 5 (1) ePD is not fit for purpose to ensure the confidentiality of human-machine communication and interpersonal communication facilitated by current human-machine communication services and similar future services. This creates a Type 3 legal problem regarding the fundamental rights to privacy and data protection.

The communication surveillance problem (Type 3)

ML, NLP and AC facilitate the surveillance of both human-machine and interpersonal communication. Major tech companies that offer human-machine communication services, such as virtual assistants, may easily intercept and otherwise process such communication. Providers of these services do not fall under the strict regime of Article 5 (1) ePD, which regulates the confidentiality of communications. This creates a significant gap in legal protection and outlines that the ePD is not fit for purpose to ensure the confidentiality of both interpersonal and human-machine communication.

Likewise, the requirement to obtain consent for the storage of information or gaining access to information already stored in the terminal equipment of a subscriber or user according to Article 5 (3) ePD as introduced in Section 3.4.3.2 is likely not applicable to virtual assistant services. With virtual assistants, the information (e.g., voice recordings) is *not* stored on the terminal equipment, nor does the service gain access to information stored on the terminal equipment. This holds true regardless of whether the virtual assistant service is embedded in a smartphone or in a smart home device such as 'Amazon Echo'. Rather, information is stored and otherwise processed within the service platform of the provider, namely, in the cloud.[1417] Regulatory guidance neglects the technical functioning of virtual assistant services such as Amazon Alexa when stating that 'consent as required by Article 5 (3) ePD would be necessary for the storing or gaining of access to information for any purpose other than executing a user request (e.g., user profiling)'.[1418] Leading virtual assistants do not store the voice recording of their users on the terminal equipment, but rather on the service platform of the provider, mostly in the cloud. [1419] Major providers of virtual assistants (e.g. Amazon, Apple, Google, Cortana) rely on cloud environments to store data processed in the context of virtual assistants.[1420] This is

[1416] Recitals 2, 5 ePD.
[1417] Daniel J Dubois et al, 'When Speakers Are All Ears: Characterizing Misactivations of IoT Smart Speakers' (2020) Iss 4 Proceedings on Privacy Enhancing Technologies 255; Lea Schönherr et al, 'Unacceptable, where is my privacy? Exploring Accidental Triggers of Smart Speakers' (2020) at 1 < https://arxiv.org/pdf/2008.00508.pdf > accessed 8 February 2024.
[1418] European Data Protection Board, 'Guidelines 02/2021 on Virtual Voice Assistants' (16 May 2011) at 29.
[1419] Tom Bolton et al, 'On the Security and Privacy Challenges of Virtual Assistants' (2021) Vol 21 Iss 7 Sensors 1-3; Allan de Barcelos Silva et al, 'Intelligent personal assistants: A systematic literature review' (2020) Vol 147 Expert Systems With Applications 1, 8.
[1420] Allan de Barcelos Silva et al, 'Intelligent personal assistants: A systematic literature review' (2020) Vol 147 Expert Systems With Applications 1, 8.

different when compared with another 'always-on' service, namely, Amazon's wearable 'Halo', which analyses voice tones to detect user emotions.[1421] According to Amazon, the recordings are never uploaded to the cloud but instead analysed on the user's device and then deleted.[1422]

For most virtual assistants, the computing performed locally on the device focusses on listening for a wake word[1423] and sampling subsequent audio information for transportation to the cloud.[1424] Computing necessary for automatic speech recognition, natural language understanding, natural language generation and ultimately speech generation[1425] are thus *not* performed or stored locally on the device used by the virtual assistant service. Most virtual assistants do not require storing information or accessing information on the user's device. Rather, by uttering the voice command, the user initiates the streaming of the voice recordings to the servers of the provider *via* the device. This does not mean that the provider retrieves the voice recording *from* the device or gains access to voice recordings *stored on* the device of the user.[1426] Moreover, virtual assistants are software applications[1427] consisting of several components[1428] and layers. They are, as such, *not* terminal equipment as referred to in Article 5 (3) ePD. Like any other software, virtual assistants rely on hardware in order to function, for example, devices like computers, smartphones, tablets or on purpose-built speaker devices.[1429] When activated by the voice command of the user, the device usually sends the speech recording directly to the service platform of the provider where it is subsequently stored.[1430] Hence, the device *solely* opens a stream to the cloud[1431] but does not store the voice recording (e.g., voice command).

[1421] Austin Carr, 'Amazon's New Wearable Will Know If I'm Angry. Is That Weird?' *Bloomberg* (New York, 31 August 2020) < https://www.bloomberg.com/news/newsletters/2020-08-31/amazon-s-halo-wearable-can-read-emotions-is-that-too-weird > accessed 8 February 2024.
[1422] Alex Hern, 'Amazon's Halo wristband: the fitness tracker that listens to your mood' *The Guardian* (London, 28 August 2020) < https://www.theguardian.com/technology/2020/aug/28/amazons-halo-wristband-the-fitness-tracker-that-listens-to-your-mood > accessed 8 February 2024.
[1423] Lea Schönherr et al, 'Unacceptable, where is my privacy? Exploring Accidental Triggers of Smart Speakers' (2020) at 2 < https://arxiv.org/pdf/2008.00508.pdf > accessed 8 February 2024.
[1424] Tom Bolton et al, 'On the Security and Privacy Challenges of Virtual Assistants' (2021) Vol 21 Iss 7 Sensors 1-3; Allan de Barcelos Silva et al, 'Intelligent personal assistants: A systematic literature review' (2020) Vol 147 Expert Systems With Applications 1, 11.
[1425] Roberto Pieraccini, *AI Assistants* (MIT Press 2021) 8.
[1426] Centre for Information Policy Leadership, 'Comments by the Centre for Information Policy Leadership on the EDPB Guidelines 02/2021 on Virtual Voice Assistants' (2021) 5 < https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl_comments_on_the_edpb_guidelines_on_virtual_voice_assistants__23_april_2021_.pdf > accessed 8 February 2024.
[1427] Matthew B Hoy, 'Alexa, Siri, Cortana, and More: An Introduction to Voice Assistants' (2018) Vol 37 No 1 Medical Reference Services Quarterly 81, 82; Tom Bolton et al, 'On the Security and Privacy Challenges of Virtual Assistants' (2021) Vol 21 Iss 7 Sensors 1.
[1428] Roberto Pieraccini, *AI Assistants* (MIT Press 2021) 7.
[1429] Matthew B Hoy, 'Alexa, Siri, Cortana, and More: An Introduction to Voice Assistants' (2018) Vol 37 No 1 Medical Reference Services Quarterly 81, 82; Tom Bolton et al, 'On the Security and Privacy Challenges of Virtual Assistants' (2021) Vol 21 Iss 7 Sensors 1.
[1430] Tom Bolton et al, 'On the Security and Privacy Challenges of Virtual Assistants' (2021) Vol 21 Iss 7 Sensors 1-3; Allan de Barcelos Silva et al, 'Intelligent personal assistants: A systematic literature review' (2020) Vol 147 Expert Systems With Applications 1, 8.
[1431] Centre for Information Policy Leadership, 'Comments by the Centre for Information Policy Leadership on the EDPB Guidelines 02/2021 on Virtual Voice Assistants' (2021) 5 < https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl_comments_on_the_edpb_guidelines_on_virtual_voice_assistants__23_april_2021_.pdf > accessed 8 February 2024.

In addition, many virtual assistants are designed as distributed web services with application services provided by different companies, organisations and developers,[1432] which indicates significant sharing of data. Furthermore, it is unclear whether the user in fact can be aware of what actually is being stored[1433] and there seems to be no accurate mechanism for the user to exercise control regarding the sharing of such stored data.[1434] Providers of virtual assistant services, along with the different actors involved in providing the service, can further process the recorded speech of their users and other data to infer a rich array of personal information without the need to obtain consent from the users. Such information includes clues about the user's biometric identity, personality, physical traits, geographical origin, level of intoxication and sleepiness, age, gender, health condition and even an individual's socioeconomic status.[1435]

Regulatory guidance implies that processing in the context of virtual assistants occurs locally on the device[1436] and that providers of virtual assistant services gain access to information stored on the user's device. However, this is not correct. The actual speech recordings, namely, the command given to the virtual assistant, is directly transmitted to the platform of the provider. Further processing, as well as storage, occurs there.[1437] For this reason, it cannot be concluded that Article 5 (3) ePD is applicable. Virtual assistant services do not store information, nor gain access to information already *stored*, *in the terminal equipment*, as is the case with cookies. Major providers of these services, such as Amazon, Apple and Google store data processed in the context of virtual assistants in the cloud, not on the device.[1438] Moreover, providers can also link speech data with other datasets (e.g. social media meta data, browsing behaviour, purchase histories) in order to draw further sensitive inferences.[1439] As explained in the communication surveillance problem, there is a loophole in the current legal framework that specifically ensures the confidentiality of human-machine communication[1440] and

[1432] Allan de Barcelos Silva et al, 'Intelligent personal assistants: A systematic literature review' (2020) Vol 147 Expert Systems With Applications 1, 8; Tom Bolton et al, 'On the Security and Privacy Challenges of Virtual Assistants' (2021) Vol 21 Iss 7 Sensors 1-3.
[1433] Tom Bolton et al, 'On the Security and Privacy Challenges of Virtual Assistants' (2021) Vol 21 Iss 7 Sensors 1, 16.
[1434] Allan de Barcelos Silva et al, 'Intelligent personal assistants: A systematic literature review' (2020) Vol 147 Expert Systems With Applications 1, 8; Tom Bolton et al, 'On the Security and Privacy Challenges of Virtual Assistants' (2021) Vol 21 Iss 7 Sensors 1, 8.
[1435] Jacob Leon Kröger, Otto Hans-Martin Lutz, Philip Raschke, 'Privacy Implications of Voice and Speech Analysis – Information Disclosure by Inference' in Michael Friedewald et al (eds) *Privacy and Identity management. Data for Better Living: AI and Privacy* (Springer 2020) 242.
[1436] European Data Protection Board, 'Guidelines 02/2021 on Virtual Voice Assistants' (16 May 2011) at 16, 29.
[1437] This is acknowledged by regulatory guidance, which also remarks that Article 5 (3) ePrivacy Directive might need to be amended in the future. See European Data Protection Board, 'Guidelines 02/2021 on Virtual Voice Assistants' (16 May 2011) at 16 and Footnote 12 on page 12; see also Tom Bolton et al, 'On the Security and Privacy Challenges of Virtual Assistants' (2021) Vol 21 Iss 7 Sensors 1-3; Allan de Barcelos Silva et al, 'Intelligent personal assistants: A systematic literature review' (2020) Vol 147 Expert Systems With Applications 1, 8.
[1438] Allan de Barcelos Silva et al, 'Intelligent personal assistants: A systematic literature review' (2020) Vol 147 Expert Systems With Applications 1, 8.
[1439] Jacob Leon Kröger, Otto Hans-Martin Lutz, Philip Raschke, 'Privacy Implications of Voice and Speech Analysis – Information Disclosure by Inference' in Michael Friedewald et al (eds) *Privacy and Identity management. Data for Better Living: AI and Privacy* (Springer 2020) 2523.
[1440] However, the GDPR regulates human-machine communications provided this relates to the processing of personal data.

prevents surveillance as well as further processing thereof without consent.[1441] Therefore, Article 5 (3) ePD is not fit for purpose to protect the fundamental rights to privacy and data protection, in particular the privacy of communications, when considering the gap of protection it creates with regard to human-machine communications. It also fails to meet the ePD's legislative aims to guarantee the confidentiality of communications,[1442] to protect natural persons concerning the automated storage and processing of data[1443] and ultimately to protect personal data and the privacy of users.[1444] This constitutes a Type 3 legal problem regarding the fundamental rights to privacy and data protection. As indicated in the communication surveillance problem, this lacuna in the current legal framework does not solely apply to human-machine communication, but also to sensitive *interpersonal communications*, including conversations about health status, domestic violence, sexual relationships, drug deals,[1445] discussions between doctors and patients and business deals.[1446]

The storage problem (Type 3)

ML, NLP and AC facilitate the provision of virtual assistant services. Providers may analyse and otherwise process human-machine and interpersonal communication without needing to obtain consent from the user. Article 5 (3) ePD does not apply to virtual assistant services as they do not store information, or gain access to information already stored, in the device of the user. This provision is not fit for purpose to protect the fundamental rights to privacy and data protection because it creates a significant gap of protection as processing of both human-machine and interpersonal communication may reveal sensitive information and that likely is to be shared with various actors.

4.10 Conclusions

Chapter 4 aimed to answer Subquestion 3, namely, what legal problems arise or may arise when the principles enshrined in the current legal framework are applied to AI. In this chapter, I have outlined that all AI disciplines as described in Section 2.2 raise or may raise legal problems when they are applied to the principles enshrined in the current legal framework as introduced in Chapter 3. Three types of legal problems were identified: (1) legal provisions that are violated, (2) legal provisions that cannot be enforced and (3) legal provisions that are not fit for purpose to protect the fundamental right at stake. These legal problems may be caused by AI disciplines *or* by the principles themselves when they are applied in the context of AI. Table 4.3 provides an overview of the legal problems identified in this chapter.

[1441] For instance, as it is the case with consent for cookies as required by Article 5 (3) ePrivacy Directive.
[1442] Recital 3 ePD.
[1443] Recital 7 ePD.
[1444] Recitals 2, 5 ePD.
[1445] Tine Munk, 'Does Online Privacy Exist in the GDPR Era? The Google Voice Assistant Case' in Tatiana-Eleni Synodiou et al (eds) *EU Internet Law in the Digital Single Market* (Springer Nature 2021) 497.
[1446] Alex Hern, 'Apple contractors regularly hear confidential details on Siri recordings' *The Guardian* (London, 26 July 2019) < https://www.theguardian.com/technology/2019/jul/26/apple-contractors-regularly-hear-confidential-details-on-siri-recordings > accessed 8 February 2024.

Problem	Principles	Type	AI Disciplines
Balancing	Lawfulness, Proportionality, Accountability	1	AR
Probability	Fairness, Accountability	1	ML, AR
Facial recognition	Fairness, Accountability	1	CV
Inaccuracy	Fairness, Accuracy, Accountability	1	ML, AC
Sensitivity	Fairness, Accountability	1	AC
Elusiveness	Fairness	2, 3	ML, NLP, CV, AC, AR
Manipulation	Fairness	3	ML, AC
Sabotage	Fairness	3	ML, NLP, CV, AC, AR
Opacity	Transparency, Accountability	1	ML, NLP, CV, AC, AR
Interpretability	Transparency	1, 2	ML
Inference	Transparency	3	ML, AC
Profiling	Transparency	3	ML, AC
Inexplicitness	Purpose limitation, Accountability	1	ML, NLP, CV, AC, AR
Function creep	Purpose limitation, Accountability	1	ML, NLP, CV, AC, AR
Restriction	Purpose limitation	3	ML, NLP, CV, AC, AR
Compatible use	Purpose limitation	3	ML
Data appetite	Data minimisation, Accountability	1	ML, NLP, CV, AC, AR
Necessity	Data minimisation, Accountability	1	ML
Verification	Data minimisation	2	ML, NLP, CV
Trade-off	Data minimisation, Accuracy, Fairness	3	ML, NLP, CV, AC, AR
Rebuttal	Accuracy, Accountability	1	AC
Common sense	Accuracy, Accountability	1	AR
Guidance	Accuracy	2, 3	ML, NLP, CV, AC, AR
Incomputability	Accuracy, Fairness	3	ML, NLP, CV, AC, AR
Emotion data	Enhanced protection for special data	3	AC
Location data	Enhanced protection for special data	3	ML
Neurodata	Enhanced protection for special data	3	ML (DL), CV, NLP
Mental data	Enhanced protection for special data	3	ML, AC
Communication surveillance	Confidentiality	3	ML, NLP, AC
Storage	Confidentiality	3	ML, NLP, AC

Table 4.3 Overview of legal problems related to the principles contained in the legal framework. The brackets surrounding DL indicate that this *specific kind* of ML causes the legal problem in question.

Table 4.3 illustrates the broad range of legal problems that arise or may arise in the context of AI. In total, 30 problems are identified.

The *lawfulness principle* does not appear to be particularly problematic when applied to AI; after all, only one legal problem relates to this principle. The reason for this is that the meaning of the lawfulness principle is substantively clear, as is further substantiated in Article 6 GDPR, which enumerates six lawful bases that can be relied upon for the processing of personal data. According to Table 4.3, only the AI discipline *AR* causes a Type 1 legal problem when applied to the lawfulness principle.

The *fairness principle* causes the most legal problems when applied to AI: 10 out of 30 problems relate to the fairness principle. In addition, it causes all three types of legal problems. When interpreted as substantive fairness to prevent adverse effects on data subjects, the principle of fairness can be violated by processing facilitated by AI. This is underscored by the fact that four legal problems regarding the fairness principle relate to Type 1 problems. The main issue with the fairness principle lies in its elusive meaning. The substantively unclear meaning of the fairness principle reduces legal certainty and makes it less likely that it will be enforced by individuals or regulators (Type 2 problem). Ultimately, this also leads to Type 3 legal problems because a substantively unclear principle is not fit for purpose to protect the fundamental right to data protection. *ML* and *AC* seem to be the most problematic AI disciplines when applied to the fairness principle.

Regarding the *transparency* principle, I have identified four legal problems of either Type 1, 2 or 3 when applied to AI. Because AI systems may be rather ubiquitous, all AI disciplines potentially clash with the transparency principle. Nevertheless, *ML* is the main driver for legal problems relating to the transparency principle: all the four legal problems relate to this AI discipline. This is mainly caused by the fact that ML is widely used to infer and derive data from existing data and because AI systems deploying DL and ANN approaches are likely to produce noninterpretable outputs.

Regarding the *purpose limitation* principle, I have identified four legal problems of either Type 1 or Type 3 when applied to AI. Generally, all AI disciplines process personal data from various sources for a plethora of purposes and are therefore in conflict with the purpose limitation principle. ML serves as a typical example: unsupervised ML processes personal data for unspecific and inexplicit purposes. Thus, the processing itself determines the purpose and future use of personal data, which causes Type 1 legal problems. The purpose limitation principle also causes Type 3 legal problems when applied to AI because it does not restrict the processing of personal data and allows further processing for compatible purposes.

Regarding the *data minimisation* principle, I have identified four legal problems of Type 1, 2 or 3 when applied to AI. Type 1 problems are mainly caused by the data appetite of AI – regardless of which discipline of AI is used to process personal data. When consequently applied in the context of AI, the data minimisation principle may create trade-offs regarding the accuracy and fairness principles, which leads to Type 3 legal problems. Here as well, *ML* is the main driver for the legal problems

with respect to the data minimisation principle: all four legal problems are caused by this single AI discipline.

Regarding the *accuracy* principle, I have identified six legal problems of Type 1, 2 or 3 when applied to AI. The main issue with the accuracy principle is caused by the fact that the required level of accuracy depends on the purpose of the processing, as suggested by relevant case law ('relative accuracy'). Such relative accuracy does not outline specific levels of accuracy that personal data processed in the context of AI must reach: there is no one-size-fits-all approach. Thus, the precise substantive requirements of the accuracy principle remain an underexplored topic in academia and case law, which is highly problematic when considering the developments in AI, causing Type 2 and 3 legal problems. *ML* and *AC* are the most problematic AI disciplines because they are likely to generate inaccurate personal data.

Regarding the principle of *enhancing protection for special categories of personal data*, I have identified four Type 3 legal problems when applied to AI. The main issue of this principle is caused by the legislators' approach to exhaustively enumerate special data in Article 9 GDPR. This exhaustive list of special personal data contained in the GDPR does not keep up with the technological developments facilitated by AI. The stringent rules concerning the processing of sensitive data do not apply to the processing of new types of sensitive personal data facilitated by AI, such as emotion data, neurodata and mental data. As apparent from Table 4.3, *ML, NLP* and *AC* are the most problematic AI disciplines in the context of this principle.

Regarding the *confidentiality of communications* principle, I have identified two Type 3 legal problems when applied to AI. Due to the restricted material scope of the ePD, the prohibition of listening, tapping, storage or other kinds of interception or surveillance does not apply in the context of virtual assistants and smart home technologies which are powered by the AI disciplines ML, NLP and AC. Likewise, the ePD does not require providers of virtual assistant services to obtain consent from their users in order to analyse and otherwise process human-machine communication because virtual assistant services typically do not store information, or gain access to information already stored, in the device of the user as required by the ePD. As apparent from Table 4.3, *ML, NLP* and *AC* are the most problematic AI disciplines in the context of this principle.

In terms of the *types of legal problems* caused by AI, Table 4.3 shows that 14 out of 30 legal problems identified within this chapter relate to *Type 3* legal problems. Thus, there is a clear mismatch between the principles enshrined in the current legal framework and the AI disciplines introduced in Chapter

2.[1447] This means that legislative measures may be needed to address said mismatch. Furthermore, almost half of the problems relate to *Type 1* legal problems. Thus, AI is likely to violate the principles enshrined in the current legal framework. *Type 2* legal problems seem to be rare: only four legal problems identified within this chapter relate to the enforcement of the provisions enshrined in the current legal framework. Therefore, more enforcement seems to be needed, both with respect to private enforcement initiated by data subjects or representative bodies and with respect to regulatory enforcement pursued by SAs.

In terms of which AI disciplines cause *how many legal problems* when applied to the principles enshrined in the current legal framework, Table 4.3 shows that ML leads to twenty-four, NLP fourteen, CV thirteen, AC nineteen and AR thirteen legal problems, respectively. The prominent role of ML is not surprising, as this AI discipline is the most widely used and often combined with other AI disciplines. In addition, AC seems to be the main driver of legal problems which only causes slightly less legal problems when compared to ML. The amounts of legal problems associated to the AI disciplines NLP, CV and AR are distributed almost equally.

[1447] This is in line with other research, e.g. Tal Z Zarsky 'Incompatible: The GDPR in the Age of Big Data' (2017) Vol 47 Iss 4 Seton Hall Law Review 995-1020.

5 Legal problems: Rights

This chapter aims to answer Subquestion 4, namely, what legal problems arise or may arise when the enforceable rights enshrined in the current EU legal framework are applied to AI. Section 5.1 introduces the approach taken to assess the legal problems. Sections 5.2 through 5.5 elaborate on the fundamental right to privacy introduced in Section 3.1 and discuss four dimensions of privacy that are derived from the elements contained in the text of the fundamental right to privacy and the corresponding case law. These four dimensions are informational privacy (Section 5.2), bodily privacy (Section 5.3), mental privacy (Section 5.4) and communicational privacy (Section 5.5). Sections 5.6 through 5.11 do the same for the fundamental right to data protection as introduced in Section 3.2. I focus on the enforceable rights that data subjects have according to the GDPR because they implement the requirements enshrined in the fundamental right to data protection.[1448] Strong[1449] and effective data subject rights[1450] constitute a prerequisite for the protection of personal data. These enforceable rights are the right of access (Section 5.6), the right to rectification (Section 5.7), the right to erasure (Section 5.8), the right to data portability (Section 5.9), the right to object (Section 5.10) and the right not to be subject to automated decision-making (Section 5.11). Section 5.12 concludes.

Note that transparency requirements according to Articles 12-14 GDPR technically do not belong to the enforceable rights of data subjects although they are listed under data subject rights. Rather, these provisions are the manifestations of the transparency principle[1451] which I discussed in Section 4.4.

5.1 Approach

The approach for assessing legal problems related to the rights enshrined in the current legal framework is the same as introduced in Section 4.1. When referring to legal problems, three types of legal problems are distinguished, namely, Type 1 (legal provisions are violated), Type 2 (legal provisions cannot be enforced) and Type 3 (legal provisions are not fit for purpose to protect the fundamental right at stake). Type 3 legal problems are discussed from the perspective of *natural persons* as the primary subject of protection envisaged by fundamental rights. These types of legal problems are identified by means of the rationales and specific aims pursued by the current legal framework as outlined in Section 4.1 (see Table 4.2 therein). To determine which type of legal problem arises or may arise due to different AI disciplines, as outlined in Chapter 2, the AI disciplines are mapped with the enforceable rights contained in the current legal framework. For each right enshrined in the current

[1448] Case C-131/12, *Google Spain* [2014] ECR I-317 para 69; Joined Cases C-141/12 & C-372/12, *YS, M* [2014] ECR I-2081, Opinion of AG Sharpston para 55.
[1449] Case C-132/21 *Nemzeti* [2023] ECR I-2 para 42.
[1450] Case C-154/21 *Österreichische Post AG* [2023] ECR I-3 para 39.
[1451] Case C-487/21, *F.F.* [2022] ECR I-1000 para 37.

legal framework, I assess whether the enforceable right at hand creates Type 1, 2 or 3 legal problems. When doing so, I follow the order of the AI disciplines outlined in Chapter 2.

Before getting started, I shall explain the focus I have chosen with respect to the fundamental right to privacy. The latter states that everyone has the right to respect for his or her private life, home and communications.[1452] Due to the broad scope of the fundamental right to privacy, I focus on four dimensions[1453] of this right that are particularly relevant in the light of AI: informational, bodily, mental and communicational. These dimensions are derived from the elements contained in the text of the fundamental right to privacy and corresponding case law.[1454] Table 5.1 maps the elements of the right to privacy derived from the text and corresponding case law with the dimensions of privacy that are discussed in this chapter.

Element of the fundamental right to privacy	**Dimension**
private life	informational privacy
private life (physical integrity)	bodily privacy
private life (mental integrity)	mental privacy
correspondence/communications	communicational privacy

Table 5.1 Mapping elements contained in the text of the fundamental right to privacy in Article 8 ECHR and corresponding case law with dimensions of the right to privacy discussed in Section 3.1.

Let me explain why I have chosen to focus on these four dimensions of privacy. First, the right to privacy provides individuals with a form of informational self-determination,[1455] which is an extremely important dimension in the context of AI because the latter relies heavily on the processing of information. I discuss informational privacy in Section 5.2. Furthermore, physical and mental integrity, two elements falling under the term 'private life' as developed in corresponding case law [1456] are particularly relevant in the context of AI. I consider these two elements to be important because some AI disciplines such as affective computing and machine learning deal with body functions and characteristics (e.g., genetic codes, biometrics, physiological information) and aim to gain access to

[1452] Art 8 ECHR, Art 7 EUCFR.
[1453] Note that I refrain from elaborating on the elements 'family life' and 'home' contained in the text of the fundamental right to privacy because these elements do not seem to be particularly relevant in the context of AI. The element 'family life' essentially relates to the right to live together so that family relationships may develop normally and those members of the family may enjoy each other's company. See *Marckx v. Belgium* App no 6833/74 (ECtHR 13 June 1979) para 31; *Olsson v. Sweden* (No. 1) App no 10465/83 (ECtHR 24 March 1988) para 59. Possible interferences with the right to respect for one's home include examples such as police entry into a person's home, including searches and seizures, and displacements from home. See *Murray v. the United Kingdom* App no 14310/88 (ECtHR 28 October 1994) para 86; *Burlya and others v. Ukraine* App no 3289/10 (ECtHR 6 February 2019) para 166.
[1454] See also Bert-Jaap Koops et al 'A Typology of Privacy' (2017) Vol. 38 Iss. 2 University of Pennsylvania Journal of International Law.
[1455] *Satakunnan Markkinapörssi Oy and Satamedia Oy v. Finland* App no 931/13 (ECtHR 27 June 2017) para 137.
[1456] *Denisov v Ukraine* App no 76639/11 (ECtHR 25 September 2018) para 95, *S. and Marper v United Kingdom* App no 30562/04 and 30566/04 (ECtHR 4 December 2008) *Pretty v United Kingdom* App no 2346/02 (ECtHR 29 April 2002) para 63.

mental states of individuals (such as thoughts, feelings, emotional states). I discuss the element of physical integrity under the dimension of bodily privacy (Section 5.3) and the element of mental integrity under the dimension of mental privacy (Section 5.4). Finally, communication also constitutes an important element in the context of AI, as AI might interfere with the right to respect confidential communication because it computes communications in various forms, for example, by means of natural language processing and machine learning. I discuss the element of communication under the dimension of communicational privacy (Section 5.5).

5.2 Informational privacy

Informational privacy refers to the idea that data and images from individuals should not be automatically available to others[1457] and that individuals may 'exercise a substantial degree of control over that data and its use'.[1458] According to ECtHR case law, the right to privacy provides individuals with a form of informational self-determination[1459] which indicates that individuals should be able to exercise control with regard to the use of their information. Informational privacy should be understood as an overarching concept[1460] rather than a separate type or form of privacy.[1461] All AI disciplines as described in Chapter 2 process various types of information. In this section, I examine how these AI disciplines may lead to legal problems when applied to informational privacy.

5.2.1 Legal problems: Type 1

Research has shown that ML models can successfully identify markers of depression by analysing photographic data from Instagram accounts, and these models even outperformed general practitioner's average diagnostic success rate for depression.[1462] This implies that sensitive information about individuals can be inferred and disclosed to others beyond the individual's control, which contradicts their right to informational self-determination.[1463]

An ML-powered system that aims to analyse customer behaviour from large volumes of customer transaction data can make accurate predictions based on patterns and correlations identified in past customer behaviour.[1464] This could reveal information an individual arguably did not want to disclose.

[1457] Rachel L. Finn, David Wright, Michael Firedewald, 'Seven Types of Privacy' in S. Gutwirth et al. (eds.), *European Data Protection: Coming of Age* (Springer 2013) 8; Bert-Jaap Koops et al 'A Typology of Privacy' (2017) Vol. 38 Iss. 2 University of Pennsylvania Journal of International Law 438, 568.
[1458] Roger Clarke, 'Introduction to Dataveillance and Information Privacy, and Definitions of Terms' (Roger Clarke's Website, 24 July 2016) < http://www.rogerclarke.com/DV/Intro.html > accessed 8 February 2024.
[1459] *Satakunnan Markkinapörssi Oy and Satamedia Oy v. Finland* App no 931/13 (ECtHR 27 June 2017) para 137.
[1460] Bert-Jaap Koops et al 'A Typology of Privacy' (2017) Vol. 38 Iss. 2 University of Pennsylvania Journal of International Law 438, 568-569.
[1461] Bart Custers, *The Power of Knowledge* (Wolf Legal Publishers 2004) 145.
[1462] Andrew G Reece, Christopher M Danforth, 'Instagram photos reveal predictive markers of depression' (2017) Vol. 6 No. 15 EPJ Data Science, 1, 8.
[1463] *Satakunnan Markkinapörssi Oy and Satamedia Oy v. Finland* App no 931/13 (ECtHR 27 June 2017) para 137.
[1464] Ethem Alpaydin, *Machine Learning: The New AI* (3rd edn MIT Press 2016) 14.

A famous example is the so-called 'pregnancy prediction' score for female customers who paid with a credit card or used a loyalty card. Based on two dozen products used as proxies, the prediction model could identify pregnant customers when analysing their past shopping cart.[1465] ML approaches, for example clustering as described in Section 2.2.1.2, may infer an individual's home and work location from widely available location metadata in public data streams like Twitter.[1466] ML approaches can also infer even more sensitive information pertaining to health, religion and nightlife from location metadata through the reconstruction of a user's location history[1467] (see also Section 4.8.3). ML models that apply dimensionality reduction (see Section 2.2.1.2) on easily accessible digital records of behaviour, for example Facebook likes, may reveal and predict highly sensitive personal attributes such as sexual orientation, ethnicity, religious and political views and personality traits.[1468] Facebook users seem to have no control to prevent that such sensitive information will subsequently be revealed at the moment they click on the like button. In addition, arguably anonymised information could identify individuals when analysed by means of ML. According to a study which deployed ML approaches, 99.98% of the population of a US state could be uniquely re-identified in any dataset using fifteen demographic attributes.[1469] This study also demonstrates that identification can be estimated with high accuracy even when the anonymised dataset is heavily incomplete, which rejects claims that re-identification is not a practical risk.[1470]

Speech recordings, if analysed by AI, can reveal not only an individual's identity, but also gender, age, native language, emotional state[1471] and information related to individual's personality traits, degree of sleepiness or intoxication and physical and mental health, as well as socioeconomic status.[1472] Individuals are often unaware of being recorded and have limited means to control what information is inferred from their recorded speech through ML and NLP approaches. For example, Amazon has patented a version of its virtual assistant Alexa that (arguably) is able to detect whether a user is ill and then subsequently offer medicine.[1473] This raises the question whether a user wants to reveal such information in the first place and how effective control can be exercised when an individual does not

[1465] Viktor Mayer-Schönberger; Kenneth Cukier, *Big Data: A Revolution That Will Transform How We Live, Work and Think* (Houghton Mifflin Harcourt 2013) 58.
[1466] Drakonakis Kostas et al, 'Please Forget Where I Was Last Summer: The Privacy Risks of Public Location (Meta) Data' (2019) 2 <https://arxiv.org/pdf/1901.00897.pdf> accessed 8 February 2024.
[1467] Ibid 1.
[1468] Michal Kosinski, David Stillwell, Thore Graepel, 'Private traits and attributes are predictable from digital records of human behaviour' (2013) Vol 110 No 15 PNAS, 5802.
[1469] Luc Rocher, Julien M Hendrickx, Yves-Alexandre de Montjoye, 'Estimating the success of re-identifications in incomplete datasets using generative models' (2019) Vol 10 Nature Communications 1.
[1470] Ibid 2.
[1471] Andreas Nautsch et al, 'Preserving privacy in speaker and speech characterisation' (2019) Vol 58 Computer Speech & Language 441, 444.
[1472] For more detailed information and related studies, see Jacob Leon Kröger, Otto Hans-Martin Lutz, Philip Raschke, 'Privacy Implications of Voice and Speech Analysis – Information Disclosure by Inference' in Michael Friedewald et al (eds) *Privacy and Identity management. Data for Better Living: AI and Privacy* (Springer 2020) 243.
[1473] James Cook, 'Amazon patents new Alexa feature that knows when you're ill and offers you medicine' *The Telegraph* (London 9 October 2018) < https://www.telegraph.co.uk/technology/2018/10/09/amazon-patents-new-alexa-feature-knows-offers-medicine/> accessed 8 February 2024.

want to reveal such information. NLP approaches embedded in virtual assistants, such as Amazon Alexa or smart home applications (e.g., smart fridges and beds), provide the technical means to track and monitor individuals in an unprecedented manner. By deploying the most recent NLP and speech recognition techniques, virtual assistants such as Siri, Google Assistant and Amazon Alexa may effortlessly recognise when a person is praying and thus reveal rather sensitive information.

Behavioural inference systems that deploy CV and ML techniques in retail spaces allow fine-grained tracking of the shoppers' behaviour and characteristics.[1474] Possibilities for shoppers to truly avoid this is difficult, if possible, at all.[1475] CV allows one to identify people based on gait. Biometric information necessary for doing so may be captured in public spaces and from a distance.[1476] For example, the police in China can identify suspects by their gait and silhouette from up to 50 metres distance, even when a person's face is covered or pretends to have a limp or hunch.[1477] The same technology can be applied in semi-public spaces such as connected retail spaces for commercial purposes. In particular, when integrated into existing surveillance systems, face recognition (see Section 2.2.3.1) and automated face analysis (AFA) systems (see Section 2.2.4.1) pose serious risks to informational privacy since they do not require the awareness or cooperation of individuals involved. The same applies to situations where AFA systems make use of digital images uploaded on the Internet, e.g. on social media, as such processing may occur without any involvement or awareness by the individuals concerned.[1478]

This is not only a theoretical risk, as the Clearview AI case clearly underscores. The company Clearview AI Inc. collected, by means of web scraping techniques, images and relevant metadata available online and further processed such biometric data in its AFA system. The Italian supervisory authority imposed a fine on the company for the violation of several provisions of the GDPR.[1479] AFA systems may not only be deployed in public or semi-public spaces. Volvo plans to install on-board cameras in their cars that can be used for identifying the driver based on face recognition systems described in Section 2.2.3.1 to automatically set climate control and seating position according to the preferences of the driver.[1480] Surveillance systems in public spaces may identify individuals participating in

[1474] In-store tracking of shoppers that are being identified based on their observable characteristics such as height, colour, width as described in a patent of 7-Eleven Inc. <https://patents.google.com/patent/US11107226B2/en> accessed 8 February 2024.
[1475] Vasilios Mavroudis, Michael Veale 'Eavesdropping Whilst You're Shopping: Balancing Personalisation and Privacy in Connected Retail Spaces' (Living in the Internet of Things Conference, London, March 2018) 4 <https://ieeexplore.ieee.org/document/8379705> accessed 8 February 2024.
[1476] See Section 2.2.3 (CV).
[1477] Chiara Giordano, 'Chinese police use surveillance technology to identify people by their walking style' *The Independent* (London 26 February 2019) <https://www.independent.co.uk/news/world/asia/china-police-walking-gait-technology-surveillance-ai-suspect-a8797836.html> accessed 8 February 2024.
[1478] Council of Europe, Consultative Committee of Convention 108, 'Guidelines on Facial Recognition' (28 January 2021) at 3 <https://rm.coe.int/guidelines-on-facial-recognition/1680a134f3> accessed 8 February 2024.
[1479] See <https://edpb.europa.eu/news/national-news/2022/facial-recognition-italian-sa-fines-clearview-ai-eur-20-million_en> accessed 8 February 2024.
[1480] 'Volvo to install in-car cameras to watch over drivers' *CAR magazine* (London 20 March 2019) <https://www.carmagazine.co.uk/car-news/tech/volvo-driver-cameras/> accessed 8 February 2024.

political activities, for example, protests, by means of CV approaches such as gait recognition and/or facial recognition.

AC can be applied in various contexts such as recruitment, casinos, restaurants, retail, hospitality and call centres and is prone to violate an individual's right to informational privacy because it gains access to the emotions of the individual beyond their control. Candidates participating in video assessments that use software that analyses their emotions based on AC techniques can hardly determine themselves to what extent their emotional state is shared with the prospective employer. Likewise, users of virtual assistant Alexa cannot control whether Amazon will use technology that enables Alexa to recognise the user's emotional state derived from the user's voice[1481] and how such information is further processed. AC provides an unprecedented means to gain access to information related to the emotional state of individuals beyond their control. It seems difficult, if possible, at all, for individuals to determine themselves whether they want, in fact, to provide access to such information.

The control problem (Type 1)

The AI disciplines discussed in Chapter 2 undermine the right to informational privacy because individuals can hardly determine to reveal certain information or not. AI can infer such information anyway, beyond the individual's control, and therefore violates the right to informational privacy.

5.2.2 Legal problems: Type 2

No specific Type 2 legal problems arise when the AI disciplines introduced in Chapter 2 are applied to informational privacy. The fundamental right to privacy has been extensively enforced, which is underscored by the wealth of case law produced by both the ECtHR and the CJEU regarding the fundamental right to privacy. According to the HUDOC database maintained by the ECtHR, at least 12,323 cases dealt with the fundamental right to privacy within the last ten years.[1482] There are no indications that the enforcement of the fundamental right to privacy will decrease in the future because of AI.

5.2.3 Legal problems: Type 3

The broad scope of the fundamental right to privacy, along with the ECtHR's refusal to define the ambit of it,[1483] enabled the ECtHR to continuously respond to modern legal dilemmas and human

[1481] Huafeng Jin, Shuo Wang 'Voice-Based Determination of Physical and Emotional Characteristics of Users' US Patent Number US 10096319 B1 (Assignee: Amazon Technologies, Inc.) October 2018 <https://patentimages.storage.googleapis.com/f6/a2/36/d99e36720ad953/US10096319.pdf> accessed 8 February 2024.
[1482] See HUDOC database accessed 8 February 2024.
[1483] Frederik Zuiderveen Borgesius, 'Improving Privacy Protection in the area of Behavioural Targeting' (Doctoral thesis, Universiteit van Amsterdam 2015) 100 <https://dare.uva.nl/search?identifier=c74bdba6-616c-4cd9-925e-33a5858935e5> accessed 8 February 2024.

rights challenges[1484] and adapt the protection to new circumstances and technological and societal developments.[1485] As introduced in Section 3.1.2, this dynamic approach to interpretation has been coined the 'living instrument doctrine'.[1486] This ensures that the fundamental right to privacy is interpreted and applied in the light of present-day conditions, thus considering, inter alia, technological developments and the issues to which these may raise.[1487] The living instrument doctrine also affects case law adopted by the CJEU.[1488] The ECtHR stressed that it will consider the extent to which 'intrusions into private life are made possible by new, more and more sophisticated technologies'[1489]. In my view, this statement addresses the technological developments facilitated by AI perfectly. Therefore, no specific Type 3 legal problems arise when AI is applied to the fundamental right to privacy.

5.3 Bodily privacy

Bodily privacy relates to the right to keep body functions and characteristics (e.g., genetic codes and biometrics) private. It specifically relates to the integrity of a person's body[1490] and physical access to it, but also encompasses the restriction and control of information about the body.[1491] Whereas traditional examples such as compulsory immunisation or blood transfusion without consent[1492] include physical and unsolicited harms to the body,[1493] examples in the context of AI shift the focus to information that is gained from a person's body and its functions without physically intruding the body, such as accessing the body by means of devices, for example, wearables that measure physiological signals. In this context, it is important to consider the distinction between informational and bodily privacy. Bodily privacy refers to access to the human body, and informational privacy relates to the observations that can be made by analysing the information gained from the human body. In other words, bodily privacy concerns the protection of the actual object of privacy which can be directly intruded, i.e. the body, and informational privacy concerns the protection of information that may be obtained by analysing the body, but not the body itself.[1494]

[1484] David Harris et al, *Law of the European Convention on Human Rights* (4th edn OUP 2018) 569, 570.
[1485] Frederik Zuiderveen Borgesius, 'Improving Privacy Protection in the area of Behavioural Targeting' (Doctoral thesis, Universiteit van Amsterdam 2015) 100 <https://dare.uva.nl/search?identifier=c74bdba6-616c-4cd9-925e-33a5858935e5> accessed 8 February 2024.
[1486] Alastair Mowbray, 'The Creativity of the European Court of Human Rights' (2005) Vol 5 Iss 1 Human Rights Law Review 57-59.
[1487] David Harris et al, *Law of the European Convention on Human Rights* (4th edn OUP 2018) 509.
[1488] Case C-400/10, *J. McB.* [2010] ECR I-582 para 53. See also Article 52 (3) EUCFR which states that the 'meaning and scope' of the rights contained in the EUCFR and ECHR shall be the same, provided that these rights 'correspond'. This holds true for Article 8 ECHR and Article 7 EUCFR.
[1489] *Köpke v Germany*, App No 420/07 (EctHR 05 October 2010) emphasis added.
[1490] Rachel L. Finn, David Wright, Michael Firedewald, 'Seven Types of Privacy' in S. Gutwirth et al. (eds.), *European Data Protection: Coming of Age* (Springer 2013) 7, 8.
[1491] Bert-Jaap Koops et al 'A Typology of Privacy' (2017) Vol. 38 Iss. 2 University of Pennsylvania Journal of International Law 438, 569.
[1492] Roger Clarke, 'What's Privacy' (Roger Clarke's Website, 7 August 2008) <http://www.rogerclarke.com/DV/Privacy.html> accessed 8 February 2024.
[1493] Bert-Jaap Koops et al 'A Typology of Privacy' (2017) Vol. 38 Iss. 2 University of Pennsylvania Journal of International Law 438, 498.
[1494] Bert-Jaap Koops et al 'A Typology of Privacy' (2017) Vol. 38 Iss. 2 University of Pennsylvania Journal of International Law 438, 555.

5.3.1 Legal problems: Type 1

Two AI disciplines are particularly relevant in the context of bodily privacy, namely, ML and AC. In what follows, I outline why these two AI disciplines cause legal problems regarding the right to bodily privacy.

Brain-computer interfaces (BCI), which are often powered by ML and DL,[1495] impact bodily privacy because they monitor physiological signals. Non-invasive BCIs, which are currently most widely used in BCI research, place sensors on the scalp to acquire electroencephalography (EEG) signals.[1496] EEG measures electrical impulses emitted by the brain.[1497] Companies develop consumer-directed wearable devices to record brain activity based on EEGs, leading to the analysis of information concerning brain activity on a large scale.[1498] BCI applications use different ML techniques for the classification of EEG signals.[1499] Neuroadaptive technologies combine AI with implantable BCIs which automatically adapt to the user's mindset without requiring explicit instructions.[1500] The company Neuralink develops a BCI system that aims to establish a direct link between the brain and everyday technology. The system records neural activity in the brain and as the user thinks about moving her arms or hands, the system decodes those intentions by means of ML and DL approaches. At a first stage, this technology is intended for individuals with paralysis and neurological disorders to regain independence by giving them the ability to control computers and mobile devices directly with their brains. Later, Neuralink intends to discover new, non-medical applications and make them available to the general population.[1501] This BCI system relies upon a small, wireless, battery-powered neural implant unseen from the outside of the body.[1502] Neuralink has already successfully implanted the device in the brains of a monkey and a pig. The company published a video showing the monkey that had been implanted with the neural device playing the video game Pong using only its mind.[1503] These approaches are invasive and physically access the body and therefore impact the physical integrity of the individuals concerned.

[1495] Mamunir Rashid et al, 'The classification of EEG Signal Using Different Machine Learning Techniques for BCI Application' in J.-H. Kim et al (Eds) *Robot Intelligence Technology and Applications* (Springer 2018) 207-221.
[1496] Hongchang Shan, 'Towards High Performance and Efficient Brain Computer Interface Character Speller: Convolutional Neural Network based Methods' Dissertation Universiteit Leiden 2020, 2.
[1497] Rachel L. Finn, David Wright, Michael Firedewald, 'Seven Types of Privacy' in S. Gutwirth et al. (eds.), *European Data Protection: Coming of Age* (Springer 2013) 7, 8.
[1498] Philipp Kellermayr, 'Big Neurodata: On the Responsible Use of Neurodata from Clinical and Consumer-Directed Neurotechnological Devices' (2018) 1, 2 <https://link.springer.com/article/10.1007/s12152-018-9371-x> accessed 8 February 2024.
[1499] Mamunir Rashid et al, 'The classification of EEG Signal Using Different Machine Learning Techniques for BCI Application' in J.-H. Kim et al (Eds) *Robot Intelligence Technology and Applications* (Springer 2018) 207-221.
[1500] Marcello Ienca, Gianclaudio Malgieri, 'Mental data protection and the GDPR' (2022) Vol 9 Iss 1 Journal of Law and the Biosciences 1, 4.
[1501] See the company's website <https://web.archive.org/web/20230331035227/https://neuralink.com/applications/> accessed 8 February 2024.
[1502] Emily Waltz, 'Elon Musk Announces Nuralink Advance Toward Syncing Our Brains With AI' *IEEE* (New York 28 August 2020) <https://spectrum.ieee.org/elon-musk-neuralink-advance-brains-ai> accessed 8 February 2024.
[1503] Rupert Neate, 'Elon Musk's brain chip firm Neuralink lines up clinical trials in humans' *The Guardian* (London 20 January 2022) <https://www.theguardian.com/technology/2022/jan/20/elon-musk-brain-chip-firm-neuralink-lines-up-clinical-trials-in-humans> accessed 8 February 2024.

Because measurable physiological changes such as changes in heart rate, galvanic skin response, muscle tension, breathing rate and electrical activity in the brain co-occur with emotions, AC technologies sense these changes and recognise emotion by detecting patterns that capture physiological responses.[1504] For example, a statistically significant increase in heart rate could be linked to the activation of the sympathetic nervous system, arguably due to the occurrence of anxiety.[1505] Therefore, AC is particularly relevant for bodily privacy. Because physiological signals cannot easily be controlled[1506] and are involuntary, they are considered to constitute a reliable method for emotion recognition.[1507] Wearables facilitate the monitoring of physiological signals in unprecedented ways and are therefore particularly suitable for emotion recognition. Such devices have the ability to detect signals from skin conductivity, skin temperature, heart rate and other emotion-related physiological parameters.[1508] Combined with ML approaches such as regression as explained in Section 2.2.1.1, wearables provide powerful means to develop emotion recognition systems.[1509] Emotion recognition systems based on physiological signals using wearables may monitor such signals in an unobtrusive manner.[1510] Research deploying ML approaches achieved high accuracy in detecting amusement and sadness by relying on an instrumented glove developed to acquire galvanic skin response signals and information about heart rate.[1511] Admittedly, the collection of bodily information through wearables and BCI as such is already problematic concerning bodily privacy. However, AI allows for inferences of bodily functions based on mere observations of the body. In this sense, AI can invade bodily integrity without touching the human body.

The two AI disciplines AC and ML (particularly DL) are highly dependent on physiological signals and body functions. In the case of body implants, physical access to the body is gained, which consequently violates the integrity of an individual's body. ML and AC technologies sense physiological signals by non-invasive means (e.g., wearables) and thus gain indirect access to the body through devices that measure physiological signals. Because bodily privacy encompasses the restriction and control of information about the body, non-invasive means also violate the right to bodily privacy since they monitor physiological signals such as changes in heart rate, galvanic skin response,

[1504] Jennifer Healey, 'Physiological Sensing of Emotion' in Rafael Calvo et al (eds), *The Oxford Handbook of Affective Computing* (OUP 2015) 204.
[1505] Francisco Lupiáñez-Villanueva et al, 'Behavioural study on unfair commercial practices in the digital environment: dark patterns and manipulative personalisation' (2022) 98 <https://op.europa.eu/o/opportal-service/download-handler?identifier=606365bc-d58b-11ec-a95f-01aa75ed71a1&format=pdf&language=en&productionSystem=cellar&part=> accessed 8 February 2024.
[1506] Lin Shu et al, 'A Review of Emotion Recognition Using Physiological Signals' (2018) Vol 18 Iss 7 Sensors 2.
[1507] Juan Antonio Domínguez-Jiménez, 'A machine learning model for emotion recognition from physiological signals' (2020) Vol 55 Biomedical Signal Processing and Control 1.
[1508] Lin Shu et al, 'A Review of Emotion Recognition Using Physiological Signals' (2018) Vol 18 Iss 7 Sensors 32.
[1509] Değer Ayata, Yusuf Yaslan, Mustafa Kamasak, 'Emotion Recognition from Multimodal Physiological Signals for Emotion Aware Healthcare Systems' (2020) Vol 40 149-157.
[1510] Juan Antonio Domínguez-Jiménez et al, 'A machine learning model for emotion recognition from physiological signals' (2020) Vol 55 Biomedical Signal Processing and Control 1.
[1511] Ibid 1, 3.

breathing rate and electrical activity in the brain. AI systems deploy approaches in ML and DL in particular to make use of information derived from the human body and its functions.

The bodily information problem (Type 1)

ML, DL and AC are highly dependent on bodily information, including its functions, by gaining physical access to the body (e.g., implants) or by non-invasive means, e.g. wearables sensing physiological signals such galvanic skin response, and electrical activity in the brain. These technologies violate the right to bodily privacy, as they invade bodily integrity by allowing for inferences of bodily functions based on observed data, either by intervening with an individual's right to keep bodily functions and characteristics private or by gaining physical access to the body.

5.3.2 Legal problems: Type 2

No specific Type 2 legal problems arise when the AI disciplines introduced in Chapter 2 are applied to bodily privacy for the same reasons as outlined in Section 5.2.2. The fundamental right to privacy has been extensively enforced in the past ten years, [1512] and there are no indications that the enforcement of the fundamental right to privacy will decrease in the future due to AI.

5.3.3 Legal problems: Type 3

No specific Type 3 legal problems arise when AI is applied to the right to bodily privacy for the same reasons as outlined in Section 5.3.2. The broad scope of the fundamental right to privacy and the 'living instrument doctrine'[1513] ensure that this right keeps up with technological developments,[1514] including AI.

5.4 Mental privacy

Mental privacy refers to controlling access to the mind and thus to information about mental processes and states.[1515] As such, mental privacy has not yet been recognised as a specific element falling under the notion of private life as enshrined in the fundamental right to privacy. However, the right to mental privacy may be derived from existing ECtHR case law on the right to privacy, in particular from the notions *psychological*[1516] and *moral integrity*[1517] covered therein. According to ECtHR case law, the

[1512] See HUDOC database accessed 8 February 2024.
[1513] Alastair Mowbray, 'The Creativity of the European Court of Human Rights' (2005) Vol 5 Iss 1 Human Rights Law Review 57-59.
[1514] David Harris et al, *Law of the European Convention on Human Rights* (4th edn OUP 2018) 509.
[1515] Abel Wajnerman Paz, 'Is Mental Privacy a Component of Personal Identity?' (2021) Vol 15 Fronties in Human Neuroscience 2.
[1516] *Botta v Italy* App no 21439/93 (EctHR 24 February 1998) para 32, *Pretty v United Kingdom* App no 2346/02 (EctHR 29 April 2002) para 61; *Tysiąc v Poland* App no 5410/03 (EctHR 24 September 2007) para 107.
[1517] *Gladysheva v Russia* App no 7097/10 (EctHR 6 December 2011) para 93, *Bensaid v United Kingdom* App no 44599/98 (EctHR 6 February 2001) para 47.

term ‘private life’ also encompasses a person’s psychological[1518] and moral integrity.[1519] In its jurisprudence, the ECtHR did not define *moral integrity*, but this term seems to be related to both dignity and freedom from coercion with respect to choices with respect to one’s own decisions or as a sense of non-invasion by outside influences.[1520] The ECtHR regards mental health as a crucial part of private life associated with the aspect of moral integrity.[1521] Neither ECtHR case law nor scholarship thoroughly examine the term *‘psychological integrity’* in the context of the right to privacy.[1522] However, harm to reputation also constitutes harm to psychological integrity,[1523] or the suffering from maltreatment without physical marks such as deprival of sleep.[1524] Thus, psychological integrity does not necessitate the suffering from mental disorders in a clinical-pathological sense.[1525] Although ‘moral’ and ‘psychological’ integrity may have slightly diverging meanings, there are no indications that they fall outside the remit of the right to privacy considering that the ECtHR repeatedly emphasised the broad interpretation of private life.[1526] Therefore, it seems likely that a right to mental privacy could be derived from or at least developed within the ECtHR’ future jurisprudence with respect to the fundamental right to privacy and particularly the notion of private life.[1527] It seems plausible that the fundamental right to privacy protects mental privacy[1528] because this fundamental right is well equipped to cover all conceivable mental privacy interests that should enjoy legal protection.[1529] This holds particularly true when considering the ECtHR’s living instrument doctrine as explained in Section 3.1.2 which requires one to apply the right to privacy in the light of present-day conditions, taking into account, inter alia, technological developments and the issues these may raise.

Mental privacy has never been considered thoroughly because, traditionally, the mind has not been conceived as an entity vulnerable to external intrusions and therefore in need of legal protection.[1530]

[1518] *Botta v Italy* App no 21439/93 (EctHR 24 February 1998) para 32, *Pretty v United Kingdom* App no 2346/02 (EctHR 29 April 2002) para 61; *Tysiąc v Poland* App no 5410/03 (EctHR 24 September 2007) para 107.
[1519] *Gladysheva v Russia* App no 7097/10 (EctHR 6 December 2011) para 93, *Bensaid v United Kingdom* App no 44599/98 (EctHR 6 February 2001) para 47.
[1520] Jill Marshall, *Personal Freedom through Human Rights Law? Autonomy, Identity and Integrity under the European Convention on Human Rights* (Martinus Nijhoff Publishers 2009) 168, 184.
[1521] *Bensaid v United Kingdom* App no 44599/98 (EctHR 6 February 2001) para 47; *Dolenec v Croatia* App no 25282/06 (EctHR 26 November 2009) para 165.
[1522] For an overview concerning relevant literature, see Footnote 57 on page 397 in Jan-Christoph Bublitz, ‘The Nascent Right to Psychological Integrity and Mental Self-Determination’ in Andreas van Arnauld, Kerstin von der Decken, Mart Susi (eds.), *The Cambridge Handbook of New Human Rights* (Cambridge University Press 2020).
[1523] *Kyriakides v Cyprus* App no 39058/05 (ECtHR 16 October 2008); *A. v Norway* App no 28070/06 (ECtHR 9 April 2009); *Axel Springer v Germany* App no 39954/08 (ECtHR 7 February 2012).
[1524] *Bati and others v Turkey* App nos 33097/96 and 57834/00 (ECtHR 3 June 2004).
[1525] Jan-Christoph Bublitz, ‘The Nascent Right to Psychological Integrity and Mental Self-Determination’ in Andreas van Arnauld, Kerstin von der Decken, Mart Susi (eds.), *The Cambridge Handbook of New Human Rights* (Cambridge University Press 2020) 396.
[1526] Ibid 395, 396.
[1527] Sjors Ligthart et al, ‘Forensic Brain-Reading and Mental Privacy in European Human Rights Law: Foundations and Challenges’ (2021) Vol 14 Neuroethics 191, 200 <https://link.springer.com/content/pdf/10.1007/s12152-020-09438-4.pdf> accessed 8 February 2024.
[1528] Thomas Douglas, Lisa Forsberg, ‘Three Rationales for a Legal Right to Mental Integrity’ in: Sjors Ligthart et al (eds) *Neurolaw Palgrave Studies in Law, Neuroscience, and Human Behavior* (Palgrave Macmillan 2021) 184.
[1529] Sjors Ligthart, ‘Freedom of thought in Europe: do advances in ‘brain-reading’ technology call for revision?’ (2020) Vol 7 Iss 1 Journal of law and the biosciences 4.
[1530] Jan Christoph Bublitz, Reinhard Merkel, ‘Crimes against Minds: On Mental Manipulations, Harms and a Human Right to Mental Self-Determination’ (2014) Vol 8 Iss 1 Criminal Law and Philosophy 51, 61; Sjors Ligthart, ‘Freedom of

For the purpose of this thesis, I interpret mental privacy broadly referring to information related to all conscious and non-conscious mental representations, events, processes and propositional attitudes, including thoughts, beliefs, emotions and moods, as well as the underlying psychological mechanisms ('mental privacy').[1531] Given the broad scope of mental privacy, it also comprises privacy of thoughts and feelings, which refers to the right of individuals not to share their feelings and thoughts or to have them revealed. This type of privacy emphasises that individuals should be able to think or feel whatever they like.[1532]

In addition, the meaning of thought must be interpreted broadly to include emotional states because research demonstrates that emotion and cognition are interrelated phenomena and that good decision-making seems to require emotional capacities.[1533] Such a broad interpretation is also in line with case law adopted by the ECtHR regarding the freedom of thought enshrined in Article 9 ECHR, which interprets this notion broadly considering the comprehensiveness of the concept of thought.[1534] However, this right is a neglected human right[1535] and has never played a decisive role in legal practice which is why its scope and meaning remain vague.[1536] Also, it is arguable that the freedom of thought protected by Article 9 ECHR relates much more to the freedom of religion and conscience than thoughts per se. It is beyond of the scope of this thesis to elaborate on this in more detail, but freedom of thought might become more relevant in the future and even provide stronger legal protection for thoughts[1537] than the fundamental right to privacy because it does not allow any interference given its absolute character.[1538]

While the body may easily be subject to domination and control by others, mental states have until recently been beyond external constraints.[1539] Advances in AI and neuroscience are changing

thought in Europe: do advances in 'brain-reading' technology call for revision?' (2020) Vol 7 Iss 1 Journal of law and the biosciences 2.

[1531] Jan-Christoph Bublitz, 'The Nascent Right to Psychological Integrity and Mental Self-Determination' in Andreas von Arnauld, Kerstin von der Decken, Mart Susi (eds) *The Cambridge Handbook of New Human Rights* (Cambridge University Press 2020) 30; Marcello Ienca, Gianclaudio Malgieri, 'Mental data protection and the GDPR' (2022) Vol 9 Iss 1 Journal of Law and the Biosciences 1, 4.

[1532] Rachel L. Finn, David Wright, Michael Firedewald, 'Seven Types of Privacy' in S. Gutwirth et al. (eds.), *European Data Protection: Coming of Age* (Springer 2013) 19.

[1533] Jan Christoph Bublitz, Reinhard Merkel, 'Crimes against Minds: On Mental Manipulations, Harms and a Human Right to Mental Self-Determination' (2014) Vol 8 Iss 1 Criminal Law and Philosophy 51, 64.

[1534] *Salonen v Finland* App no 27868/95 (ECtHR 2 July 1997).

[1535] Sjors Ligthart, 'Freedom of thought in Europe: do advances in 'brain-reading' technology call for revision?' (2020) Vol 7 Iss 1 Journal of law and the biosciences 3.

[1536] Jan Christoph Bublitz, Reinhard Merkel, 'Crimes against Minds: On Mental Manipulations, Harms and a Human Right to Mental Self-Determination' (2014) Vol 8 Iss 1 Criminal Law and Philosophy 51; Leonard M Hammer, *The international human right to freedom of conscience: some suggestions for its development and application* (Ashgate 2001).

[1537] Sjors Ligthart et al, 'Forensic Brain-Reading and Mental Privacy in European Human Rights Law: Foundations and Challenges' (2021) Vol 14 Neuroethics 191, 200 <https://link.springer.com/content/pdf/10.1007/s12152-020-09438-4.pdf> accessed 8 February 2024.

[1538] Article 9 (1) EUCHR which does not allow for any interferences, as opposed to the right to the right to privacy according to Article 8 EUCHR.

[1539] Marcello Ienca, Roberto Andorno, 'Towards new human rights in the age of neuroscience and neurotechnology. Life Sciences, Society and Policy' (2017) Vol 13 Life Sciences, Society and Policy 1 <https://lsspjournal.biomedcentral.com/articles/10.1186/s40504-017-0050-1> accessed 8 February 2024.

traditional boundaries of the mind and yield information from the brain that enables to draw inferences about particular mental states of individuals and thus to some extent enable 'brain-reading'.[1540] Although the mind and mental states were insusceptible or irresistible to interference in the past, this seems no longer to be the case[1541] considering the progress in AI and neuroscience, in particular involving the AI disciplines ML (especially DL), CV as well as NLP and AC.

5.4.1 Legal problems: Type 1

Developments in AI raise legal problems regarding mental privacy, especially concerning the interpretation of neural activity patterns aiming to determine what an individual is thinking.[1542] These concerns partially overlap with the legal problems with respect to the processing of neurodata and mental data as discussed in Section 4.8.3. Brain-computer interfaces (BCIs) translate brain signals into computer commands and enable the communication between the human brain and devices.[1543] Such BCIs are often powered by ML and DL approaches.[1544] Measuring an individual's brain activity by means of electroencephalography (EEG) or functional magnetic resonance imaging (fMRI) in the form of BCI systems deploy ML and DL approaches and facilitate the drawing of inferences about particular mental properties, such as a person's emotions and memory.[1545]

Notably, the developments in neuro-AI may circumvent the cognitive process of filtering and selectively sharing information that humans typically perform to control the flow of information about them (e.g. thoughts and feelings). Thus, information that humans have considered and decided not to share may become available to entities[1546] anyway by interpreting neural activity and decoding it in order to determine those individual's thoughts, powered by ML and DL approaches as well as feature extraction techniques from the AI discipline CV[1547] that adaptively decode neurodata.[1548] Researchers have achieved to translate brain activity into text by means of ML and ANN approaches.[1549]

[1540] Sjors Ligthart, 'Freedom of thought in Europe: do advances in 'brain-reading' technology call for revision?' (2020) Vol 7 Iss 1 Journal of law and the biosciences 1, 2.
[1541] Thomas Douglas, Lisa Forsberg, 'Three Rationales for a Legal Right to Mental Integrity' in: Sjors Ligthart et al (eds) *Neurolaw Palgrave Studies in Law, Neuroscience, and Human Behavior* (Palgrave Macmillan 2021) 194.
[1542] Abel Wajnerman Paz, 'Is Mental Privacy a Component of Personal Identity?' (2021) Vol 15 Fronties in Human Neuroscience 2.
[1543] Hongchang Shan, 'Towards High Performance and Efficient Brain Computer Interface Character Speller: Convolutional Neural Network based Methods' Dissertation Universiteit Leiden 2020, 1.
[1544] Mamunir Rashid et al, 'The classification of EEG Signal Using Different Machine Learning Techniques for BCI Application' in J.-H. Kim et al (Eds) *Robot Intelligence Technology and Applications* (Springer 2018) 207-221.
[1545] Sjors Ligthart, 'Freedom of thought in Europe: do advances in 'brain-reading' technology call for revision?' (2020) Vol 7 Iss 1 Journal of law and the biosciences 1, 2.
[1546] Abel Wajnerman Paz, 'Is Mental Privacy a Component of Personal Identity?' (2021) Vol 15 Fronties in Human Neuroscience 2.
[1547] Mark Nixon, Alberto Aguado, *Feature Extraction & Image Processing for Computer Vision* (3rd edn Elsevier 2012).
[1548] Stephen Rainey et al, 'Is the European Data Protection Regulation sufficient to deal with emerging data concerns relating to neurotechnology?' (2020) Journal of Law and the Biosciences 3 < https://academic.oup.com/jlb/advance-article/doi/10.1093/jlb/lsaa051/5864051?searchresult=1> accessed 8 February 2024.
[1549] Joseph G Makin, David A Moses, Edward F Chang, 'Machine translation of cortial activity to text with an encoder-decoder framework' (2020) Vol 23 Nature Neuroscience 575.

Developments in neurotechnology, powered by ML and DL approaches, have partially unlocked the human brain and made it readable under scientific lenses.[1550]

Whereas current applications of ML, DL and ANN in the context of neuroscience are being used in restricted scientific settings, such approaches will be used in a broader context in the future. It does not seem unlikely that the upcoming decades will see neurotechnology becoming pervasive and embedded in numerous aspects of human lives. Take, for example, a BCI system that records neural activity in the brain, and as the user thinks about moving an arm or a hand, the system decodes those intentions by means of ML and DL approaches. Whereas this system is initially intended to be used in a medical context, the provider of the system announced that it intends to discover new, non-medical applications allowing to control computers directly with the brain and make them available to the general population.[1551] In fact, there are already commercial brain-reading devices available to consumers,[1552] such as EEG sensor headsets for gaming, self-monitoring and entertainment.[1553] The right to privacy protects individuals from unwanted intrusions into their private lives, including intrusions into processes that occur solely inside one's brain, [1554] for instance thoughts that are not being communicated to others. These developments can violate mental privacy simply because they provide access to mental processes and states themselves as well as further information about mental states and information derived thereof.[1555] In addition to the infringement of mental privacy caused by the mere access to mental states and processes themselves (and information inferred thereof), the fact that individuals are unable to control access to mental processes and states violates mental privacy. Individuals are deprived of the opportunity to not share their feelings and thoughts or disclose them. Consequently, individuals are also unable to think or feel whatever they like.[1556]

Additionally, such approaches in AI may become increasingly effective in modulating the neural correlates of human psychology and behaviour.[1557] Neurotools such as BCIs allow interventions into

[1550] Marcello Ienca, Roberto Andorno, 'Towards new human rights in the age of neuroscience and neurotechnology. Life Sciences, Society and Policy' (2017) Vol 13 Life Sciences, Society and Policy 5 <https://lsspjournal.biomedcentral.com/articles/10.1186/s40504-017-0050-1> accessed 8 February 2024.
[1551] See the company's website <https://web.archive.org/web/20230331035227/https://neuralink.com/applications/> accessed 8 February 2024.
[1552] Sjors Ligthart, 'Freedom of thought in Europe: do advances in 'brain-reading' technology call for revision?' (2020) Vol 7 Iss 1 Journal of law and the biosciences 3.
[1553] Marcello Ienca, Pim Haselager, Ezekiel J Emanuel, 'Brain Leaks and Consumer Technology' (2018) Vol 36 Iss 9 Nature Biotechnology 805-815.
[1554] Jill Marshall, *Personal Freedom through Human Rights Law? Autonomy, Identity and Integrity under the European Convention on Human Rights* (Martinus Nijhoff Publishers 2009) 3.
[1555] It might be argued that access to mental states and processes in fact refers to informational privacy as described in Section 5.2. However, I do see mental states and processes themselves as the source and thus object worthy of protection. Information about mental states such as concrete thoughts might then be protected under both mental and informational privacy.
[1556] Rachel L. Finn, David Wright, Michael Firedewald, 'Seven Types of Privacy' in S. Gutwirth et al. (eds.), *European Data Protection: Coming of Age* (Springer 2013) 19.
[1557] Marcello Ienca, Roberto Andorno, 'Towards new human rights in the age of neuroscience and neurotechnology. Life Sciences, Society and Policy' (2017) Vol 13 Life Sciences, Society and Policy 5 <https://lsspjournal.biomedcentral.com/articles/10.1186/s40504-017-0050-1> accessed 8 February 2024.

minds changing desires and beliefs without inflicting pain, harming bodily integrity or the need to indoctrinate persons over extended periods of time.[1558] Neuroenhancement, closed-loop brain interventions with on-chip ML[1559] and digital influences of the brain, such as nudging and other persuasive concepts, may allow interventions into minds and thus change desires and beliefs.[1560] Nudges are 'interventions that steer people in particular directions but that also allows them to go their own way'.[1561] Information about mental states and processes, including thoughts, provides powerful means to exhibit external influences, such as manipulation of individuals and their decision-making processes. As outlined in Section 4.3.3, manipulation perverts the way a person reaches decisions, forms preferences or adopts goals.[1562] It has been argued that case law does not provide hints as to whether mind-interventions such as manipulation of decision-making, fall within the ambit of mental integrity.[1563] In my view, manipulations violate what the ECtHR considers to constitute moral integrity.[1564] The latter covers non-invasion by outside influences.[1565] Therefore, I take the view that such manipulations may violate moral integrity which forms part of the broad concept of private life as elaborated by the ECtHR.

The AI discipline AC[1566] aims to detect emotional states and thus raises legal problems regarding mental privacy because it arguably renders emotional states and emotions machine-readable. AC violates mental privacy, simply because it detects and discloses emotions, moods and feelings of individuals.[1567] Systems that deploy AC and NLP approaches affecting mental privacy include automated border control systems aimed at detecting whether an individual lies, virtual assistants that detect the user's emotional state, video-based job assessments and wristbands that tell managers whether employees are unhappy.[1568] A notably EU funded automated border control system called

[1558] Jan Christoph Bublitz, Reinhard Merkel, 'Crimes against Minds: On Mental Manipulations, Harms and a Human Right to Mental Self-Determination' (2014) Vol 8 Iss 1 Criminal Law and Philosophy 51, 61.
[1559] See Bingzhao Zhu, Uisub Shin, Masha Shoaran, 'Closed-Loop Neural Prostheses with On-Chip Intelligence: A Review and A Low-Latency Machine Learning Model for Brain State Detection' (2021) <https://www.epfl.ch/labs/inl/wp-content/uploads/2021/09/2109.058482.pdf> accessed 8 February 2024.
[1560] Sjors Ligthart, 'Freedom of thought in Europe: do advances in 'brain-reading' technology call for revision?' (2020) Vol 7 Iss 1 Journal of law and the biosciences 2.
[1561] Cass R Sunstein, 'The Ethics of Nudging' (2015) Vol 32 Yale Journal of Regulation 413, 417.
[1562] Joseph Raz, *The Morality of Freedom* (OUP 1986) 377; Cass R Sunstein, 'The Ethics of Nudging' (2015) Vol 32 Yale Journal of Regulation 413, 444.
[1563] Jan-Christoph Bublitz, 'The Nascent Right to Psychological Integrity and Mental Self-Determination' in Andreas van Arnauld, Kerstin von der Decken, Mart Susi (eds.), *The Cambridge Handbook of New Human Rights* (Cambridge University Press 2020) 397.
[1564] *Gladysheva v Russia* App no 7097/10 (ECtHR 6 December 2011) para 93, *Bensaid v United Kingdom* App no 44599/98 (ECtHR 6 February 2001) para 47.
[1565] Jill Marshall, *Personal Freedom through Human Rights Law? Autonomy, Identity and Integrity under the European Convention on Human Rights* (Martinus Nijhoff Publishers 2009) 168, 184.
[1566] See the approaches in AC as discussed in Chapter 2.2.4.
[1567] It might be argued that access to emotions in fact refers to informational privacy as described in Section 5.2. However, given that emotions constitute srather sensitive information, I take the view that emotional states constitute a new object worthy of its own dedicated protection in the context of the right to privacy, namely under mental privacy.
[1568] For the latter, see Suzanne Bearne, 'A wristband that tells your boss if you are unhappy' *BBC* (London, 18 January 2021) <https://www-bbc-com.cdn.ampproject.org/c/s/www.bbc.com/news/amp/business-55637328> accessed 31 January 2021.

IBORDERCTRL 'analyses the micro-gestures of travellers to figure out if the interviewee is lying'.[1569] HireVue video interview software claims to be able to evaluate a candidate's employability, including personality traits, in under 30 minutes[1570] by means of on-demand video interviews where job candidates record responses to structured interview questions.[1571] The software detects and analyses the emotions a candidate portrays during the video assessment[1572] based on AC and AFA components. Amazon patented technology that enables its virtual assistant Alexa to recognise the users emotional state derived from the user's voice[1573] by combining AC and NLP approaches. Thus, systems that incorporate the discipline AC, sometimes[1574] combined with NLP, provide access to the emotional states and feelings of individuals.

The mere access to this sensitive information violates mental privacy. Furthermore, access to emotional states and feelings of individuals occurs beyond the control of the individuals concerned. AC deprives individuals of the opportunity not to share their feelings and emotional states because these disciplines may detect such information by non-invasive means anyway, such as by analysing facial expressions, gestures, physiological sensors and speech when combined with NLP. Individuals are also unable to feel whatever they like[1575] considering that their emotional states and feelings may be detected by non-invasive means and beyond their control. By means of revealing the emotional states of individuals, AC provides the necessary information needed to effectively manipulate decision-making of individuals, which arguably violates what the ECtHR considers to constitute moral integrity[1576] aiming to protect from undue external influences.[1577] Emotions play an important role in the elicitation of autonomous motivated behaviour.[1578] According to research in behavioural sciences, especially psychology, emotions constitute powerful, pervasive and predictable drivers of decision-making.[1579] Emotions can have significant effects on economic transactions and play a powerful role

[1569] European Commission, 'Smart lie-detection system to tighten EU's busy borders' (24 October 2018) < https://ec.europa.eu/research-and-innovation/en/projects/success-stories/all/smart-lie-detection-system-tighten-eus-busy-borders > accessed 8 February 2024.
[1570] Nathan Mondragon, Clemens Aicholzer, Kiki Leutner, 'The Next Generation of Assessments' (HireVue 2019) <http://hrlens.org/wp-content/uploads/2019/11/The-Next-Generation-of-Assessments-HireVue-White-Paper.pdf> accessed 8 February 2024.
[1571] See <https://www.hirevue.com/demo> accessed 8 February 2024.
[1572] Nathan Mondragon, Clemens Aicholzer, Kiki Leutner, 'The Next Generation of Assessments' (HireVue 2019) <http://hrlens.org/wp-content/uploads/2019/11/The-Next-Generation-of-Assessments-HireVue-White-Paper.pdf> accessed 8 February 2024.
[1573] Huafeng Jin, Shuo Wang 'Voice-Based Determination of Physical and Emotional Characteristics of Users' US Patent Number US 10096319 B1 (Assignee: Amayon Technologies, Inc.) October 2018 <https://patentimages.storage.googleapis.com/f6/a2/36/d99e36720ad953/US10096319.pdf> accessed 8 February 2024.
[1574] When emotional states are derived from speech.
[1575] Rachel L. Finn, David Wright, Michael Firedewald, 'Seven Types of Privacy' in S. Gutwirth et al. (eds.), *European Data Protection: Coming of Age* (Springer 2013) 19.
[1576] *Gladysheva v Russia* App no 7097/10 (ECtHR 6 December 2011) para 93, *Bensaid v United Kingdom* App no 44599/98 (ECtHR 6 February 2001) para 47.
[1577] Jill Marshall, *Personal Freedom through Human Rights Law? Autonomy, Identity and Integrity under the European Convention on Human Rights* (Martinus Nijhoff Publishers 2009) 168, 184.
[1578] Leen Vandercammen et al, 'On the Role of Specific Emotions in Autonomous and Controlled Behaviour' (2014) Vol 28 Iss 5 European Journal of Personality 413, 445.
[1579] Jennifer S Lerner et al, 'Emotion and Decision Making' (2015) Vol 66 Annual Review of Psychology 799, 802.

in everyday economic choices.[1580] The powerful insights that AC provides can be used to influence individuals, e.g. emotional states and feelings, which can violate mental privacy.

Although the mind and mental states were insusceptible or irresistible to interference, this seems no longer to be the case when considering the developments in AI, in particular BCI systems powered by ML, DL and CV,[1581] as well as approaches from the AI discipline AC (alone or combined with NLP).[1582] These AI disciplines may violate mental privacy in a yet unknown and unprecedented manner which constitutes a Type 1 legal problem simply because they enable mere access to mental states themselves as well as information that might be inferred or derived thereof. Furthermore, these disciplines violate mental privacy because individuals cannot control access to mental states and are deprived of the opportunity to not share such information. Consequently, individuals are also unable to think or feel whatever they like.[1583] Additionally, these developments in AI become increasingly relevant for the purpose of manipulating individuals, which arguably violates what, according to the ECtHR, constitutes moral integrity.[1584]

The mental information problem (Type 1)

Except for AR, all AI disciplines introduced in Chapter 2 facilitate access to mental states and information that might be inferred or derived thereof. Consequently, mental states and related information are no longer insusceptible or irresistible to interference. The AI disciplines ML, CV, NLP and AC are therefore prone to violate mental privacy.

5.4.2 Legal problems: Type 2

No specific Type 2 legal problems arise when the AI disciplines introduced in Chapter 2 are applied to mental privacy for the same reasons as outlined in Section 5.2.2. The fundamental right to privacy has been extensively enforced in the past ten years, [1585] and there are no indications that the enforcement of the fundamental right to privacy will decrease in the future due to AI.

[1580] Jennifer S Lerner, Deborah A Small, George Loewenstein, 'Heart Strings and Purse Strings' (2004) Vol 15 No 5 American Psychology Society 337-340.
[1581] Mark Nixon, Alberto Aguado, *Feature Extraction & Image Processing for Computer Vision* (3rd edn Elsevier 2012).
[1582] Thomas Douglas, Lisa Forsberg, 'Three Rationales for a Legal Right to Mental Integrity' in: Sjors Ligthart et al (eds) *Neurolaw Palgrave Studies in Law, Neuroscience, and Human Behavior* (Palgrave Macmillan 2021) 194.
[1583] Rachel L. Finn, David Wright, Michael Firedewald, 'Seven Types of Privacy' in S. Gutwirth et al. (eds.), *European Data Protection: Coming of Age* (Springer 2013) 19.
[1584] *Gladysheva v Russia* App no 7097/10 (ECtHR 6 December 2011) para 93, *Bensaid v United Kingdom* App no 44599/98 (ECtHR 6 February 2001) para 47.
[1585] See HUDOC database accessed 8 February 2024.

5.4.3 Legal problems: Type 3

If the broad interpretation of mental privacy[1586] in the context of AI does not hold true[1587] and the ECHR will not recognise mental privacy in such a broad way, it must be concluded that there is a Type 3 legal problem. In this case, the fundamental right to privacy, and particularly the broad concept of private life, is not fit for purpose to protect mental privacy. Nevertheless, and as outlined in Section 5.4, the broad scope of the fundamental right to privacy and the 'living instrument doctrine'[1588] are well equipped to ensure that the fundamental right to privacy keeps up with technological developments.[1589] I am therefore confident that the fundamental right to privacy will recognise and protect mental privacy considering the developments facilitated by AI that enable access to mental information.

Moreover, some interferences with mental privacy caused by AI may simultaneously also infringe bodily privacy (see Section 5.3.1). This might be the case with AC that detects emotions based on physiological signals or ML and DL approaches that use neuro implants to record neural activity and decode the intentions and thoughts of the individual concerned.

5.5 Communicational privacy

The right to communicational privacy as part of the fundamental right to privacy aims to avoid unsolicited interception of communication. Typical violations include eavesdropping or intercepting communication,[1590] including mere access to stored communication.[1591] Communication is to be understood broadly and includes telephone and wireless communication, as well as mail and email and, in line with the living instrument doctrine, future means of communication. Possible infringements also entail the interception of communication by means of bugs, microphones or other sensors.[1592] Communicational privacy is typified by an individual's interest in restricting access to communications or controlling the use of information communicated to third parties.[1593] According to ECtHR case law,

[1586] Sjors Ligthart et al, 'Forensic Brain-Reading and Mental Privacy in European Human Rights Law: Foundations and Challenges' (2021) Vol 14 Neuroethics 191, 200 <https://link.springer.com/content/pdf/10.1007/s12152-020-09438-4.pdf> accessed 8 February 2024; Abel Wajnerman Paz, 'Is Your Neural Data Part of Your Mind? Exploring the Conceptual Basis of Mental Privacy' (2022) Vol 32 Minds and Machines 395, 399.
[1587] For instance, Ienca and Andorno, which argue that the right to privacy is insufficient to protect mental privacy. See Marcello Ienca, Roberto Andorno, 'Towards new human rights in the age of neuroscience and neurotechnology. Life Sciences, Society and Policy' (2017) Vol 13 Life Sciences, Society and Policy 15 <https://lsspjournal.biomedcentral.com/articles/10.1186/s40504-017-0050-1> accessed 8 February 2024.
[1588] Alastair Mowbray, 'The Creativity of the European Court of Human Rights' (2005) Vol 5 Iss 1 Human Rights Law Review 57-59.
[1589] David Harris et al, *Law of the European Convention on Human Rights* (4th edn OUP 2018) 509.
[1590] Roger Clarke, 'What's Privacy?' (2006) <http://www.rogerclarke.com/DV/Intro.html > accessed 8 February 2024.
[1591] Rachel L. Finn, David Wright, Michael Firedewald, 'Seven Types of Privacy' in S. Gutwirth et al. (eds.), *European Data Protection: Coming of Age* (Springer 2013) 8.
[1592] Rachel L. Finn, David Wright, Michael Firedewald, 'Seven Types of Privacy' in S. Gutwirth et al. (eds.), *European Data Protection: Coming of Age* (Springer 2013) 8.
[1593] Bert-Jaap Koops et al 'A Typology of Privacy' (2017) Vol. 38 Iss. 2 University of Pennsylvania Journal of International Law 438, 567.

the form and content of the communication are irrelevant to the question of interference.[1594] Moreover, the right to privacy aims to protect the confidentiality of communication in a wide range of situations and technologies. The living instrument doctrine explained in Section 3.1.2 enables the fundamental right to privacy to keep up with technological developments. This doctrine is also helpful for new methods of communication,[1595] arguably including methods involving AI. For example, I take the view that human-machine communication occurring in the context of virtual assistants and similar services as discussed in Section 4.9.3 is protected by the right to communicational privacy. The ECtHR anticipated new technological developments and emphasised that it will consider the extent to which 'intrusions into private life are made possible by new, more and more sophisticated technologies'[1596] (see also Section 5.5.3).

5.5.1 Legal problems: Type 1

Three AI disciplines are particularly relevant regarding communicational privacy. Speech-based emotion recognition systems that combine approaches from the AI disciplines ML and AC rely on the processing of personal communication, speech signals in particular. NLP requires the analysis of communication because it concerns the understanding and generation of natural language.

Approaches that implement AC and ML measure and quantify the emotions of individuals by observing the speech signals of these individuals. Supervised ML algorithms are at the heart of many emotion recognition efforts[1597] and methods applied to emotion recognition from speech also involve DL approaches.[1598] As explained in Section 2.2.4.2, effects of emotion tend to be present in acoustic signal features such as average pitch, pitch range and pitch changes, speech rate and articulation.[1599] ML maps the input, namely, the automatically derived acoustic features, to emotion labels that represent the characteristics for a given emotion category.[1600] For example, the detected acoustic feature of a high speech rate is typically associated with the emotional state of anger or fear.[1601] Virtual assistants as introduced in Section 4.9.1 deploy AC approaches to detect a user's emotional state, which allows them to modify their behaviour accordingly.[1602]

[1594] *A. v France* App no 14838/89 (ECtHR 23 November 1993) paras 35-37; *Frérot v France* App no 70204/01 (ECtHR 12 June 2007) para 54.
[1595] David Harris et al, *Law of the European Convention on Human Rights* (4th edn OUP 2018) 509.
[1596] *Köpke v Germany*, App No 420/07 (ECtHR 05 October 2010) emphasis added.
[1597] Chi-Chun Lee et al, 'Speech in Affective Computing' in Rafael Calvo et al (eds), *The Oxford Handbook of Affective Computing* (OUP 2015) 177.
[1598] Haytham M Fayek, Margaret Lech, Lawrence Cavedon, 'Evaluating deep learning architectures for Speech Emotion Recognition' (2017) Vol 92 Neural Networks 60.
[1599] Rosalind W Picard, *Affective Computing* (MIT Press 1997) 179, 180.
[1600] Chi-Chun Lee et al, 'Speech in Affective Computing' in Rafael Calvo et al (eds), *The Oxford Handbook of Affective Computing* (OUP 2015) 173, 177.
[1601] Rosalind W Picard, *Affective Computing* (MIT Press 1997) 179.
[1602] Giorgo Manfredi, Claudio Gribaudo, Virtual Assistant with real-time emotions, WIPO Patent WO 2008/049834 A2, Laurence Goasduff 'Emotion AI Will Personalize Interactions' (Gartner, 22 January 2018) <https://www.gartner.com/smarterwithgartner/emotion-ai-will-personalize-interactions/> accessed 8 February 2024.

For example, Amazon has been granted a patent for 'Voice-based determination of physical and emotional characteristics of users'. According to this patent, the system may detect emotions such as happiness, joy, anger, sorrow, sadness, fear, disgust, boredom and other emotional states based on analysis of acoustic features such as pitch or speech rate, as determined from the processing of the voice data. The patent specifically refers to the AI disciplines NLP and ML, including ANN approaches.[1603] Following the claims of this patent, virtual assistant Alexa is able to detect a user's emotional or physical state. This enables Alexa to intuitively suggest specific products based on the user's current emotional state or offer medicine if it detects a cough when a user makes a request.[1604] Spotify patented a virtual assistant that improves the way a machine processes and generates a response to a human's emotion based on an utterance (human vocalisation) from a user containing both a command and an emotion.[1605] The virtual assistant is designed for a 'media playback device' and can recognise when a user sounds sad and is able to offer encouragement by 'cheering' the user up.[1606] Apart from sadness, other detectable emotions enlisted in the patent are surprise, anger, fear, anxiety, disgust and joy.[1607] As mentioned in Section 2.2.4.1, these six 'basic emotions'[1608] are the most common ones used in emotion research.[1609] According to the patent, emotions are derived from a variety of cues associated with user's utterance. In the case of a command, such cues may be the tone, cadence, volume, pitch and pace of the user's speech.[1610] These cues resemble the acoustic signal features typically used in speech-based emotion recognition systems as outlined in Section 2.2.4.2. They are often related to prosody which considers the intonational and rhythmic aspects of language.[1611] Typical examples are pitch and energy of speech,[1612] including voice level and speech rate.[1613] Where the user's utterance

[1603] Huafeng Jin, Shuo Wang, 'Voice-based Determination of Physical and Emotional Characteristics of Users' US Patent Number US 10096319B1 (Assignee: Amazon Technologies, Inc.) October 2018 <https://patentimages.storage.googleapis.com/f6/a2/36/d99e36720ad953/US10096319.pdf>, accessed 8 February 2024.
[1604] James Cook, 'Amazon patents new Alexa feature that knows when you're ill and offers you medicine' *The Telegraph* (London 9 October 2018) < https://www.telegraph.co.uk/technology/2018/10/09/amazon-patents-new-alexa-feature-knows-offers-medicine/> accessed 8 February 2024.
[1605] Daniel Bromand et al, 'Systems and Methods for Enhancing Responsiveness to Utterances Having Detectable Emotion' US Patent Number US 10566010 B2 (Assignee: Spotify AB) February 2020 11 < https://patentimages.storage.googleapis.com/2a/9d/2d/926b58a2bd956f/US10566010.pdf >, accessed 8 February 2024.
[1606] Josh Mandell, 'Spotify Patents A Voice Assistant That Can Read Your Emotions' *Forbes* (New York, 12 March 2020) < https://www.forbes.com/sites/joshmandell/2020/03/12/spotify-patents-a-voice-assistant--that-can-read-your-emotions/ > accessed 8 February 2024.
[1607] Daniel Bromand et al, 'Systems and Methods for Enhancing Responsiveness to Utterances Having Detectable Emotion' US Patent Number US 10566010 B2 (Assignee: Spotify AB) February 2020 12 < https://patentimages.storage.googleapis.com/2a/9d/2d/926b58a2bd956f/US10566010.pdf >, accessed 8 February 2024.
[1608] These six emotions refer to research conducted by psychologists in the early seventies that developed the methodology of 'basic emotions'; see Paul Ekman, Wallace v Friesen, 'Constants across cultures in the face and emotion' (1971) Vol 17 (2) Journal of Personality and Social Psychology 124.
[1609] Lisa Feldman Barrett et al. 'Emotional Expressions Reconsidered' (2019) Vol 20 (1) Psychological Science in the Public Interest 1, 4.
[1610] Daniel Bromand et al, 'Systems and Methods for Enhancing Responsiveness to Utterances Having Detectable Emotion' US Patent Number US 10566010 B2 (Assignee: Spotify AB) February 2020 12 < https://patentimages.storage.googleapis.com/2a/9d/2d/926b58a2bd956f/US10566010.pdf >, accessed 8 February 2024.
[1611] Daniel Jurafsky, James H Martin, *Speech and Language Processing* (2 edn, Pearson Education Limited 2014) 238.
[1612] Ricardo A. Calix, Leili Javadpour, Gerald M. Knapp, 'Detection of Affective States From Text and Speech For Real-Time Human-Computer Interaction' (2012) Vol 54 No 4 Human Factors and Ergonomics Society 530, 531.
[1613] Christina Sobn and Murray Alpert, 'Emotion in Speech: The Acoustic Attributes of Fear, Anger, Sandess, and Joy' (1999) Vol 28 No 4 Journal of Psycholinguistic Research, 347.

contains no words from which a command can be extracted, (e.g. 'Ugh'), emotions are derived from the tone of this utterance.[1614]

Virtual assistants are not the only domain in which speech emotion recognition systems could be implemented. Speech emotion recognition may be used in various areas, such as call centres, smart devices or cars.[1615] In fact, they are already used in practice. A real-world application of AC aiming to derive emotional states from speech is Amazon's wearable 'Halo' that analyses voice tones to detect user emotions.[1616] A Hungarian bank used an AI system with the aim to detect and measure emotions of customers that called the bank's customer service.[1617] In order to identify customer dissatisfaction, the AI system deployed by the bank relied on acoustic signal features introduced in Section 2.2.4.2, namely, speed, volume and pitch of speech.[1618]

Speech-based emotion recognition systems combine approaches from the AI disciplines ML, AC and NLP. Because such systems are highly dependent on speech analysis, they violate communicational privacy. Speech falls under the term 'communication' according to the right to communicational privacy: the form and content of the communication is irrelevant to the question of interference.[1619] Individuals concerned cannot control the further use of such communication and might not even be aware of the fact that communication is analysed to detect their emotional state, let alone be aware of what information can be derived from analysing speech. Speech-based emotion recognition systems therefore violate communicational privacy, which leads to a Type 1 legal problem.

The speech analysis problem (Type 1)

By combining approaches from ML, AC and NLP, speech-based emotion recognition systems are highly dependent on the processing of communication (speech) to detect the emotional states of the individual concerned. These systems intercept, analyse and otherwise process communications in various contexts, including virtual assistants, call centres and cars. Individuals cannot control the further use of such communication. This violates communicational privacy.

[1614] Daniel Bromand et al, 'Systems and Methods for Enhancing Responsiveness to Utterances Having Detectable Emotion' US Patent Number US 10566010 B2 (Assignee: Spotify AB) February 2020 13 < https://patentimages.storage.googleapis.com/2a/9d/2d/926b58a2bd956f/US10566010.pdf >, accessed 8 February 2024.
[1615] See services of the company audeering: https://www.audeering.com/.
[1616] Alex Hern, 'Amazon's Halo wristband: the fitness tracker that listens to your mood' *The Guardian* (London, 28 August 2020) < https://www.theguardian.com/technology/2020/aug/28/amazons-halo-wristband-the-fitness-tracker-that-listens-to-your-mood > accessed 8 February 2024; Austin Carr, 'Amazon's New Wearable Will Know If I'm Angry. Is That Weird?' *Bloomberg* (New York, 31 August 2020) < https://www.bloomberg.com/news/newsletters/2020-08-31/amazon-s-halo-wearable-can-read-emotions-is-that-too-weird > accessed 8 February 2024.
[1617] Sebastião Barros Vale, Gabriela Zanfir-Fortuna, 'Automated Decision-Making Under the GDPR: Practical Cases from Courts and Data Protection Authorities' (2022) 48 <https://fpf.org/wp-content/uploads/2022/05/FPF-ADM-Report-R2-singles.pdf> accessed 8 February 2024.
[1618] Cesar Manso-Sayao, Summary of Hungarian SA Decision NAIH-85-3/2022 < https://gdprhub.eu/NAIH_(Hungary)_-_NAIH-85-3/2022 > accessed 8 February 2024.
[1619] *A. v France* App no 14838/89 (ECtHR 23 November 1993) paras 35-37; *Frérot v France* App no 70204/01 (ECtHR 12 June 2007) para 54.

NLP develops novel practical applications to facilitate the interactions between computers and humans,[1620] including the generation and understanding of natural language.[1621] Developments in the discipline NLP have led to the integration of AI technologies in daily life. Nowadays, individuals routinely communicate with virtual assistants[1622] such as Alexa, Siri or Google Assistant. Such interactions are expected to increase even more in the future.[1623] For example, car manufacturers already offer in-vehicle virtual assistants.[1624] A study concerning Amazon Alexa's ecosystem revealed that a user's activities can be reconstructed due to the large amount of data with timestamps.[1625]

Most of the virtual assistant's processing occurs on a remote server and every transaction and recording is kept by the company which provides the service.[1626] Contrary to what was claimed in the terms, a study revealed that Amazon Alexa records speech even if the wake word is not spoken: 91% of the study participants had instances of unintended voice recordings, i.e. recordings occurring without mentioning the wake word. Study participants reported that such unintended recordings contained sensitive conversations.[1627] This means that users do not have complete control over what is recorded, transmitted and stored in the cloud environment of the virtual assistant's provider.[1628] Unintended recordings may contain sensitive recordings of speech[1629] given the broad range of applications of virtual assistants, which are used at home, in cars and at any given location in case the virtual assistant service is used on a mobile phone. A whistle-blower who used to work for Apple revealed that he had listened to hundreds of recordings every day, often including unintentional recordings, for quality control purposes ('grading of Apple's virtual assistant'). According to the whistle-blower, these recordings concerned sensitive communications such as discussions between doctors and patients, business deals, seemingly criminal acts and sexual encounters.[1630] Such recordings are also interesting for law enforcement agencies.[1631]

[1620] Deng Li and Liu Yang, 'A Joint Introduction to Natural Language Processing and Deep Learning' in Deng Li and Liu Yang (eds) *Deep learning in natural language processing* (Springer 2018) 1.
[1621] Stan Franklin, 'History, motivations, and core themes' in Frankish Keith and Ramsey William M. (eds) *The Cambridge Handbook of Artificial Intelligence* (2014) 26.
[1622] Refer to Section 4.9.3 to learn more how virtual assistants work.
[1623] Andrea L Guzman, Seth C Lewis, 'Artificial intelligence and communication: A Human-Machine Communication agenda' (2020) Vol 22 Iss 1 New Media & Society 70, 71.
[1624] For instance, 'Hey Mercedes', which is able to understand different accents and will adjust to the driver over time; see <https://www.mercedes-benz.co.uk/passengercars/mercedes-benz-cars/models/eqc/comfort.pi.html/mercedes-benz-cars/models/eqc/comfort/standard-equipment/mbux> accessed 8 February 2024.
[1625] Hyunji Chung, Jungheum Park, Sangjin Lee, 'Digital forensic approaches for Amazon Alexa ecosystem' (2017) Vol 22 Digital Investigation 15, 18.
[1626] Tom Bolton et al, 'On the Security and Privacy Challenges of Virtual Assistants' (2021) Vol 21 Iss 7 Sensors 1-2.
[1627] Yousra Javed, Shashank Sethi, Akshay Jadoun, 'Alexa's Voice Recording Behavior: A Survey of User Understanding and Awareness' (ARES '19, Canterbury 26-29 August 2019) 7 <https://dl.acm.org/doi/10.1145/3339252.3340330> accessed 8 February 2024.
[1628] Hyunji Chung et al, 'Alexa, Can I Trust You?' (2017) Vol 50 Iss 9 Computer 100, 103.
[1629] Yousra Javed, Shashank Sethi, Akshay Jadoun, 'Alexa's Voice Recording Behavior: A Survey of User Understanding and Awareness' (ARES '19, Canterbury 26-29 August 2019) 2 <https://dl.acm.org/doi/10.1145/3339252.3340330> accessed 8 February 2024.
[1630] Alex Hern, 'Apple contractors regularly hear confidential details on Siri recordings' *The Guardian* (London, 26 July 2019) < https://www.theguardian.com/technology/2019/jul/26/apple-contractors-regularly-hear-confidential-details-on-siri-recordings > accessed 8 February 2024.
[1631] Hyunji Chung, Jungheum Park, Sangjin Lee, 'Digital forensic approaches for Amazon Alexa ecosystem' (2017) Vol 22 Digital Investigation 2.

Generating and understanding natural language in NLP requires analysis of personal communication. Virtual assistants such as Siri or Alexa unintentionally intercept and record personal communication from their users and other people such as relatives, children and friends. Developments in NLP provide means to listen to private communications and also facilitate the identification of the individual who is speaking. For example, Microsoft's Speaker Recognition Application Programming Interface (SAPI)[1632] allows one to identify individual speakers within a group and can be easily deployed.[1633] Because they are highly dependent on the processing of communication, these NLP empowered systems violate communicational privacy. This affects the confidentiality of communication in a wide range of situations and technologies regardless of the form and content of the communication.[1634] This constitutes a Type 1 legal problem.

The interception and identification problem (Type 1)

Generating and understanding natural language in NLP requires the processing of communication. Virtual assistants unintendedly intercept and record personal communication of their users and other individuals such as relatives, children, and friends. Developments in NLP such as Speaker Recognition APIs facilitate the identification of individual speakers within a group. This violates communicational privacy.

Keyword determination systems are based on the AI discipline NLP. They are highly problematic in the context of communication privacy. Such systems aim to detect keywords from recorded speech and use them for targeted advertising. Users suspected their smartphones to be secretly eavesdropping on them, and many reports[1635] have claimed that private conversations occurring in the presence of smartphones consequently resulted in targeted online advertisements. Advertisements referred to in these reports relate to a broad range of product categories matching either an overall discussion topic or a specific brand or product mentioned in a preceding face-to-face conversation.[1636] For example, 20 employees of the research and advisory firm Forrester reported that some of their 'real-life' conversations seemingly resulted in ads and sponsored posts on Facebook without having searched for the item advertised after the conversations took place.[1637]

[1632] A set of functions and procedures allowing the creation of applications that access features or data of an operating system, application or other service; see <https://www.dictionary.com/browse/api> accessed 8 February 2024.
[1633] <https://azure.microsoft.com/en-us/services/cognitive-services/speaker-recognition/> accessed 8 February 2024.
[1634] *A. v France* App no 14838/89 (ECtHR 23 November 1993) paras 35-37; *Frérot v France* App no 70204/01 (ECtHR 12 June 2007) para 54.
[1635] Jacob Leon Kröger, Philip Raschke, Is My Phone Listening in? On the Feasibility and Detectability of Mobile Eavesdropping' in: Simon N Foley (eds) *Data and Applications security and Privacy XXXIII* (Springer 2019) 102, 103.
[1636] Ibid.
[1637] Fatemeh Khatibloo, 'Is Facebook Listening (And So What If They Are)?' *Forbes* (New York, 17 March 2017) <https://www.forbes.com/sites/forrester/2017/03/17/is-facebook-listening-and-so-what-if-they-are/> accessed 8 February 2024.

Some consider the fear that private companies could target their ads based on eavesdropped conversation as baseless and paranoid.[1638] For example, a former product manager of Facebook stated that alleged eavesdropping would be economically and technically unfeasible, referring to CPU,[1639] battery and data storage limitations.[1640] The technological and economic feasibility argument has been rebutted in research, however.[1641] Smartphone-based eavesdropping can be deployed efficiently and scalable by means of keyword detection instead of full speech recognition. Keyword detection only recognises a predefined vocabulary of spoken words and runs on devices with much lower computational power than smartphones. It allows one to search for trigger words indicating a person's interest, such as 'love' or 'enjoy', to identify relevant sections of a private conversation instead of searching for millions or perhaps billions of targetable keywords.[1642]

Amazon's US patent 'Keyword Determinations from Voice Data'[1643] indicates that the technology for such advertisements is already available. The patent, which relies on NLP, describes a system that captures voice content when a user speaks into or near the device (e.g., Alexa), notably without activating the virtual assistant by mentioning the 'wake word' (e.g., 'hey Alexa'). Sniffer algorithms identify trigger words that indicate statements of preference (such as 'like' or 'love') and translate them into keywords. The identified keywords are subsequently transmitted to a location accessible to advertisers, who then use the keywords to select content that is likely relevant to the user.[1644] Amazon has denied that it uses voice recordings for advertising at the moment and claimed that the patent might never actually come to the market.[1645] This statement seems to be contradictory to a journalist's report that suspects Amazon to have listened to a private conversation between herself and her husband. The conversation involved a very specific kitchen gadget. She suspects that Alexa snooped into the conversation, as she has subsequently received an ad for that kitchen gadget on Amazon.[1646] When considering the capabilities of Amazon's keyword determination system, this does not seem to be an unrealistic or far-fetched claim. The Amazon patent clearly shows that the technical

[1638] Jacob Leon Kröger, Philip Raschke, Is My Phone Listening in? On the Feasibility and Detectability of Mobile Eavesdropping' in: Simon N Foley (eds) *Data and Applications security and Privacy XXXIII* (Springer 2019) 103.
[1639] Central Processing Unit (CPU), sometimes also called main processor, constitutes the physical heart of the entire computer system and is generally composed of the main memory, control unit, and arithmetic-logic unit; see <https://www.britannica.com/technology/central-processing-unit> accessed 8 February 2024.
[1640] Antonio García Martínez, 'Facebook's Not Listening Through Your Phone. It Doesn't Have To' *Wire* (New York, 18 November 2017) <https://www.wired.com/story/facebooks-listening-smartphone-microphone/> accessed 8 February 2024.
[1641] Jacob Leon Kröger, Philip Raschke, Is My Phone Listening in? On the Feasibility and Detectability of Mobile Eavesdropping' in: Simon N Foley (eds) *Data and Applications security and Privacy XXXIII* (Springer 2019) 112.
[1642] Ibid.
[1643] Edara Kiran, 'Key Word Determinations From Voice Data' US Patent Number US 8798995B1 (Assignee: Amazon Technologies, Inc.) August 2014 <https://patentimages.storage.googleapis.com/bd/ed/2b/c4c67cc5a9f1ab/US8798995.pdf>, accessed 8 February 2024.
[1644] Ibid.
[1645] Griffin Andrew, 'Amazon files for Alexa patent to let it listen to people all the time and work out what they want' *The Independent* (London, 11 April 2018) <https://www.independent.co.uk/life-style/gadgets-and-tech/news/amazon-alexa-patent-listening-to-me-facebook-phone-talking-ads-a8300246.html> accessed 8 February 2024.
[1646] Morgan Blake, 'Are Digital Assistants Always Listening?' *Forbes* (New York, 5 February 2018) <https://www.forbes.com/sites/blakemorgan/2018/02/05/are-digital-assistants-always-listening/#2f000e1a4eeb> accessed 8 February 2024.

means for such eavesdropping are available and could be used for targeted advertisement. A marketing team within media giant Cox Media Group claims it can listen to ambient conversations of consumers through embedded microphones in smartphones, smart TVs, and other devices to gather data and use it to serve targeted ads.[1647] Hence, advances in NLP such as keyword determination systems for targeted advertising may violate communicational privacy because they are designed to intercept and analyse communication with the aim of subsequently using the information for targeted advertising. This violates communicational privacy and constitutes a Type 1 legal problem.

The keyword problem (Type 1)

Keyword determination systems powered by approaches in NLP identify trigger words that indicate statements of preference (such as 'like' or 'love') from recorded speech and translate these into keywords. These keywords are then used by advertisers to select content that is likely relevant to the user. Such systems intercept and analyse communications, which violates the right to communicational privacy.

5.5.2 Legal problems: Type 2

No specific Type 2 legal problems arise when the AI disciplines introduced in Chapter 2 are applied to communicational privacy for the same reasons as outlined in Section 5.2.2. The fundamental right to privacy has been extensively enforced in the past ten years,[1648] and there are no indications that the enforcement of the fundamental right to privacy will decrease in the future due to AI.

5.5.3 Legal problems: Type 3

As already discussed in Sections 5.5 and 4.9.3, the developments in AI require protection of human-machine communication under the remit of communicational privacy. Historically, communication has been conceptualised as a human process potentially mediated by technology.[1649] Case law of the ECtHR refers to the historic conception of communication, i.e. communication between humans. Therefore, it might be argued that human-machine communications, such as between the user and its virtual assistant, do not neatly fall within the scope of communicational privacy. However, I do not think such an argument is valid. First, the ECtHR stressed that it will consider the extent to which 'intrusions into private life are made possible by new, more and more sophisticated technologies'.[1650] Second, the living instrument doctrine as described in Section 3.1.2 proved to be very effective to address issues at the forefront of technology. Third, the ECtHR interprets the confidentiality of

[1647] Joseph Cox, 'Marketing Company Claims That It Actually Is Listening to Your Phone and Smart Speakers to Target Ads' *404 Media* (United States, 14 December 2023) <Marketing Company Claims That It Actually Is Listening to Your Phone and Smart Speakers to Target Ads (404media.co)> accessed 8 February 2024.
[1648] See HUDOC database accessed 8 February 2024.
[1649] Andrea L Guzman, Seth C Lewis, 'Artificial intelligence and communication: A Human-Machine Communication agenda' (2020) Vol 22 Iss 1 New Media & Society 70 -68.
[1650] *Köpke v Germany*, App No 420/07 (ECtHR 05 October 2010) emphasis added.

communication broadly, regardless of the *form* and *content* of the communication.[1651] Therefore, I take the view that communicational privacy as enshrined in the fundamental right to privacy not only covers communication between individuals, but also communication between humans and machines. Therefore, no Type 3 legal problems arise. However, if my broad interpretation of communicational privacy does not hold and the ECtHR will refrain from considering human to machine communications to fall under communicational privacy, it must be concluded that there is a Type 3 legal problem.

5.6 Access

In many cases, the right of access is the point of departure for the data subject in exercising control over his or her personal data. The right of access allows the data subject to verify the lawfulness[1652] of processing and enables the data subject to obtain, depending on the circumstances, the rectification, erasure or blocking of personal data by the controller.[1653] The right of access must be considered a conditio sine qua non for exercising other data subject rights and restrictions on or around this right cause a knock-on effect on the entire data protection law regime.[1654] The CJEU repeatedly stressed the importance of the right of access as a prerequisite to other data protection rights.[1655] Given the important role of the right of access, the analysis in this section will be more extensive than for other data subject rights.

The right of access is not an absolute right, which means that this right may be restricted. Indeed, the right of access may be restricted in to ways, namely, in line with the provisions contained in Article 23 GDPR and in accordance with Article 15 (4) GDPR. Both provisions refer to the rights and freedoms of others, which particularly encompasses *trade secrets* or IP rights, including copyrights protecting the software.[1656] Restrictions under Article 15 (4) GDPR differ from restrictions possible under Article 23 GDPR. Article 15 (4) *exclusively* applies to the right to obtain a copy of the personal data undergoing processing and allows restrictions on a *case-by-case* basis, whereas restrictions according to Article 23 GDPR need to be laid down in Member State or Union law. According to Custers and Hijne, both the tools used for data analysis (AI systems) and the resulting knowledge (output of the AI system) fall within the scope of IP, trade secrets or other rights of the controller deserving protection.[1657] This is particularly relevant when analysing the right of access in the light of AI. In Sections

[1651] *A. v France* App no 14838/89 (ECtHR 23 November 1993) paras 35-37; *Frérot v France* App no 70204/01 (ECtHR 12 June 2007) para 54.
[1652] Recital 63 GDPR.
[1653] Joined Cases C-141/12 & C-372/12, *YS, M* [2014] ECR I-2081 para 44.
[1654] Jef Ausloos, Michael Veale, René Mahieu, 'Getting Data Subject Rights Right' (2019) Vol 10 Iss 3 JIPITEC 283, 285.
[1655] Case C-579/21, *Pankki S* [2023] ECR I-501 paras 56-58; Case C-487/21, *F.F.* [2022] ECR I-1000 paras 34-35; Case C-434/16, *Nowak* [2017] ECR I-994 para 57; Joined Cases C-141/12 & C-372/12, *YS, M* [2014] ECR I-2081 para 44; Case C-553/07 *Rijkeboer* [2009] ECR I-03889, para 51.
[1656] Recital 63 GDPR.
[1657] Bart Custers, Anne-Sophie Heijne, 'The right of access in automated decision-making: The scope of article 15(1)(h) GDPR in theory and practice' (2022) Vol 46 Computer Law & Security Review 1, 10

5.6.1 – 5.6.3, I outline that the relationship between the trade secrets directive ('TSD')[1658] and the GDPR is particularly problematic.[1659]

The scope of protection of the TSD covers AI itself, including the technical method used to process and obtain information. This protection applies to all AI disciplines,, as introduced in Chapter 2. Trade secrets are broadly defined in the TSD. To qualify as a trade secret according to Article 2 TSD, the information must (i) be secret, (ii) have commercial value due to its secrecy and (iii) be subject to reasonable steps to keep it secret.

Requirement (i), i.e. secrecy, is already met when the information is not generally known among or readily accessible to persons within the circles that normally deal with the kind of information in question.[1660] Protection offered by the TSD might be sought for AI technology including the technical methods used to obtain and process information and thus algorithms, training data (created or selected) and methods to create and select training data and output data (for example, the detected emotional state of an individual).[1661] The TSD lists a diverse range of information that is protectable.[1662] According to Recital 2 TSD, trade secrets protect a wide range of know-how and business information. It comprises information such as business practices, information on or knowledge about customers, personal data inferred or predicted by controllers and personal data analytics itself.[1663] Recital 14 TSD specifically includes 'technological information' in the definition of trade secrets. Arguably, the definition of a trade secret is so broad to include nearly any data handled by a commercial entity, such as shopping habits and history of customers,[1664] information about a customer's behaviour (creditworthiness, lifestyle, reliability, etc.),[1665] customer lists and profiles,[1666] algorithms,[1667] personalised marketing plans (e.g. pricing) or forecasts about customer's future life based on probabilistic studies

[1658] Directive (EU) 2016/943 of the European Parliament and of the Council of 8 June 2016 on the protection of undisclosed know-how and business information (trade secrets) against their unlawful acquisition, use and disclosure [2016] OJ L157/1 (TSD).
[1659] Teresa Trallero Ocaña, *The Notion of Secrecy* (Nomos 2021) 312.
[1660] Recital 14 Trade Secrets Directive; Teresa Trallero Ocaña, *The Notion of Secrecy* (Nomos 2021) 307; Thomas Hoeren, 'The EU Directive on the Protection of Trade Secrets and its Relation to Current Provisions in Germany' (2018) Vol 9 Iss 2 JIPITEC 140 <https://www.jipitec.eu/issues/jipitec-9-2-2018/4725> accessed 8 February 2024.
[1661] Ana Nordberg, 'Trade secrets, big data and artificial intelligence innovation: a legal oxymoron?' in Jens Schovsbo, Timo Minssen and Thomas Riis (eds) *The Harmonization and Protection of Trade Secrets in the EU* (Edward Elgar Publishing 2020) 197, 201.
[1662] Rochelle Cooper Dreyfusss, Mireille van Eechoud 'Choice of law in EU trade secrecy cases' in Jens Schovsbo, Timo Minssen and Thomas Riis (eds) *The Harmonization and Protection of Trade Secrets in the EU* (Edward Elgar Publishing 2020) 177.
[1663] Claudio Malgieri, Giovanni Comandé, 'Why a Right to Legibility of Automated Decision-Making Exists in the General Data Protection Regulation' (2017) Vol 7 No 4 International Data Privacy Law 243, 262.
[1664] Inge Graef, Martin Husovec, Nadezhda Purtova 'Data Portability and Data Control: Lessons for an Emerging Concept in EU Law' (2018) Vol 19 Iss 6 German Law Journal 1359, 1381.
[1665] Gianclaudio Malgieri, 'Trade Secrets v. Personal Data: A Possible Solution for Balancing Rights' (2016) Vol 6 No 2 International Data Privacy Law 102, 113-114.
[1666] Teresa Trallero Ocaña, *The Notion of Secrecy* (Nomos 2021) 312; Nadezhda Prtova, 'Do property rights in personal data make ense after the Big Data turn?' (2017) Vol 10 No 2 Journal of Law & Economic Regulation 64, 71.
[1667] Guido Noto La Diega, 'Against the Dehumanisation of Decision-Making: Algorithmic Decisions at the Crossroads of Intellectual Property, Data Protection, and Freedom of Information' (2018) Iss 1 Vol 9 JIPITEC 3, 4, 26, 28.

(life expectancy, estimated advancements in career, etc.).[1668] In addition, information or knowledge does not necessarily need to be correct or complete in order to enjoy protection under the TSD.[1669]

Protected information or knowledge has commercial value according to requirement (ii), if its unlawful acquisition, use or disclosure is likely to harm the interests of the person lawfully controlling it, in the sense that it undermines that person's business or financial interests, strategic position or ability to compete.[1670] Under the TSD, commercial value includes both *potential* or *actual* value. The latter seems to indicate that individual data may also be eligible for protection and, therefore, the notion of commercial value should be interpreted broadly. It refers to any harm to the scientific and technical capacity as well as the economic interests of the trade secret holder resulting from the disclosure of (secret) information, including the ability to compete in a broad sense.[1671]

Criterion (iii), i.e. the trade secret holder taking 'reasonable steps' to keep the protected information secret, is arguably the most tangible for businesses to demonstrate. To satisfy this requirement, companies may adopt non-disclosure agreements, include clauses banning reverse engineering into their licencing agreements or limit the number of possible licences altogether to not undermine secrecy.[1672] The threshold for this requirement seems to be rather low. It does not require trade secret holders to conclude individual confidentiality agreements with each third party to whom the trade secret is conveyed. In the absence of explicit non-disclosure agreements, even an implied duty of confidence might be sufficient to meet criterion (iii), for example, between the employer and employee.[1673]

AI is particularly valuable for companies because it may be used to derive or infer data, such as statistical inferences about a multitude of subjects, a given arrangement of a list of information and technical information related to a product or process.[1674] Trade secrets are extensively used by most types of companies. A study conducted by the EU Intellectual Property Office (EUIPO) in 2017[1675] demonstrated that the use of trade secrets is higher than the use of patents by most types of company,

[1668] Gianclaudio Malgieri, 'Trade Secrets v. Personal Data: A Possible Solution for Balancing Rights' (2016) Vol 6 No 2 International Data Privacy Law 102, 113-114; Sandra Wachter, Brent Mittelstadt 'A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI' (2019) No 2 Columbia Business Law Review 495, 607.
[1669] Ana Nordberg, 'Trade secrets, big data and artificial intelligence innovation: a legal oxymoron?' in Jens Schovsbo, Timo Minssen and Thomas Riis (eds) *The Harmonization and Protection of Trade Secrets in the EU* (Edward Elgar Publishing 2020) 200.
[1670] Recital 14 TSD; Jens Schovsbo, 'The Directive on trade secrets and its background' in Jens Schovsbo, Timo Minssen and Thomas Riis (eds) *The Harmonization and Protection of Trade Secrets in the EU* (Edward Elgar Publishing 2020) 14.
[1671] Teresa Trallero Ocaña, *The Notion of Secrecy* (Nomos 2021) 311, 411.
[1672] Nazrin Huseinzade, 'Algorithm Transparency: How to Eat the Cake and Have it Too' *European Law Blog* (27 January 2021) <https://europeanlawblog.eu/2021/01/27/algorithm-transparency-how-to-eat-the-cake-and-have-it-too/> accessed 8 February 2024.
[1673] Teresa Trallero Ocaña, *The Notion of Secrecy* (Nomos 2021) 315, 412.
[1674] Ana Nordberg, 'Trade secrets, big data and artificial intelligence innovation: a legal oxymoron?' in Jens Schovsbo, Timo Minssen and Thomas Riis (eds) *The Harmonization and Protection of Trade Secrets in the EU* (Edward Elgar Publishing 2020) 197
[1675] EUIPO 'Protecting Innovation Through Trade Secrets and Patents: Determinants for European Union Firms' (2017) <https://euipo.europa.eu/ohimportal/en/web/observatory/news/-/action/view/3704420> accessed 8 February 2024.

in most economic sectors and in all Member States.[1676] A very high prevalence of trade secrets has been observed in the sectors of computer programming, consultancy and related services.[1677] AI systems and their underlying algorithms[1678] may undoubtedly[1679] fall under the broad term of trade secrets and are likely to be treated as such. As a result, these algorithms will rarely be disclosed to the public or individuals affected by it.[1680] In fact, most of the complex algorithms including the algorithms of Google or Facebook are proprietary and shielded as trade secrets while only a negligible minority of algorithms are open source.[1681] Amazon's recommendation system, the Instagram algorithm for publication diffusion and Google's search engine are among the most well-known examples of trade secrets.[1682]

The scope of protection of the TSD is broad and protects not only AI and its underlying algorithms, but also input data (including training data) and selection methods, as well as output data, which constitute personal data. This applies to *all* AI disciplines as introduced in Chapter 2 and thus allows restrictions of all data subject rights and principles introduced in Sections 3.3.4 and 3.3.3 provided that such restrictions comply with Article 23 GDPR. The wording contained in Article 23 GDPR concerns the ability of Member States to impose restrictions on data subject rights and principles by means of legislative measures and expressly refers to Union law. Thus, EU legislation may adopt, by legislative measures, any restriction on the rights and principles contained in the GDPR.[1683] In fact, the TSD constitutes such Union law and provides controllers with the possibility to restrict data subject rights, for example, the right of access, to protect their trade secrets. Such restrictions must respect both the fundamental right to data protection and trade secrets simultaneously.[1684] In addition, trade secrets may be protected by the right to property according to Article 17 EUCFR.[1685] AI may also be

[1676] This seems logical since the protected information under trade secrets is much broader than compared to patents where the patentability thresholds need to be met. Furthermore, there are no formal registration requirements as it is the case with IP laws.
[1677] EUIPO 'Protecting Innovation Through Trade Secrets and Patents: Determinants for European Union Firms' (2017) 8-9 37 <https://euipo.europa.eu/ohimportal/en/web/observatory/news/-/action/view/3704420> accessed 8 February 2024
[1678] Which arguably constitutes 'technological information' according to Recital 14 TSD.
[1679] Teresa Trallero Ocaña, *The Notion of Secrecy* (Nomos 2021) 409; Maja Brkan, Grégory Bonnet, 'Legal and Technical Feasibility of the GDPR's Quest for explanation of Algorithmic Decisions: of Black Boxes, White Boxes and Fata Morganas (2020) Vol 11 Iss 1 European Journal of Risk Regulation 18, 39.
[1680] Gintarė Surblytė-Namavičienė, *Competition and Regulation in the Data Economy* (Edward Elgar Publishing 2020) 243.
[1681] Nazrin Huseinzade, 'Algorithm Transparency: How to Eat the Cake and Have it Too' *European Law Blog* (27 January 2021) < https://europeanlawblog.eu/2021/01/27/algorithm-transparency-how-to-eat-the-cake-and-have-it-too/> accessed 8 February 2024.
[1682] Maja Brkan, Grégory Bonnet, 'Legal and Technical Feasibility of the GDPR's Quest for explanation of Algorithmic Decisions: of Black Boxes, White Boxes and Fata Morganas (2020) Vol 11 Iss 1 European Journal of Risk Regulation 18, 39.
[1683] Dominique Moore, Commentary of Article 23 in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020) 552.
[1684] Recital 34 TSD states that the TSD 'respects the fundamental rights….[ ]…notably the right to protection of personal data.... [ ]…while respecting business secrecy'.
[1685] Case C-1/11 *Interseroh Scrap* [2012] ECR I-194 para 43; Case T-189/14 *Deza* [2017] para 163.

protected by intellectual property rights[1686] such as patents or copyrights[1687] - alone or in combination with trade secrets.[1688] Because trade secrets are more widely used than IP rights[1689] and easier for companies to rely on,[1690] I focus on trade secrets.

5.6.1 Legal problems: Type 1

Article 15 (1) lit h GDPR grants the right to data subjects to receive 'meaningful information about the logic involved' in automated decision-making (ADM). It is not yet clear what 'meaningful information' and the 'logic involved' mean when put into practice. In the view of AG Pikamäe, information about the 'logic involved' particularly includes the factors taken into account in the decision-making process and their weighting at an aggregate level.[1691] As indicated in Section 4.4.1, I interpret meaningful information according to Article 15 (1) lit h GDPR as information that is useful and/or has practical value for data subjects to (i) become aware of processing relating to ADM, (ii) enforce their data subject rights and (iii) exercise control over the processing of their personal data. AG Pikamäe stresses that such information must be useful for data subjects, so they can challenge 'decisions' within the meaning of Article 22 (1) of the GDPR.[1692] This is also in line with the CJEU's focus on intelligibility regarding Article 12 (1) GDPR, which ensures that the data subject fully understands the information provided to it.[1693] According to AG Pitruzella, Article 12 (1) GDPR aims to allow the data subject to effectively exercise the right of access and other data subject rights.[1694] In view of the AG, information should be provided in a manner that enables the data subject to familiarise itself with it fully, easily and without difficulty. Controllers do not comply with Articles 12 (1) and 15 GDPR if they provide information in a way that makes it 'extremely difficult or burdensome' for the data subject to be acquainted with that information.[1695] The emphasis on intelligibility is further justified by CJEU case law relating to the right of access.[1696]

There is limited understanding of how each data point impacts an ML model used for ADM.[1697] This holds true in case of complex models based on DL and ANNs and the problems described in Section

[1686] Ana Nordberg, 'Trade secrets, big data and artificial intelligence innovation: a legal oxymoron?' in Jens Schovsbo, Timo Minssen and Thomas Riis (eds) *The Harmonization and Protection of Trade Secrets in the EU* (Edward Elgar Publishing 2020) 198.
[1687] Sandra Wachter, Brent Mittelstadt 'A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI' (2019) No 2 Columbia Business Law Review 495, 600-604.
[1688] Recital 2 TSD.
[1689] EUIPO 'Protecting Innovation Through Trade Secrets and Patents: Determinants for European Union Firms' (2017) <https://euipo.europa.eu/ohimportal/en/web/observatory/news/-/action/view/3704420> accessed 8 February 2024.
[1690] Because it is not required to undergo the burdensome process of obtaining a patent, for instance. See also Bart Custers, Anne-Sophie Heijne, 'The right of access in automated decision-making: The scope of article 15(1)(h) GDPR in theory and practice' (2022) Vol 46 Computer Law & Security Review 1, 10.
[1691] Case C-634/21, *SCHUFA Holding AG* [2023] ECR I-220, Opinion AG Pikamäe para 58.
[1692] Ibid.
[1693] Case C-487/21, *F.F.* [2022] ECR I-1000 paras 37-38, also Opinion of AG Pitruzella paras 55-56.
[1694] Case C-487/21, *F.F.* [2022] ECR I-1000, Opinion of AG Pitruzella paras 55-56.
[1695] Ibid para 76.
[1696] Cases C/141/12 and C-372/12, *YS* [2014] ECR I-2081 paras 57, 60.
[1697] Lucas Bourtoule et al, 'Machine Unlearning' (IEEE Symposium on Security and Privacy, San Jose, May 2021) 1 and 3 <https://arxiv.org/abs/1912.03817> accessed 8 February 2024.

2.2.1.4. More specifically, it seems impossible to understand what happened in the intermediate (hidden) layers of an ANN.[1698] Most of the current DL models lack reasoning and explanatory capabilities, which makes them vulnerable to produce unexplainable outcomes. In addition, ML becomes increasingly opaque, and even if the underlying principles of ML models are understood, they lack explicit declarative knowledge.[1699] Understanding the causes and correlations of algorithmic decisions currently constitutes one of the major challenges of computer science.[1700] There are some methods to facilitate comprehension of ADM logic.[1701] For example, external explanation systems aim to analyse an AI system and propose explanations by means of two approaches: the white-box approach analyses the code, and the black-box approach is used to probe the ADM by simulating different input and observing the results if no knowledge of the code is available. Both approaches have advantages and drawbacks. Due to technical constraints, the explanation might be limited in case of external black-box approaches, which cannot explain the different steps of an ADM process: only the output, i.e. the final step of the ADM, is explained. How the input is used to produce internal representations remains unknown. External white-box approaches need access to the source code and do not provide explanations in itself but only show some general properties of an ADM system.[1702] It remains unclear whether these methods are helpful for laypersons.[1703] It has been argued that they fall short in providing optimal granularity of explanation for non-experts.[1704] In particular, in ML which is often used for ADM, an affected individual may hardly have any concrete sense of how or why a particular classification results from input.[1705]

Even if an AI system in the future will be able to list all factors that have influenced the ADM process and rank them according to their statistical relevance, it is likely that such information exceeds a data subject's capacity to process such information, resulting in the provision of information that is meaningless rather than meaningful.[1706] Due to these technological shortcomings, controllers cannot

[1698] Ethem Alpaydin, *Machine Learning: The New AI* (3rd edn MIT Press 2016) 155.
[1699] Andreas Holzinger, 'From Machine Learning to Explainable AI' (IEEE DISA Conference, Kosice, August 2018) <https://www.aholzinger.at/wordpress/wp-content/uploads/2020/07/For-Students-HOLZINGER-2018.pdf> accessed 8 February 2024.
[1700] Maja Brkan, Grégory Bonnet, 'Legal and Technical Feasibility of the GDPR's Quest for explanation of Algorithmic Decisions: of Black Boxes, White Boxes and Fata Morganas' (2020) Vol 11 Iss 1 European Journal of Risk Regulation, 18.
[1701] Lee A Bygrave, 'Machine Learning, Cognitive Sovereignty and Data Protection Rights with Respect to Automated Decisions' (2021) University of Oslo Faculty of Law Legal Studies Research Paper Series No. 202-35 at 19 <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3721118> accessed 8 February 2024.
[1702] Maja Brkan, Grégory Bonnet, 'Legal and Technical Feasibility of the GDPR's Quest for explanation of Algorithmic Decisions: of Black Boxes, White Boxes and Fata Morganas (2020) Vol 11 Iss 1 European Journal of Risk Regulation 18, 34, 37-38.
[1703] Thomas Wischmeyer, 'Artificial Intelligence and Transparency: Opening the Black Box' in Thomas Wischmeyer and Timo Rademacher (eds) *Regulating Artificial Intelligence* (Springer Nature 2020) 90.
[1704] Lee A Bygrave, 'Machine Learning, Cognitive Sovereignty and Data Protection Rights with Respect to Automated Decisions' (2021) University of Oslo Faculty of Law Legal Studies Research Paper Series No. 202-35 at 19 <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3721118> accessed 8 February 2024.
[1705] Jenna Burrel, 'How the machine 'thinks': understanding opacity in machine learning algorithms' (2016) Vol 3 Iss 1 Big Data Society 1-12 < https://journals.sagepub.com/doi/pdf/10.1177/2053951715622512 > accessed 8 February 2024.
[1706] Thomas Wischmeyer, 'Artificial Intelligence and Transparency: Opening the Black Box' in Thomas Wischmeyer and Timo Rademacher (eds) *Regulating Artificial Intelligence* (Springer Nature 2020) 89.

comply with the legal obligation imposed on them to provide meaningful information about the *logic* involved in ADM.[1707] According to Article 12 (1) GDPR, information must be intelligible, allowing the data subject to familiarise itself with it fully, easily and without difficulty.[1708] However, current approaches to explain the logic involved in ADM are hardly helpful for laypersons[1709] because they fall short in providing optimal granularity of explanation for non-experts[1710] such as data subjects. Rather, such information makes it 'extremely difficult or burdensome' for the data subject to be acquainted with that information.[1711] This leads to a Type 1 legal problem because meaningful information about the logic involved in ADM cannot be provided in an intelligible manner, which violates both Article 15 (1) lit h and Article 12 (1) GDPR. In fact, empirical research on the matter confirms this conclusion: controllers do not routinely comply with Article 15 (1) lit h GDPR in practice.[1712]

The meaningless information problem (Type 1)

With complex models based on DL and ANNs, it seems impossible to understand what happened in the intermediate (hidden) layers of an ANN when used for ADM. Even if future AI systems will be able to list all factors that have influenced an ADM process, it is likely that such information exceeds a data subject's capacity to understand it, resulting in the provision of meaningless, rather than meaningful information. This violates Articles 12 (1) and 15 (1) lit h GDPR.

5.6.2 Legal problems: Type 2

A Type 2 legal problem is caused by the non-absolute nature of the right of access combined with the broad scope of protection for AI as trade secrets. This Type 2 legal problem with respect to the enforcement of the right of access is twofold. As outlined in Section 5.6, the right of access may be restricted in two ways, i.e. in line with the provisions contained in Article 23 GDPR and in accordance with Article 15 (4) GDPR. Restrictions under Article 15 (4) GDPR differ from restrictions possible under Article 23 GDPR. Article 15 (4) *exclusively* applies to the right to obtain a copy of the personal data enshrined in Article 15 (3) GDPR and allows restrictions on a *case-by-case* basis, whereas restrictions according to Article 23 GDPR need to be laid down in Member State or Union law. Thus, trade secret protection allows controllers to restrict the right to obtain a copy of the personal data in line with Article 15 (4) GDPR, as well as to restrict access to information about processing according

[1707] Maja Brkan, Grégory Bonnet, 'Legal and Technical Feasibility of the GDPR's Quest for explanation of Algorithmic Decisions: of Black Boxes, White Boxes and Fata Morganas (2020) Vol 11 Iss 1 European Journal of Risk Regulation 18, 39.
[1708] Case C-487/21, *F.F.* [2022] ECR I-1000, Opinion of AG Pitruzella para 76.
[1709] Thomas Wischmeyer, 'Artificial Intelligence and Transparency: Opening the Black Box' in Thomas Wischmeyer and Timo Rademacher (eds) *Regulating Artificial Intelligence* (Springer Nature 2020) 90.
[1710] Lee A Bygrave, 'Machine Learning, Cognitive Sovereignty and Data Protection Rights with Respect to Automated Decisions' (2021) University of Oslo Faculty of Law Legal Studies Research Paper Series No. 202-35 at 19 <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3721118> accessed 8 February 2024.
[1711] Case C-487/21, *F.F.* [2022] ECR I-1000, Opinion of AG Pitruzella para 76.
[1712] Bart Custers, Anne-Sophie Heijne, 'The right of access in automated decision-making: The scope of article 15(1)(h) GDPR in theory and practice' (2022) Vol 46 Computer Law & Security Review 1, 11.

to Article 15 (1) GDPR,[1713] provided this occurs in accordance with Article 23 GDPR. I begin with the latter.

Access to information

Many of the fundamental components required to understand AI systems and ensure accountability are barely subject to scrutiny because they are hidden by trade secrets or IP laws.[1714] According to an AI now report, 'one significant barrier to accountability is the culture of industrial and legal secrecy that dominates AI development.'[1715] In fact, Recital 14 TSD specifically includes 'technological information' in the definition of trade secrets, which is significant in the context of AI. Technological information related to an AI system is protected if there is a legitimate interest in maintaining the confidentiality of such information and if there is also a legitimate expectation in the preservation of such confidentiality.[1716] Technological information includes both information about the development and production of the product concerned, as well as information about its actual configuration and functionalities.[1717] In the context of AI systems used to process and generate personal data, technological information is protected in the form of the algorithm[1718] as well as the system's internal components expressed in source code format,[1719] its functionality[1720] and other system artefacts. Put simply, an algorithm is 'the sum of logic and control that has its origins in ancient mathematics'[1721] and is typically a numerical process that consists of a sequence of well-defined steps leading to the solution of a particular type of problem.[1722] The source code is a set of human readable computer commands written in high-level programming languages.[1723]

In order to thoroughly evaluate compliance with applicable legal provisions such as the fairness principle[1724] or ADM,[1725] access to the source code and algorithms at the heart of the AI systems would be required.[1726] For example, to assess potentially discriminatory outcomes of ADM, information regarding comparison groups would be needed. However, particular information about the functionality of algorithms is often poorly accessible[1727] and falls under the scope of trade secret protection within the

[1713] Most importantly information about the logic involved in ADM according to Article 15 (1) lit h GDPR.
[1714] Alex Campolo et al, 'AI Now Report' (2018) 11 < https://ainowinstitute.org/publication/ai-now-2018-report-2 > accessed 8 February 2024.
[1715] Ibid.
[1716] Recital 14 TSD.
[1717] Teresa Trallero Ocaña, *The Notion of Secrecy* (Nomos 2021) 346.
[1718] Ibid 72, 308.
[1719] Noam Shemtov, Beyond the Code: Protection of Non-Textual Features of Software (Oxford University Press 2017) 71.
[1720] Teresa Trallero Ocaña, *The Notion of Secrecy* (Nomos 2021) 346.
[1721] Andrew Goffey, 'Algorithm' in Matthew Fuller (ed) *Software Studies: A Lexicon* (MIT Press 2008).
[1722] Yadolah Dodge, 'Algorithm' in: *The Concise Encyclopedia of Statistics* (Springer New York 2006) 1-2.
[1723] Joasia Krysa, Grzesiek Sedek, 'Source Code' in Matthew Fuller (ed) *Software Studies: A Lexicon* (MIT Press 2008).
[1724] Art 5 (1) lit a GDPR.
[1725] Art. 22 GDPR, and applicable requirements regarding transparency according to Art 13 (2) lit f GDPR.
[1726] Danielle Citron Keats, Frank Pasquale, 'The scored society: Due process for automated predictions' (2014) Vol 89 Iss 1 Washington Law Review, 1, 14.
[1727] Brent Daniel Mittelstadt et al, 'The ethics of algorithms: Mapping the debate' (2016) Vol 3 Iss 2 Big Data & Society 1, 6.

EU.[1728] This makes it difficult for supervisory authorities (SAs) and individuals concerned to verify compliance with the existing legal framework. The legislator anticipated the need for data subjects to obtain information with respect to ADM by requiring controllers to inform them about 'meaningful information about the logic involved' when they enforce their right of access.[1729] Information about the logic involved could fall under the scope of trade secrets because protected technological information includes both information about the development and production of a product and information about its actual configuration and functionalities.[1730] Applied to AI systems and products, the protection offered is broad and comprises the technical method and tools[1731] used to process and obtain information[1732] and thus arguably also how the AI system achieved its automated decision.

For example, in a case dealing with the creation of score values concerning the creditworthiness of individuals, the German Federal Court of Justice ruled that the abstract method of the calculation of the score value, comprising of i) general operandi such as statistical values used, ii) the weighing of specific elements within the calculation of the probability value and iii) the creation of comparison groups do not have to be disclosed because it falls within the scope of trade secrets.[1733] One case[1734] pending at the CJEU specifically addresses the tension between trade secrets and the right of access enshrined in the GDPR. It concerns the German credit agency that automatedly calculated a credit score for a data subject. The data subject exercised her right to access according to Article 15 GDPR and requested the credit agency to provide 'meaningful information about the logic involved' with respect to the ADM to which she was subject (the automated calculation of the credit score). The CJEU is supposed to provide an answer to the question whether Article 15 (1) lit h GDPR obliges the controller to disclose the information which is essential for enabling the comprehensibility of the result of the ADM in the individual case, if necessary while maintaining an existing trade secret.[1735] One of the questions referred to the CJEU is of significant importance in the context of AI and trade secrets, namely, whether meaningful information about the logic involved requires the controller to disclose parts of the algorithm on which the ADM is based for achieving comprehensibility of the

[1728] Directive (EU) 2016/943 of the European Parliament and of the Council of 8 June 2016 on the protection of undisclosed know-how and business information (trade secrets) against their unlawful acquisition, use and disclosure [2016] OJ L157/1 (Trade Secrets Directive).
[1729] Article 15 (1) lit h GDPR.
[1730] Teresa Trallero Ocaña, *The Notion of Secrecy* (Nomos 2021) 346.
[1731] Bart Custers, Anne-Sophie Heijne, 'The right of acces in automated decision-making: The scope of article 15(1)(h) GDPR in theory and practice' (2022) Vol 46 Computer Law & Security Review 1, 10.
[1732] Ana Nordberg, 'Trade secrets, big data and artificial intelligence innovation: a legal oxymoron?' in Jens Schovsbo, Timo Minssen and Thomas Riis (eds) *The Harmonization and Protection of Trade Secrets in the EU* (Edward Elgar Publishing 2020) 197, 201.
[1733] VI/ZR 156/13, BGH (German Federal Court of Justice), judgement of 28 January 2014 [27].
[1734] Case C-203/22 *Dun & Bradstreet Austria.*
[1735] Among other, the CJEU needs to answer whether the data subject exercising its right of access in the context of an ADM must be provided with a) information outlining in which manner personal data are processed, b) input data used for profiling, c) parameters and input variables used in the assessment determination, d) the influence of theses parameters and input variables on the calculated rating, e) information on how the parameters or input variables were arrived at and f) explanations on why the data subject was assigned to a certain evaluation result and presentations of the statement associated with this evaluation, enumeration of the profile categories and explanation of which evaluation statement is associated with each of the profile categories. Article 15 (1) lit h GDPR.

ADM.[1736] In my view, it is unlikely that the CJEU answers this question in the affirmative because the partial disclosure of an algorithm is not intelligible as required by Article 12 (1) GDPR. According to Article 12 (1) GDPR, information must be intelligible, allowing the data subject to familiarise itself with it fully, easily and without difficulty.[1737] This criterion will not be met if the controller provides the data subject with the algorithm or a part of it. AG Pikamäe agrees. He notes that Article 12 (1) GDPR precludes the provision of highly complex information, such as the algorithm used to calculate a score value.[1738]

Information according to Article 15 (1) lit h GDPR may be restricted provided that the requirements set out in Article 23 GDPR are complied with. Article 23 GDPR allows for restrictions of the rights enshrined in Articles 12 to 22 GDPR if (i) provided for in EU or Member State law applying to the controller, (ii) the restriction respects the essence of the fundamental rights and freedoms and (iii) is a necessary and proportionate measure to safeguard, among others, the rights and freedoms of others. Thus, in the situations listed in Article 23 GDPR, private interests can limit the scope of the rights conferred on data subject as introduced in Section 3.3.4 and the corresponding obligations imposed on controllers mentioned in Section 3.3.3.[1739] This holds true regardless of the AI discipline used because trade secret protection applies to all AI disciplines. As outlined in Section 5.6, the term 'rights and freedoms of others' includes trade secrets and IP rights. The TSD and its national laws implementing it constitute EU or Member State law in the sense of the first requirement (i).

With regard to requirement (iii), Malgieri and Comandé argue that there is a legal preference for data protection rights when the latter clash with trade secrets and that the GDPR has intensified this preference.[1740] Among others, they derive this prevalence from Recital 35 TSD which states that the latter should *not affect* the fundamental right to data protection, particularly the right of access and other rights enshrined in the GDPR,[1741] whereas Recital 63 GDPR states that data protection rights should not *adversely* affect trade secrets. According to the authors, the adverb 'adversely' contained in the GDPR reveals that trade secrets can never affect data protection rights, while the right of access can affect trade secrets, but not 'adversely'.[1742] However, that EU law provides greater priority to the

[1736] Case C-203/22 *Dun & Bradstreet Austria* p 2 <https://www.ris.bka.gv.at/Dokumente/Lvwg/LVWGT_WI_20220211_VGW_101_042_791_2020_44_00/LVWGT_WI_20220211_VGW_101_042_791_2020_44_00.pdf> accessed 8 February 2024.
[1737] Case C-487/21, *F.F.* [2022] ECR I-1000, Opinion of AG Pitruzella para 76.
[1738] Case C-634/21, *SCHUFA Holding AG* [2023] ECR I-220, Opinion AG Pikamäe para 57.
[1739] Case C-620/19 *Land Nordrhein Westfalen* [2020] ECR I-1011 paras 42, 46; Case C-620/19 *Land Nordrhein Westfalen* [2020] ECR I-649, Opinion of AG Bobek, para 81.
[1740] Claudio Malgieri, Giovanni Comandé, 'Why a Right to Legibility of Automated Decision-Making Exists in the General Data Protection Regulation' (2017) Vol 7 No 4 International Data Privacy Law 243, 262, 264.
[1741] Note that Recital 35 TDS refers to the Data Protection Directive which was replaced by the GDPR.
[1742] Claudio Malgieri, Giovanni Comandé, 'Why a Right to Legibility of Automated Decision-Making Exists in the General Data Protection Regulation' (2017) Vol 7 No 4 International Data Privacy Law 243, 263.

fundamental right to data protection than safeguarding commercial interests covered by trade secrets is not an accurate claim.[1743]

First, the TSD itself clearly does *not* contain a priority for the fundamental right to data protection. Article 5 TSD limits the scope of trade secret protection,[1744] among others, by restricting trade secret protection in situations of conflicts with the fundamental right to freedom of expression and information (Article 11 EUCFR). If the intention of the EU legislator was to grant the fundamental right to data protection priority over trade secret protection, it would have referred to Article 8 EUCFR in the text of Article 5 TSD, as it did with the fundamental right to freedom of expression and information.

Second, the claim of prevalence for the fundamental right to data protection neglects relevant case law adopted by the CJEU in the context of balancing the fundamental right to data protection with IP rights. Case law indicates that the protection of IP rights may prevail over the protection of personal data: The CJEU considered that the obligation to communicate personal data to private persons in civil proceedings was likely, in principle, to ensure a fair balance between the protection of IP rights and the protection of personal data.[1745] This requirement affirms the rule of non-prevalence in line with other CJEU case law and also rejects arguments made in academia that trade secrets generally prevail over the interests of data subjects and their right of access.[1746] The CJEU stressed the need to reconcile the requirements of the protection of different fundamental rights, such as the fundamental right to privacy and data protection, on the one hand, and the fundamental right to property (including IP and trade secrets[1747]) on the other hand.[1748] According to the CJEU, a 'fair balance' must be struck between the various fundamental rights protected by the EU legal order and any restriction on those rights must comply with the principle of proportionality.[1749] More specifically, Article 23 (1) GDPR seeks to strike a fair balance between the data subjects fundamental right to data protection and the need to safeguard other legitimate interests. This necessitates weighing the fundamental right to data protection conferred on natural persons against the interests that those restrictions are intended to preserve.[1750]

[1743] Lee A Bygrave, 'Machine Learning, Cognitive Sovereignty and Data Protection Rights with Respect to Automated Decisions' (2021) University of Oslo Faculty of Law Legal Studies Research Paper Series No. 202-35 at 19 <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3721118> accessed 8 February 2024.
[1744] Ana Nordberg, 'Trade secrets, big data and artificial intelligence innovation: a legal oxymoron?' in Jens Schovsbo, Timo Minssen and Thomas Riis (eds) *The Harmonization and Protection of Trade Secrets in the EU* (Edward Elgar Publishing 2020) 211.
[1745] See C-461/10 *Bonnier Audio AB* [2012] paras 57-60.
[1746] See, for instance, Paul B de Laat, 'Algorithmic decision-making employing profiling: will trade secrecy protection render the right to explanation toothless?' (2022) Vol 24 Iss 1 Ethics and Information Technology 1, 17.
[1747] Case C-1/11 *Interseroh Scrap* [2012] ECR I-194 para 43; Case T-189/14 *Deza* [2017] para 163.
[1748] Dominique Moore, Commentary of Article 23 in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020) 548.
[1749] Case C-275/06 *Promusicae* [2008] ECR I-00271 paras 65, 68.
[1750] Case C-620/19 *Land Nordrhein Westfalen* [2020] ECR I-1011 para 48; Case C-620/19 *Land Nordrhein Westfalen* [2020] ECR I-649, Opinion of AG Bobek, paras 86 and 88.

Third, the TSD also indicates that the fundamental right to data protection and trade secrets must be respected simultaneously.[1751] Fourth, AG Pikamäe stresses that the legislator clearly did not contemplate sacrificing the fundamental right to intellectual property for the benefit of the fundamental right to data protection, or the other way around. Rather, the legislator intended to strike an appropriate balance between these two rights.[1752]

Therefore, it is possible that meaningful information about the logic involved in ADM might be restricted in accordance with Article 23 GDPR. Regulatory guidance adopted by the EDPB acknowledges this possibility by providing the example that a controller is not bound to reveal any part of the technical operating of software as long as such information can be regarded as a trade secret.[1753] Ultimately, the CJEU will provide clarity on how to proceed when the information to be provided according to Article 15 (1) lit h GDPR is classified as a trade secret within the meaning of Article 2 TSD.[1754] It seems clear that access to certain information is required in order to accurately evaluate compliance with applicable legal provisions such as the fairness principle[1755] or to evaluate ADM[1756] and enforce other data protection rights.

In the case pending at the CJEU, the technical expert appointed by the referring court concluded that specific information is required to ensure the comprehensibility of the calculated credit score. The expert argued that to make the concrete arithmetic operation used to calculate the credit score comprehensible, the detailed mathematical formula used needs to be disclosed next to the processed data. In addition, the expert concluded that comprehensibility is only given if the part of the algorithm is disclosed that was *actually used* by the controller for the calculation of the concrete credit score.[1757] If access to such information is denied in accordance with Article 23 GDPR, the individual concerned will have no opportunity to accurately assess compliance with applicable data protection rules and subsequently enforce other data protection rights, such as the right to rectification or erasure of personal data.[1758] This is because the AI system itself and the technical methods used to process and obtain information might be protected as a trade secret and/or as an IP right.[1759] As a consequence, Article 15 (1) lit h GDPR cannot be enforced, which constitutes a Type 2 legal problem. In fact, empirical research on Article 15 (1) lit h GDPR shows that most of the information required by this

[1751] Recital 34 TSD states that the TSD 'respects the fundamental rights….[ ]…notably the right to protection of personal data.... [ ]…while respecting business secrecy'.
[1752] Case C-634/21, *SCHUFA Holding AG* [2023] ECR I-220, Opinion AG Pikamäe para 55.
[1753] European Data Protection Board, 'Guidelines 01/2022 on data subject rights – Right of access Version 2.0' (28 March 2023) at 173.
[1754] Case C-203/22 *Dun & Bradstreet Austria*.
[1755] Art 5 (1) lit a GDPR.
[1756] Art. 22 GDPR, and applicable requirements regarding transparency according to Art 13 (2) lit f GDPR.
[1757] Case C-203/22 *Dun & Bradstreet Austria* p 12-14 <https://www.ris.bka.gv.at/Dokumente/Lvwg/LVWGT_WI_20220211_VGW_101_042_791_2020_44_00/LVWGT_WI_20220211_VGW_101_042_791_2020_44_00.pdf> accessed 8 February 2024.
[1758] Case C-553/07 *Rijkeboer* [2009] ECR I-03889, paras 51-52.
[1759] About IP rights, see Paul B de Laat, 'Algorithmic decision-making employing profiling: will trade secrecy protection render the right to explanation toothless?' (2022) Vol 24 Iss 1 Ethics and Information Technology 1-20.

provision is rarely or not at all provided in practice because controllers invoke trade secret protection to block or restrict such access requests.[1760]

The information restriction problem (Type 2)

Trade secret protection under the TSD covers AI itself including the technical methods used to process and obtain information and arguably also the particular way how the AI system achieved its ADM. Meaningful information about the logic involved in ADM according to Article 15 (1) lit h GDPR could therefore fall under trade secret protection, allowing controllers to restrict or refuse the provision of such information if this complies with the requirements set out in Article 23 GDPR. Consequently, data subject cannot enforce Article 15 (1) lit h GDPR.

It has been argued that even though algorithms may be protected by trade secrets, explaining the ADM based on that algorithm would not necessarily disclose the trade secret, for example, if only the main factor influencing a decision is required to explain the ADM. Alternatives proposed that do not involve the unlawful disclosure of trade secrets are probing the algorithm by a court or reverse engineering of the protected algorithm in the public domain.[1761] ADM is often influenced by more than one main factor. Probing the algorithm by a court does not seem to be a practical solution for the data subjects and would be in contravention with the law. The GDPR imposes the duty to explain the logic involved in ADM on the controller – it is not the data subject's task to invest time and financial resources to obtain such information. Article 12 (1) GDPR stipulates that information must be intelligible, meaning that it should be understandable for a data subject.[1762] Information should be provided in a manner that enables the data subject to familiarise herself with it fully, easily and without difficulty. Asking the data subject to probe the algorithm by a court would make it 'extremely difficult or burdensome' for the data subject to be acquainted with the information[1763] according to Article 15 (1) lit h GDPR. The same conclusion applies to reverse engineering, a technique to understand how a product was designed and operates[1764] (see also Section 6.5.2). In addition, the data subject should not have to extensively seek out information,[1765] as it must be 'easily accessible'[1766] and the verb 'provide' implies that the data subject is not required to actively search for information covered by Article 15 GDPR.[1767]

[1760] Bart Custers, Anne-Sophie Heijne, 'The right of access in automated decision-making: The scope of article 15(1)(h) GDPR in theory and practice' (2022) Vol 46 Computer Law & Security Review 1-16.
[1761] Maja Brkan, Grégory Bonnet, 'Legal and Technical Feasibility of the GDPR's Quest for explanation of Algorithmic Decisions: of Black Boxes, White Boxes and Fata Morganas (2020) Vol 11 Iss 1 European Journal of Risk Regulation 18, 41.
[1762] Art 29 Working Party, 'Guidelines on transparency under Regulation 2016/679' (WP 260 rev.01, 11 April 2018) at 9.
[1763] Case C-487/21, *F.F.* [2022] ECR I-1000, Opinion of AG Pitruzella para 76.
[1764] Noam Shemtov, Beyond the Code: Protection of Non-Textual Features of Software (Oxford University Press 2017) 71.
[1765] Art 29 Working Party, 'Guidelines on transparency under Regulation 2016/679' (WP 260 rev.01, 11 April 2018) at 11.
[1766] Article 12 (1) GDPR.
[1767] European Data Protection Board, 'Guidelines 01/2022 on data subject rights – Right of access Version 2.0' (28 March 2023) at 130; Art 29 Working Party, 'Guidelines on transparency under Regulation 2016/679' (WP 260 rev.01, 11 April 2018) at 33.

The transparency principle and particularly Article 12 GDPR is intertwined with the right of access. Article 12 GDPR is an expression of the transparency principle and aims to ensure that the data subject fully understands the information provided, allowing to effectively exercise the right of access and other data subject rights.[1768] Controllers must inform data subjects under Article 15 (1) GDPR in a way that enables complete access to the requested information.[1769] The two alternatives proposed, namely, probing by the Court and reverse engineering in the public domain, would conclusively lead to a Type 1 legal problem. The alternatives violate the modalities to provide information as required by Article 12 (1) GDPR, because that information must be intelligible and easily accessible.

Obtaining a copy of personal data

As mentioned in Section 3.3.4.1, the concept of copy is not defined in the GDPR. The CJEU ruled that a 'copy' refers to 'faithful reproduction or transcription' of an original. A purely general description of the data undergoing processing or a reference to categories of personal data does not correspond to that definition.[1770] In addition, the right to obtain a copy also includes information resulting from the processing of personal data, for example, a credit score.[1771] Faithful means 'true and accurate; not changing anything'[1772] and/or 'true or not changing any of the details, facts, style, etc. of the original'.[1773] The copy must enable the data subject to effectively exercise the right of access in full knowledge of all personal data undergoing processing, including personal data *generated* by the *controller*.[1774] The latter makes crystal clear that personal data generated by the controller with the support of AI systems or applications do fall within the scope of the right to obtain a copy of the personal data undergoing processing. However, Article 15 (3) GDPR does not require the provision of a copy of the document, but a copy of the personal data. The CJEU found that there is no right to obtain a copy of the document containing the personal data.[1775] In addition, Article 15 (3) GDPR does not provide the data subject with a right to obtain information regarding the criteria, models, rules or internal procedures (whether or not computational) used for processing the personal data.[1776]

Article 15 (4) GDPR states that the right to obtain a copy of personal data undergoing processing according to Article 15 (3) GDPR 'should not adversely affect the rights or freedoms of others'. Recital 63 GDPR clarifies that this refers to trade secrets, intellectual property and copyright protecting the software. By directly referring to Article 15 (3) GDPR, this limitation of the right of access

[1768] Case C-487/21, *F.F.* [2022] ECR I-1000, Opinion of AG Pitruzella paras 55-56.
[1769] European Data Protection Board, 'Guidelines 01/2022 on data subject rights – Right of access Version 2.0' (28 March 2023) at 130.
[1770] Case C-487/21, *F.F.* [2022] ECR I-1000 para 21.
[1771] Ibid, para 26.
[1772] See <https://www.oxfordlearnersdictionaries.com/definition/english/faithful?q=faithful> and < accessed 8 February 2024.
[1773] See < https://dictionary.cambridge.org/dictionary/english/faithful > accessed 8 February 2024.
[1774] Case C-487/21, *F.F.* [2022] ECR I-1000 paras 26, 39; see also the opinion of AG Pitruzella paras 45, 70.
[1775] Cases C/141/12 and C-372/12, *YS* [2014] ECR I-2081 paras 58-59.
[1776] Case C-487/21, *F.F.* [2022] ECR I-1000, Opinion of AG Pitruzella para 52.

only applies to the right to obtain a copy of personal data undergoing processing, but not to Article 15 (1) GDPR.[1777]

As outlined in Section 5.6, Article 15 (4) GDPR allows controllers to restrict the right to obtain a copy of the personal data on a case-by-case basis. This gives controllers more leeway and flexibility in restricting access requests concerning Article 15 (3) GDPR, because these restrictions do not have to be enshrined in EU or MS law.[1778] The restriction of the right to obtain a copy of the personal data enshrined in Article 15 (4) GDPR is particularly relevant if the information protected as a trade secret constitutes personal data. As outlined in Section 5.6, trade secret protection covers training and output data,[1779] information[1780] on or knowledge about customers, information about a customer's behaviour (creditworthiness, lifestyle)[1781] and predictions such as a customer's future life (life expectancy, estimated advancements in career, etc.).[1782] More generally, any output generated by an AI system constituting personal data, such as a data subject's detected emotional state could fall under the trade secret protection.[1783] Therefore, the exception enshrined in Article 15 (4) GDPR is particularly problematic in the context of AI.

Imagine, for instance, an AI system that intends to detect the emotional state of an individual powered by the discipline AC. The data subject enforces her right of access by specifically requesting a copy of the personal data undergoing processing[1784] to determine what emotional state the system has discovered. Then, the controller refers to trade secret protection and argues that he is not obliged to disclose the detected emotional state (e.g., fear or anger) even though such information constitutes sensitive personal data. In the sketched situation, the data subject cannot gain access to her own personal data generated by AI (AC).

The same applies to the AI system introduced in Section 4.4.3 which uses unsupervised ML techniques to automatically predict the life expectancy of insurance companies' clients based on relatively simple personal data, such as the gender and place of residence of the clients. Also, here, the controller may refuse to disclose the life expectancy predictions due to trade secret protection.

[1777] European Data Protection Board, 'Guidelines 01/2022 on data subject rights – Right of access Version 2.0' (28 March 2023) at 169. However, such restriction might be possible under Article 23 GDPR.
[1778] As it is the case of restrictions made on the basis of Article 23 GDPR.
[1779] Ana Nordberg, 'Trade secrets, big data and artificial intelligence innovation: a legal oxymoron?' in Jens Schovsbo, Timo Minssen and Thomas Riis (eds) *The Harmonization and Protection of Trade Secrets in the EU* (Edward Elgar Publishing 2020) 197, 201.
[1780] Teresa Trallero Ocaña, *The Notion of Secrecy* (Nomos 2021) 312.
[1781] Gianclaudio Malgieri, 'Trade Secrets v. Personal Data: A Possible Solution for Balancing Rights' (2016) Vol 6 No 2 International Data Privacy Law 102, 113-114.
[1782] Sandra Wachter, Brent Mittelstadt 'A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI' (2019) No 2 Columbia Business Law Review 495, 607; Gianclaudio Malgieri, 'Trade Secrets v. Personal Data: A Possible Solution for Balancing Rights' (2016) Vol 6 No 2 International Data Privacy Law 102, 113-114.
[1783] Gianclaudio Malgieri, 'Trade Secrets v. Personal Data: A Possible Solution for Balancing Rights' (2016) Vol 6 No 2 International Data Privacy Law 102, 113-114.
[1784] Article 15 (3) GDPR

Output produced by AI that constitute personal data does not need to be correct to fall under trade secret protection: information or knowledge protected under the TSD may be very well incorrect or incomplete.[1785] Therefore, if a data subject enforces the right of access in order to 'be aware of, and verify, the lawfulness of the processing',[1786] controllers are likely to argue that disclosure of output produced by an AI system constituting personal data infringes their trade secrets or IP rights.[1787] Companies may merely provide restricted information, such as naming the category 'emotion data' or 'life expectancy prediction' instead of disclosing the predicted life expectancy or detected emotional state. This approach would be in line with Recital 63 GDPR, which states that considerations with respect to trade secrets should not result in 'a refusal to provide all information to the data subject'. This will likely be considered acceptable by supervisory authorities (SAs). Regulatory guidance concerning transparency[1788] simply requires controllers to inform data subjects about the 'categories of the inferred data processed'.[1789] Both emotional states as well as life expectancy predictions are inferred data defined as 'the product of probability-based processes'.[1790] Furthermore, a data subject cannot request a copy of the *document* containing the personal data undergoing processing, such as the report generated by AC system HireVue.[1791] Article 15 (3) GDPR does not require the controller to provide a copy of the document containing personal data.[1792] Indeed, the CJEU confirmed that there is no right to receive a copy of the document containing the personal data undergoing processing.[1793] Likewise, Article 15 (3) GDPR does not provide the data subject with a right to obtain information regarding the criteria, models, rules or internal procedures (whether or not computational) used for processing the personal data.[1794] Again, instead of disclosing the detected emotional state or predicted life expectancy, the controller may just indicate the category of personal data, such as 'emotion data' or 'life expectancy prediction', in order to protect its trade secrets.

[1785] Ana Nordberg, 'Trade secrets, big data and artificial intelligence innovation: a legal oxymoron?' in Jens Schovsbo, Timo Minssen and Thomas Riis (eds) *The Harmonization and Protection of Trade Secrets in the EU* (Edward Elgar Publishing 2020) 200.
[1786] Recital 63
[1787] Bart Custers, Anne-Sophie Heijne, 'The right of acces in automated decision-making: The scope of article 15(1)(h) GDPR in theory and practice' (2022) Vol 46 Computer Law & Security Review 1, 10.
[1788] Dealing with the transparency principle and Articles 12-14 GDPR.
[1789] Art 29 Working Party, 'Guidelines on transparency under Regulation 2016/679' (WP 260 rev.01, 11 April 2018) in footnone 30 at page 14.
[1790] OECD Working Party on Security and Privacy in the Digital Economy JT03357584 (2014) <https://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=dsti/iccp/reg(2014)3&doclanguage=en> accessed 8 February 2024.
[1791] Nathan Mondragon, Clemens Aicholzer, Kiki Leutner, 'The Next Generation of Assessments' (HireVue 2019) <http://hrlens.org/wp-content/uploads/2019/11/The-Next-Generation-of-Assessments-HireVue-White-Paper.pdf> accessed 8 February 2024.
[1792] Note however that this depends on local guidance and local case law, arguably leading to 'unharmonized' results across the EU; Gabriela Zanfir-Fortuna, Commentary of Article 15 in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020) 464.
[1793] Cases C-141/12 and C-372/12, *YS* [2014] ECR I-2081 paras 58-59. Note that the CJEU relativated this to some extent. It might be needed to provide the reproduction of extracts from documents or even entire documents ore extracts from databases containing personal data to ensure the copy provided is intelligible. See Case C-487/21, *F.F.* [2022] ECR I-1000 para 41.
[1794] Case C-487/21, *F.F.* [2022] ECR I-1000, Opinion of AG Pitruzella para 52.

As a result of this, the individual concerned has no means of accessing the specific emotional state detected by the AI system or the predicted life expectancy. This is particularly relevant because the right of access is decisive for other data protection rights and enables the data subject to obtain, depending on the circumstances, rectification, erasure or blocking of his or her personal data by the controller. This leads to a significant loophole because the data subject cannot verify the accuracy of the emotion data detected by the AI system. I use the term 'loophole' because in my view, data subjects should be able to see which emotion the machine recognises, in particular when considering the sensitive nature of emotion data.[1795] Without that knowledge, an individual will hardly be able to obtain rectification of inaccurate data because it is the individual that must demonstrate the inaccuracy of personal data (see Section 5.7). Note that life expectancy predictions or emotional states detected by the AI system may be protected under the TSD even if they are incorrect.[1796]

Because the TSD provides extensive protection for input data and output data in all AI disciplines and because trade secrets are widely used, it will hardly be possible for individuals concerned to accurately assess compliance with the GDPR and enforce other data subject rights such as rectification or erasure. Controllers are likely to invoke trade secret protection to deny full or partial access to personal data undergoing processing.[1797] Already in 2011, Facebook denied a data subject access to his personal data because such disclosures 'would adversely affect trade secrets'.[1798] Trade secret protection hampers the thorough enforcement of the right to obtain a copy of the personal data processed, which constitutes a Type 2 legal problem.

The trade secrets problem (Type 2)

Trade secret protection under the TSD covers AI itself as well as output generated by the AI system, including personal data relating to emotional states and life expectancy predictions. When data subjects invoke their right to obtain a copy of personal data undergoing processing according to Article 15 (3) GDPR, controllers are likely to argue that disclosure of the output generated by the AI system infringes their trade secrets and restrict access to such personal data in accordance with Article 15 (4) GDPR. Consequently, data subjects cannot enforce their right to obtain a copy of their personal data.

[1795] I derive this requirement from the underlying ideas of the transparency *and* fairness principle.
[1796] Information or knowledge protected under the TSD does not have to be correct or complete.
[1797] Bart Custers, Anne-Sophie Heijne, 'The right of acces in automated decision-making: The scope of article 15(1)(h) GDPR in theory and practice' (2022) Vol 46 Computer Law & Security Review 1, 15.
[1798] See < http://www.europe-v-facebook.org/FB_E-Mails_28_9_11.pdf > accessed 8 February 2024.

5.6.3 Legal problems: Type 3

The trade secrets problem explained in Section 5.6.2, i.e. that controllers may deny data subjects copies of personal data undergoing processing, also leads to a Type 3 legal problem. Because data subjects cannot gain access to the personal data processed by a controller to verify the lawfulness of processing[1799] and to obtain the rectification, erasure or blocking of personal data,[1800] Article 15 (3) GDPR is not fit for purpose to effectively[1801] protect the fundamental right to data protection. The right of access is a conditio sine qua non for exercising other data subject rights and restrictions on or around this right cause a knock-on effect on the entire data protection law regime.[1802] The CJEU stressed the importance of ensuring that data subject rights granted by the GDPR are effective.[1803] Article 15 (3) GDPR is not effective because it allows controllers to extensively restrict this right based on Article 15 (4) GDPR. Controllers may easily invoke this provision by arguing that the disclosure of personal data generated by means of AI violates their trade secret protection. In such cases, the data subject must initiate legal proceedings against the controller[1804] or lodge a complaint with the competent SA[1805] to challenge the controller's restriction of Article 15 (3) GDPR. The lack of sufficient resources for SAs[1806] and the EDPB[1807] is widely known, which causes delay of regulatory enforcement. According to a report published by the EDPB in 2021, it took the Irish SA an average of 16 months to formally decide on purely *national cases* and 23 months for cases subject to the *cooperation procedure*.[1808] The Irish SA is the lead supervisory authority for most of the 'big tech' companies, including Meta Platforms Ireland Limited, Google Ireland Limited, WhatsApp Ireland Limited, Airbnb Ireland UC, Twitter International Company, Microsoft Ireland Operations Limited, LinkedIn Ireland UC and Apple Distribution International.[1809] In terms of private enforcement, according to Article 79 GDPR, the timeframe to obtain a final decision is even longer, when considering that 'big tech' companies may be willing to exhaust all possible legal remedies and that such cases raise new points of law and ultimately end up at the CJEU. Thus, when taking the broad exception

[1799] Recital 63 GDPR.
[1800] Joined Cases C-141/12 & C-372/12, *YS, M* [2014] ECR I-2081 para 44.
[1801] Recital 11 GDPR.
[1802] Jef Ausloos, Michael Veale, René Mahieu, 'Getting Data Subject Rights Right' (2019) Vol 10 Iss 3 JIPITEC 283, 285.
[1803] Case C-154/21 *Österreichische Post AG* [2023] ECR I-3 para 39.
[1804] Article 79 GDPR.
[1805] Article 77 GDPR.
[1806] EDPB, 'Overview on resources made available by Member States to the Data Protection Supervisory Authorities' (2022) at 5 <https://edpb.europa.eu/system/files/2022-09/edpb_overviewresourcesmade_availablebymemberstates-tosas2022_en.pdf> accessed 8 February 2024.
[1807] The EDPB and EDPS have jointly sent an open letter to the European Parliament and European Council expressing concerns about the budget for 2023; see <https://edps.europa.eu/system/files/2022-09/22-09-12_edps-edpb-open-letter-budget-2022_en.pdf> accessed 8 February 2024.
[1808] EDPB, 'Overview on resources made available by Member States to the Data Protection Supervisory Authorities' (2021) at 21 < https://edpb.europa.eu/system/files/2021-08/edpb_report_2021_overviewsaressourcesandenforcement_v3_en_0.pdf > accessed 8 February 2024.
[1809] See < https://www.dataprotection.ie/sites/default/files/uploads/2022-03/DPC%20statistical%20report%20on%20OSS%20cross-border%20complaints.pdf > accessed 8 February 2024.

according to Article 15 (4) GDPR and the long enforcement timeframes into account, Article 15 (3) GDPR is not an effective right.

Articles 15 (3) and 15 (4) GDPR also fail to achieve the GDPR's legislative aim to strengthen the rights of data subjects.[1810] As noted by the CJEU, effective protection of personal data requires the strengthening of the rights of data subjects, which is emphasised by Recital 11 GDPR.[1811] Article 15 (3) GDPR specifically aims to strengthen the position of the data subject.[1812] Instead of strengthening the right of access, the broad scope of restrictions possible under Article 15 (4) GDPR weakens this right and thus fails to achieve the GDPR's legislative aim. Additionally, these newly introduced provisions do not achieve the GDPR's goal that 'natural persons should have control of their own personal data',[1813] although this was one of the main reasons for the data protection reform.[1814] As outlined in Section 4.4.3, one of the main mechanisms for data subjects to exercise control over the processing of their personal data under the GDPR are data subject rights. Control in the sense of the GDPR is rather limited from a conceptual point of view. In a preliminary ruling, the AG acknowledged that 'the scope for individual action is limited' and 'confined to the exercise of those rights in specified circumstances'.[1815] Article 15 (4) GDPR further restricts the already limited mechanism for data subjects to exercise control over the processing of their personal data.

This is especially true for AC, which can generate inaccurate personal data (see Section 4.7.1). Without access to the specific emotional state detected by the AI system deploying AC approaches, a data subject cannot verify the accuracy of the output data and subsequently request the rectification or erasure of such personal data. The same applies to ML, which generates predictions and establishes correlations that are probabilistic and thus constitute uncertain knowledge, which may lead to inaccurate evaluations and representations of data subjects. Article 15 (3) GDPR is the last resort for data subjects to obtain the specific emotional state detected by AC or the exact prediction or correlation generated by ML to subsequently enforce other rights of the data subject, such as the right to rectification or erasure. Due to the broad scope of protection provided by the TSD for *all AI* disciplines as introduced in Chapter 2, this Type 3 legal problem constitutes a general problem and relates to all AI disciplines discussed in Chapter 2.

It remains unclear how a controller must, in fact, respond to an access right request according to Article 15 (1) GDPR and what information must be included in such a response. The standard adopted by the GDPR requires that information must be provided to data subjects in a 'concise, intelligible

[1810] Recital 11 GDPR.
[1811] Case C-132/21 *Nemzeti* [2023] ECR I-2 para 42.
[1812] Case C-487/21, *F.F.* [2022] ECR I-1000 para 33; see also the opinion of AG Pitruzella para 69.
[1813] Recital 7 GDPR.
[1814] Case C-300/21 *UI* [2022] ECR I-756 Opinion AG Manuel Campos Sánchez-Bordona paras 69, 70.
[1815] Case C-300/21 *UI* [2022] ECR I-756 Opinion AG Manuel Campos Sánchez-Bordona para 72.

and easily accessible form, using clear and plain language'.[1816] It is unclear what this means in the context of the right of access, particularly when it concerns meaningful information about the *logic* involved in ADM. Research[1817] suggests that such information refers to a description of the technologies used rather than access to the code or software itself. The dearth of corresponding literature underscores that the matter has not received much attention in academia or practice,[1818] nor in regulatory guidance.

With regard to information according to Article 15 (1) lit h, the EDPB takes the view that such information could be based on the privacy notice of a controller subject to being 'updated and tailored' to the data subject making the request[1819] and should, *if possible*, 'be more specific in relation to the reasoning that lead to specific decisions concerning the data subject who asked for access'.[1820] Regulatory guidance also states that such information does not necessarily entail complex information of the algorithms used or disclosure of the algorithm.[1821] Instead of providing a complex mathematical explanation about how algorithms and AI used for ADM work, controllers should provide general information such as factors taken into account for the ADM process and their respective weight on an aggregated level. In addition, controllers should disclose the categories of data that have been or will be used for ADM, why these categories are pertinent, how any profile used in the ADM process is built, including any statistics used in the analysis, why this profile is relevant to the ADM process and how it is used for a decision with respect to the data subject.[1822] In addition, regulatory guidance seems to indicate that Article 15 (1) lit h GDPR does not oblige controllers to explain *particular decisions* to data subjects, but rather to oblige them to provide information about the envisaged consequences of the processing. In view of the EDPB, the right to receive meaningful information about the logic involved in ADM does *not* seem to entail a right for data subjects to obtain explanation of particular decisions because Article 15 (1) lit h GDPR entitles data subjects to obtain the *same information* as required under Articles 13 (2) lit f and 14 (2) lit g GDPR.[1823]

Views in scholarship diverge on what information controllers must provide under Article 15 (1) lit h GDPR. There is a vivid debate whether or not the GDPR provides a right to explanation of specific ADM or not.[1824] With regard to the information to be provided specifically under Article 15 (1) lit h

[1816] Article 12 (1) GDPR.
[1817] Bart Custers, Anne-Sophie Heijne, 'The right of acces in automated decision-making: The scope of article 15(1)(h) GDPR in theory and practice' (2022) Vol 46 Computer Law & Security Review 1, 16.
[1818] Ibid.
[1819] European Data Protection Board, 'Guidelines 01/2022 on data subject rights – Right of access Version 2.0' (28 March 2023) at 20, 113.
[1820] Ibid at 119.
[1821] Art 29 Working Party 'Guidelines on Automated Individual Decision-Making and Profiling for the Purposes of Regulation 2016/679', (WP251rev.01, 6 February 2018) at 25.
[1822] Ibid 31; see also Case C-634/21, *SCHUFA Holding AG* [2023] ECR I-220, Opinion AG Pikamäe para 58.
[1823] Art 29 Working Party 'Guidelines on Automated Individual Decision-Making and Profiling for the Purposes of Regulation 2016/679', (WP251rev.01, 6 February 2018) at 26, 27.
[1824] Thomas Wischmeyer, 'Artificial Intelligence and Transparency: Opening the Black Box' in Thomas Wischmeyer and Timo Rademacher (eds) *Regulating Artificial Intelligence* (Springer Nature 2020) 75-101; Sandra Wachter, Brent

GDPR, Malgieri and Comandé argue that such information must adhere to the standard of legibility, which requires that the information provided is both transparent and comprehensible and that such information must go 'beyond the mere mathematical functionality of an algorithm' and consider contextual use, expected and actual impact, rationales and purposes.[1825] Wachter, Mittelstadt and Floridi take the view that the right of access only grants an explanation of the logic and general functionality of an ADM system, but not the rationale and circumstances of specific decisions. Additionally, 'meaningful information' would not entail an obligation to disclose the algorithm, but only the provision of 'basic information' about its logic.[1826] Finally, empirical research on Article 15 (1) lit h GDPR suggests interpreting it as information that is useful and/or has practical value for data subjects.[1827] Obviously, this interpretation has a contextual component. It refers to useful and practical information for data subjects to (i) become aware of processing relating to ADM, (ii) enforce their data subject rights (e.g. contesting to ADM) and thus (iii) exercise control over the processing of their personal data. This interpretation is also in line with the requirement of intelligibility as enshrined in Article 12 (1) GDPR. This provision ensures that the data subject fully understands the information provided,[1828] enabling *effectively* exercise of the right of access and other data subject rights.[1829]

However, as pointed out in Section 5.6.2, in a CJEU case relating to explanation of the logic involved in ADM, the technical expert appointed by the referring court concluded that, in order to comprehend the logic involved and evaluate the ADM at hand, at least the disclosure of a part of the algorithm would be required, together with other detailed information. The latter include the concrete factors and mathematical formula used, the concrete value assigned to the data subject and the disclosure of the intervals within which different data on the same factor are assigned to the same value.[1830] It is unlikely that the CJEU will accept this interpretation. As AG Pikamäe notes, the requirement of intelligibility enshrined in Article 12 (1) GDPR precludes the provision of highly complex information, such as the algorithm or the mathematical formula used.[1831] As outlined in Section 5.6.2, this information is meaningless rather than meaningful for most data subjects.

Mittelstadt, Luciano Floridi, 'Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation' (2017) Vol 7 Iss 2 IDPL 76-99; Giancludio Malgieri, Giovanni Comandé, 'Why a Right to Legibility of Automated Decision-Makig Exists in the General Data Protection Regulation' (2017) Vol 7 Iss 4 IDPL 243-265.

[1825] Giancludio Malgieri, Giovanni Comandé, 'Why a Right to Legibility of Automated Decision-Making Exists in the General Data Protection Regulation' (2017) Vol 7 Iss 4 IDPL 243, 245, 257, 258.

[1826] Sandra Wachter, Brent Mittelstadt, Luciano Floridi, 'Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation' (2017) Vol 7 Iss 2 IDPL 76, 84, 90.

[1827] Bart Custers, Anne-Sophie Heijne, 'The right of access in automated decision-making: The scope of article 15(1)(h) GDPR in theory and practice' (2022) Vol 46 Computer Law & Security Review 1, 14.

[1828] Case C-487/21, *F.F.* [2022] ECR I-1000 paras 37, 38.

[1829] Case C-487/21, *F.F.* [2022] ECR I-1000, Opinion of AG Pitruzella paras 55-56.

[1830] Case C-203/22 *Dun & Bradstreet Austria;* see page 12 <https://www.ris.bka.gv.at/Dokumente/Lvwg/LVWGT_WI_20220211_VGW_101_042_791_2020_44_00/LVWGT_WI_20220211_VGW_101_042_791_2020_44_00.pdf> accessed 8 February 2024.

[1831] Case C-634/21, *SCHUFA Holding AG* [2023] ECR I-220, Opinion AG Pikamäe para 57.

Another highly relevant question is whether controllers must provide meaningful information about the logic involved with regard to a *particular decision*. Regulatory guidance[1832] as well as the scholars Wachter, Mittelstadt and Floridi[1833] suggest answering this question negatively. I do not agree. If information according to Article 15 (1) lit h GDPR does not relate to a particular decision, it cannot be useful or meaningful for data subjects. To determine what information could be useful and/or of practical value ('meaningful') for data subjects, it is worth considering what is typically required if humans are asked for an explanation of a specific decision. What humans usually want to know is whether and how certain input factors affected the final decision or outcome.[1834] Such causal explanation helps individuals to modify their behaviour or consider which factors they must challenge in order to change the decision.[1835] Thus, in order to be meaningful for data subjects, information according to Article 15 (1) lit h GDPR needs to explain how certain input factors affected the final ADM.[1836] Causal explanation relating to a *specific* automated decision would enable data subjects to determine which factors they must challenge in order to change the ADM,[1837] by obtaining human intervention, expressing their point of view and contesting the decision as foreseen in Article 22 (3) GDPR. AG Pikamäe seems to agree. With regard to the automated establishment of a score value, controllers must provide 'sufficiently detailed explanations of the method for calculating the score value and the reasons that led to a *particular* result.'[1838]

However, neither the GDPR and its corresponding recitals nor regulatory guidance seem to suggest such an interpretation of meaningful information about the logic involved in a specific automated decision. The opinion of AG Pikamäe is not legally binding, and the CJEU completely ignored this point in the corresponding ruling. Thus, data subjects do not know the input factors that affected a specific automated decision and cannot effectively enforce their right to contest ADM according to Article 22 (3) GDPR. Therefore, information according to Article 15 (1) lit h GDPR is not useful and/or of practical value for data subjects. In addition, empirical legal research on Article 15 (1) lit h GDPR concludes that the right of access, particularly Article 15 (1) lit h GDPR, does not function adequately in practice.[1839]

[1832] Art 29 Working Party 'Guidelines on Automated Individual Decision-Making and Profiling for the Purposes of Regulation 2016/679', (WP251rev.01, 6 February 2018) at 26, 27.
[1833] Sandra Wachter, Brent Mittelstadt, Luciano Floridi, 'Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation' (2017) Vol 7 Iss 2 IDPL 76, 84, 90.
[1834] Finale Doshi-Velez et al, 'Accountability of AI Under the Law: The Role of Explanation' (2017) Berkman Klein Center Working Group on Explanation and the Law Working Paper 1 <https://dash.harvard.edu/bitstream/handle/1/34372584/2017-11_aiexplainability-1.pdf> accessed 8 February 2024.
[1835] Thomas Wischmeyer, 'Artificial Intelligence and Transparency: Opening the Black Box' in Thomas Wischmeyer and Timo Rademacher (eds) *Regulating Artificial Intelligence* (Springer Nature 2020) 88.
[1836] Finale Doshi-Velez et al, 'Accountability of AI Under the Law: The Role of Explanation' (2017) Berkman Klein Center Working Group on Explanation and the Law Working Paper 1 <https://dash.harvard.edu/bitstream/handle/1/34372584/2017-11_aiexplainability-1.pdf> accessed 8 February 2024.
[1837] Thomas Wischmeyer, 'Artificial Intelligence and Transparency: Opening the Black Box' in Thomas Wischmeyer and Timo Rademacher (eds) *Regulating Artificial Intelligence* (Springer Nature 2020) 88.
[1838] Case C-634/21, *SCHUFA Holding AG* [2023] ECR I-220, Opinion AG Pikamäe para 58.
[1839] Bart Custers, Anne-Sophie Heijne, 'The right of access in automated decision-making: The scope of article 15(1)(h) GDPR in theory and practice' (2022) Vol 46 Computer Law & Security Review 1, 16.

Therefore, Article 15 (1) lit h GDPR is not fit for purpose to effectively protect the fundamental right to data protection[1840] and strengthen data subject rights as envisaged by the GDPR.[1841] The CJEU emphasised that effective protection of personal data requires the strengthening of the rights of data subjects[1842] and also stressed the importance of ensuring that data subjects rights granted by the GDPR are effective.[1843] A right that fails to provide data subjects with information that is useful and/or of practical value with regard to other data subject rights enshrined in the GDPR is ineffective. Consequently, it also fails to strengthen the rights of the data subject. Likewise, this provision fails to achieve the GDPR's legislative goals to enhance legal and practical certainty for data subjects and to provide data subjects with control over the processing of their own personal data.[1844] As outlined in Section 4.4.3, control in the sense of the GDPR is limited to two main mechanisms, namely, consent and data subject rights. Regarding the latter, even AG Campos Sánchez-Bordona acknowledged that 'the scope for individual action is limited' and 'confined to the exercise of those rights in specified circumstances'.[1845] Article 15 (1) lit h GDPR further restricts the mechanism for data subjects to exercise control, in particular regarding their right to contest to ADM according to Article 22 (3) GDPR. Due to the lack of causal explanation relating to a *specific* automated decision, it may be difficult for data subjects to determine which factors they must challenge to change the ADM,[1846] by obtaining human intervention, expressing their point of view and contest the decision as foreseen in Article 22 (3) GDPR. This leads to a Type 3 legal problem occurring regardless of which AI discipline is used for ADM. The problem is caused by the wording enshrined in Article 15 (1) lit h GDPR, which does not impose an obligation on controllers to provide data subjects with a causal explanation about a specific automated decision. It is therefore a general problem and relates to all AI disciplines as introduced in Chapter 2.

The logic and causal explanation problem (Type 3)

The right to obtain meaningful information about the logic involved in ADM according to Article 15 (1) lit h GDPR does not seem to require controllers to provide causal information about specific ADM, i.e. how input factors affected the final decision. Consequently, data subjects cannot determine which factors they must challenge when contesting ADM according to Article 22 (3) GDPR. Therefore, Article 15 (1) lit h GDPR is not fit for purpose to protect the fundamental right to data protection.

[1840] Case C-645/19 [2021] *Facebook Ireland* ECR I-483 para 45; Case C-319/20, *Meta Platforms Ireland Limited* [2022] ECR I-322 para 73.
[1841] Recital 11 GDPR.
[1842] Case C-132/21 *Nemzeti* [2023] ECR I-2 para 42.
[1843] Case C-154/21 *Österreichische Post AG* [2023] ECR I-3 para 39.
[1844] Recital 7 GDPR.
[1845] Case C-300/21 *UI* [2022] ECR I-756 Opinion AG Manuel Campos Sánchez-Bordona para 72.
[1846] Thomas Wischmeyer, 'Artificial Intelligence and Transparency: Opening the Black Box' in Thomas Wischmeyer and Timo Rademacher (eds) *Regulating Artificial Intelligence* (Springer Nature 2020) 88.

5.7 Rectification

Both the EUCFR and the GDPR provide individuals with a right to have personal data rectified.[1847] Article 16 GDPR is more specific and grants data subjects a right to obtain the rectification of inaccurate personal data with respect to him or her or to have incomplete personal data completed. The right to rectification constitutes a key element of the fundamental right to data protection.[1848] Its significance has been emphasised by both the ECtHR[1849] and CJEU.[1850] It applies to false, inaccurate and incomplete information.[1851] Neither the GDPR itself, nor CJEU case law nor regulatory guidance yield details about the standard of proof applying to the rectification of personal data. It remains unclear which requirements data subjects must meet concerning the accuracy or completeness of the personal data designated to *replace* the personal data currently processed by the controller when they exercise their right to rectification.

A case relating to the request to erasure of inaccurate personal data and the freedom of expression according to Article 17 (3) lit a GDPR provides some insight about the standard of proof to be met in order to establish the inaccuracy of personal data processed. According to the CJEU, the data subject bears the burden of proof to establish the manifest inaccuracy of the information in question.[1852] To avoid an excessive burden, the data subject must provide evidence that can reasonably be required. It must submit '*relevant* and *sufficient* evidence capable of substantiating his or her request and of establishing the *manifest inaccuracy* of the information'.[1853] Apparently, the CJEU did not follow the opinion of AG Pitruzella, who suggested a lower evidence threshold. In his view, the data subject must provide 'prima facie evidence of the false nature of the content'.[1854] However, the context of this case must be taken into account. It relates to the weighing of the fundamental rights to privacy and the protection of personal data on the one hand and the fundamental right to freedom of expression and information on the other. Arguably, the CJEU might establish a lower standard when balancing the rights and freedoms of data subjects against those of controllers. Therefore, I do not give the standard of 'manifest inaccuracy' much weight. Rather, I rely on a PNR opinion issued by the CJEU which suggests that rectification somehow relates to the notion of verification because it used the terms 'verified' and 'unverified' personal data in the opinion.[1855] The CJEU has pointed to the significant 'margin of error' that may result from the automated processing of personal data, in particular

[1847] Article 8 (2) EUCFR and Article 16 GDPR.
[1848] Cécile de Terwangne, Commentary of Article 16 in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020) 473.
[1849] *Leander v Sweden,* App No 9248/81 (ECtHR 26 March 1987) para 48; *Rotaru v Romania,* App No 28341/95 (ECtHR 4 May 2000) para 46.
[1850] Case C-434/16, *Nowak* [2017] ECR I-994 para 49; Case C-362/14 *Schrems v Data Protection Commissioner* [2015] ECR I-650 para 95; Case C-553/07 *Rijkeboer* [2009] ECR I-03889 para 51.
[1851] *Cemalettin Canli v Turkey* App No 22427/04 (ECtHR 18 February 2009) para 37 and 42; Case C-131/12 Google Spain [2014] ECR I-317 para 70.
[1852] Case C-460/20, *TU* [2022] ECR I-962 para 68.
[1853] Case C-460/20, *TU* [2022] ECR I-962 paras 68, 72.
[1854] Case C-460/20, *TU* [2022] ECR I-962, Opinion AG Pitruzella para 50.
[1855] Opinion 1/15 CJEU [2017] ECR I-592 paras 131, 169.

if such processing is carried out on the basis of '*unverified* personal data […] and pre-established *models* and criteria'.[1856]

According to the ECtHR, natural persons should adduce 'objectively verifiable evidence' for having personal data relating to them changed.[1857] Case law on the right to rectification in the Netherlands seems to apply a similar standard: inaccuracies in personal data to be rectified must be 'easily' and 'objectively' verifiable.[1858] In Germany, the standard concerning the right to rectification amounts to 'objective reality': correct data reflect reality, and data are incorrect if not corresponding with reality.[1859] Differences in local case law with respect to the standard of proof are caused by the principle of national procedural autonomy. In the absence of EU procedural law, Member States may set up the procedural system as they deem fit.[1860] Thus, the manner of regulating procedural law is generally considered a matter of Member State autonomy, as long as it satisfies the minimum principles of effectiveness and equivalence[1861] (see Section 5.7.1). Unfortunately, there are no standards that define the required degree of accuracy[1862] that could serve as a benchmark when a data subject wishes to rectify personal data (see Section 4.7.2).

Because this thesis relates to EU law, I introduce a distinct 'EU' standard. From ECtHR case law[1863] as well as the CJEU's PNR opinion,[1864] it can be concluded that the right to rectification relies on the notion of verification. Thus, when data subjects dispute the accuracy or completeness of personal data processed by the controller ('current data'), they must provide verifiable evidence that the 'new' personal data envisaged to replace the current data are accurate. I call this 'the objective verifiability standard'. The latter is seemingly met with ease when the personal data in question is verifiable by nature (such as a name, date of birth, email address).[1865] In what follows, I explain that this is not the case regarding personal data processed in the context of AI. Personal data generated by AI is often unverifiable by nature. This applies particularly to inferred personal data (including predictions) produced by means of ML and emotion data generated by AC.

[1856] Opinion 1/15 CJEU [2017] ECR I-592 paras 169, 170 emphasis added.
[1857] *Ciubotaru v Moldov* App No 27138/04 (ECtHR 27 July 2010) para 59.
[1858] Raad van State, ECLI:NL:RVS:2021:1020, 20 February 2019 para 5.1.
[1859] BVerwG – 6 C 7.20 [2022], ECLI:DE:BVerwG:2022:020322U6C7.20, para 32.
[1860] Bart Krans, Anna Nylund, 'Aspects of Procedural Autonomy' in Bart Krans, Anna Nylund (eds) *Procedural Autonomy Across Europe* (Intersentia 2020) 1.
[1861] Anna Wallerman, 'Towards an EU law doctrine on the exercise of discretion in national courts? The Member States' self-imposed limits on national procedural autonomy' (2016) Vol 53 Iss 2 Common Market Law Review 339-360.
[1862] Diana Dimitrova, 'The rise of the personal data quality principle: is it legal and does it have an impact on the right to rectification?' (2021) Vol 12 No 3 European Journal of Law and Technology 25.
[1863] *Ciubotaru v Moldov* App No 27138/04 (ECtHR 27 July 2010) para 59.
[1864] Opinion 1/15 CJEU [2017] ECR I-592 paras 131, 169.
[1865] Sandra Wachter and Brent Mittelstadt, 'A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI' (2019) Vol 2 Columbia Business Law Review 494, 548.

5.7.1 Legal problems: Type 1

When applied to AI-generated personal data, the right to rectification could be violated due to the procedural law applicable in a given Member State ('MS'). More specifically, the right to rectification may be violated when the procedural law and/or judicial practice of a Member State does not meet the minimum principles of equivalence and effectiveness as elaborated by CJEU case law. These principles appear in numerous cases[1866] and are, together with the principle of effective judicial protection, the most widely recognised limits on national procedural autonomy.[1867] The principle of equivalence essentially amounts to the law of remedies concerning the general principle of non-discrimination.[1868] More importantly in the context of this thesis, the minimum principle of *effectiveness* demands that procedural rules applicable in any given MS must not render the exercise of rights conferred to individuals by EU law '*virtually impossible or excessively difficult*'. [1869] In a landmark ruling, the CJEU found that any provision, legislative, administrative or judicial practice that 'might prevent, even temporarily, Community rules from having full force and effect'[1870] is incompatible with the very essence of EU law.[1871]

One case in Germany dealing with the right to rectification further illustrates this problem. The German Federal Administrative Court ruling[1872] mentioned in Section 5.7 arguably violates the right to rectification according to the GDPR because the judicial practice and national procedural law make it *excessively difficult* for data subjects to enforce their right to rectification conferred to them by Article 16 GDPR. In the dispute of this ruling, the Republic of Turkey issued a new passport for the data subject containing a corrected date of birth (01.01.1953 'new date'), following the ruling of a Turkish district court that declared the data subject's date of birth currently registered (01.01.1958 'current date') to be incorrect. Consequently, the data subject requested that the entry of his date of birth contained in the German population register (current date) be changed in accordance with the newly issued Turkish passport containing the new date of birth.[1873]

[1866] Case 33–76, *Rewe-Zentralfinanz eG and Rewe-Zentral AG* [1976] European Court Reports 1976-01989; Joined cases C-430/93 and C-431/93, *Jeroen van Schijndel et al* [1995] ECR I-4705 para 17; Case C-312/93 Peterbroeck [1995] ECR I-437; Case C-126/97, *Eco Swiss China Time Ltd* [1999] ECR I-269; see Bart Krans, Anna Nylund, 'Aspects of Procedural Autonomy' in Bart Krans, Anna Nylund (eds) *Procedural Autonomy Across Europe* (Intersentia 2020) in Footnote 5 at page 3 for more cases.
[1867] Anna Wallerman, 'Towards an EU law doctrine on the exercise of discretion in national courts? The Member States' self-imposed limits on national procedural autonomy' (2016) Vol 53 Iss 2 Common Market Law Review 339, 342.
[1868] Koen Lenaerts, 'National Remedies for Private Parties in the Light of the EU Law Principles of Equivalence and Effectiveness' (2011) Vol 46 Irish Jurist 13, 14.
[1869] Case C-353/20, Skeyes [2022] ECR I-423 para 52; Case C-497/20, *Randstad Italia SpA* [2021] ECR I-1037 para 58, Joined Cases C-222/05 and C-225/05, *Van der Weerd* [2007] ECR I-4233 para 28; *Jeroen van Schijndel et al* [1995] ECR I-4705 para 17.
[1870] Case C-213/89, Factortame [1990] ECR I-527 para 20.
[1871] Bart Krans, Anna Nylund, 'Aspects of Procedural Autonomy' in Bart Krans, Anna Nylund (eds) *Procedural Autonomy Across Europe* (Intersentia 2020) 3.
[1872] BVerwG – 6 C 7.20 [2022], ECLI:DE:BVerwG:2022:020322U6C7.20.
[1873] BVerwG – 6 C 7.20 [2022], ECLI:DE:BVerwG:2022:020322U6C7.20, paras 1-3.

Although the data subject provided the newly issued passport as evidence for the rectification of his personal data, the German Federal Administrative Court concluded that the controller cannot be obliged to change the registered date of birth in the German population register as, in accordance with Germany's code of civil procedure concerning the evidentiary value of public documents,[1874] the 'correctness of the date of birth as "01.01.1953" *[new date]* does not follow from the entry in the plaintiff's Turkish passport'.[1875] The Court referred to the accountability principle according to Article 5 (2) GDPR that puts the burden of proof on the controller to demonstrate compliance with the accuracy principle. Considering this burden of proof, the controller cannot be required to rectify the current date and instead process the new date of birth of which the accuracy cannot be determined, in particular, where the data subject *fails to prove the correctness* of the new date as required by applicable procedural law. According to the Court, the burden of proof regarding the accuracy of the new data lies on the data subject. The data subject's inability to prove the correctness of the new data is at the data subject's expense.[1876] Hence, the data subject cannot exercise the right to rectification according to Article 16 GDPR if it cannot establish the accuracy of personal data designated to replace the current personal data processed by the controller with sufficient certainty.[1877]

In my view, the judicial practice adopted by the German Court, as well as the procedural laws in Germany, render it 'virtually impossible or excessively difficult'[1878] for data subjects to exercise their right to rectification according to EU data protection law. Ultimately, this contradicts the minimum principle of effectiveness and thus, in itself, may violate EU law. In addition, the judicial practice is contrary to the GDPR's objectives to ensure that the level of protection is *equivalent* in all MS,[1879] strengthening data subject rights,[1880] and particularly providing the same level of legally enforceable data subject rights.[1881] Also, in my view, a newly issued passport containing the correct date of birth should be considered to meet the objective verifiability standard as introduced in Section 5.7. Furthermore, the Court's ruling appears to adopt a prevalence for 'current' data processed by the controller, making it excessively difficult, if not impossible, for data subjects to obtain rectification of such personal data and opens the door for controllers to easily reject rectification requests. In addition, it should be kept in mind that this case concerned personal data whose accuracy appears to be easy to verify, as opposed to personal data generated by AI (see Sections 5.7.2 and 5.7.3). When the judicial practice adopted by the German court as well as the German procedural laws are applied to the rectification of unverifiable and highly subjective personal data generated by AI, it will be virtually

[1874] Zivilprozessordnung (ZPO) § 418 Beweiskraft öffentlicher Urkunden mit anderem Inhalt.
[1875] BVerwG – 6 C 7.20 [2022], ECLI:DE:BVerwG:2022:020322U6C7.20, para 7, emphasis added by the author.
[1876] BVerwG – 6 C 7.20 [2022], ECLI:DE:BVerwG:2022:020322U6C7.20, paras 9, 52.
[1877] BVerwG – 6 C 7.20 [2022], ECLI:DE:BVerwG:2022:020322U6C7.20, paras 9, 51.
[1878] Joined Cases C-222/05 and C-225/05, *Van der Weerd* [2007] ECR I-4233 para 28; *Jeroen van Schijndel et al* [1995] ECR I-4705 para 17.
[1879] Recital 10 GDPR.
[1880] Recital 11 GDPR, Case C-132/21 *Nemzeti* [2023] ECR I-2 para 42.
[1881] Recital 13 GDPR.

impossible or excessively difficult for data subjects to rectify inaccurate personal data. This constitutes a Type 1 legal problem.

> ***The procedural autonomy problem (Type 1)***
> *Due to the principle of national procedural autonomy, Member States (MS) may set up their own procedural laws as they deem fit. This may lead to the violation of the right to rectification when the procedural law and/or judicial practice of a MS renders it virtually impossible or excessively difficult for data subjects to exercise their right to rectification according to Article 16 GDPR. This problem applies particularly to the rectification of unverifiable and highly subjective personal data generated by the AI disciplines ML and AC as discussed in Sections 5.7.2 and 5.7.3.*

5.7.2 Legal problems: Type 2

ML as introduced in Section 2.2.1 is particularly eligible to generate inferred data, defined as 'the product of probability-based processes' used to create predictions of behaviour[1882] (see also Sections 4.4.1 and 4.4.3). ML applies data-driven methods, combining fundamental concepts in computer science with approaches from statistics, probability and optimisation[1883] and is used for classification as well as the detection of patterns and predictions. Therefore, ML constitutes a powerful tool of computational methods using experience to make predictions.[1884] Due to its probabilistic approach, ML is closely related to the field of statistics and is particularly helpful to handle ambiguous cases.[1885] Given that predictions produced by ML, such as life expectancy, score value ratings and career perspectives are probabilistic by nature, ML poses the risk that personal data generated by it might be inaccurate, wrong or incomplete. Essentially, ML-based predictions or classifications constitute 'educated guesses or bets, based on large amounts of data'.[1886] ML systems that aim to predict the future behaviour of individuals cannot achieve absolute accuracy due to the predictive nature of the generated output and the lack of a baseline truth for comparison.[1887] Thus, as outlined in Section 4.7.1, ML generates output that constitutes uncertain knowledge because it is probabilistic by nature and not based on human reasoning. Therefore, such an output can be inaccurate.

[1882] OECD Working Party on Security and Privacy in the Digital Economy JT03357584 (2014) <https://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=dsti/iccp/reg(2014)3&doclanguage=en> accessed 8 February 2024.
[1883] Mehryar Mohri, Afshin Rostamizadeh, Ameet Talwalkar, *Foundations of Machine Learning* (MIT Press 2012) 1.
[1884] Mehryar Mohri, Afshin Rostamizadeh, Ameet Talwalkar, *Foundations of Machine Learning* (MIT Press 2012) 1.
[1885] Kevin P Murphy, *Machine Learning: A Probabilistic Perspective* (MIT Press 2012) 1, 4.
[1886] Teresa Scantaburlo, Andrew Charleswoth, Nello Cristianini, 'Machine Decisions and Human Consequences' in Karen Yeung and Martin Lodge (eds) *Algorithmic Regulation* (OUP 2019) 57; Lee A Bygrave, 'Machine Learning, Cognitive Sovereignty and Data Protection Rights with Respect to Automated Decisions' (2021) University of Oslo Faculty of Law Legal Studies Research Paper Series No. 202-35 at 5 <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3721118> accessed 8 February 2024.
[1887] Diana Dimitrova, 'The rise of the personal data quality principle: is it legal and does it have an impact on the right to rectification?' (2021) Vol 12 No 3 European Journal of Law and Technology 21.

With the output generated by ML, it is rather difficult or even impossible to meet the objective verifiability standard. The main reason for this is that the predictions generated by ML relate to *future behaviour* that has not yet happened. Examples of such output generated by ML are predictions about a customer's future life, including estimated advancements in career,[1888] credit risk scores, life expectancy or likelihood of future health outcomes.[1889] An individual's phone-charging habit is currently used as a relevant factor for determining individual creditworthiness. AI, in particular when powered by ML, assesses data points such as phone-charging habits that would commonly not be considered when determining someone's creditworthiness. For example, Smart Finance disclosed that customers who regularly let their phone batteries drop below 12% are not considered good prospects. Another FinTech company called Lenddo considers hyper well-maintained smartphone batteries as a red flag because such a phone-charging habit seems to be robotic or not human enough.[1890] In fact, research suggests that behaviour revealed in mobile phone usage can predict the likelihood of credit repayment. By means of ML, the likelihood of repayment was predicted using behavioural features derived from mobile phone usage.[1891]

Often, predictions or correlations are essentially considered *facts*, although the output generated by ML is probabilistic and can relate to conduct that has not yet happened. Such inferred data can be used by controllers for decision-making with respect to data subjects, whether automated or not. Output generated by ML is not only problematic due to the possible impact they may have for the data subject concerned, but also because such output may be fed back into the AI system and influence future decisions and predictions which could lead to discrimination.[1892] Difficulties concerning the provision of objectively verifiable evidence are particularly problematic when considering the highly subjective nature of predictive inference techniques such as ML:[1893] predictions generated by ML are essentially educated guesses based on large amounts of data.[1894] Inferred data generated by ML may also ascribe attributes to people using ML techniques such as regression, classification (see Section

[1888] Gianclaudio Malgieri, 'Trade Secrets v. Personal Data: A Possible Solution for Balancing Rights' (2016) Vol 6 No 2 International Data Privacy Law 102, 113-114; Sandra Wachter, Brent Mittelstadt 'A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI' (2019) No 2 Columbia Business Law Review 495, 607.
[1889] OECD Working Party on Security and Privacy in the Digital Economy JT03357584 (2014) <https://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=dsti/iccp/reg(2014)3&doclanguage=en> accessed 8 February 2024.
[1890] Tanya Goodin, 'The battery life of your phone could affect your loan application' (2022) <https://tanya-goodin.com/2022/08/credit-rating-algorithmic-transparency/> accessed 8 February 2024.
[1891] Daniel Björkegren, Darrell Grissen, 'Behavior Revealed in Mobile Phone Usage Predicts Credit Repayment' (2020) Vol 34 Iss 3 The World Bank Economic Review 618, 623.
[1892] Solon Barocas, Andrew D Selbst 'Big Data's disparate impact' (2016) Vol. 104 California Law Review 671, 681, 726; Bart Custers, 'Profiling as inferred data. Amplifier effects and positive feedback loops' in Emre Bayamlıoğlu, Irina Baraluic, Liisa Janssens and Mireille Hildebrandt (eds), *Being Profiled: Cogitas Ergo Sum* (Amsterdam University Press 2018) 113.
[1893] Jef Ausloos, Michael Veale, René Mahieu, 'Getting Data Subject Rights Right' (2019) Vol 10 Iss 3 JIPITEC 283, 302.
[1894] Teresa Scantaburlo, Andrew Charleswoth, Nello Cristianini, 'Machine Decisions and Human Consequences' in Karen Yeung and Martin Lodge (eds) *Algorithmic Regulation* (OUP 2019) 57; Lee A Bygrave, 'Machine Learning, Cognitive Sovereignty and Data Protection Rights with Respect to Automated Decisions' (2021) University of Oslo Faculty of Law Legal Studies Research Paper Series No. 202-35 at 5 <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3721118> accessed 8 February 2024.

2.2.1.1) or clustering (Section 2.2.1.2) and thus amount to profiling as defined in Article 4 (4) GDPR (see also Section 4.4.3). Attributes ascribed to data subjects are quite imprecise (e.g., inferred from Facebook likes) and constitute estimates rather than factual information. Therefore, profiles are simply new inferred personal data.[1895]

In terms of accuracy, personal data can be divided into three categories: i) factual data that accurately reflect a known reality about an individual, ii) counter-factual data that inaccurately reflect a known reality about an individual and iii) data that cannot be described as completely falling under the former or the latter.[1896] I call the last category 'unverifiable personal data'. According to CJEU case law, facts are susceptible to proof.[1897] Unverifiable personal data, e.g. inferred personal data such as ML predictions or subjective emotion data are not susceptible to proof because they do not constitute factual nor counter-factual data.

Inferred data, including estimates or predictions generated by AI systems and other output generated by ML, fall into the category of unverifiable personal data. For example, life expectancy and estimated advancements in career may prove to be wrong or true in the future, but in essence they are probabilistic and not verifiable at the time when they are generated. Data subjects cannot meet the objective verifiability standard when they intend to enforce their right to rectify the output generated by ML. This is mainly due to the fact that such data relates to the future, its highly probabilistic nature and the lack of a baseline truth for comparison.[1898] In addition, it is generally impossible for individuals to prove that personal data inferred by means of AI is inaccurate without having access to the tools used to infer the data.[1899] As outlined in Section 5.6, such tools, including specific technological information, are likely to be subject to trade secret protection which hinders individuals from proving the inaccuracy of inferred personal data.[1900]

Therefore, it seems extremely difficult, if not impossible, for data subjects to meet the objective verifiability standard regarding unverifiable data inferred by ML. Possibly, this leads to serious consequences for data subjects as inferred data may propagate existing biased patterns, leading to disparate

[1895] Bart Custers, 'Profiling as inferred data. Amplifier effects and positive feedback loops' in Emre Bayamlıoğlu, Irina Baraluic, Liisa Janssens and Mireille Hildebrandt (eds), *Being Profiled: Cogitas Ergo Sum* (Amsterdam University Press 2018) 112.
[1896] Dara Hallinan, Frederik Zuiderveen Borgesius 'Opinions can be incorrect (in our opinion)! On data protection law's accuracy principle' (2020) Vol. 10 No. 1 IDPL 1, 4-5.
[1897] Case C-460/20, *TU* [2022] ECR I-962 para 66.
[1898] Art 29 Working Party 'Guidelines on Automated Individual Decision-Making and Profiling for the Purposes of Regulation 2016/679', (WP251rev.01, 6 February 2018) at 17-18; Diana Dimitrova, 'The rise of the personal data quality principle: is it legal and does it have an impact on the right to rectification?' (2021) Vol 12 No 3 European Journal of Law and Technology 21.
[1899] Bart Custers, 'Profiling as inferred data. Amplifier effects and positive feedback loops' in Emre Bayamlıoğlu, Irina Baraluic, Liisa Janssens and Mireille Hildebrandt (eds), *Being Profiled: Cogitas Ergo Sum* (Amsterdam University Press 2018) 114.
[1900] The situation is even more difficult in the case of emotion data because it is questionable if and how such data in fact can be verified.

impacts.[1901] Additionally, it is highly problematic when unverifiable data are essentially considered as *facts*, although they are not. The personal data generated by ML are probabilistic and relate to future behaviour that has not yet occurred. Actions taken based on probabilistic predictions and correlations may have real impact on human interests[1902] (e.g., to receive a loan or to be employed). Regulatory guidance indicates that data subjects cannot rectify inferred personal data such as a prediction if this *may* be factually correct, even if the prediction *never materialises*. If, according to this guidance, a computer system puts the data subject into the group that 'most likely will develop heart disease', the data subject cannot request the rectification of the inferred personal data because the prediction solely states that the data subject is more likely to develop heart disease. This might be factually correct as a matter of statistics, even if the data subject will never suffer from heart disease.[1903] Because output generated by ML, including inferred data, represents unverifiable personal data, data subjects cannot meet the objective verifiability standard when they enforce their right to rectification. This constitutes a Type 2 legal problem.

The unverifiable data problem (Type 2)

ML generates probabilistic output concerning a data subject's future life such as credit risk and life expectancy scores, or future health, constituting uncertain knowledge. Due to the lack of truth serving as a verification mechanism and the lack of access to the tools used to generate them, this output represents unverifiable personal data. Consequently, data subjects cannot meet the objective verifiability standard when enforcing their right to rectification.

The difficulty to meet the objective verifiability standard applicable the right of rectification also occurs regarding emotion data generated by the AI discipline AC. To illustrate this problem in more detail, I take the example of emotion data inferred by an AI system that relies on the AI discipline AC combined with other AI disciplines (e.g., CV for facial movements or NLP). Emotion data are subjective by nature and, therefore, are not objectively verifiable.

Naturally, emotion data can only be a known reality for the natural person that has these emotions (and not for other parties or entities) because every individual has its own personal experience of emotion.[1904] Thus, emotion data are not objectively verifiable due to the subjective perception of emotion. Rather, it is subjectively verifiable: Emotion data can uniquely be verified by the individual

[1901] Bart Custers, 'Profiling as inferred data. Amplifier effects and positive feedback loops' in Emre Bayamlıoğlu, Irina Baraluic, Liisa Janssens and Mireille Hildebrandt (eds), *Being Profiled: Cogitas Ergo Sum* (Amsterdam University Press 2018) 115.
[1902] Brent Daniel Mittelstadt et al, 'The ethics of algorithms: Mapping the debate' (2016) Vol 3 Iss 2 Big Data & Society 1, 5; Solon Barocas, 'Data Mining and the Discourse on Discrimination' (2014) <https://dataethics.github.io/proceedings/DataMiningandtheDiscourseOnDiscrimination.pdf> accessed 8 February 2024.
[1903] Art 29 Working Party 'Guidelines on Automated Individual Decision-Making and Profiling for the Purposes of Regulation 2016/679', (WP251rev.01, 6 February 2018) at 18.
[1904] Jennifer Healey, 'Physiological Sensing of Emotion' in Rafael Calvo et al (eds), *The Oxford Handbook of Affective Computing* (OUP 2015) 213, 214.

experiencing the emotional state in question. Emotion data derived from AC-powered applications represent unproven and factually uncertain information about the emotional states of individuals. As described in Section 4.7.1, it is likely that emotion data detected by AC systems is inaccurate. For example, imagine that an AC-powered automated video assessment wrongfully detects that the job applicant was angry while performing the automated video assessment. Because the data subject in fact was surprised by an unexpected question posed during the automated video assessment, he seeks the rectification of the inaccurate emotional state of anger to be replaced by the emotional state 'surprise'. Because emotion data are subjective by nature, it is impossible for the data subject to meet the objective verifiability standard. Emotional states cannot be verified objectively because they are by definition subjective as every individual has its own, personal, experience of emotion.[1905] Because emotion data are subjective by nature, data subjects cannot meet the objective verifiability standard when enforcing their right to rectification to correct inaccurate emotion data. This leads to a Type 2 legal problem.

The subjectivity problem (Type 2)

Scientific research suggests that AC powered systems are likely to generate inaccurate emotional data. Emotional data are highly subjective because every individual has its own personal experience of emotion. Due to this inherently subjective nature, data subjects cannot meet the objective verifiability standard when they seek the rectification of inaccurate emotional data.

The right to rectification enshrined in Article 16 GDPR also allows data subjects to provide a 'supplementary statement'. This amounts to adding missing elements rather than rectifying inaccurate personal data.[1906] It seems unclear what specific obligations such a supplementary statement imposes on the controller.[1907] Regulatory guidance simply states that Article 16 GDPR contains a right for the data subject to complement the personal data with additional information.[1908] Thus, the right to have incomplete personal data completed may not be particularly helpful in the context of AI because it does not solve the problem of inaccurate data. Even if the data subject could prove that personal data generated by AI is inaccurate, similar issues arise in the context of the right to erasure (Section 5.8.1). Such issues concern the practical consequences for controllers, e.g. whether and how they should rectify the personal data contained in the ML model, for example, by means of machine unlearning.

[1905] Jennifer Healey, 'Physiological Sensing of Emotion' in Rafael Calvo et al (eds), *The Oxford Handbook of Affective Computing* (OUP 2015) 213, 214.
[1906] Cécile de Terwangne, Commentary of Article 16 in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020) 473.
[1907] Diana Dimitrova, 'The rise of the personal data quality principle: is it legal and does it have an impact on the right to rectification?' (2021) Vol 12 No 3 European Journal of Law and Technology 27.
[1908] Art 29 Working Party 'Guidelines on Automated Individual Decision-Making and Profiling for the Purposes of Regulation 2016/679', (WP251rev.01, 6 February 2018) at 18.

5.7.3 Legal problems: Type 3

Two views have been presented that significantly restrict the scope of the right to rectification in Article 16 GDPR. First, it has been argued that inferred data 'cannot be rectified under data protection law and can only be contested if there is a procedure in place to contest the evaluation'.[1909] According to this view, the right to rectification is limited to assess the accuracy and completeness of the input data, but excludes the output data generated by means of AI, including opinions.[1910] Second, AG Sharpston takes the view that 'only information relating to *facts* about an individual can be personal data'.[1911] Consequently, only factual personal data can be rectified under the right to rectification. In the Netherlands, there is established case law restricting the right to rectification to *factual* personal data.[1912] Accordingly, the right to rectification is in principle not applicable to impressions, assessments and conclusions relating to the data subject.[1913] If only factual personal data fall under the scope of this right, inferred data cannot be rectified because such data represent unproven and factually uncertain knowledge relating to the future, rather than facts. The view that only *input data* and *factual* personal data fall within the scope of Article 16 GDPR unduly limits the right to rectification. When applied to personal data inferred by AI-powered systems such as ML predictions or emotional states inferred by AC approaches, such personal data cannot be rectified at all. However, this narrow interpretation of the right to rectification not only contradicts regulatory guidance[1914] but also the CJEU's teleological approach to interpret data subject rights.[1915]

In my view, the problem is not the scope of Article 16 GDPR, but the objective verifiability standard. According to CJEU case law, facts are susceptible to proof.[1916] Since unverifiable data are neither factual nor counter-factual data, it is extremely difficult if not impossible to provide evidence that they are inaccurate. As outlined in the unverifiable data and subjectivity problems discussed in Section 5.7.2, data subjects cannot rectify unverifiable and subjective personal data generated by AI when the objective verifiability standard is applied. Regarding both unverifiable and subjective personal data, the question arises of what information the data subject can adduce in order to meet the objective verifiability standard.

[1909] Sandra Wachter, Brent Mittelstadt 'A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI' (2019) No 2 Columbia Business Law Review 1, 550-551; see also 549, 590.
[1910] Sandra Wachter, Brent Mittelstadt 'A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI' (2019) No 2 Columbia Business Law Review 1, 550-590.
[1911] Joined Cases C-141/12 & C-372/12, *YS, M* [2014] ECR I-2081, Opinion of AG Sharpston para 56.
[1912] Raad van State, ECLI:NL:RVS:2011:BR6338, 31 August 2011 para 2.3; Raad van State, ECLI:NL:RVS:2019:520, 20 February 2019 para 7.2; Rechtbank Den Haag, ECLI:NL:RBDHA:2022:2432, 25 February 2022 para 7.3.
[1913] Raad van State, ECLI:NL:RVS:2019:520, 20 February 2019 para 7.2; Rechtbank Den Haag, ECLI:NL:RBDHA:2022:2432, 25 February 2022 para 7.3.
[1914] Art 29 Working Party 'Guidelines on Automated Individual Decision-Making and Profiling for the Purposes of Regulation 2016/679' (WP251rev.01, 6 February 2018) at 8–9 and 17-18.
[1915] Case C-434/16, *Nowak* [2017] ECR I-994 para 53.
[1916] Case C-460/20, *TU* [2022] ECR I-962 para 66.

In terms of subjective emotion data, I refer to the example mentioned in Section 5.7.2. An automated video assessment powered by AC wrongfully detects that the applicant was angry while conducting the assessment, although the job applicant was surprised. The data subject can put forward a simple statement indicating he has experienced another emotional state, It is difficult to imagine objectively verifiable evidence for this, however. Ultimately, only the data subject can determine the accuracy of a detected emotional state because emotion data are inherently subjective as every individual has his or her own, personal, experience of emotion.[1917] Emotions can only be a known reality for the natural person that has these emotions and not for other parties or entities. There is simply no such thing as objectively verifiable evidence that a data subject may adduce to rectify inaccurate emotion data. Consequently, the data subject cannot request the controller to replace the emotional state detected by the AI system (sadness) with the correct emotional state (surprise).

In case of inferred personal data generated by ML, the data subject cannot request the rectification of the inferred personal data because there are *no facts* available to prove inaccuracy. The prediction simply states that the data subject is more likely to develop heart disease, which might be correct as a matter of statistics, even if the data subject in fact will never suffer from heart disease.[1918] As pointed out in the unverifiable data problem in Section 5.7.2, ML can generate probabilistic output with respect to the data subject's future life. Examples are estimated career advancements, credit risk scores, life expectancy scores or the likelihood of future health outcomes, constituting uncertain knowledge. This output represents unverifiable personal data because it relates to future behaviour that has not (yet) happened and cannot be considered as facts, even if such output is based on mathematical calculations.[1919] It is unverifiable because it is probabilistic. There is a lack of truth serving as a verification mechanism, and data subjects cannot access the tools used to generate the output. Consequently, a data subject is unable to meet the objective verifiability standard and provide the corresponding factual evidence outlining that a prediction is wrong. This is problematic when considering that inferred personal data might have adverse consequences for the data subject, in particular when considered and treated as facts, despite their probabilistic nature. This occurs, for example, when a data subject seeks to obtain health care insurance or a loan. Likewise, inaccurate emotion data can have adverse consequences for data subjects when used by controllers, for example, in an employment context or when such data are used to influence or manipulate the data subject (see Section 4.3.3).

Due to the objective verifiability standard, data subjects cannot enforce their right to rectification regarding personal data generated by AI. Such data are unverifiable and/or subjective, and factual data eligible to prove inaccuracy are absent. Therefore, the right to rectification is not fit for purpose

[1917] Jennifer Healey, 'Physiological Sensing of Emotion' in Rafael Calvo et al (eds), *The Oxford Handbook of Affective Computing* (OUP 2015) 213, 214.
[1918] Art 29 Working Party 'Guidelines on Automated Individual Decision-Making and Profiling for the Purposes of Regulation 2016/679', (WP251rev.01, 6 February 2018) at 18.
[1919] A chance of something based on e.g. mathematical/statistical calculations can also be considered factual data.

to achieve the GDPR's legislative aim to strengthen the rights of data subjects.[1920] As noted by the CJEU, effective protection of personal data requires the strengthening of the rights of data subjects, which is emphasised by Recital 11 GDPR.[1921] A right that cannot be enforced with regard to unverifiable and subjective personal data generated by AI systems is not suitable to strengthen the rights of data subjects. Furthermore, the objective verifiability standard hampers the GDPR's aim to improve the legal and practical protection of data subjects (Recital 7). It remains unclear how data subjects can enforce their right to rectification regarding unverifiable or highly subjective personal data generated by AI systems.

According to the CJEU, it is important that the data subject rights granted by the GDPR are effective.[1922] However, this is not the case with the right to rectification. Data subjects can hardly enforce this right regarding unverifiable or highly subjective personal data generated by AI due to the objective verifiability standard. Article 16 GDPR does not achieve the GDPR's goal that 'natural persons should have control of their own personal data',[1923] although this was one of the main reasons for the data protection reform.[1924] As outlined in Section 4.4.3, one of the main mechanisms for data subjects to exercise control over the processing of their personal data under the GDPR are data subject rights. Control in the sense of the GDPR is rather limited from a conceptual point of view. In his opinion concerning a preliminary ruling, AG Campos Sánchez-Bordona acknowledged that 'the scope for individual action is limited' and 'confined to the exercise of those rights in specified circumstances'.[1925] The objective verifiability standard further restricts the mechanism for data subjects to exercise control because data subjects cannot rectify arguably inaccurate personal data generated by means of AI. This leads to a Type 3 legal problem.

The verifiability standard problem (Type 3)

Data subjects need to meet the objective verifiability standard to have output generated by ML and AC powered systems rectified. Output generated by means of ML may constitute unverifiable personal data. Emotion data are by nature highly subjective. Therefore, data subjects cannot provide evidence that meets the objective verifiability standard. Thus, the right to rectification is not fit for purpose to protect the fundamental right to data protection, as this standard hinders data subjects from exercising their right.

[1920] Recital 11 GDPR.
[1921] Case C-132/21 *Nemzeti* [2023] ECR I-2 para 42.
[1922] Case C-154/21 *Österreichische Post AG* [2023] ECR I-3 para 39.
[1923] Recital 7 GDPR.
[1924] Case C-300/21 *UI* [2022] ECR I-756 Opinion AG Manuel Campos Sánchez-Bordona paras 69, 70.
[1925] Case C-300/21 *UI* [2022] ECR I-756 Opinion AG Manuel Campos Sánchez-Bordona para 72.

5.8 Erasure

The right to erasure of personal data, as currently understood by courts and regulators, relies on conceptions of how human memories work and how they 'forget'. However, the seemingly easy request to erase or 'forget' personal data poses various practical problems in the context of AI, arguably on the edge of impossibility.[1926]

5.8.1 Legal problems: Type 1

ML models are trained with historical personal data to make predictions and inferences about the future.[1927] Data deletion[1928] in the context of AI is very complex, and machines could be considered as unable to 'forget' because they must be able to go back to an older state of the system in order to be compliant with technical requirements, for example compliance with provisions with respect to databases.[1929] ML models can remember data they have been trained on or - in some cases - simply store it as part of the models.[1930] ANNs unintendedly memorise training data, which is convincingly demonstrated in experiments conducted by researchers at the Berkeley Artificial Intelligence Research Centre.[1931] With a generative text model trained on a data set including one piece of personal data in the form of a credit card number, it is possible to extract the latter from the model itself completely. Thus, where predictive ML models are trained with personal data of users, the models can unexpectedly disclose such personal data, in the case of an ANN in particular. The ANN quickly memorises data contained in the training set, even when these values are rare and the models do not overfit in the traditional sense[1932] (see also Section 4.7.1).

Personal data used as training data for an ML system might, in some cases, be reconstructed from an ML model, for example, by means of model inversion. This permits the training data to be estimated. Membership-inference recovers information to figure out whether or not a particular data subject was

[1926] Eduard Fosch Villaronga, Peter Kieseberg, Tiffany Li 'Humans forget, machines remember: Artificial intelligence and the Right to Be Forgotten' (2018) Vol 34 Iss 2 Computer Law & Security Review 304, 305, 313.
[1927] Asia Biega, Michèle Finck, 'Reviving Purpose Limitation and Data Minimisation in Data-Driven Systems' (2021) Technology and Regulation 44, 60 <https://arxiv.org/abs/2101.06203> accessed 8 February 2024.
[1928] I use the term 'deletion' as a synonym for erasure.
[1929] Eduard Fosch Villaronga, Peter Kieseberg, Tiffany Li 'Humans forget, machines remember: Artificial intelligence and the Right to Be Forgotten' (2018) Vol 34 Iss 2 Computer Law & Security Review 304, 305, 313.
[1930] Jef Ausloos, Michael Veale, René Mahieu, 'Getting Data Subject Rights Right' (2019) Vol 10 Iss 3 JIPITEC 283, 295-296; Michael Veale et al, 'Algorithms that Remember: Model Inversion Attacks and Data Protection Law' (2018) A 376 Philosophical Transactions of the Royal Society A 376.
[1931] Nicholas Carlin, 'Evaluating and Testing Unintended Memorization in Neural Networks' (*Berkeley Artificial Intelligence Research Blog*, 13 August 2019) < https://bair.berkeley.edu/blog/2019/08/13/memorization/ > accessed 8 February 2024
[1932] Nicholas Carlini et al, 'The Secret Sharer: Evaluating and Testing Unintended Memorization in Neural Networks' (USENIX Security Symposium, Santa Clara, August 2019) <https://arxiv.org/abs/1802.08232> accessed 8 February 2024; Nicholas Carlin, 'Evaluating and Testing Unintended Memorization in Neural Networks' (*Berkeley Artificial Intelligence Research Blog*, 13 August 2019) < https://bair.berkeley.edu/blog/2019/08/13/memorization/ > accessed 8 February 2024.

in the training set.[1933] Making ML systems forget is a difficult challenge.[1934] It is not even straightforward to detect that a training algorithm attempts to memorise personal data within the ML model although there are several techniques and places for encoding such information.[1935] Having ML models forget necessitates knowledge of exactly how individual training points contributed to model parameter updates. This is possible when the algorithm queries the data in a previously defined order. However, when the data are queried adaptively, the divergence induced is bounded only for relatively simple models, which require a small number of iterations for learning. However, efficient approaches for complex models such as ANNs introduced in Section 2.2.1.4 do not yet exist.[1936] If individuals request the deletion of their personal data initially used as training data for the ML model, there are basically two ways for the erasure of personal data *and* what the ML model has learnt from it. These are re-training or amending the ML model by means of machine *un*learning.[1937]

For most of the standard ML models, the only way to completely remove an individual's personal data is to retrain the whole model from scratch on the remaining data.[1938] From a computational perspective, re-training the affected ML models is inefficient and typically also requires one to re-access the original training data and redeploy the retrained model. [1939] Such re-training is considered to constitute a naïve way to have ML models provably forget due to the large computational and time overhead associated with it.[1940] It leads to significant efforts in terms of costs, time, labour and energy consumption and is therefore a rather burdensome task for the controller.[1941] Re-training is computationally often not practical because large-scale algorithms can take weeks to train and learning algorithms known to support fast data deletion operations are scarce.[1942] Ultimately, requiring a controller to retrain a prediction model could create a vicious circle, in particular when many data subjects want

[1933] Michael Veale et al, 'Algorithms that Remember: Model Inversion Attacks and Data Protection Law' (2018) A 376 Philosophical Transactions of the Royal Society A 376, 2 and 4.
[1934] Yinzhi Cao, Junfeng Yang, 'Towards Makign Systems Forget with Machine Unlearning' (IEEE Symposium on Security and Privacy, San Jose, May 2015) 464 <https://www.ieee-security.org/TC/SP2015/papers-archived/6949a463.pdf> accessed 8 February 2024.
[1935] Congzheng Song, Thomas Ristenpart, Vitaly Shmatikov, 'Machine Learning Models that Remember Too Much' (2017) in Bhavani Thuraisingham et al (eds) Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security, CCS 2017 Dallas US, 587, 598.
[1936] Lucas Bourtoule et al, 'Machine Unlearning' (IEEE Symposium on Security and Privacy, San Jose, May 2021) 1 <https://arxiv.org/abs/1912.03817> accessed 8 February 2024.
[1937] Michael Veale et al, 'Algorithms that Remember: Model Inversion Attacks and Data Protection Law' (2018) A 376 Philosophical Transactions of the Royal Society A 376, 9.
[1938] Antonio Ginart et al, 'Making AI Forget You: Data Deletion in Machine Learning', Advances in Neural Information Processing Systems (2019) 1 <https://proceedings.neurips.cc/paper/2019/file/cb79f8fa58b91d3af6c9c991f63962d3-Paper.pdf > accessed 8 February 2024.
[1939] Sebastian Schelter, 'Amnesia – A Selection of Machine Learning Models That Can Forget User Data Very Fast' (Conference on Innovative Data Systems, Amsterdam, January 2020) <http://cidrdb.org/cidr2020/papers/p32-schelter-cidr20.pdf> accessed 8 February 2024.
[1940] Lucas Bourtoule et al, 'Machine Unlearning' (IEEE Symposium on Security and Privacy, San Jose, May 2021) 1 <https://arxiv.org/abs/1912.03817> accessed 8 February 2024.
[1941] Michael Veale et al, 'Algorithms that Remember: Model Inversion Attacks and Data Protection Law' (2018) A 376 Philosophical Transactions of the Royal Society A 376, 9.
[1942] Antonio Ginart et al, 'Making AI Forget You: Data Deletion in Machine Learning', Advances in Neural Information Processing Systems (2019) 2 <https://proceedings.neurips.cc/paper/2019/file/cb79f8fa58b91d3af6c9c991f63962d3-Paper.pdf> accessed 8 February 2024.

to erase their personal data. It would lead to less training data and consequently lower accuracy[1943] which ultimately negatively affects the accuracy principle as outlined in Section 3.3.3.6.

The second approach, 'machine unlearning', might be described as the process to revert the effects of the training data on the extracted features and models.[1944] Because ML models might memorise personal data used for training (training data), these models must unlearn what they have learnt from data that must be deleted. Machine unlearning assures that the ML model is no longer trained using the personal data to be deleted.[1945] It has been argued that machine unlearning is rarely possible in modern systems and that methods currently available cannot be retrofitted onto existing systems.[1946] Any ML model trained with personal data may have memorised it and having ML models unlearn is notoriously difficult. First, there is a rather limited understanding of how each data point impacts the ML model because work that measures the influence of a particular training point on the parameters of a model is scarce if not to say non-existent.[1947] This argument particularly applies to complex models based on DL and ANNs and the problems described in Section 2.2.1.4, as it seems impossible to understand what happened in the intermediate (hidden) layers of the ANN.[1948] The second reason is stochasticity. This refers to the lack of any predictable order or plan in the training methods for complicated models such as DL and ANNs. Third, training is an incremental process in which any given update reflects all updates that have occurred previously. For example, if a model is updated based on a particular training data point at a particular time, all subsequent model updates will depend implicitly on this training point.[1949] Approaches for quick 'machine unlearning' are relatively unexplored and not ready for deployment.[1950] Thus machine unlearning seems not to be readily available due to technological difficulties. It is, however, subject to ongoing research.[1951]

There seems to be a disconnect between the right to erasure and the technical reality[1952] in the context of AI and particularly ML. Approaches to remove personal data from ML models do not seem to be

[1943] The more data are fed into the algorithm, the better the performance of the algorithm, namely the accuracy rate of the prediction model.
[1944] Yinzhi Cao, Junfeng Yang, 'Towards Making Systems Forget with Machine Unlearning' (IEEE Symposium on Security and Privacy, San Jose, May 2015) 464 <https://www.ieee-security.org/TC/SP2015/papers-archived/6949a463.pdf> accessed 8 February 2024.
[1945] Lucas Bourtoule et al, 'Machine Unlearning' (IEEE Symposium on Security and Privacy, San Jose, May 2021) 1 <https://arxiv.org/abs/1912.03817> accessed 8 February 2024.
[1946] Michael Veale et al, 'Algorithms that Remember: Model Inversion Attacks and Data Protection Law' (2018) A 376 Philosophical Transactions of the Royal Society A 376, 9.
[1947] Lucas Bourtoule et al, 'Machine Unlearning' (IEEE Symposium on Security and Privacy, San Jose, May 2021) 1 and 3 <https://arxiv.org/abs/1912.03817> accessed 8 February 2024.
[1948] Ethem Alpaydin, *Machine Learning: The New AI* (3rd edn MIT Press 2016) 155.
[1949] Lucas Bourtoule et al, 'Machine Unlearning' (IEEE Symposium on Security and Privacy, San Jose, May 2021) 3 <https://arxiv.org/abs/1912.03817> accessed 8 February 2024.
[1950] Michael Veale et al, 'Algorithms that Remember: Model Inversion Attacks and Data Protection Law' (2018) A 376 Philosophical Transactions of the Royal Society A 376, 9.
[1951] For an overview see Asia Biega, Michèle Finck, 'Reviving Purpose Limitation and Data Minimisation in Data-Driven Systems' (2021) Technology and Regulation 44, 60 <https://arxiv.org/abs/2101.06203> accessed 8 February 2024.
[1952] Eduard Fosch Villaronga, Peter Kieseberg, Tiffany Li 'Humans forget, machines remember: Artificial intelligence and the Right to Be Forgotten' (2018) Vol 34 Iss 2 Computer Law & Security Review 304, 305, 313.

technically or economically feasible at present. Challenges concerning the reliable deletion of personal data are not constrained to ML models, but extend to the entire data management lifecycle, including data replication when run in cloud environments.[1953] From a computational perspective, re-training seems to be impractical due to the significant effort needed in terms of time, labour and energy consumption. Also, amending ML models after training seems to be technically unfeasible because research is still ongoing in this area, and the scarce approaches are arguably not yet ready for deployment. Ultimately, this violates the right to erasure and leads to a Type 1 legal problem.

The training data problem (Type 1)

When data subjects submit requests to delete their personal data used for the purpose of training ML models, it will in most cases technically not be feasible for the controller to delete such personal data, re-train or unlearn the ML model in question or alternatively anonymise the personal data. The right to erasure enshrined in Article 17 GDPR will then be violated.

At first sight, this problem might be solved by rendering the personal data to be erased anonymous. Recital 26 GDPR describes anonymous data as 'information which does not relate to an identified or identifiable natural person or to personal data rendered anonymous in such a manner that the data subject is not or no longer identifiable'. In the view of the CJEU, anonymisation hinges on whether identification is 'practically impossible on account of the fact that it requires a disproportionate effort in terms of time, cost and man-power, so that the risk of identification appears in reality to be insignificant'.[1954] Recital 26 outlines what must be considered to determine whether means are 'reasonably likely' to be used for identification: all objective factors, such as costs, the amount of time required for identification, available technology at the time of the processing and technological developments.[1955] In particular, technological developments and related research indicate that there is no solid technical basis for assuming de-identification that will be effective in the long run.[1956] Perfect anonymisation is often unfeasible if not impossible[1957] and computer scientists already warned more than a decade ago that de-identification of personal data constitutes an 'unattainable goal'.[1958] In light of the technological developments, many data formats simply cannot be anonymised, which particularly

[1953] Sebastian Schelter, 'Amnesia – A Selection of Machine Learning Models That Can Forget User Data Very Fast' (Conference on Innovative Data Systems, Amsterdam, January 2020) 9 <http://cidrdb.org/cidr2020/papers/p32-schelter-cidr20.pdf> accessed 8 February 2024.
[1954] Case C-582/14 *Breyer* [2016] ECR I-779 para 46.
[1955] See also Article 29 Working Party, 'Opinion 05/2014 on Anonymisation Techniques' (10 April 2014).
[1956] Jef Ausloos, Michael Veale, René Mahieu, 'Getting Data Subject Rights Right' (2019) Vol 10 Iss 3 JIPITEC 283, 294; Arvind Narayanan et al, 'A Precatuionary Approach to Big Data Privacy' in Serge Gutwirth et al (eds) *Data Protection on the Move: Current Developments in ICT and Privacy/Data Protection* (Springer Netherlands 2014); Solon Barocas, Helen Nissenbaum, 'Big Data's End Run around Anonimity and Consent' in *Privacy, Big Data, and the Public Good: Frameworks for Engagement* (Cambridge University Press 2014).
[1957] Michèle Finck, Frank Pallas 'They who must not be identified- distinguishing personal from non-personal data under the GDPR' (2020) Vol 10 No 1 International Data Privacy Law 12.
[1958] Arvind Narayanan and Vitaly Shmatikov, 'Myths and Fallacies of Personally Identifiable Information' (2010) 53 Communications of the ACM 24, 26.

holds true in the case of ML models. These can remember data on which they have been trained or in some cases simply store it as part of their models.[1959]

5.8.2 Legal problems: Type 2

The training data problem outlined in Section 5.8.1 automatically leads to a Type 2 legal problem. If controllers cannot erase personal data used to train ML models, the right to erasure cannot be enforced. This constitutes a Type 2 legal problem.

Article 17 (1) lit d GDPR allows data subjects to request the erasure of their personal data if these have been unlawfully processed. This provision constitutes a general clause for data subjects to request the erasure of their personal data if the processing thereof does not comply with the GDPR in a broad sense.[1960] Based on Article 17 (1) lit d GDPR, data subjects may obtain the erasure of inaccurate personal data. According to the CJEU, the accuracy of personal data constitutes one of the 'conditions of lawfulness'.[1961] Also, Recital 65 GDPR supports this interpretation by stating that 'the data subject has the right to have his or her personal data erased […] where the processing of his or her data does not *otherwise comply* with this Regulation'.

Article 17 (1) lid d GDPR is closely intertwined with the right of access according to Article 15 GDPR, which enables the data subject to verify the lawfulness[1962] and allows one to obtain the rectification, erasure or blocking of its personal data by the controller.[1963] In Section 5.6, I have outlined that input data as well as output data produced by AI, including personal data generated by it, is likely to fall under trade secrets protection and that controllers can therefore restrict access to such personal data. This has a knock-on effect on the entire data protection law regime[1964] and particularly regarding the enforcement of data subject rights such as the right to erasure. The CJEU repeatedly stressed the importance of the right of access as a prerequisite to other data protection rights.[1965] Limitations on the right of access have significant consequences for the right to erasure, because it will be hardly possible for data subjects concerned to assess compliance with the GDPR and subsequently request the erasure of personal data in case of detected non-compliance. Non-compliance is likely to occur as indicated by the various Type 1 legal problems discussed in Chapter 4. For example, the principle of

[1959] Jef Ausloos, Michael Veale, René Mahieu, 'Getting Data Subject Rights Right' (2019) Vol 10 Iss 3 JIPITEC 283, 295-296.
[1960] Herke Kranenborg, Commentary of Article 17 in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020) 481.
[1961] Case C-136/17, *GC and Others* [2019] ECR I-773 para 64; see also Case C-460/20, *TU* [2022] ECR I-962 Opinion AG Pitruzella para 32.
[1962] Recital 63 GDPR.
[1963] Joined Cases C-141/12 & C-372/12, *YS, M* [2014] ECR I-2081 para 44.
[1964] Jef Ausloos, Michael Veale, René Mahieu, 'Getting Data Subject Rights Right' (2019) Vol 10 Iss 3 JIPITEC 283, 285.
[1965] Case C-434/16, *Nowak* [2017] ECR I-994 para 57; Joined Cases C-141/12 & C-372/12, *YS, M* [2014] ECR I-2081 para 44; Case C-553/07 *Rijkeboer* [2009] ECR I-03889, para 51.

fairness (Section 4.3.1) and accuracy (Section 4.7.1) is likely to be violated by the AI disciplines ML and AC. Consequently, data subjects cannot enforce their right to erasure and request the controller to delete their personal data unlawfully processed by means of AI systems. This constitutes a Type 2 legal problem.

The erasure problem (Type 2)

Access requests to personal data generated and otherwise processed by means of AI can be denied due to trade secret protection. This has a knock-on effect for the right to erasure, because data subjects cannot verify the lawfulness of such processing. As indicated by the various Type 1 legal problems identified in Chapter 1, non-compliance is likely to occur when personal data are processed by AI systems. Consequently, data subjects cannot request the erasure of personal data unlawfully processed as enshrined in Article 17 (1) lit d GDPR.

5.8.3 Legal problems: Type 3

No Type 3 legal problems arise when the right to erasure is applied to the AI disciplines introduced in Chapter 2. This is mainly due to the broad wording contained in Article 17 (1) lit d GDPR,[1966] which allows data subjects to request the erasure of personal data that 'have been unlawfully processed'. Data subjects may enforce their right to erasure according to Article 17 (1) lit d GDPR regarding all Type 1 legal problems identified in this thesis, provided that the violation in question concerns the GDPR and no exception enshrined in Article 17 (3) GDPR applies. However, there is one important caveat. As mentioned in Section 5.7, the data subject bears the burden of proof to establish the manifest inaccuracy of the information in question. The CJEU seems to place the emphasis on *factual* evidence. The CJEU ruled that facts, in particular, are susceptible to provable evidence. [1967] In Section 5.7.2, I have outlined that it is extremely difficult, not to say impossible, for data subjects to provide factual evidence for unverifiable personal data generated by means of AI (e.g. predictions or emotion data).

5.9 Portability

The right to data portability enshrined in Article 20 GDPR enables data subjects to transfer personal data among controllers[1968] and to the data subject's own systems. Recital 68 GDPR emphasises its strong connection with the legislative objective to strengthen the data subjects' control over their own personal data.[1969] As outlined in Section 4.4.3, the main mechanism for data subjects to exercise control over the processing of their personal data under the GDPR are data subject rights. The right to

[1966] Herke Kranenborg, Commentary of Article 17 in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020) 481.
[1967] Case C-460/20, *TU* [2022] ECR I-962 para 66.
[1968] Inge Graef, Martin Husovec, Nadezhda Purtova 'Data Portability and Data Control: Lessons for an Emerging Concept in EU Law' (2018) Vol 19 Iss 6 German Law Journal 1359, 1364.
[1969] Recitals 7 and 68 GDPR.

data portability empowers data subjects to exercise control as it facilitates to move, copy or transmit personal data easily from one IT environment to another, regardless of whether this refers to the data subject's own systems or the systems of others (e.g., other controllers).[1970] The wording of Article 20 (1) GDPR indicates that the right is twofold meaning that the data subject has the right to receive the personal data '*and*' the right to transmit those to another controller. Article 20 (2) GDPR states that the data subject may request the controller to transfer the personal data *directly* to another controller, which would be obsolete if Article 20 (1) GDPR would mean to exclude the possibility to have the data transferred to the data subject's own system.

5.9.1 Legal problems: Type 1

No Type 1 legal problems arise when the right to data portability is applied to the AI disciplines introduced in Chapter 2. As will be outlined in Sections 5.9.2 and 5.9.3, the right to data portability is particularly problematic with regard to its enforcement and scope.

5.9.2 Legal problems: Type 2

As outlined in the copy problem discussed in Section 5.6, trade secret protection under the TSD covers AI itself as well as output generated by the AI system, including emotional states and life expectancy predictions. Like the right of access (Section 5.6.2), the right to data portability contains a provision that enables the controller to restrict the right to data portability on a case-by-case basis. Article 20 (4) GDPR states that the right to data portability 'shall not adversely affect the rights and freedoms of others'. This gives controllers more leeway and flexibility in restricting data portability requests by means of Article 20 (4) GDPR, because these restrictions do not have to be enshrined in EU or MS law.[1971] Therefore, controllers could argue that the transmission of personal data constituting the output of AI systems from one IT environment to another (thus to the data subject or another controller) infringes their trade secrets and refuse to transmit such data. Consequently, data subjects cannot enforce their right to data portability, which constitutes a Type 2 legal problem. Because the broad scope of protection under the TSD applies to *all AI* disciplines as introduced in Chapter 2, this Type 3 legal problem constitutes a general problem and relates to all AI disciplines discussed in Chapter 2.

The transmission problem (Type 2)

Due to the broad scope of trade secrets protection in the EU, controllers are likely to argue that the transmission of personal data constituting outputs generated by AI systems from one IT system to another infringes their trade secrets. Consequently, data subjects cannot enforce their right to data portability.

[1970] Article 29 Working Party, 'Guidelines on the right to data portability' (WP 242rev.01, 5 April 2017) at 4.
[1971] As it is the case of restrictions made on the basis of Article 23 GDPR.

5.9.3 Legal problems: Type 3

The text of Article 20 (1) GDPR limits the scope of the right to data portability in two ways. First, the right applies only to processing of personal data based on the lawful bases of consent or performance of a contract and therefore not to processing which is based on a controller's legitimate interest.[1972] Second, the scope of the right is limited to personal data that are 'provided by' the data subject[1973] which only refers to personal data actively and knowingly disclosed by the data subject. Examples mentioned in regulatory guidance include the email address or user name submitted via online forms, the photos and videos uploaded on social media and personal data *observed* by the controller. The latter, according to the regulator, includes raw personal data observed in the context of the use of the service or device, for example, search history, traffic data, location data and heartbeat, all tracked by a wearable device.[1974] According to both regulatory guidance and the European Commission, [1975] observed data constitutes 'raw data' and *excludes* personal data generated by the controller.[1976] Regulatory guidance specifically mentions that data generated by the controller, such as a user profile created by analysis of raw data collected by the controller, does *not* fall under the notion of personal data 'provided by the data subject'. Thus, regulatory guidance explicitly excludes *inferred* and *derived* personal data from the scope of the right to data portability.[1977] As I outline in the following paragraphs, this limitation is significant regarding processing of personal data in the context of AI.

Regulatory guidance does not further explain the two terms 'inferred' and 'derived' personal data but indicates that this may include 'algorithmic results'.[1978] It seems that the regulatory guidance relies on a paper published by the OECD which introduces a data taxonomy distinguishing between four categories: provided, observed, derived and inferred data.[1979] Derived data are described as 'data generated from other data, after which they become new data elements related to a particular individual' created by simple reasoning and basic mathematics to detect patterns and create classifications (e.g. detection of common attributes among profitable customers used for classification). Inferred data are defined as 'the product of probability-based processes' and used, for instance, to create predictions of behaviour deployed to categorise individuals.[1980] Unlike derived data, inferred data are based on probabilistic reasoning and may include 'statistical data' (e.g., credit risk scores, life expectancy scores) and

[1972] Art. 6 (1) lif f GDPR.
[1973] Art. 20 (1) GDPR, Recital 68.
[1974] Article 29 Working Party, 'Guidelines on the right to data portability' (WP 242rev.01, 5 April 2017) at 10.
[1975] See letter from Member of the European Commission Věra Jourová to Chairman of WP29 (2017) < https://zwenneblog.weblog.leidenuniv.nl/files/2018/06/Letter-Cssr-Jourova-to-Falque-Pierrotin.pdf > accessed 8 February 2024.
[1976] Article 29 Working Party, 'Guidelines on the right to data portability' (WP 242rev.01, 5 April 2017) at 10.
[1977] Article 29 Working Party, 'Guidelines on the right to data portability' (WP 242rev.01, 5 April 2017) at 10.
[1978] Ibid.
[1979] OECD Working Party on Security and Privacy in the Digital Economy JT03357584 (2014) 5 <https://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=dsti/iccp/reg(2014)3&doclanguage=en> accessed 8 February 2024.
[1980] Ibid.

'advanced analytical data' (e.g., likelihood of future health outcomes based on analysis of extensive medical data sets).[1981]

Considering the AI discipline ML as introduced in Section 2.2.1, I take the view that ML-generated data most likely constitutes both derived and/or inferred personal data. ML applies data-driven methods, combining fundamental concepts in computer science with approaches from statistics, probability and optimisation[1982] and is used for classification and the detection of patterns and predictions. Unsupervised ML detects patterns by means of clustering and dimensionality reduction techniques. Supervised ML uses classification and regression techniques. Recommendations or predictions generated by ML, for example, personalised suggestions for Netflix users[1983] or predictions concerning one's sexual orientation[1984] either constitute derived or inferred personal data. Additionally, the AI discipline CV applies basic mathematics and might produce derived data. In particular, face recognition systems rely on the mathematical concept convolution, which is considered a specialised kind of linear operation (see Section 2.2.3.2). Models combining CV and ML disciplines and applying convolutional ANNs and regression techniques were able to predict sexual orientation from dating profile photographs.[1985] In addition, personal data generated by systems relying on any other discipline of AI combined with ML approaches might constitute derived or inferred personal data that falls outside the scope of application of the right to data portability.

Consider, for example, emotion data generated by an AI system using AC and ML. Regulatory guidance states that data generated by the controller's algorithms, including derived or inferred profiles and the outcome of an assessment, personalisation or recommendation process, are excluded from the scope of the right to data portability. According to the regulator, this limitation also applies to inferred or derived personal data which relate to special categories of personal data, for example data concerning health.[1986] Thus, personal data derived and inferred by means of the AI disciplines CV, AC, ML and potentially any other AI discipline combined with ML does not fall within the scope of the right enshrined in Article 20 GDPR.[1987] This also holds true if the personal data generated by the controller with the help of AI constitutes special data according to Article 9 (1) GDPR.

[1981] OECD Working Party on Security and Privacy in the Digital Economy JT03357584 (2014) 5 <https://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=dsti/iccp/reg(2014)3&doclanguage=en> accessed 8 February 2024.
[1982] Mehryar Mohri, Afshin Rostamizadeh, Ameet Talwalkar, *Foundations of Machine Learning* (MIT Press 2012) 1.
[1983] See <https://research.netflix.com/research-area/recommendations> accessed 8 February 2024.
[1984] John Leuner, 'A Replication Study: Machine Learning Models Are Capable of Predicting Sexual Orientation From Facial Images' (2018) <https://arxiv.org/pdf/1902.10739.pdf> accessed 8 February 2024.
[1985] Ibid 52.
[1986] Article 29 Working Party, 'Guidelines on the right to data portability' (WP 242rev.01, 5 April 2017) at 10, 11.
[1987] Sandra Wachter, Brent Mittelstadt, 'A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI' (2019) Issue 2 Columbia Business Law Review 494, 519.

Arguably, the right to data portability was drafted without considering AI systems that can generate derived or inferred personal data. Consequently, individuals have limited control over their personal data, which is contrary to the legislative aim of the right, as emphasised in Recital 68 GDPR. This recital stresses the strong connection of the right to portability of data with the legislative objective of strengthening the control of data subjects over their own personal data as propagated in Recital 7 GDPR. The right to data portability intends to empower data subjects by facilitating them to move, copy or transmit personal data easily from one IT environment to another, including their own systems.[1988] Because personal data generated by AI systems, including inferred or derived special personal data such as predictions concerning sexual orientation and mental health, do not fall under the scope of the right to data portability, individuals have no control with regard to such data. This particularly holds true when considering that data subjects even cannot obtain access to such data by means of Article 15 GDPR due to the trade secrets problem discussed in Section 5.6.2. In other words, the right of access cannot close this gap, although it is precisely the right of access that is supposed to do so. It is acknowledged that the scope of the right to data portability is intentionally limited when compared to Article 15 GDPR, as indicated by the European Commission.[1989]

The limited scope of the right to portability about personal data inferred and/or derived by the controller ultimately leads to a Type 3 legal problem. This right is not fit for purpose to achieve the GDPR's goal that 'natural persons should have control of their own personal data'.[1990] It was one of the main reasons for the data protection reform[1991] and was *specifically intended* to 'further strengthen the control over his or her own data'.[1992] As outlined in Section 4.4.3, one of the main mechanisms for data subjects to exercise control over the processing of their personal data under the GDPR are data subject rights. Due to the limited scope with regard to inferred and / or derived personal data, the right to data portability does not achieve the goal of strengthening the control of the data subject over the processing of personal data. Likewise, Article 20 GDPR fails to strengthen the rights of data subjects as envisaged by the GDPR.[1993] As noted by the CJEU, effective protection of personal data requires the strengthening of the rights of data subjects, which is emphasised by Recital 11 GDPR.[1994] A right of which the scope excludes personal data inferred and/or derived by the controller fails to strengthen the data subject's rights, in particular when considering that the right to obtain a copy of the personal data undergoing processing allows for restrictions due to trade secret protection. The CJEU has stressed the importance of ensuring that data subject rights granted by the GDPR are

[1988] Article 29 Working Party, 'Guidelines on the right to data portability' (WP 242rev.01, 5 April 2017) at 10, 11.
[1989] See letter from Member of the European Commission Věra Jourová to Chairman of WP29 (2017) page 2 < https://zwenneblog.weblog.leidenuniv.nl/files/2018/06/Letter-Cssr-Jourova-to-Falque-Pierrotin.pdf > accessed 8 February 2024.
[1990] Recital 7 GDPR.
[1991] Case C-300/21 *UI* [2022] ECR I-756 Opinion AG Manuel Campos Sánchez-Bordona paras 69, 70.
[1992] Recital 68 GDPR.
[1993] Recital 11 GDPR.
[1994] Case C-132/21 *Nemzeti* [2023] ECR I-2 para 42.

effective.[1995] However, this is not the case for the right to data portability due to the severely restricted scope concerning personal data generated by AI. This Type 3 legal problem occurs regardless of which AI discipline has been used to infer or derive personal data because the problem is caused by the restricted scope of Article 20 GDPR. It is therefore a general problem and potentially relates to all AI disciplines introduced in Chapter 2.

The restricted scope problem (Type 3)
The right to data portability excludes personal data derived and/or inferred by AI from its scope and thus fails to enhance the data subjects' control over their own personal data. The right to data portability is therefore not fit for purpose to protect the fundamental right to data protection.

5.10 Objection

As introduced in Section 3.3.4.5, Article 21 (1) GDPR provides the data subject a right to object to processing 'on grounds relating to his or her particular situation'. Simultaneously, it imposes a duty on the controller to cease processing unless it can demonstrate 'compelling legitimate grounds for the processing' which override the interests, rights and freedoms of the data subject *or* for the establishment, exercise or defence of legal claims.[1996] The right to object *exclusively* applies to processing based on the legal ground 'performance of a task carried out in the public interest' according to Article 6 (1) lit e and legitimate interest according to Article 6 (1) lit f GDPR. Data subjects do not have a right to object to processing if controllers rely on legal grounds other than those mentioned.

5.10.1 Legal problems: Type 1

As described in Section 4.2.1, AI has the potential to determine why and how to process personal data due to its autonomous and adaptive characteristics. The balancing problem explained in Section 4.2.1 also applies to the right to object because processing based on the legal basis of 'legitimate interest' constitutes one of the two grounds on which data subjects can exercise this right. In essence, the balancing problem refers to the incapability of autonomous AI systems to appropriately balance the fundamental rights and freedoms of the parties involved in accordance with the Legitimate Interest Assessment (LIA) and the proportionality principle (Sections 4.2.1 and 3.3.2 respectively). This is caused by the reasoning deficiencies in the AI discipline AR. The said problem also applies to the balance of interests that a controller must perform in order to demonstrate its 'compelling legitimate ground' according to Article 21 (1) GDPR.

[1995] Case C-154/21 *Österreichische Post AG* [2023] ECR I-3 para 39.
[1996] Article 21 (1) GDPR.

If a data subject objects to processing based on Article 21 (1) GDPR and the controller does not intend to cease such processing, it must be able to demonstrate its compelling legitimate ground for processing overrides the interests, rights and freedoms of the data subject.[1997] As explained in Section 3.3.4.5, the burden of proof that the conditions in Article 21 (1) are met lies with the controller.[1998] However, current AI systems have been called to be clueless[1999] to understand cause and effect and to be devoid of common sense.[2000] Common sense reasoning constitutes a major challenge in AI,[2001] particularly in the discipline of automated reasoning (see Sections 2.2.5, 4.3.1, 4.4.1 and 4.7.1). Apparently, there is not one AI system today which has a semblance of common sense or has capabilities such as human cognition. Hence, AI systems are unable to think on par with human thinking[2002] and are therefore not able (at least not in the near future) to appropriately weigh the fundamental rights and freedoms of the parties involved as required by the 'compelling legitimate ground' balancing according to Article 21 (1) GDPR.

Data processing is likely to continue after a data subject enforced the right to object. AI systems autonomously perform processing activities meaning that the AI system makes its own decisions and executes tasks on the controller's behalf.[2003] When a data subject exercises the right to object, whether successful or not, the controller must *immediately* restrict the processing pursuant to Article 18 (1) lit d GDPR.[2004] It is unlikely that a controller immediately restricts the processing of personal data. In addition, there is arguably not 'one' command that the controller can execute that immediately restricts all relevant processing activities that occur in the complex environment of AI systems. Take, for example, a supermarket chain that processes personal data of its customers by means of an ML-powered system to obtain valuable insights about the personal aspects of the customers based on purchase history. The supermarket relies on its legitimate interest according to Article 6 (1) lit f GDPR as the legal ground for such processing. Based on two dozen products used as proxies, the powerful ML prediction model identifies pregnant customers. After becoming aware, one customer objected to such processing according to Article 21 (1) GDPR. The processing performed by the ML-powered

[1997] Article 21 (1) GDPR.
[1998] Gabriela Zanfir-Fortuna, Commentary of Article 21 in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020) 517.
[1999] Brian Bergstein, 'What AI still can't do' MIT Technology Review (Cambridge 31 January 2020) <https://www.technologyreview.com/2020/01/31/304844/ai-common-sense-reads-human-language-ai2/> accessed 8 February 2024.
[2000] Cade Metz, 'Paul Allen Wants to Teach Machines Common Sense' *The New York Times* (New York, 28 February 2018) <https://www.nytimes.com/2018/02/28/technology/paul-allen-ai-common-sense.html> accessed 09 November 2019.
[2001] Shoham Yoav et al, 'The AI Index 2018 Annual Report' (AI Index Steering Committee Stanford University 2018) 64 <https://hai.stanford.edu/sites/default/files/2020-10/AI_Index_2018_Annual_Report.pdf > accessed 8 February 2024.
[2002] Lance Eliot, 'AI Ethics And The Quagmire Of Whether You Have A Legal Right To Know Of AI Inferences About You, Including Those Via AI-Based Self-Driving Cars' *Forbes* (New York, 25 May 2022) < https://www-forbes-com.cdn.ampproject.org/c/s/www.forbes.com/sites/lanceeliot/2022/05/25/ai-ethics-and-the-quagmire-of-whether-you-have-a-legal-right-to-know-of-ai-inferences-about-you-including-those-via-ai-based-self-driving-cars/amp/> accessed 8 February 2024.
[2003] Eduardo Alonso, 'Actions and agents' in Frankish Keith and Ramsey William M. (eds) *The Cambridge Handbook of Artificial Intelligence* (2014) 235, 236.
[2004] Gabriela Zanfir-Fortuna, Commentary of Article 21 in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020) 518.

system is complex and entangled, and the customer's personal data are incorporated into the system's models. As mentioned in Section 5.8.1, ML could store personal data as a part of its models.[2005] But the system processes also personal data of all other customers of the supermarket, who did not object to such processing. It seems rather unlikely that the supermarket chain shuts down the whole ML system, simply because one customer objected to such processing. Instead, the supermarket may consider to retrain the ML models to cease the processing of personal data relating to the customer who successfully objected to it. However, such re-training is computationally burdensome because large-scale algorithms can take weeks to train.[2006] In any case, the controller will not be able to *immediately* restrict processing pursuant to Article 18 (1) lit d GDPR.[2007]

Due to the balancing problem explained in Section 4.2.1, AI systems cannot balance interests to demonstrate the controller's 'compelling legitimate ground' according to Article 21 (1) GDPR. This is mainly caused by the reasoning deficiencies in the AI discipline AR explained in Section 4.3.1. Processing is likely to continue after the data subject enforced its right to object because AI systems autonomously perform processing activities.[2008] Controllers are unable to *immediately* restrict the processing pursuant Article 18 (1) lit d GDPR because processing performed by AI is complex. In addition, ML systems process personal data of various data subjects, and it seems unlikely that controllers shut down a whole system simply because only one data subject enforced its right to object. Therefore, processing of personal data does not cease, but continues. This violates the right to object according to Article 21 (1) GDPR. This Type 1 legal problem applies to all AI disciplines as introduced in Chapter 2 because the ability to autonomously make decisions and execute tasks on the designer's behalf[2009] constitutes a key element of AI (see Section 2.1).

The continuance problem (Type 1)

The balancing problem introduced in Section 4.2.1 also applies to the 'compelling legitimate ground' balancing required by Article 21 (1) GDPR allowing data subjects to object to processing performed by autonomous AI systems. Because AI systems make their own decisions and execute tasks independently, processing of personal data can continue after the data subject has enforced its right to object. Because processing performed by AI systems is highly entangled and complex, controllers cannot immediately restrict processing as required by Article 18 (1) lit d GDPR. Consequently, the right to object is violated.

[2005] Jef Ausloos, Michael Veale, René Mahieu, 'Getting Data Subject Rights Right' (2019) Vol 10 Iss 3 JIPITEC 283, 295-296.
[2006] Antonio Ginart et al, 'Making AI Forget You: Data Deletion in Machine Learning', Advances in Neural Information Processing Systems (2019) 2 <https://proceedings.neurips.cc/paper/2019/file/cb79f8fa58b91d3af6c9c991f63962d3-Paper.pdf> accessed 8 February 2024.
[2007] Gabriela Zanfir-Fortuna, Commentary of Article 21 in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020) 518.
[2008] Eduardo Alonso, 'Actions and agents' in Frankish Keith and Ramsey William M. (eds) *The Cambridge Handbook of Artificial Intelligence* (2014) 235, 236.
[2009] Ibid.

5.10.2 Legal problems: Type 2

No Type 2 legal problems arise when the right to object is applied to the AI disciplines introduced in Chapter 2. This is mainly because the controller must demonstrate a compelling legitimate ground if the controller intends to continue with the processing of personal data after the data subject has enforced its right to object. Thus, the burden of proof is imposed on the controller. In addition, data subjects may object to processing for direct marketing unconditionally: no conditions are attached to effectively enforce this right. The data subject simply needs to object to processing for direct marketing purposes to be successful.[2010]

5.10.3 Legal problems: Type 3

AI provides powerful tools to infer and otherwise generate personal data. Such data provides controllers with valuable insights about data subjects, their personal aspects in particular. Controllers may use AI for profiling as defined in the GDPR. Article 4 (4) GDPR defines profiling as 'any form of automated processing of personal data consisting of the use of personal data to evaluate certain personal aspects relating to a natural person'. Inferred data generated by ML may ascribe attributes to individuals using ML techniques such as regression, classification (see Section 2.2.1.1) or clustering (Section 2.2.1.2) and thus amount to profiling as defined in the GDPR. ML infers personal data by detecting patterns and correlations and making predictions, such as likelihood of pregnancy, life expectancy or credit risks (see Section 4.4.1). AC as introduced in Section 2.2.4 generates personal data which indicates the emotional state of a data subject. Processing through AC amounts to profiling as defined in the GDPR because it evaluates a particular aspect of the data subject, namely, his emotional state exhibited during a given activity (for example, during the data subject's conversation with its virtual assistant). When the right to object according to Article 21 (1) GDPR is applied to profiling, problems arise regarding the subsequent erasure of inferred personal data in cases in which the rights and interests of the data subject prevail.

Let me explain this through the supermarket's ML-powered system introduced in Section 5.10.1, which infers valuable information about the personal aspects of its customers based on their purchase history. After identifying pregnant customers through the powerful ML system, the supermarket sends them a targeted email announcement and offers vouchers for baby food. One of the customers concerned, a 21-year-old student still living at home, is rather upset and considers the marketing communication of the supermarket very intrusive. She is also very concerned that her parents will learn about her unexpected pregnancy because the family shares a common account with the

[2010] Gabriela Zanfir-Fortuna, Commentary of Article 21 in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020) 517.

supermarket.[2011] Curious about her data protection rights, the customer consults the supermarket's privacy notice and objects to the processing of her personal data for marketing purposes according to Article 21 (2) GDPR. According to Articles 12 (3) and 21 (3) GDPR, the supermarket confirms to the data subject by email that it does not process information about her pregnancy for marketing purposes. Understandably, the customer assumes that the supermarket has erased this sensitive information.

However, the conclusion drawn by the customer is incorrect. Following a successful objection according to Article 21 (2) GDPR, the personal data are not erased from the supermarket's systems. On the contrary, the wording contained in Article 21 (3) GDPR points to the possibility of processing for *other* purposes[2012] because the provision states that 'personal data shall no longer be processed for such [direct marketing] purposes'. To do so, the supermarket needs to comply with all the requirements of the GDPR, in particular the data protection principles introduced in Section 3.3.3. Nevertheless, as already outlined in Section 4.5.3, if controllers make an effort to define purposes with sufficient specificity and can demonstrate that such purposes are legitimate, any purpose is a valid purpose under the GDPR.[2013] This holds particularly true given the lack of judicial guidance with respect to the relevant criteria for determining the precision of the purpose.[2014] Thus, it is not unlikely that controllers will successfully fiddle about a new purpose. To have her personal data concerning pregnancy deleted, the customer must submit a separate erasure request based on Article 17 (1) lit c GDPR. However, even then, the supermarket may opt to only erase the pregnancy-related personal data from a dedicated list or database kept for direct marketing purposes and continue with processing for other purposes.[2015] Then, the supermarket can argue that Article 17 (1) lit a GDPR does not apply because processing is still necessary for these other purposes. This provision requires controllers to erase personal data that 'are no longer necessary in relation to the purposes for which they were collected or otherwise processed'.

The outcome is the same with respect to the profile and the personal data inferred by AC. Take, as an example, Amazon's patent introduced in Section 5.5.1, which specifically refers to AI disciplines NLP and ML (particularly ANN as applied in DL). Following the claims of this patent, Amazon's virtual assistant Alexa is able to detect a user's emotional state such as happiness, joy, anger, sorrow,

[2011] Example taken from Viktor Mayer-Schönberger; Kenneth Cukier, *Big Data: A Revolution That Will Transform How We Live, Work and Think* (Houghton Mifflin Harcourt 2013) 58.
[2012] Helena U Vrabec, 'Uncontrollable: Data Subject Rights and the Data-driven Economy' (Dissertation, Leiden University 2019) 180; Gabriela Zanfir-Fortuna, Commentary of Article 21 in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020) 518.
[2013] Asia Biega, Michèle Finck, 'Reviving Purpose Limitation and Data Minimisation in Data-Driven Systems' (2021) Technology and Regulation 44, 49 <https://arxiv.org/abs/2101.06203> accessed 8 February 2024.
[2014] Maximilian von Grafenstein, *The Principle of Purpose Limitation in Data Protection Laws* (Nomos 2017) 232, 233, 244.
[2015] Gabriela Zanfir-Fortuna, Commentary of Article 21 in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020) 518.

sadness or fear based on analysis of acoustic features as determined from the user's speech.[2016] This enables Alexa to intuitively advertise specific products based on the user's current emotional state.[2017] As in the supermarket example, if the data subject objects to the processing of its emotion data for direct marketing purposes, Amazon is not required to entirely erase such emotion data, even if the data subject hands in a separate erasure request. Amazon may simply erase such data from a dedicated list or database kept for direct marketing purposes and further process emotion data for other purposes. Obviously, such further processing for other purposes requires a corresponding assessment of the controller.

As an alternative to object to processing for direct marketing purposes according to Article 21 (2) GDPR, the customer may also object to the processing of her personal data based on Article 21 (1) GDPR. The customer could argue that on grounds relating to her particular situation, namely, that her pregnancy constitutes rather sensitive information and her parents are not yet aware of it, the controller must cease the processing of the personal data for all *conceivable or envisaged* purposes. It is unlikely that the supermarket in this case can demonstrate 'compelling legitimate grounds for processing', which override the interests, rights and freedoms of the customer. It has been argued that, if the objection of the data subject has merit (like in this particular case), the controller cannot retain the personal data in question but must erase it[2018] without undue delay.[2019] According to this view, the controller cannot retain personal data subsequent to a successful objection because storage constitutes a form of processing defined in Article 4 (2) GDPR, and, when interpreted together with Article 17 (1) lit c GDPR, imposes the obligation on the controller to erase the personal data in question, without requiring the data subject to submit a separate erasure request according to Article 17 (1) lit c GDPR.

In my view, it must be added that particularly the storage limitation principle as introduced in Section 3.3.3.7 obliges the controller to erase the personal data in question. This principle requires controllers to not store personal data longer than necessary in relation to the purpose of processing. When applied to the supermarket case, the supermarket must erase the personal data concerning the pregnancy of the customer. Processing is no longer necessary in relation to all conceivable processing purposes because the customer's rights and interests prevail. However, the view that controllers are obliged to erase personal data after a successful objection request, without a separate erasure request according to Article 17 (1) lit c GDPR, is by no means supported by CJEU case law. On the basis of a

[2016] Huafeng Jin, Shuo Wang, 'Voice-based Determination of Physical and Emotional Characteristics of Users' US Patent Number US 10096319B1 (Assignee: Amazon Technologies, Inc.) October 2018 <https://patentimages.storage.googleapis.com/f6/a2/36/d99e36720ad953/US10096319.pdf>, accessed 8 February 2024.
[2017] James Cook, 'Amazon patents new Alexa feature that knows when you're ill and offers you medicine' *The Telegraph* (London 9 October 2018) < https://www.telegraph.co.uk/technology/2018/10/09/amazon-patents-new-alexa-feature-knows-offers-medicine/> accessed 8 February 2024.
[2018] Gabriela Zanfir-Fortuna, Commentary of Article 21 in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020) 518.
[2019] Article 17 (1) lit c GDPR.

teleological interpretation, the CJEU could confirm this interpretation, but has not done so yet. However, as demonstrated by the training data problem contained in Section 5.8.1, in most cases it will technically not be feasible for the controller to delete such personal data, retrain or unlearn the ML model or alternatively anonymise the personal data.

The AI disciplines AC, NLP and ML provide controllers with powerful means for profiling and allow them to infer and otherwise generate personal data. If controllers rely on their legitimate interest for profiling and infer personal data by means of these AI disciplines, and if data subjects successfully object to this, personal data generated by AI systems will not be automatically erased and may be further processed for *other* purposes. The outcome of an objection according to Article 21 (1) and (2) GDPR varies regarding the subsequent erasure of the personal data in question. If the data subject opts to object to the processing for direct marketing purposes, the personal data inferred or otherwise generated by means of AC, NLP and ML approaches will not necessarily be entirely erased by the controller, if the latter specified another purpose for processing. If the data subject objects based on paragraph 1 instead of paragraph 2 of Article 21 GDPR, the personal data must be erased by the controller if the teleological interpretation of Articles 21 and 17 GDPR is affirmed by the CJEU. In any case, it is highly unlikely that the data subjects are aware of these legal nuances when objecting to the processing of their personal data. Data subjects are arguably more likely to rely on paragraph 2 of Article 21 GDPR because there are no conditions attached to enforce this right.[2020]

Therefore, the right to object is not fit for purpose to effectively[2021] protect the fundamental right to data protection. In its case law, the CJEU has repeatedly stressed that EU data protection law aims to effectively protect[2022] the data subject's personal data against risk of misuse.[2023] Such risk of misuse seems likely to occur when controllers are not obliged to erase highly sensitive personal data because data subjects chose paragraph 2 instead of paragraph 1 when objecting to processing according to Article 21 GDPR. Examples of such sensitive data generated by means of AI are emotion data derived by means of AC or pregnancy predictions facilitated by ML. Similarly, the legal nuances contained in Article 20 GDPR fail to achieve the GDPR's aim of enhancing the legal and practical certainty for data subjects (Recital 7). In addition, Article 21 GDPR does not achieve the GDPR's goal that 'natural persons should have control of their own personal data',[2024] although this was one of the main reasons for the data protection reform.[2025] As outlined in Section 4.4.3, enforceable rights are one of the main

[2020] Gabriela Zanfir-Fortuna, Commentary of Article 21 in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020) 518.
[2021] Recital 11 GDPR.
[2022] Case C-645/19 [2021] *Facebook Ireland* ECR I-483 para 45; Case C-319/20, *Meta Platforms Ireland Limited* [2022] ECR I-322 para 73.
[2023] Joined Cases C-293/12 and C-594/12 *Digital Rights Ireland* [2014] ECR-I 238 paras 54 and 66; Joined Cases C-203/15 and C-698/15 [2016] *Tele 2 Sverige* ECR-1 970 paras 109 and 122; Case C-362/14 *Schrems v Data Protection Commissioner* [2015] ECR I-650 para 91.
[2024] Recital 7 GDPR.
[2025] Case C-300/21 *UI* [2022] ECR I-756 Opinion AG Manuel Campos Sánchez-Bordona paras 69, 70.

mechanisms for data subjects to exercise control over the processing of their personal data. However, due to the complex legal nuances in Article 21 GDPR, data subjects cannot really exercise control over the processing of their personal data. When the data subject objects based on Article 21 (2) GDPR, personal data will not necessarily be erased, and it can be further processed for purposes other than direct marketing.

The erasure after objection problem (Type 3)
ML, NLP and AC provide controllers with powerful means for profiling. When data subjects object to such profiling, controllers are not necessarily required to erase the generated personal data because erasure depends on legal nuances of which data subjects are most likely not aware. This right is not fit for purpose to protect the fundamental right to data protection, as it fails to effectively protect data subjects from misuse and to provide data subjects with control concerning processing of profiling outcomes generated by AI for purposes other than direct marketing.

5.11 Automated decision-making

As outlined in Section 3.3.4.6, Article 22 (1) GDPR grants individuals the right 'not to be subject to a decision based only on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her'. Preparatory documents on the drafting of Article 22 GDPR provide little explanation about its rationale. It seems that the rationale is rooted in the predecessor of Article 22 GDPR, namely, Article 15 DPD. Article 15 DPD aimed to address the potential weakening of the ability of individuals to exercise influence over decision-making processes that significantly affect them considering the growth of automated decision-making (ADM) and concerns about the quality of ADM. Other concerns are the fear that ADM will cause humans to take the validity of ADM for granted, thereby reducing own responsibility to investigate the matters involved, and the concern to uphold human dignity by ensuring that humans keep their autonomy. The same concerns arguably also apply to Article 22 GDPR in addition to harms related to profiling, on which the preparatory documents of the GDPR mainly focus.[2026] This would also match with the rationale of Article 22 GDPR identified by the CJEU: protecting individuals effectively against the particular risks associated with the automated processing of personal data, including profiling.[2027] In AG Pikamäe's opinion, Article 22 GDPR aims to safeguard human dignity. It also prevents data subjects from being subject to ADM without any human intervention, which monitors whether ADM has been taken properly, fairly and without discrimination.[2028]

[2026] Isak Mendoza, Lee A Bygrave, 'The Right Not to be Subject to Automated Decisions Based on Profiling' in Tatiana-Eleni Synodinou et al (eds) *EU Internet Law: Regulation and Enforcement* (Springer International 2017) 83-84.
[2027] Case C-634/21, *SCHUFA Holding AG* [2023] ECR I-957 para 57.
[2028] Case C-634/21, *SCHUFA Holding AG* [2023] ECR I-220, Opinion AG Pikamäe para 19.

AI contributes significantly to ADM. As outlined in Section 2.2.1, ML uses data-driven methods, combining fundamental concepts in computer science with approaches from statistics, probability and optimisation in order to achieve its main goal, which is to generate accurate predictions for unseen data and to design efficient algorithms to produce these predictions.[2029] ADM may be facilitated by ML alone or in combination with other AI disciplines. In fact, ML may be fused with other AI disciplines in dedicated systems, for example, emotion detection systems which, depending on the system at hand, combine the disciplines ML, CV, NLP and AC in order to produce automated decisions concerning the data subject.

Article 22 GDPR suffers from significant weaknesses[2030] and the ambiguity and complexity of the right makes it difficult to apply in practice.[2031] The complexity also relates to the mechanics of Article 22 GDPR: The first paragraph provides for a right not to be subject to ADM, and the second paragraph provides exceptions to that right, while the third paragraph qualifies two of those exceptions by adding requirements to them ('suitable safeguards'). Finally, the fourth paragraph introduces a further qualification to all the exceptions provided in paragraph 2, i.e. a prohibition on ADM based on special categories of personal data but simultaneously provides some exceptions to this prohibition.[2032]

5.11.1 Legal problems: Type 1

As outlined in Sections 4.2.1 and 5.10.1, AI has the potential to decide itself why and how to process personal data due to its autonomous characteristics. Current AI systems have been called to be clueless[2033] to understand cause and effect and to be devoid of common sense.[2034] Common sense reasoning constitutes a major challenge in AI,[2035] particularly in the discipline AR (see Sections 2.2.5, 4.3.1, 4.4.1 and 4.7.1). Because AI systems make their own autonomous decisions[2036] about the processing of personal data and lack cognitive skills on par with human thinking,[2037] they are prone to violate the

[2029] Mehryar Mohri, Afshin Rostamizadeh, Ameet Talwalkar, *Foundations of Machine Learning* (MIT Press 2012) 1.
[2030] Lee A Bygrave, 'Minding the machine v2.0: The EU General Data Protection Regulation and Automated Decision Making' in Karen Yeung and Martin Lodge (eds) *Algorithmic Regulation* (OUP 2019) 249.
[2031] Paul De Hert, Guillermo Lazcoz, 'Radical rewriting of Article 22 GDPR on machine decisions in the AI era' *European Law Blog* (13 October 2021) <https://europeanlawblog.eu/2021/10/13/radical-rewriting-of-article-22-gdpr-on-machine-decisions-in-the-ai-era/> accessed 8 February 2024.
[2032] Isak Mendoza, Lee A Bygrave, 'The Right Not to be Subject to Automated Decisions Based on Profiling' in Tatiana-Eleni Synodinou et al (eds) *EU Internet Law: Regulation and Enforcement* (Springer International 2017) 85.
[2033] Brian Bergstein, 'What AI still can't do' MIT Technology Review (Cambridge 31 January 2020) <https://www.technologyreview.com/2020/01/31/304844/ai-common-sense-reads-human-language-ai2/> accessed 8 February 2024.
[2034] Cade Metz, 'Paul Allen Wants to Teach Machines Common Sense' *The New York Times* (New York, 28 February 2018) <https://www.nytimes.com/2018/02/28/technology/paul-allen-ai-common-sense.html> accessed 8 February 2024.
[2035] Brandon Bennet, Anthony G Cohn, 'Automated Common-sense Spatial Reasoning: Still a Hughe Challenge' in Stephen Muggleton, Nicholas Chater (eds) Human-Like Machine Intelligence (Oxford University Press 2021) 405; Gary Marcus, Ernest Davis, *Rebooting AI: Buidling Artificial Intelligence we can trust* (Pantheon Books 2019); Shoham Yoav et al, 'The AI Index 2018 Annual Report' (AI Index Steering Committee Stanford University 2018) 64 <https://hai.stanford.edu/sites/default/files/2020-10/AI_Index_2018_Annual_Report.pdf > accessed 8 February 2024.
[2036] Eduardo Alonso, 'Actions and agents' in Frankish Keith and Ramsey William M. (eds) *The Cambridge Handbook of Artificial Intelligence* (2014) 235, 236.
[2037] Lance Eliot, 'AI Ethics And The Quagmire Of Whether You Have A Legal Right To Know Of AI Inferences About You, Including Those Via AI-Based Self-Driving Cars' *Forbes* (New York, 25 May 2022) < https://www-forbes-com.cdn.ampproject.org/c/s/www.forbes.com/sites/lanceeliot/2022/05/25/ai-ethics-and-the-quagmire-of-whether-you-

prohibition of ADM enshrined in Article 22 (1) GDPR. The balancing problem explained in Section 4.2.1 outlines that autonomous AI systems cannot balance the fundamental rights and freedoms of the parties involved due to the reasoning deficiencies in the AI discipline AR. Due to exactly these reasoning deficiencies, autonomous AI systems are also not capable of assessing whether ADM produces legal or similarly significant effects for the data subjects concerned. Consequently, autonomous AI systems can produce ADM with legal or similarly significant effects for data subjects despite the prohibition contained in Article 22 (1) GDPR. Due to these reasoning deficiencies, it is unlikely that these systems can determine which exception to the prohibition according to Article 22 (2) GDPR applies to a particular case, that is, whether ADM is (i) necessary to enter or perform a contract, (ii) authorised by EU or MS law and (iii) based on the consent of the data subject. This Type 1 legal problem applies to all AI disciplines as introduced in Chapter 2 because the ability to make autonomous decisions and execute tasks on the designer's behalf[2038] constitutes a key element of AI (see Section 2.1).

The autonomous ADM problem (Type 1)

Autonomous AI systems could make their own decisions on how and why to process personal data. Due to the reasoning deficiencies in the AI discipline AR, such systems are likely to generate automated decisions that have legal or similarly significant effects for data subjects, even in cases in which the prohibition of ADM takes effect and none of the exceptions applies. This violates Article 22 (1-2) GDPR.

5.11.2 Legal problems: Type 2

Provided that all cumulative requirements mentioned in Article 22 (1) GDPR are met, Article 22 (3) GDPR provides the data subject with the right to obtain human intervention on the part of the controller. The corresponding Recital 71 does not further elaborate on what is required for such human intervention. To be effective, it has been argued that human intervention must be meaningful[2039] - and this is a rightful claim. Regulatory guidance explains that the human reviewer should undertake a thorough assessment of all the relevant data, including additional information provided by the data subject.[2040]

In my view, the human reviewer seems to have an almost unachievable task when taking the problems with respect to the interpretability of AI systems into account. As outlined in Section 4.4.1, most

have-a-legal-right-to-know-of-ai-inferences-about-you-including-those-via-ai-based-self-driving-cars/amp/> accessed 8 February 2024.

[2038] Eduardo Alonso, 'Actions and agents' in Frankish Keith and Ramsey William M. (eds) *The Cambridge Handbook of Artificial Intelligence* (2014) 235, 236.

[2039] Paul De Hert, Guillermo Lazcoz, 'Radical rewriting of Article 22 GDPR on machine decisions in the AI era' *European Law Blog* (13 October 2021) <https://europeanlawblog.eu/2021/10/13/radical-rewriting-of-article-22-gdpr-on-machine-decisions-in-the-ai-era/> accessed 8 February 2024.

[2040] Article 29 Working Party, 'Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679' (WP 251rev.01, 6 February 2018) at 27.

current DL models lack reasoning and explanatory capabilities, making them vulnerable to produce unexplainable outcomes. In particular, DL methods based on ANNs generally lack interpretability[2041] due to the hierarchical and nonlinear structure of ANNs. There is limited understanding of how each data point impacts the ML model and the ADM produced by it. Methods to measure the influence of a particular training point on the parameters of a model are scarce and subject to ongoing research.[2042] I take the view that in the case of complex AI systems, for example, involving DL and ANNs, obtaining meaningful human interventions in ADM is currently hardly possible due to the lack of interpretability of the AI systems and the ADM deployed by them. An additional factor is the incapacity of humans to grasp the logic of multidimensional ML algorithms. Typically, humans will struggle even more than machines with decisions produced by the ML algorithms currently used simply because humans cannot handle such an array of operational factors.[2043] Therefore, the right to obtain human intervention as enshrined in Article 22 (3) GDPR – if it shall be meaningful – cannot be enforced. This constitutes a Type 2 legal problem.

The intervention problem (Type 1)

AI systems deploying DL and ANN approaches are likely to produce output that is not interpretable for humans. When used in the context of ADM, meaningful human intervention as required by Article 22 (3) is impossible. Consequently, the data subject's right to obtain human intervention cannot be enforced.

Even if issues concerning interpretability can be overcome, it seems questionable whether humans are, in fact, able to assess the quality of output generated by means of AI correctly. There is experimental evidence suggesting that humans are not, although the concept of human oversight (intervention) rests on the assumption that humans are able to do so.[2044] Thus, the concept of human intervention seems to be flawed, which could also lead to a Type 3 legal problem.

5.11.3 Legal problems: Type 3

Article 22 GDPR creates three Type 3 legal problems when applied to AI. These three legal problems are the cumulativeness, opaque ADM and procedural safeguard problems.

[2041] Deng Li and Liu Yang, 'A Joint Introduction to Natural Language Processing and Deep Learning' in Deng Li and Liu Yang (eds) *Deep learning in natural language processing* (Springer 2018) 11, 12.
[2042] Lucas Bourtoule et al, 'Machine Unlearning' (IEEE Symposium on Security and Privacy, San Jose, May 2021) 1 and 3 <https://arxiv.org/abs/1912.03817> accessed 8 February 2024.
[2043] Lilian Edwards, Michael Veale, 'Slave to the Algorithm: Why a 'Right to Explanation' is Probably not the Remedy You are Looking for' (2017) Vol 16 Iss 1 Duke Law & Technology Review 19, 51.
[2044] Jan Biermann, John Horton, Johannes Walter, 'Algorithmic Advice as a Credence Good' (2022) Centre for European Economic Research Discussion Paper No 22-071 at 14, 17 < https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4326911 > accessed 8 February 2024.

Narrow and ambiguous scope

As outlined in Section 3.3.4.6, Article 22 (1) GDPR rests on three cumulative conditions to apply: (i) a decision is made that is (ii) based only on automated processing, including profiling, and (iii) has either legal or similarly significant effects.[2045]

Regarding condition (i), serious difficulties exist in determining precisely when a decision has been made, in particular in ML contexts.[2046] Apart from Recital 71 GDPR, which states that a decision 'may include a measure', and the first case on Article 22 GDPR referred to the CJEU,[2047] there is little guidance on what constitutes a 'decision' as mentioned in Article 22 (1) GDPR. In the first case dealing with Article 22 GDPR, the CJEU ruled that the automated establishment of a probability value concerning the ability of a data subject to service a loan ('score value')[2048] adopted by the credit agency SCHUFA in itself constitutes a solely-automated decision in the sense of Article 22 (1) GDPR.[2049] In this scenario, that score value is transmitted to a third party controller (financial institution), which then enters into or refrains from entering into contractual relationships with the data subject strongly drawing on that score value.[2050] However, it could be argued that a score value in itself does not represent a decision in the sense of Article 22 (1) GDPR. It rather constitutes a prediction of the data subject's future behaviour and/or the result of profiling that evaluates personal aspects about the data subject which *could subsequently* be used for decision-making (whether automated or not).[2051] Bygrave suggests that a decision in the sense of Article 22 (1) GDPR covers a large range of situations and should be viewed in a fairly generic sense, provided it is formalised and can be distinguished from other stages that prepare, support or complement decision-making.[2052] A decision in this sense usually requires some degree of binding effect which follows from the very concept of a decision.[2053] It can be argued that this binding effect is absent in this specific case because it is *another controller*, i.e. the financial institution, that takes the decision by applying the score value when determining whether the data subject receives the loan. However, the CJEU and AG Pikamäe reject such an interpretation. Following AG Pikamäe's opinion, [2054] the CJEU interprets the notion of a

[2045] Lee A. Bygrave, Commentary of Article 22 in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020) 532.
[2046] Lee A Bygrave, 'Minding the machine v2.0: The EU General Data Protection Regulation and Automated Decision Making' in Karen Yeung and Martin Lodge (eds) *Algorithmic Regulation* (OUP 2019) 253.
[2047] Case C-634/21, *SCHUFA Holding AG* [2023] ECR I-957.
[2048] Based on personal data of the data subject.
[2049] Case C-634/21, *SCHUFA Holding AG* [2023] ECR I-957 para 73.
[2050] Case C-634/21, *SCHUFA Holding AG* [2023] ECR I-957 paras 14-21.
[2051] Regulatory guidance names the example that where a human decides to agree the loan based on a profile based by purely automated means constitutes decision-making based on profiling; see Article 29 Working Party, 'Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679' (WP 251rev.01, 6 February 2018) at 6, 7.
[2052] Lee A. Bygrave, Commentary of Article 22 in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020) 532.
[2053] Lee A Bygrave, 'Automated Profiling, minding the machine: Article 15 of the EC Data Protection Directive and Automated Profiling' (2001) Vol 17 No. 1 Computer & Law Security Report 1, 18-19; Andreas Häuselmann, 'Profiling and the GDPR: Harmonised Confusion' (2018) Jusletter 13 <https://jusletter.weblaw.ch/juslissues/2018/924/profiling-in-the-gdp_3b8e8a124f.html__ONCE&login=false> accessed 8 February 2024.
[2054] Case C-634/21, *SCHUFA Holding AG* [2023] ECR I-220, Opinion AG Pikamäe paras 42, 47, 52.

decision broadly.[2055] According to the CJEU, such a broad interpretation is needed to prevent a circumvention of Article 22 GDPR and to avoid the resulting lacuna in legal protection.[2056] In the view of the CJEU, this interpretation also serves the purposes and objectives pursued by the GDPR. In addition, it reinforces the effective protection which Article 22 GDPR aims to achieve.[2057]

Even when interpreting the notion of a decision broadly, the question is whether ML-powered systems actually produce decisions in the sense of Article 22 GDPR. In many cases, ML will generate output that lacks the binding effect. ML merely generates predictions, which is one of its core goals.[2058] Thus, the output of an ML system constitutes something which *may* be used for decision-making, whether automated or not. ML models mostly generate classifications or uncertain estimations as they are incapable of synthesising the estimation and relevant uncertainties into a decision for action.[2059] Therefore, the output generated by ML, notably predictions concerning the future behaviour of data subjects, does arguably not constitute decisions in the sense of Article 22 (1) GDPR. Such output lacks the degree of binding effect required by the very concept of a decision. Instead, they prepare, support or complement decision-making. Predictions may have a binding effect once they are *applied towards* the data subject. Whereas obvious cases, such as the automated establishment of a score value constitute decisions in the sense of Article 22 (1), this is less clear in the context of AI. Decision-making processes with several stages[2060] are more complex, making it difficult to determine when and how a decision is made. Think, for example, of all the actors involved in targeted advertisement online.

Requirement (ii), i.e. the decision must be based 'solely' on automated processing, excludes AI systems that only provide decisional support for decision-making from the scope of Article 22 GDPR.[2061] When there is a 'human in the loop', which is the case when the automated processing functions solely as decisional support, Article 22 GDPR is not applicable.[2062] According to regulatory guidance, Article 22 (1) GDPR cannot be circumvented by 'fabricating' human intervention in the decision process so that the decision is no longer 'solely' automated.[2063] Thus, the crucial question concerning requirement (ii) is whether the processing of personal data involves human intervention and if so, what the extent of such intervention is. In fact, the first case on ADM referred to the CJEU for a

[2055] Case C-634/21, *SCHUFA Holding AG* [2023] ECR I-957 paras 44-46.
[2056] Case C-634/21, *SCHUFA Holding AG* [2023] ECR I-957 para 61.
[2057] Case C-634/21, *SCHUFA Holding AG* [2023] ECR I-957 paras 51 and 60.
[2058] See Section 2.2.1.
[2059] Lilian Edwards, Michael Veale, 'Slave to the Algorithm: Why a "Right to Explanation" is Probably not the Remedy You are Looking for' (2017) Vol 16 Iss 1 Duke Law & Technology Review 19, 46.
[2060] For an overview see Ruben Binns, Michael Veale, 'Is that your final decision? Multi-stage profiling, selective effects, and Article 22 of the GDPR' (2021) Vol 11 No 4 International Data Privacy Law 319-332.
[2061] Lee A Bygrave, 'Minding the machine v2.0: The EU General Data Protection Regulation and Autmated Decision Making' in Karen Yeung and Martin Lodge (eds) *Algorithmic Regulation* (OUP 2019) 253.
[2062] Lee A Bygrave, 'Machine Learning, Cognitive Sovereignty and Data Protection Rights with Respect to Automated Decisions' (2021) University of Oslo Faculty of Law Legal Studies Research Paper Series No. 202-35 at 20 <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3721118> accessed 8 February 2024.
[2063] Article 29 Working Party, 'Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679' (WP 251rev.01, 6 February 2018) at 27.

preliminary ruling addresses this issue. The establishment of the score value adopted by the controller SCHUFA meets requirement (ii) as such processing constitutes profiling and is thus 'based solely on automated processing'.[2064]

However, the *SCHUFA* ruling did not address the question what type of human involvement renders Article 22 (1) GDPR inapplicable, meaning processing is not 'solely automated' anymore. In many cases, this will be the decisive question regarding the applicability of Article 22 GDPR. Due to the lack of judgements at the CJEU level, it is worth considering case law at the level of the Member State ('MS'). Cases at MS level have tended to result in findings that the automated processing at issue was not fully automated.[2065] In fact, a report assessing ADM in light of the GDPR concludes that 'Courts across the EU have found that some (often limited) degree of human involvement...[..] was enough to set aside the application' of Article 22 GDPR.[2066] One case[2067] in the Netherlands specifically addressed the question what constitutes 'solely' automated processing according to Article 22 (1) GDPR. In this case, the data subjects (Uber drivers) contested the arguably fully automated deactivation of their Uber Driver account resulting from potential fraud signals detected by Uber's algorithm intended to prevent and detect fraud.[2068] However, Uber argued that its 'risk team' ultimately takes the decision to deactivate the Uber account of the drivers.[2069] The Amsterdam district Court accepted Uber's argumentation and ruled that there were no fully automated decisions. Consequently, the Court also denied the drivers' right to obtain meaningful information about the logic involved according to Article 15 (1) lit h GDPR with respect to the processing performed by Uber.[2070] This strongly underscores the problem regarding condition (ii). The Court of Appeal overturned the district Court's ruling. In the opinion of the Court of Appeal, Uber failed to sufficiently substantiate actual human intervention.[2071] Although Uber claimed that one or more members of Uber's risk team carried out manual investigations in each deactivation case, it failed to make this sufficiently plausible. In view of the Court of Appeal, Uber did not in any way demonstrate that the actions performed by the members of Uber's risk team was much more than merely a token gesture[2072] as mentioned in

[2064] Case C-634/21, *SCHUFA Holding AG* [2023] ECR I-957 para 47.
[2065] Lee A Bygrave, 'Machine Learning, Cognitive Sovereignty and Data Protection Rights with Respect to Automated Decisions' (2021) University of Oslo Faculty of Law Legal Studies Research Paper Series No. 202-35 at 20 <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3721118> accessed 8 February 2024.
[2066] Sebastião Barros Vale, Gabriela Zanfir-Fortuna, 'Automated Decision-Making Under the GDPR: Practical Cases from Courts and Data Protection Authorities' (2022) 8 <https://fpf.org/wp-content/uploads/2022/05/FPF-ADM-Report-R2-singles.pdf> accessed 8 February 2024.
[2067] Amsterdam District Court 13 March 2021, ECLI:NL:RBAMS:2021:1018.
[2068] Ibid paras 2.4, 3.1, 3.2.
[2069] Ibid para 4.19.
[2070] Ibid para 4.26; Raphaël Gellert, Marvin van Bekkum, and Frederik Zuiderveen Borgesius, 'The Ola & Uber judgments: for the first time a court recognises a GDPR right to an explanation for algorithmic decision-making' *EU Law Analysis* (28 April 2021) accessed 8 February 2024.
[2071] Amsterdam Court of Appeal 4 April 2023, ECLI:NL:GHAMS:2023:793 para 3.24; Amsterdam Court of Appeal 4 April 2023, ECLI:NL:GHAMS:2023:796 para 3.37.
[2072] Amsterdam Court of Appeal 4 April 2023, ECLI:NL:GHAMS:2023:793 para 3.24.

regulatory guidance.[2073] A decisive factor for this was the lack of any personal conversation between members of Uber's risk team and the drivers affected by the deactivations. In the only deactivation case involving such a personal conversation, the Court of Appeal ruled that there was indeed sufficient human intervention.[2074] Thus, a personal conversation seems to satisfy the requirements of actual human intervention, at least in view of the Amsterdam Court of Appeal. To me, this seems to be a rather low threshold. Ultimately, the ruling reaffirms the conclusion of a report assessing ADM in light of the GDPR: 'Courts across the EU have found that some (often limited) degree of human involvement...[..] was enough to set aside the application' of Article 22 GDPR.[2075]

In the context of AI, the requirement (i) that Article 22 (1) GDPR applies exclusively to decisions 'solely' based on automated processing creates a significant loophole because the output generated by AI is often used to support nonautomated decision-making. For example, the AC-powered HireVue software analyses the emotions a job candidate portrays during the video assessment[2076] and automatically assigns the candidate with an average rating (score) and recommendation whether or not to be employed. Subsequently, the recruiter has the discretion to decide, i.e. to select one of the recommended candidates. In such a scenario, Article 22 (1) GDPR does not apply because the decision-making process is not 'solely' automated. This is different with the automated establishment of a credit score adopted by a credit agency, which occurs without any human involvement. Also, credit scores are proven to play a pivotal role in the bank's decision to grant a loan.[2077] Requirement (i) is also problematic regarding decision-making processes involving multiple stages[2078] and multiple processing activities and controllers. AG Pikamäe acknowledges the difficulty in identifying the ultimately relevant decision, particularly[2079] when processing in the context of ADM involves several actors.

Requirement (iii) states that the decision produces legal effects concerning the data subject or 'significantly affects' him or her. Recital 71 GDPR names only two examples: automatic refusal of an online credit application or e-recruiting practices without any human intervention. Legal effects are effects that are able to alter or determine a person's rights or duties.[2080] An automated court decision

[2073] Article 29 Working Party, 'Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679' (WP 251rev.01, 6 February 2018) at 21.
[2074] Amsterdam Court of Appeal 4 April 2023, ECLI:NL:GHAMS:2023:793 para 3.25.
[2075] Sebastião Barros Vale, Gabriela Zanfir-Fortuna, 'Automated Decision-Making Under the GDPR: Practical Cases from Courts and Data Protection Authorities' (2022) 8 <https://fpf.org/wp-content/uploads/2022/05/FPF-ADM-Report-R2-singles.pdf> accessed 8 February 2024.
[2076] Nathan Mondragon, Clemens Aicholzer, Kiki Leutner, 'The Next Generation of Assessments' (HireVue 2019) <http://hrlens.org/wp-content/uploads/2019/11/The-Next-Generation-of-Assessments-HireVue-White-Paper.pdf> accessed 8 February 2024.
[2077] Case C-634/21, *SCHUFA Holding AG* [2023] ECR I-957 para 50.
[2078] See for an overview: Ruben Binns, Michael Veale, Is that your final decision? Multi-stage profiling, selective effects, and Article 22 of the GDPR' (2021) Vol 11 No 4 International Data Privacy Law 319-332.
[2079] Case C-634/21, *SCHUFA Holding AG* [2023] ECR I-220, Opinion AG Pikamäe para 40.
[2080] Lee A. Bygrave, 'Automated Profiling, minding the machine: Article 15 of the EC Data Protection Directive and Automated Profiling' 2001 Vol 17 No 1 Computer & Law Security Report 19.

is an example of a decision with legal effects.[2081] Regulatory guidance names as examples the cancellation of a contract, entitlement or denial of social benefits or refused admission to a country or denial of citizenship.[2082] The real ambiguity of requirement (iii) lies within the wording 'significantly affects'. This appears to be rather vague, and it is difficult to determine what should be considered 'sufficiently significant' to meet the threshold, which is even acknowledged in regulatory guidance.[2083] AG Pikamäe sheds some light on this notion. In his view, these significant effects may be of economic and/or social nature and relate to severe consequences for the data subject's freedoms and autonomy. They include adverse effects resulting from a negative score value, if it significantly restricts the data subject in exercising its freedoms or even stigmatises the data subject.[2084] In its decision in *SCHUFA*, the CJEU ruled that the automated establishment of a credit score by a credit agency significantly affects the data subject in the sense of requirement (iii). An insufficient credit score leads, in almost all cases, to the bank refusing to grant the loan applied for.[2085]

The ambiguity surrounding the notion of significant effects is quite unfortunate when considering that requirement (iii) constitutes one of the three decisive components that determines whether Article 22 (1) GDPR is applicable or not. Indeed, Belgium, Germany, Ireland, Italy, Finland, Poland and the UK stated during the law-making process that this wording is unclear and needs further clarification.[2086] Italy mentioned that 'it should be specified that this expression covers, for example, the application of network analysis instruments, user behaviour tracking, the creation of movement profiles via portable applications and the creation of personal profiles through social networking sites'.[2087] Poland argued that the vague term may lead to abuses by entities using profiling techniques.[2088] Finally, the EDPB's predecessor WP29 doubted in its opinion on the data protection reform proposals if the approach taken is sufficient to reflect the issues of creating and using profiles, an online environment in particular. Further need for clarification was mentioned by promoting that the term 'significantly affects' also 'covers the application of, for example, web analysis tools, tracking for assessing user behaviour, the creation of location profiles by mobile applications, or the creation of personal profiles by social networks'.[2089]

[2081] Frederik Zuiderveen Borgesius, 'Improving Privacy Protection in the area of Behavioural Targeting' (Doctoral thesis, Universiteit van Amsterdam 2015) 375 <https://dare.uva.nl/search?identifier=c74bdba6-616c-4cd9-925e-33a5858935e5> accessed 8 February 2024.
[2082] Article 29 Working Party, 'Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679' (WP 251rev.01, 6 February 2018) 21.
[2083] Ibid 22.
[2084] Case C-634/21, *SCHUFA Holding AG* [2023] ECR I-220, Opinion AG Pikamäe paras 38, 39, 42, 43.
[2085] Case C-634/21, *SCHUFA Holding AG* [2023] ECR I-957 paras 48-50.
[2086] Belgium p 12, Germany p 48, Ireland p 129, 137 Italy p 137, 172 Poland p 172, Finland p 189, UK pa 237 see < https://data.consilium.europa.eu/doc/document/ST-14147-2012-INIT/en/pdf > accessed 8 February 2024.
[2087] Ibid 137.
[2088] Ibid 172.
[2089] Article 29 Working Party, 'Opinion 1/2012 on the data protection reform proposals' (WP 191, 23. March 2012) at 14.

Despite these numerous requests for clarification, the term 'significantly affects' was not further specified, not even in the corresponding Recital 71 GDPR. Typically, targeted advertising based on profiling does not meet this threshold according to regulatory guidance. However, this might be different due to the intrusiveness of the profiling process, the expectations of the data subjects, the way the advertisement is delivered or when using knowledge of the vulnerabilities of the targeted data subject.[2090] AI is very well suited to facilitate such intrusive profiling. Take, for example, Amazon's US patent 'Keyword Determinations from Voice Data'[2091] introduced in Section 4.5.1. The patent relies on the AI discipline NLP and describes a system that can capture voice content when a user speaks into or near the device (e.g., Alexa), notably without activating the virtual assistant by mentioning the 'wake word' (e.g., 'hey Alexa'). Sniffer algorithms attempt to identify trigger words that indicate statements of preference (such as like or love) and translate them into keywords. The identified keywords are then transmitted to a location accessible to advertisers, who can use the keywords to select content that is likely relevant to the user.[2092] Amazon has denied that it uses voice recordings for advertising at the moment and mentioned that the patent might never actually come to the market.[2093] In any case, it is questionable whether controllers and Courts will agree that such kind of intrusive advertisement significantly affects the data subjects in the sense of requirement (iii). Neither the GDPR nor its preparatory documents provide substantive guidance on the threshold that must be met in this regard, which ultimately leads to legal uncertainty.

It is problematic when life decisions about a person[2094] such as being hired or receiving a loan are influenced by or based on possibly inaccurate data (see Section 4.7.1) automatically *generated* by AI. The relatively narrow scope of Article 22 GDPR and the cumulative requirements that must be met to render it applicable actually provide far less support for data subjects seeking control over ADM involving automated processing facilitated by AI than initially expected.[2095] In my view, this holds true despite the CJEU's broad interpretation of a decision in *SCHUFA*[2096] because conditions (ii) and (iii) must be met simultaneously. In many cases, processing is not 'solely automated' as required by condition (ii). In addition, the vagueness in terms of the required effects foreseen by condition (iii) comes into play often.

[2090] Article 29 Working Party, 'Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679' (WP 251rev.01, 6 February 2018) 22.
[2091] Edara Kiran, 'Key Word Determinations From Voice Data' US Patent Number US 8798995B1 (Assignee: Amazon Technologies, Inc.) August 2014 <https://patentimages.storage.googleapis.com/bd/ed/2b/c4c67cc5a9f1ab/US8798995.pdf>, accessed 8 February 2024.
[2092] Ibid.
[2093] Griffin Andrew, 'Amazon files for Alexa patent to let it listen to people all the time and work out what they want' *The indipendent* (London, 11 April 2018) <https://www.independent.co.uk/life-style/gadgets-and-tech/news/amazon-alexa-patent-listening-to-me-facebook-phone-talking-ads-a8300246.html> accessed 8 February 2024.
[2094] Tim Lewis, 'AI can read your emotions. Should it?' *The Guardian* (London 17 August 2019) <https://www.theguardian.com/technology/2019/aug/17/emotion-ai-artificial-intelligence-mood-realeyes-amazon-facebook-emotient> accessed 8 February 2024.
[2095] Lilian Edwards, Michael Veale, 'Slave to the Algorithm: Why a "Right to Explanation" is Probably not the Remedy You are Looking for' (2017) Vol 16 Iss 1 Duke Law & Technology Review 19, 46.
[2096] Case C-634/21, *SCHUFA Holding AG* [2023] ECR I-957 paras 45, 60.

MS-level case law concerning Article 22 GDPR places upon data subjects the onus of showing that all the cumulative requirements are fulfilled. Often, this may be difficult to satisfy considering that the AI systems used for ADM utilise non-transparent logic and come with covert consequences.[2097] Thus, the right of data subjects not to be subject to ADM creates a Type 3 legal problem. This right is not fit for purpose to strengthen the rights of data subjects.[2098] As noted by the CJEU, effective protection of personal data requires the strengthening of the rights of data subjects, which is emphasised by Recital 11 GDPR.[2099] With its cumulative and vague requirements determining the applicability of Article 22 GDPR, this right does not effectively contribute to the GDPR's aim to strengthen data subject rights. The CJEU has stressed the importance of ensuring that data subject rights granted by the GDPR are effective.[2100] Thus, Article 22 GDPR fails to *effectively* protect individuals against the particular risks associated with the automated processing of personal data, which is the aim of this provision according to the CJEU.[2101] Controllers are likely to exploit the ambiguousness of the requirements enshrined in Article 22 GDPR to argue that this right does not apply.[2102] For example, a report assessing ADM in light of the GDPR concludes: 'Courts across the EU have found that some (often limited) degree of human involvement...[..] was enough to set aside the application' of Article 22 GDPR.[2103] A right with vague cumulative requirements cannot be considered effective.

In addition, Article 22 GDPR fails to protect data subjects against risk of misuse[2104] and from concerns relating to ADM which the GDPR aims to address. These include, among others, (i) potential weakening of the ability of individuals to exercise influence over ADM and (ii) concerns over the quality of ADM.[2105]

The ability to exercise influence over ADM (i) is intertwined with the GDPR's goal that 'natural persons should have control of their own personal data'.[2106] As outlined in Section 4.4.3, enforceable rights are one of the main mechanisms for data subjects to exercise control over the processing of

[2097] Lee A. Bygrave, Commentary of Article 22 in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary – 2021 Update* (OUP 2021) 100.
[2098] Recital 11 GDPR.
[2099] Case C-132/21 *Nemzeti* [2023] ECR I-2 para 42.
[2100] Case C-154/21 *Österreichische Post AG* [2023] ECR I-3 para 39.
[2101] Case C-634/21, *SCHUFA Holding AG* [2023] ECR I-957 paras 57, 60.
[2102] Lee A Bygrave, 'Machine Learning, Cognitive Sovereignty and Data Protection Rights with Respect to Automated Decisions' (2021) University of Oslo Faculty of Law Legal Studies Research Paper Series No. 202-35 at 20 <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3721118> accessed 8 February 2024.
[2103] Sebastião Barros Vale, Gabriela Zanfir-Fortuna, 'Automated Decision-Making Under the GDPR: Practical Cases from Courts and Data Protection Authorities' (2022) 8 <https://fpf.org/wp-content/uploads/2022/05/FPF-ADM-Report-R2-singles.pdf> accessed 8 February 2024.
[2104] Joined Cases C-293/12 and C-594/12 *Digital Rights Ireland* [2014] ECR-I 238 paras 54 and 66; Joined Cases C-203/15 and C-698/15 [2016] *Tele 2 Sverige* ECR-1 970 paras 109 and 122; Case C-362/14 *Schrems v Data Protection Commissioner* [2015] ECR I-650 para 91.
[2105] Recitals 4 and 71 GDPR; see also the rationales mentioned in COM(92) 422 final—SYN 287 at page 26 and COM(90) 314 final—SYN 287 at page 29 relating to Article 22 GDPR's predecessor DPD which remain valid for the GDPR as convincingly outlined by Isak Mendoza, Lee A Bygrave, 'The Right Not to be Subject to Automated Decisions Based on Profiling' in Tatiana-Eleni Synodinou et al (eds) *EU Internet Law: Regulation and Enforcement* (Springer International 2017) 83-84.
[2106] Recital 7 GDPR.

their personal data. Due to the narrow scope and the cumulative criteria enshrined in Article 22 GDPR, this right is in many cases not applicable to personal data automatedly processed by AI systems. Control in the sense of the GDPR is rather limited as acknowledged by AG Campos Sánchez-Bordona, stating that 'the scope for individual action is limited' and 'confined to the exercise of those rights in specified circumstances'.[2107] Article 22 (1) GDPR further restricts the already limited means for data subjects to exercise control with respect to the automated processing by means of AI systems and therefore fails to achieve this goal. Data subjects cannot obtain human intervention, express their point of view and contest the decision because Article 22 GDPR is not applicable due to the restricted scope and cumulative criteria that must be met.

Article 22 GDPR fails to protect data subjects from issues relating to the quality of ADM and 'the particular risks to their rights and freedoms associated with the automated processing of personal data, including profiling' which is the rationale of Article 22 GDPR according to the CJEU.[2108] As explained in the inaccuracy and rebuttal problems discussed in Section 4.7.1, ML and AC may *automate the generation of* inaccurate personal data. Such inaccurate data might be used for partially automated decision-making with significant effects for data subjects, like the decision to receive a loan, job offer or to be allowed to pass border control. This is also problematic with respect to Recital 4 GDPR, which states that 'processing of personal data should be designed to serve mankind'. In the examples mentioned, automated processing performed by AI serves the interest of controllers, rather than those of natural persons who want to obtain a loan, seek employment or cross a border. Thus, Article 22 GDPR fails to safeguard *human dignity*, which is another rationale of this provision, as noted by AG Pikamäe.[2109] In conclusion, Article 22 GDPR fails to achieve its aim, which is, according to the CJEU, *effective protection* against the particular risks associated with the automated processing of personal data, including profiling.[2110]

This Type 3 legal problem occurs regardless of which AI discipline has been used for ADM because the problem is caused by the cumulativeness requirement enshrined in Article 22 GDPR. It is therefore a general problem and potentially relates to all AI disciplines, as introduced in Chapter 2.

The cumulativeness problem (Type 3)
The cumulative and vague requirements in Article 22 GDPR render it inapplicable to many decisions enabled, taken by or generated with the support of AI. Therefore, Article 22 GDPR is not fit for purpose to effectively protect data subjects from the particular risks associated with the automated processing of personal data, which is the main rationale of this provision according to the CJEU.

[2107] Case C-300/21 *UI* [2022] ECR I-756 Opinion AG Manuel Campos Sánchez-Bordona para 72.
[2108] Case C-634/21, *SCHUFA Holding AG* [2023] ECR I-957 para 57.
[2109] Case C-634/21, *SCHUFA Holding AG* [2023] ECR I-220, Opinion AG Pikamäe para 19.
[2110] Case C-634/21, *SCHUFA Holding AG* [2023] ECR I-957 paras 57, 60.

Information about the logic involved in ADM

As described in Section 3.3.4.6, Article 22 GDPR applies only if all three cumulative requirements are met simultaneously. Only then can the data subject enforce its right to obtain meaningful information about the logic involved in ADM and the significance and the envisaged consequences of such processing. If Article 22 GDPR is not applicable, for example, because the decision is not *solely* automated, the result will be that the data subject cannot enforce its right according to Article 15 (1) lit h GDPR. This interpretation is confirmed by regulatory guidance. Article 15 (1) lit h GDPR is discussed under Chapter IV of the guidelines on ADM, which 'explains the specific provisions that *only* apply to solely automated individual decision-making, including profiling'.[2111]

However, research on Article 15 (1) lit h GDPR suggests that 'meaningful information about the logic involved and the significance and consequences for data subjects can also be invoked where decision-making processes are only partially (rather than completely) automated'.[2112] Whereas this interpretation is certainly welcome from the data subject's perspective, it does not stand when applying the grammatical (literal) and systematic method of interpretation. The wording contained in Article 15 (1) lit h GDPR obliges controllers to inform data subjects about 'the existence of automated decision-making, including profiling, referred to in *Article 22(1)* and (4) and, *at least in those cases*, meaningful information about the logic involved, as well as the significance and the envisaged consequences of such processing for the data subject'.[2113] Due to the wording '*at least in those cases*', controllers are *not* legally required to inform data subjects about decision-making which is only partially automated. This follows from a grammatical (literal) interpretation of Article 15 (1) lit h (see also Section 4.4.3). The result is the same when applying the method of systematic interpretation. Systematically, Article 15 (1) lit h GDPR explicitly refers to Article 22 (1) GDPR, which outlines that decisions must be based *solely* on automated processing to fall within the scope of this right. Consequently, the data subjects concerned are not entitled to receive meaningful information about the logic involved in the output generated by AI systems simply because the decision taken is not fully automated.

The HireVue software and similar services[2114] aim to detect the emotional states portrayed during the automated video assessment. It will be difficult for applicants to assess the accuracy of emotion data detected by this software without having access to additional information concerning the logic involved in the processing performed by the AI system. Within the iBorderCtrl system, an 'automatic

[2111] Adding, in Footnote 3 of the guidelines 'as defined in Article 22 (1) GDPR'; see Article 29 Working Party, 'Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679' (WP 251rev.01, 6 February 2018) at 10.
[2112] Bart Custers, Anne-Sophie Heijne, 'The right of access in automated decision-making: The scope of article 15(1)(h) GDPR in theory and practice' (2022) Vol 46 Computer Law & Security Review 1, 5.
[2113] Emphasis added by the author.
[2114] HumeAI which provides AI-powered tools helping recruiters to assess personality traits as well as emotional states of candidates, see < https://hume.ai/products/facial-expression-model/ > and < https://gethume.com/blog5/artificial-intelligence-for-recruiting > accessed 8 February 2024.

deception detection system' quantifies the probability of deceit in interviews by analysing interviewees' non-verbal micro-gestures.[2115] There is an inherent risk of inaccuracy, namely, false positives that wrongly identify the interviewee as being deceptive, which might lead to a stigmatisation or prejudice against the interviewee, for example when talking to the human border guard.[2116] Because the final decision will be made by a human border guard, Article 22 GDPR is not applicable.[2117] Therefore, interviewees do not have to be informed about the logic and functionality of the iBorderCtrl system.[2118] In addition, the human border guard taking the final decision could be unduly influenced by the possibly inaccurate output of the iBorderCtrl system.[2119]

Individuals might be subject to decisions enabled or supported by AI, but do not have the means to verify whether the relevant legal provisions were respected. They face difficulties with regard to effective access to justice in case such decisions negatively affect them.[2120] Individuals cannot obtain information about the logic involved in the processing performed by the AI system because one of the requirements enshrined in Article 22 (1) GDPR is not met. This leads to a Type 3 legal problem for the same reasons as outlined in the cumulativeness problem. Article 22 (1) GDPR is not fit for purpose to strengthen the rights of data subjects and ensure that they are effective.[2121] It also fails to facilitate that data subjects can exercise control[2122] regarding the processing of their personal data processed by AI systems. As outlined in Section 4.4.3, enforceable rights are one of the main mechanisms for data subjects to exercise control over the processing of their personal data. Due to the narrow scope and cumulative criterion enshrined in Article 22 GDPR, this right is in many cases not applicable to personal data automatedly processed by means of AI systems. This is in stark contrast to what Article 22 GDPR aims to achieve according to the CJEU: *effective protection* against risks associated with the automated processing of personal data.[2123] Article 22 (1) GDPR further restricts the already limited means for data subjects to exercise control[2124] concerning the automated processing by means of AI systems and therefore fails to achieve the GDPR's legislative goal. Because one of the cumulative requirements enshrined in Article 22 (1) GDPR is not met, data subjects cannot obtain meaningful information about processing concerning ADM when enforcing their right of access. Therefore, they have no effective means to exercise control, for example, enforcing other data subject

[2115] See <https://www.iborderctrl.eu/Technical-Framework/> accessed 8 February 2024.
[2116] See <https://www.iborderctrl.eu/Frequently-Asked-Questions/> accessed 8 February 2024.
[2117] If the border guard does not blindly follow the system and 'rubber-stamp' its decision.
[2118] Tim Lewis, 'AI can read your emotions. Should it?' *The Guardian* (London 17 August 2019) <https://www.theguardian.com/technology/2019/aug/17/emotion-ai-artificial-intelligence-mood-realeyes-amazon-facebook-emotient> accessed 8 February 2024.
[2119] See <https://www.iborderctrl.eu/Technical-Framework/> accessed 8 February 2024.
[2120] Commission, 'White Paper on Artificial Intelligence - A European approach to excellence and trust' COM (2020) 65 final 12 <https://commission.europa.eu/document/d2ec4039-c5be-423a-81ef-b9e44e79825b_en> accessed 8 February 2024.
[2121] Recital 11 GDPR; Case C-132/21 *Nemzeti* [2023] ECR I-2 para 42; Case C-154/21 *Österreichische Post AG* [2023] ECR I-3 para 39.
[2122] Recital 7 GDPR.
[2123] Case C-634/21, *SCHUFA Holding AG* [2023] ECR I-957 para 60.
[2124] Case C-300/21 *UI* [2022] ECR I-756 Opinion AG Manuel Campos Sánchez-Bordona para 72.

rights, regarding decisions supported and enabled by AI that may negatively affect them. This Type 3 legal problem occurs regardless of which AI discipline has been used for ADM because it is caused by the cumulativeness requirement enshrined in Article 22 GDPR. It is therefore a general problem and potentially relates potentially to all AI disciplines as introduced in Chapter 2.

The opaque ADM problem (Type 3)

The cumulativeness problem renders Article 22 GDPR inapplicable to many decisions taken by or generated with the support of AI. Consequently, data subjects cannot obtain meaningful information about the logic involved in decisions taken by or generated with the support of AI. Data subjects are not effectively protected and have no means to exercise control regarding decisions supported and enabled by AI that may negatively affect them.

Contesting to ADM

In case all the cumulative requirements enshrined in Article 22 (1) GDPR are indeed met, the right to contest the ADM as enshrined in Article 22 (3) GDPR provides the data subject with an effective remedy with respect to ADM,[2125] at least from a preliminary point of view. The scarce literature in academia suggests that the term 'contest' means a right of appeal and therefore more than simply a right to object or oppose to ADM. To be meaningful, the right to contest shall at least oblige the controller to hear and consider the merits of an appeal made by the data subject. To be fair, the appeal process shall carry a qualified obligation to provide the data subject with reasons for the ADM.[2126]

Although these claims are valid, it seems that the right to contest ADM mostly offers a procedural safeguard rather than meaningful protection against ADM and personal data automatedly processed by AI systems. In fact, it is unlikely that a company deploying ADM will actually revise such decisions when an individual invokes her right to contest under Article 22 (3) GDPR unless sector-specific decision-making standards or other provisions of data protection law are violated.[2127]

This holds particularly true for types of ADM which determine whether to conclude a contract with the data subject. The freedom of contract is a cornerstone of EU contract law and grants parties the legal freedom to enter into a contract (or not) and agree on its content.[2128] According to the CJEU, freedom of contract is covered by the freedom to conduct a business enshrined in Article 16

[2125] Isak Mendoza, Lee A Bygrave, 'The Right Not to be Subject to Automated Decisions Based on Profiling' in Tatiana-Eleni Synodinou et al (eds) *EU Internet Law: Regulation and Enforcement* (Springer International 2017) 93.
[2126] Ibid 93-94.
[2127] Sandra Wachter, Brent Mittelstadt 'A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI' (2019) No 2 Columbia Business Law Review 570, 571.
[2128] Olha O Cherednychenko, 'Fundamental Freedoms, Fundamental Rights, and the Many Faces of Freedom of Contract in the EU' in Mads Andenas, Tarjei Bekkedal, Luca Pantaleo (eds) *The Reach of Free Movement* (Springer 2017) 273, 276.

EUCFR.[2129] A company may decide on its own discretion whether and how to conclude a contract with the data subject, provided that this complies with EU and MS law. Consider, for example, a data subject who applies for a loan at a bank. The bank has a highly sophisticated AI system in place that automatically decides whether the loan will be granted. The system deploys approaches from the AI disciplines ML and DL in particular and analyses all personal data provided by the data subject, including behaviour related to mobile phone usage, to determine the creditworthiness of the data subject. The AI system decides to not grant the loan to the data subject because the likelihood of repayment was predicted negatively due to behavioural features derived from the data subject's mobile phone usage[2130] (see Section 4.4.3).

In this scenario, all requirements enshrined in Article 22 (1) GDPR are met: There is a decision (i) which is fully automated (ii) and significantly affects the data subject (iii) because the latter will not receive the loan to buy its own apartment. In addition, the prohibition on ADM is lifted because it is necessary to assess and determine the creditworthiness of the data subject, from the bank's perspective, to enter a contract with the data subject. The data subject may very well invoke its right to contest the ADM, but the bank is by no means obliged to revert its decision. The freedom of contract grants the bank legal freedom not to enter into a contract with the data subject. Imagine a second scenario, in which an employer relies on the AC-powered HireVue software to analyse the emotions a job candidate portrays during the video assessment,[2131] automatically assigns an average score and selects the candidate with the highest score. Here as well, candidates who have been rejected may invoke their right to contest the ADM, but the employer is under no requirement to change the decision.

The right to contest to ADM is a procedural safeguard rather than a right which allows data subjects to exercise real influence over ADM that legally or significantly affect them. This leads to a Type 3 legal problem. The right not to be subject to ADM is not fit for purpose to strengthen the rights of data subjects and ensure that they are effective.[2132] This is in stark contrast to what Article 22 GDPR aims to achieve according to the CJEU: *effective protection* against risks associated with the automated processing of personal data.[2133] A right that merely provides procedural safeguards but no meaningful influence on the ADM facilitated or supported by AI systems cannot be effective, nor can it strengthen the rights of data subjects.

[2129] Case C-426/11, *Alemo-Herron* [2013] ECR I-521 para 32; Case C-283/11, *Sky Österreich* [2013] ECR-28 paras 42, 43.
[2130] Daniel Björkegren, Darrell Grissen, 'Behavior Revealed in Mobile Phone Usage Predicts Credit Repayment' (2020) Vol 34 Iss 3 The World Bank Economic Review 618, 623.
[2131] Nathan Mondragon, Clemens Aicholzer, Kiki Leutner, 'The Next Generation of Assessments' (HireVue 2019) <http://hrlens.org/wp-content/uploads/2019/11/The-Next-Generation-of-Assessments-HireVue-White-Paper.pdf> accessed 8 February 2024.
[2132] Recital 11 GDPR; Case C-132/21 *Nemzeti* [2023] ECR I-2 para 42; Case C-154/21 *Österreichische Post AG* [2023] ECR I-3 para 39.
[2133] Case C-634/21, *SCHUFA Holding AG* [2023] ECR I-957 para 60.

The ability to influence ADM is intertwined with the GDPR's goal that 'natural persons should have control of their own personal data'.[2134] As outlined in Section 4.4.3, enforceable rights are one of the main mechanisms for data subjects to exercise control. Because the right to contest is only a procedural safeguard, it fails to achieve the GDPR's goal that data subjects be able to control the processing of their personal data related to ADM. The right to contest will not really change the controller's ADM, as it is in many cases not obliged to revert its decision due to the freedom of contract. This Type 3 legal problem occurs regardless of which AI discipline has been used for ADM because it is caused by the fact that the right to contest according to Article 22 (3) GDPR is solely a procedural safeguard. It is therefore a general problem and potentially relates to all AI disciplines introduced in Chapter 2.

The procedural safeguard problem (Type 3)

The right to contest ADM as enshrined in Article 22 (3) GDPR is a procedural safeguard rather than a right that allows data subjects to exercise influence over ADM that significantly affects them. If a data subject contests ADM generated by means of AI and based on Article 22 (2) lit a GDPR, the controller is by no means required to change the outcome of the decision due to the freedom of contract. The right to contest fails to provide data subjects with effective protection and meaningful influence over ADM based on personal data.

5.12 Conclusions

This chapter addressed Subquestion 4, i.e. what legal problems arise or may arise when the enforceable rights enshrined in the current legal framework are applied to AI. I have outlined that all AI disciplines as described in Section 2.2 may raise legal problems when they are applied to the enforceable rights enshrined in the current legal framework discussed in Chapter 3. Three types of legal problems were identified, i.e. that (1) legal provisions are violated, (2) that legal provisions cannot be enforced and (3) that legal provisions are not fit for purpose to protect the fundamental right at stake. These legal problems may be caused by the AI disciplines *or* by the enforceable rights themselves when applied in the context of AI. Table 5.2 provides an overview of the legal problems identified in this chapter. The table illustrates the broad range of legal problems that arise or may arise in the context of AI. In total, twenty-five problems are identified.

[2134] Recital 7 GDPR.

Problem	Right	Type	AI Disciplines
Control	Informational privacy	1	ML, NLP, CV, AC, AR
Bodily information	Bodily privacy	1	ML (DL), AC
Mental information	Mental privacy	1, (3)	ML (DL), NLP, CV, AC
Speech analysis	Communicational privacy	1	ML, NLP, AC
Interception and identification	Communicational privacy	1	NLP
Keyword	Communicational privacy	1	NLP
Meaningless information	Access	1	ML (DL)
Information restriction	Access	2	ML, NLP, CV, AC, AR
Trade secrets	Access	2, 3	ML, NLP, CV, AC, AR
Logic and causal explanation	Access	3	ML, NLP, CV, AC, AR
Procedural autonomy	Rectification	1	ML, AC
Unverifiable data	Rectification	2	ML
Subjectivity	Rectification	2	AC
Verifiability standard	Rectification	3	ML, AC
Training data	Erasure	1, 2	ML
Erasure	Erasure	2	ML, AC
Transmission	Portability	2	ML, NLP, CV, AC, AR
Restricted scope	Portability	3	ML, NLP, CV, AC, AR
Continuance problem	Object	1	ML, NLP, CV, AC, AR
Erasure after objection	Object	3	ML, NLP, CV, AC, AR
Autonomous ADM	Automated decision-making	1	ML, NLP, CV, AC, AR
Intervention	Automated decision-making	2	ML (DL)
Cumulativeness	Automated decision-making	3	ML, NLP, CV, AC, AR
Opaque ADM	Automated decision-making	3	ML, NLP, CV, AC, AR
Procedural safeguard	Automated decision-making	3	ML, NLP, CV, AC, AR

Table 5.2 Overview of legal problems, enforceable rights concerned, type of legal problem (1, 2, 3) and AI disciplines concerned. The brackets surrounding DL indicate that this *specific kind* of ML causes the legal problem.

Regarding the *right to informational privacy*, I have identified one Type 1 legal problem when applied to AI. This problem constitutes an overarching issue. *All AI disciplines* introduced in Chapter 2 process various types of information beyond the control of the individuals concerned. It thus attacks the core of informational privacy which is to provide individuals with a form of informational self-determination, allowing them to exercise control over the collection, dissemination and use of their information. No Type 2 or 3 legal problems arise due to the broad scope of the fundamental right to privacy and the living instrument doctrine adopted by the ECtHR, which considers technological developments such as AI and the issues to which they may give rise.

Regarding the *right to bodily privacy*, I have identified one Type 1 legal problem when applied to AI. This problem relates to the AI disciplines *ML (particularly DL)* and *AC* which are highly dependent on bodily information, including its functions and either gain physical access to the body (e.g., implants) or derive information from it through non-invasive means (e.g., wearables sensing neural activity in the brain). Due to the broad scope of the fundamental right to privacy and the living instrument doctrine adopted by the ECtHR, no Type 2 or 3 legal problems arise.

Regarding the *right to mental privacy*, I have identified one Type 1 legal problem when applied to AI. This problem constitutes a major issue. *All AI disciplines (*except AR) facilitate access to mental states and information that might be derived from this, which means that the mind is no longer insusceptible to interferences. As such, the right to mental privacy is not yet recognised as a specific element falling under the notion of private life as enshrined in the fundamental right to privacy. However, the existence of the right to mental privacy could be derived from existing case law or developed in future ECtHR jurisprudence due to the broad scope of this right and the doctrine of living instruments. If not, there will also be a Type 3 legal problem which is indicated by the brackets surrounding the Type 3 problem as illustrated in Table 5.2.

Regarding the *right to communicational privacy*, I have identified three Type 1 legal problems when applied to AI. *NLP* is the main driver: All three legal problems relate to this AI discipline. This is not surprising because *NLP* aims to give machines the ability to process human language, which unavoidably involves the processing of communications. Two other AI disciplines, i.e. *ML* and *AC*, give rise to one Type 1 legal problem. Due to the broad scope of the fundamental right to privacy and the living instrument doctrine adopted by the ECtHR, no Type 2 or 3 legal problems arise.

Regarding the *right of access*, I have identified four legal problems of either Type 1, 2 or 3 when applied to AI. Table 5.2 shows that *all AI disciplines* are associated with these legal problems. This is mainly caused by the non-absolute nature of the right of access and trade secret protection for AI under the EU trade secrets directive (TSD). The broad scope of protection for AI under the TSD and restrictions to the right of access have severe effects on the entire data protection law regime because this right constitutes a conditio sine qua non for exercising other data subject rights.

Regarding the *right to rectification*, I have identified four legal problems of Type 1, 2 or 3 when applied to AI. All these problems relate to the AI disciplines ML and/or AC. This is mainly due to the unverifiable and subjective nature of the personal data generated by these two AI disciplines and the close connection with the right to rectification. Both ML and AC can generate inaccurate personal data, and the right to rectification grants data subjects the right to rectify inaccurate personal data.

Regarding the *right to erasure*, I have identified two Type 1 and/or 2 legal problems when applied to AI. ML is the main driver: All three legal problems relate to this AI discipline. As such, no Type 3 legal problems arise when the right to erasure is applied to the AI disciplines. This is mainly due to the broad wording contained in Article 17 (1) lit d GDPR[2135] which allows data subjects to request the erasure of personal data that 'have been unlawfully processed'.

Regarding the right to *data portability*, I have identified two legal problems when applied to AI: Types 2 and 3. Table 5.2 shows that *all AI disciplines* are associated with these legal problems. Both legal problems occur regardless of which *AI discipline* is applied to the right to data portability because the problems relate to the broad scope of protection for AI under the TSD and the restricted scope of this right. As such, no Type 1 legal problems arise when the right to data portability is applied to AI.

Regarding the *right to object*, I have identifed two legal problems when applied to AI: Types 1 and 3. Both legal problems occur regardless of which *AI discipline* is applied to the right to object. No Type 2 legal problems arise because data subjects can easily enforce their right to object, and the burden of proof is imposed on the controller if the latter intends to continue processing.

Regarding the *right not to be subject to ADM*, I have identified five legal problems when applied to AI: either Type 1, 2 or 3. Table 5.2 shows that *all AI disciplines* are associated with these legal problems, except for the Type 2 legal problem, which only relates to ML or, more specifically, to *DL*. All other legal problems are not caused by AI, but rather by the right itself: The right not to be subject to ADM suffers from significant flaws. The ambiguity and complexity of this right make it difficult to apply in practice.

In terms of the *types of legal problems* identified in this chapter, Table 5.2 shows that the total number of legal problems per type is almost evenly distributed. In total, there are eleven Type 1 legal problems, eight Type 2 legal problems and nine Type 3 legal problems. The almost equal distribution per type of legal problem underscores that the problems caused by AI are diverse, leading to situations in which the fundamental rights to privacy and data protection are violated, cannot be enforced or are not fit for purpose.

In terms of the *fundamental right to privacy*, a clear trend can be observed regarding the types of legal problem identified within this chapter: Only Type 1 legal problems occur. The fundamental right to privacy appears to be well equipped to protect privacy from the challenges and risks posed by AI.

[2135] Herke Kranenborg, Commentary of Article 17 in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020) 481.

This is mainly due to the broad scope of the right and the living instrument doctrine adopted by the ECtHR.

In terms of the *fundamental right to data protection* and the enforceable rights enshrined in the GDPR, no clear trend can be observed regarding the types of legal problems. There are five Type 1 legal problems, eight Type 2 legal problems and eight Type 3 legal problems. However, the finding that Types 2 and 3 legal problems occur just as often indicates two things: that there is an enforcement problem and that legislative measures and judicial action are needed to overcome the shortcomings of the current legal framework.

In terms of which AI disciplines cause *how many legal problems* when applied to the enforceable rights enshrined in the current legal framework, Table 5.2 shows that ML leads to twenty-two, NLP sixteen, CV thirteen, AC nineteen and AR twelve legal problems. The prominent role of ML is not surprising, as this AI discipline is the most widely used and often combined with other AI disciplines. In addition, AC seems to be the main driver of legal problems, as it causes only slightly less legal problems than ML. The total amount of legal problems associated to the AI disciplines NLP, CV and AR are almost evenly distributed.

6 Addressing the legal problems

This chapter aims to answer Subquestion 5, i.e. how should the incompatibilities of the current legal framework identified in Chapters 4 and 5 be addressed. Chapters 4 and 5 strongly emphasise the difference between the law in books and the law in action by unveiling, in total, 55 legal problems when the current legal framework is applied to AI. This chapter discusses how the gaps between the law in books and the law in action can be addressed by means of legal solutions.

This chapter is structured as follows. Section 6.1 starts by introducing the selected legal problems, the selection criteria and the approach taken to address the legal problems. Section 6.2 discusses the elusiveness problem, Section 6.3 the mental data problem, Section 6.4 the communication surveillance problem, Section 6.5 the trade secret problem, Section 6.6 the verifiability standard problem and Section 6.7 the cumulativeness problem. Section 6.8 concludes.

6.1 Approach

In Chapters 4 and 5, 55 legal problems were identified when the current legal framework is applied to the AI disciplines introduced in Chapter 2. Because it is impossible to address all of them in sufficient depth in this chapter, I focus on six selected legal problems, as shown in Table 6.1.

Problem	Principle / Right	Type	AI Disciplines
Elusiveness	Fairness	2, 3	ML, NLP, CV, AC, AR
Mental data	Exhaustive enumeration	3	ML, AC
Communication surveillance	Confidentiality	3	ML, NLP, AC
Trade secrets	Access	2, 3	ML, NLP, CV, AC, AR
Verifiability standard	Rectification	3	ML, AC
Cumulativeness	Automated decision-making	3	ML, NLP, CV, AC, AR

Table 6.1 Overview of legal problems addressed in this chapter, principle/right concerned, type of legal problem (1, 2, 3) and AI disciplines concerned.

The decision to focus on the six selected legal problems contained in Table 6.1 is based on three selection criteria: effectiveness, urgency and novelty. I have chosen these selection criteria because I want to focus on the problems unique to AI that are most urgent and seem to have the highest impact, either by their weight (influencing several other problems) or by their sensitive nature. Choosing isolated legal problems such as the storage, verification and restriction problem would not be very effective because they are not closely intertwined with other legal problems, as is the case with the six selected legal problems. Solving these six legal problems would address simultaneously eight highly related legal problems, i.e. the manipulation, sabotage, emotion data, location data, neurodata, information restriction, unverifiable data and subjectiveness. In terms of urgency, some of the

remaining legal problems are less pressing. This applies to the transmission and restricted scope problem. The right to data portability, to which these two legal problems relate, is not a classic data protection right as it mainly aims to facilitate the transfer of personal data from one controller to another. Thus, this right stimulates competition and innovation in data-driven markets and does not entirely align with the nature of the fundamental right to data protection.[2136] In terms of novelty as a selection criterion, some legal problems are not 'new' but well known for quite a while, such as opacity, interpretability or training data problems.

Let me explain why I discuss exactly these six legal problems. First, the *elusiveness problem* is important to solve as it relates to the fairness principle, which is under great pressure considering that ten legal problems relate to this principle. In addition, the elusiveness problem raises two other legal problems, namely, the manipulation and sabotage problem. A substantively sound fairness principle may address these three problems together and could also prove helpful for other potential challenges caused by AI. Second, the *mental data problem* is very pressing due to the highly sensitive nature of mental data as it relates to the core of an individual's private sphere. Finding a solution for the mental data problem might simultaneously solve the emotion data, neurodata and location data problem, as these problems essentially arise due to the principle of enhancing protection for special data and the approach taken to enumerate such data exhaustively. Third, the *communication surveillance* problem reveals that virtual assistant services are able to intercept, analyse and otherwise process both human-machine and interpersonal communication which is problematic in terms of communicational privacy. Fourth, the *trade secrets* problem is particularly pressing as it allows controllers to restrict access to personal data, which prevents individuals from enforcing other data subject rights. This is problematic because the right of access constitutes a conditio sine qua non for the enforcement of other data subject rights, for example the right to rectification or erasure. Solving the trade secrets problem simultaneously address the inherently related information restriction problem. Fifth, the *verifiability standard* problem deserves particular attention because some AI disciplines are prone to generate inaccurate personal data, which is both problematic regarding the right to rectification and the accuracy principle. An effective solution is needed for people to seek the rectification of inaccurate personal data, as the processing of such data might be harmful to the individuals concerned. Solving the verifiability standard problem might also address two closely-related legal problems, namely, the unverifiable data and subjectivity problems. Sixth, the *cumulativeness* problem should be solved because there is a need for protection against ADM facilitated by AI. Important decisions about individuals are increasingly influenced by personal data generated through AI. Controllers increasingly rely on algorithmic tools to support their decision-making.[2137] Such data might be

[2136] Inge Graef, Martin Husovec, Nadezhda Purtova, 'Data portability and data control: Lessons for an emerging concept in EU law' (2018) Vol 19 No 6 German Law Journal 1360-1398.
[2137] Jan Biermann, John Horton, Johannes Walter, 'Algorithmic Advice as a Credence Good' (2022) Centre for European Economic Research Discussion Paper No 22-071 1, 2 < https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4326911 > accessed 8 February 2024.

inaccurate, which could lead to detrimental effects for individuals (e.g. in an employment or financial context).

As shown in Table 6.1, in this chapter, Type 1 legal problems will not be discussed. The solution for such problems is obvious: violations of provisions contained in the current legal framework must be enforced through data subjects and/or representative bodies going to court ('private enforcement') and through supervisory authorities ('regulatory enforcement'). Without oversimplifying the issues at stake, the first step towards improved regulatory enforcement may be harmonisation of some procedural aspects of regulatory GDPR enforcement. In 2023, the EDPB sent a wish list to the European Commission, which points to current weaknesses in terms of cross-border cooperation between SAs.[2138] After receiving this wish list, the European Commission launched an initiative to adopt a proposal in the form of a regulation to specify and harmonise procedural rules relating to the *regulatory* enforcement of the GDPR.[2139] This initiative aims to harmonise some aspects of the administrative procedures the national SAs apply in cross-border cases and to support a smooth functioning of the GDPR cooperation and dispute resolution mechanisms. Another step towards improved regulatory enforcement could be to provide SAs with more financial resources. The latter seems to be needed both on EU and Member State level.[2140]

After discarding the 23 Type 1 legal problems and the eight related legal problems that can be addressed by solving the elusiveness, verifiability standard, trade secrets and mental data problem,[2141] eighteen legal problems remain that will not be discussed. However, these remaining eighteen legal problems are not necessarily less relevant. They simply do not appear on the top of the list when applying the selection criterion effectiveness, urgency and novelty.

Sections 6.2 through 6.7 discuss how the gap (i.e. the identified legal problems) between technology (AI) and the law (legal framework) might be closed. Essentially, these gaps might be closed by either changing the technology or the law (or both). Thus, two types of solutions may address the six selected legal problems: technological solutions and legal solutions. The former refers to solutions relating to the design of and applications of AI or techniques used for it. The latter refers to new or revised legislation as well as detailing existing legislation through policies or re-interpretation by courts. The

[2138] See < https://edpb.europa.eu/system/files/2022-10/edpb_letter_out2022-0069_to_the_eu_commission_on_procedural_aspects_en_0.pdf > accessed 8 February 2024.
[2139] See <https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/13745-Further-specifying-procedural-rules-relating-to-the-enforcement-of-the-General-Data-Protection-Regulation_en> accessed 8 February 2024.
[2140] The EDPB and EDPS have jointly sent an open letter to the European Parliament and European Council expressing concerns about the budget for 2023; see <https://edps.europa.eu/system/files/2022-09/22-09-12_edps-edpb-open-letter-budget-2022_en.pdf> accessed 8 February 2024; EDPB, 'Overview on resources made available by Member States to the Data Protection Supervisory Authorities' (2022) <https://edpb.europa.eu/system/files/2022-09/edpb_overviewresourcesmade_availablebymemberstatestosas2022_en.pdf> accessed 8 February 2024.
[2141] As outlined in the previous paragraph, namely manipulation, sabotage, emotion data, location data, neurodata, information restriction, unverifiable data and subjectiveness problems.

focus is on legal solutions, although it is important to stress that nonlegal solutions, technological solutions in particular, may exist or can be developed. In terms of technological solutions, approaches such as randomisation techniques,[2142] secure multiparty computation,[2143] homomorphic encryption,[2144] differential privacy,[2145] synthetic data[2146] or knowledge-infused learning[2147] should be further explored.

When referring to legal solutions, I mean (i) new interpretations of existing provisions through policies and courts, (ii) amending existing provisions or (iii) introducing new provisions that may 'solve' the selected legal problems. The verb 'solve' in the latter sense refers to suggestions and recommendations that can contribute to actual solutions to the selected legal problems. In some cases, it might be sufficient to simply interpret existing provisions in a new manner (i). This applies, for example, to the fairness principle. Re-interpreting legislation should ideally occur through judicial action performed by courts, i.e. the CJEU. New interpretations should also be reflected in regulatory guidance, for example, in guidelines established by the EDPB. In other cases, it might be necessary to tweak or completely redraft existing provisions (ii). This applies to the approach to exhaustively enumerate special categories of personal data, the right to rectification, the exceptions mentioned in the TSD and the right not to be subject to ADM. If a new interpretation or redrafting of existing provisions is not sufficient to solve the legal problem, it might be necessary to introduce new provisions (iii). This applies to the communication surveillance problem.

Let me briefly explain how I proceed when discussing solutions that could solve the selected legal problems. For each legal problem, I first set the scene and then propose concrete legal solutions to solve it. As a first step, I further examine the selected Type 2 and 3 legal problems and introduce additional analysis and interpretations that may be helpful in addressing these problems. In the second step, I provide concrete legal solutions for each legal problem discussed, namely, by proposing (i) a

[2142] Durga Prasad, Adi Narayana Reddy, Devara Vasumathi, 'Privacy-Preserving Naive Bayesian Classifier for Continuous Data and Discrete Data' in Raju Surampudi Bapi et al (eds) *First International Conference on Artificial Intelligence and Cognitive Computing* (Springer Nature 2019) 289-299; Ling Guo, 'Randomization Based Privacy Preserving Categorical Data Analysis' (DPhil thesis, University of North Carolina 2010) < http://csce.uark.edu/~xintaowu/publ/DissertationLing.pdf > accessed 8 February 2024; Klaus Jansen et al (eds), *Approximation, Randomization, and Combinatorial Optimization* (Springer 2004).
[2143] Peter Laud, Liina Kamm (eds), *Applications of Secure Multiparty Computing* (IOS Press BV 2015); Ronald Cramer, Ivan Bjerre Damgård, Jesper Buus Nielsen, *Secure Multiparty Computation and Secret Sharing* (Cambridge University Press 2015).
[2144] Justin Zhan, 'Using Homomorphic Encryption For Privacy-Preserving Collaborative Decision Tree Classification' (IEEE Symposium on Computational Intelligence and Data Mining, Honolulu 2007) <https://ieeexplore.ieee.org/document/4221360> accessed 8 February 2024; Zhiqiang Yang, Sheng Zhong, Rebecca N Wright, 'Privacy-Preserving Classification of Customer Data without Loss of Accuracy' (2005) <https://www.cs.columbia.edu/~rwright/Publications/sdm05.pdf > accessed 8 February 2024.
[2145] Cynthia Dwork, Aaron Roth, *The Algorithmic Foundations of Differential Privacy* (Now Publishers Inc 2014); Cynthia Dwork, 'Differential Privacy' in Michele Bugliesi et al (eds) *Automata, Languages and Programming* (Springer 2006) 1-12.
[2146] Sergey I Niolenko, 'Synthetic Data for Deep Learning' (2019) <https://arxiv.org/pdf/1909.11512.pdf> accessed 8 February 2024; Khaled El Emam, Lucy Mosquera, Richard Hoptroff, *Practical Synthetic Data Generation* (O'Reilly Media Inc 2020).
[2147] Manas Gaur et al, 'Knowledge-Infused Learning: A Sweet Spot in Neuro-Symbolic AI' (2022) Vol 26 Iss 4 IEE Internet Computing, 5-11; Ugur Kursuncu, Manas Gaur, Amit Sheth, 'Knowledge Infused Learning (K-IL): Towards Deep Incorporation of Knowledge in Deep Learning' (2020) <https://arxiv.org/pdf/1912.00512.pdf > accessed 8 February 2024.

new interpretation of the relevant provision, where possible. Thus, preference is given to new interpretations of existing legislation through judicial action by the CJEU or through guidelines. If this is impossible, I suggest (ii) amendments of existing provisions or (iii) entirely new provisions. The third step wraps up by means of a short conclusion.

6.2 Fairness principle – the elusiveness problem

The elusiveness problem (Type 2)

AI systems are likely to process personal data in a way that would typically be considered as unfair. The elusive role and meaning of the fairness principle reduces legal certainty and makes it difficult for data subjects to challenge the fairness of processing enabled by AI systems and enforce the fairness principle accordingly.

6.2.1 Setting the scene

As indicated[2148] in Section 4.3, scholars distinguish two types of fairness, i.e. procedural and substantive fairness. *Procedural fairness* refers to formal or process-oriented requirements[2149] focussing on whether the data have been obtained or processed through unfair means, e.g. by deception or without the knowledge of the individual concerned.[2150] Eskens as well as Wachter and Mittelstadt interpret fairness as a mere proxy for transparency[2151] which essentially falls under procedural fairness as it merely focusses on formal transparency requirements. According to their views, fairness does not merit an independent meaning because it solely relates to transparency, it is not defined in the GDPR and it only appears in the context of lawfulness or transparency.[2152] Eskens interpretation of fairness as mere transparency is backed by the argument that 'fair processing' is never mentioned in the GDPR.[2153]

[2148] Parts of Section 4.3 and Section 6.2 resulted in a publication see Andreas Häuselmann, Bart Custers, 'Substantive fairness in the GDPR: Fairness Elements for Article 5.1a GDPR' (2024) Vol 52 Computer Law & Security Review 105942.
[2149] Inge Graef, Damien Clifford, Peggy Valcke, 'Fairness and enforcement: bridging competition, data protection, and consumer law' (2018) Vol 8 No 3 International Data Privacy Law 200, 203.
[2150] Cecile de Terwangne, Commentary of Article 5 in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020) 314.
[2151] Sarah Johanna Eskens, 'Profiling the European Citizen in the Internet of Things: How Will the General Data Protection Regulation Apply to this Form of Personal Data Processing, and How Should It? (2016) Master thesis, University of Amsterdam 27 <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2752010> accessed 8 February 2024; Sandra Wachter, Brent Mittelstadt 'A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI' (2019) No 2 Columbia Business Law Review 494, 581-582.
[2152] Sarah Johanna Eskens, 'Profiling the European Citizen in the Internet of Things: How Will the General Data Protection Regulation Apply to this Form of Personal Data Processing, and How Should It? (2016) Master thesis, University of Amsterdam 27 <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2752010> accessed 8 February 2024; Sandra Wachter, Brent Mittelstadt 'A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI' (2019) No 2 Columbia Business Law Review 494, 582.
[2153] Sarah Johanna Eskens, 'Profiling the European Citizen in the Internet of Things: How Will the General Data Protection Regulation Apply to this Form of Personal Data Processing, and How Should It? (2016) Master thesis, University of Amsterdam 27 <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2752010> accessed 8 February 2024.

I think such an interpretation of fairness is not convincing. First, none of the terms mentioned in the data protection principles are defined as such in Article 4 GDPR. Rather, *some* of these principles are further substantiated in the GDPR. Article 6 GDPR, for example, implements the lawfulness principle by enumerating six legal grounds for processing. Articles 12-14 GDPR further substantiate the transparency principle by imposing specific information obligations on controllers. Other principles, such as accuracy and data minimisation, are not further substantiated in the GDPR. Second, the fact that Article 5 (1) lit a GDPR mentions fairness together with lawfulness, and transparency does not imply that these notions mean the same. If so, the legislator would not have introduced these three distinct notions and mentioned in Recital 39 GDPR that 'any processing shall be lawful *and* fair'. Of course, recitals do not have binding legal value in EU law, but they are helpful to determine the nature of a provision and expand an ambiguous provision's scope.[2154] Fourth, the claim that 'fair processing' is never mentioned in the GDPR is simply wrong. Article 5 (1) lit a GDPR literally states that 'personal data shall be processed [..], *fairly*', which is another linguistic form of expressing 'fair processing'. Fifth, regulatory enforcement at the EU level confirms that the fairness principle has an independent meaning.[2155]

Thus, the interpretation of fairness as merely procedural fairness is not convincing. Principles are open norms. They allow judges to adjust the law to changing circumstances when approaching contemporary problems. As open norms, principles are well suited to adjust data protection legislation to changing technological circumstances to achieve the goals set by the fundamental right to data protection, including legislative goals pursued by the GDPR. The latter particularly aims to achieve a consistent and high level of protection for personal data (Recitals 6 and 10), a strong and coherent data protection framework (Recital 7) and effective protection[2156] (Recital 11). The fairness principle's breadth of scope and its open texture[2157] make it a particularly suitable candidate to host normative parameters beyond transparency and to prevent data subjects from unwarranted discrimination, power imbalance and risk of vulnerability.[2158] *Substantive fairness* is more promising and suitable to solve fairness issues concerning the processing of personal data by AI systems. It aims at preventing adverse

[2154] Tadas Klimas, Jflrate Vaitiukait, 'The Law of Recitals in European Community Legislation' (2008) Vol 15 No 1 ILSA Journal of International & Comparative Law 61, 63.
[2155] Binding Decision 3/2022 on the dispute submitted by the Irish SA on Meta Platforms Ireland Limited and its Facebook service (Art. 65 GDPR), adopted on 5 December 2022 paras 22, 477; Binding Decision 4/2022 on the dispute submitted by the Irish SA on Meta Platforms Ireland Limited and its Instagram service (Art. 65 GDPR), adopted on 5 December 2022 para 226, 444; Binding Decision 5/2022 on the dispute submitted by the Irish SA on Meta Platforms Ireland Limited and its WhatsApp service (Art. 65 GDPR), adopted on 5 December 2022.
[2156] Case C-645/19 [2021] *Facebook Ireland* ECR I-483 para 45; Case C-319/20, *Meta Platforms Ireland Limited* [2022] ECR I-322 para 73.
[2157] Lee A Bygrave, 'Minding the machine v2.0: The EU General Data Protection Regulation and Autmated Decision Making' in Karen Yeung and Martin Lodge (eds) *Algorithmic Regulation* (OUP 2019) 260.
[2158] Lee A Bygrave, 'Machine Learning, Cognitive Sovereignty and Data Protection Rights with Respect to Automated Decisions' (2021) University of Oslo Faculty of Law Legal Studies Research Paper Series No. 202-35 at 22, 23 <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3721118> accessed 8 February 2024.

effects in concrete circumstances, in particular when conflicting interests need to be balanced.[2159] Also, EU primary sources seem to refer to a substantive conception of fairness.[2160] Interpreting fairness as substantial fairness complies with the CJEU's approach to favour the interpretation of a provision which is the most effective. According to settled case law, if a provision of EU law is open to several interpretations, preference must be given to the interpretation that ensures and maintains the effectiveness of the provision in question.[2161] Both regulatory guidance[2162] and regulatory enforcement at the EU level[2163] point to substantive fairness by mentioning *reasonable expectations* of the data subjects, possible *adverse consequences* of processing and effects of *power imbalance* as relevant factors of the fairness principle. Therefore, I suggest that fairness, in addition to procedural fairness covered by transparency obligations, be interpreted as *substantive fairness.* I further explain this concept in Section 6.2.2. Before doing so, I quickly elaborate on how the notion of fairness is interpreted in two other fields of EU law, namely, consumer protection and competition law. These two areas of law are particularly relevant because they deal with notions of fairness. This might provide helpful information to further substantiate this notion under data protection law. Fairness under these areas of law could therefore inform the principle of fairness under data protection law.[2164]

In consumer protection law, fairness focusses on the decision capacity of consumers. Fairness acts as the substantive standard against which the legality of contractual terms and commercial practices are tested.[2165] Under the Unfair Terms Directive (UTD),[2166] 'good faith' and 'no significant imbalance' are components of fairness that must be examined together. The principle of good faith has its roots in Roman law[2167] under the term 'bona fides'. Applying the principle of good faith in the context of consumer law requires the contracting parties to take each other's interests into account in order to achieve a fair balance.[2168] A contractual term is unfair if, contrary to the requirement of good faith, it

[2159] Gianclaudio Malgieri, 'The concept of Fairness in the GDPR' (FAT* '20: Conference on Fairness, Accountability, and Transparency, Barcelona, January 2020) 2, 3 <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3517264> accessed 8 February 2024.
[2160] Giulia Gentile, 'Two Strings to One Bow? Article 47 of the EU Charter of Fundamental Rights in the EU Competition Case Law: Between Procedural and Substantive Fairness' (2020) Vol 4 No 2 Market and Competition Law Review 169, 177.
[2161] Case C-31/17 *Cristal Union* [2018] ECR I-168 para 41; Case C-517/07 *Afton Chemical* [2008] ECR I-751 para 43; Case C-152/13 *Holger Forstmann Transporte* [2014] ECR I-2184 para 26.
[2162] European Data Protection Board, 'Guidelines on Article 6(1)(b) GDPR' (Guidelines 2/2019, 8 October 2019), at 6.
[2163] Binding Decision 3/2022 on the dispute submitted by the Irish SA on Meta Platforms Ireland Limited and its Facebook service (Art. 65 GDPR), adopted on 5 December 2022 paras 219-220; Binding Decision 4/2022 on the dispute submitted by the Irish SA on Meta Platforms Ireland Limited and its Instagram service (Art. 65 GDPR), adopted on 5 December 2022 paras 223-224; Binding Decision 5/2022 on the dispute submitted by the Irish SA on Meta Platforms Ireland Limited and its WhatsApp service (Art. 65 GDPR), adopted on 5 December 2022.
[2164] Milda Mačėnaitė, 'Protecting Children Online: Combining the Rationale and Rules of Personal Data Protection Law and Consumer Protection Law' in Mor Bakhoum et al (eds) *Personal Data in Competition, Consumer Protection and Intellectual Property Law* (Springer Nature 2018) 361.
[2165] Inge Graef, Damien Clifford, Peggy Valcke, 'Fairness and enforcement: bridging competition, data protection, and consumer law' (2018) Vol 8 No 3 International Data Privacy Law 200, 204.
[2166] Articles 3-5 of the Council Directive 93/13/EEC of 5 April 1993 on unfair terms in consumer contracts [1993] OJ 01993L0013 further on UTD.
[2167] Hugh Collins, 'Good Faith in European Contract Law' (1994) Vol 14 No 2 Oxford Journal of Legal Studies 229, 250.
[2168] Mahmoud Fayyad, 'Measures of the Principle of Good Faith in European Consumer Protection and Islamic Law, a Comparative Analysis' (2014) Vol 28 Arab Law Quarterly 205, 208; Martin Schermaier, 'Bona Fides in Roman Contract

causes a significant imbalance to the detriment of the consumer.[2169] In order to pass the fairness test under the UTD, a term must not necessarily have been individually negotiated, it must be contrary to good faith and cause a significant imbalance in the contracting parties' rights and obligations to the detriment of the consumer. In addition, when assessing good faith, particular regard should be given to the strength of the bargaining positions of the parties.[2170] From the UTD, I define *good faith* as preventing *imbalances* between the interests of the seller and consumer that are to the detriment of the consumer as a component of fairness. In the Unfair Commercial Practices Directive (UCPD),[2171] fairness focusses on the average consumer's capacity to make informed autonomous decisions.[2172] A commercial practice is unfair if it is contrary to professional diligence and distorts or is likely to distort the consumer's economic behaviour,[2173] causing the consumer to act transactionally in a way he would have otherwise not done.[2174]

Article 5 UCPD divides fairness into three levels. The UCPD protects from misleading and aggressive commercial practices and contains a blacklist of practices that are deemed de facto unfair.[2175] Aggressive practices prohibit coercion and undue influence.[2176] The prohibition of misleading practices protects consumers from taking transactional decisions that they would not have taken in the absence of false or untruthful information provided by the trader.[2177] Thus, from the UCPD I derive undue *interferences* with a consumer's *autonomy* as a component of fairness. What also follows from the concept of fairness under EU consumer law is the rationale to protect the *weaker party* (i.e. a consumer) vis-à-vis the *stronger party* (i.e. trader).[2178]

The exact meaning of fairness in EU competition law is controversial,[2179] and it is not clear what constitutes 'fair' or 'unfair' behaviour.[2180] This is, among other reasons, due to the fact that fairness depends on the context as the legality of practices under competition, law is evaluated on the basis of

Law' in Reinhard Zimmermann, Simon Whittaker (eds) *Good Faith in European Contract Law* (Cambridge University Press 2000) 65.
[2169] Article 3 (1) UTD.
[2170] Pinar Akman, *The Concept of Abuse in EU Competition Law* (Hart Publishing Ltd 2012) 177.
[2171] Directive 2005/29/EC of the European Parliament and of the Council of 11 May 2005 concerning unfair business-to-consumer commercial practices in the internal market OJ L 149/22 furtheron 'UCPD'.
[2172] Inge Graef, Damien Clifford, Peggy Valcke, 'Fairness and enforcement: bridging competition, data protection, and consumer law' (2018) Vol 8 No 3 International Data Privacy Law 200, 204.
[2173] Pinar Akman, *The Concept of Abuse in EU Competition Law* (Hart Publishing Ltd 2012) 180.
[2174] Sarah Brown, 'European regulation of consumer credit: enhancing consumer confidence and protection from a UK perspective?' in James Devenney et al (eds) *Consumer credit, debt and investment in Europe* (Cambridge University Press 2012) 74.
[2175] Inge Graef, Damien Clifford, Peggy Valcke, 'Fairness and enforcement: bridging competition, data protection, and consumer law' (2018) Vol 8 No 3 International Data Privacy Law 200, 204.
[2176] Article 8 UCPD.
[2177] Article 6 UCPD.
[2178] Pinar Akman, *The Concept of Abuse in EU Competition Law* (Hart Publishing Ltd 2012) 182.
[2179] Giulia Gentile, 'Two Strings to One Bow? Article 47 of the EU Charter of Fundamental Rights in the EU Competition Case Law: Between Procedural and Substantive Fairness' (2020) Vol 4 No 2 Market and Competition Law Review 169, 170.
[2180] Pinar Akman, *The Concept of Abuse in EU Competition Law* (Hart Publishing Ltd 2012) 146.

anticompetitive nature or effects in the specific circumstances of a case.[2181] However, anticompetitive effects are considered unfair because they ultimately deprive consumers of the power to arbitrate the marketplace, which underscores the social rationale of EU competition policy.[2182] In EU competition law, Article 102 TFEU prohibits certain unfair behaviour as abuse of a dominant position.[2183] Such abuse consists, for example, of imposing unfair purchase or selling prices as well as other unfair trading conditions, limiting production, markets or technical development to the prejudice of consumers.[2184] Ultimately, Article 102 TFEU aims at regulating the abuse of market power, and it has been argued that unfairness in the context of competition law simply means exploitation.[2185] In competition law, fairness is pivotal for a pluralistic market in which companies shall not exploit dominant positions and consumers can efficiently use their financial resources.[2186] Exploitation presupposes power inequalities between the parties concerned. In this context, power relates to the ability of private parties to influence one another to their respective preferred outcomes.[2187] In case of power inequalities, one party uses its stronger position vis-à-vis the weaker party to obtain outcomes that it could not have achieved without that disparity in power.[2188] Thus, from EU competition law, I derive two components of fairness: i) non-exploitation of dominant positions and ii) recalibrating power inequalities.

Table 6.2 lists the components of (un)fairness according to EU consumer protection and competition law. As will be illustrated in Section 6.2.2, these components are also helpful to substantiate the fairness principle under EU data protection law.

Components of (un)fairness	**Area of EU law**
Preventing unfair imbalances between the parties to the detriment of *the consumer by means of the concept of good faith*	Consumer protection
Exercising undue influence on the consumer's *autonomy*	Consumer protection
Protecting the weaker party (consumer) from the stronger party (trader)	Consumer protection
Non-exploitation of dominant positions	Competition law
Recalibrating *power inequalities*	Competition law

Table 6.2 Components of 'un'fairness according to EU consumer protection and competition law.

[2181] Inge Graef, Damien Clifford, Peggy Valcke, 'Fairness and enforcement: bridging competition, data protection, and consumer law' (2018) Vol 8 No 3 International Data Privacy Law 200, 204.
[2182] Damien Gerard, 'Fairness in EU Competition Policy: Significance and Implications' (2018) Vol 9 No 4 Journal of European Competition Law & Practice 211-212.
[2183] Pinar Akman, *The Concept of Abuse in EU Competition Law* (Hart Publishing Ltd 2012) 146.
[2184] Article 102 TFEU.
[2185] Pinar Akman, *The Concept of Abuse in EU Competition Law* (Hart Publishing Ltd 2012) 184.
[2186] Giulia Gentile, 'Two Strings to One Bow? Article 47 of the EU Charter of Fundamental Rights in the EU Competition Case Law: Between Procedural and Substantive Fairness' (2020) Vol 4 No 2 Market and Competition Law Review 169, 177.
[2187] Daniel D Barnhizer, 'Inequality of bargaining power' (2005) Vol 76 Iss 1 University of Colorado Law Review 139,159.
[2188] Pinar Akman, *The Concept of Abuse in EU Competition Law* (Hart Publishing Ltd 2012) 173.

6.2.2 Solution: interpretation including substantive fairness

Fairness relates to both procedural and substantive fairness. The provisions in the GDPR and the corresponding recitals mostly refer to procedural fairness and provide clarity and protection in that respect. Procedural fairness contributes to fairness by elevating the controller's accountability duty to ensure effective compliance with data protection principles in the concrete case at stake. However, the lack of clarity regarding the substantive meaning of fairness creates the elusiveness problem. In this section, I argue that including substantive fairness can solve this problem. Substantive fairness, as suggested here, has two main elements.

First, substantive fairness focusses on the *outcome* or *consequences* of a process[2189] as opposed to procedural fairness which examines the fairness of the procedure by which that outcome was reached.[2190] To focus on the *outcome* or *consequence* of a certain processing activity in the context of the fairness principle neatly aligns with other provisions in the GDPR. For example, if a controller intends to further process personal data for purposes other than those for which the personal data have been initially collected, the possible consequences of such further processing must be considered.[2191] Articles 13 (2) lit f and 14 (2) lit g GDPR[2192] oblige controllers to inform data subjects about the envisaged consequences of ADM and profiling. Article 35 (1) GDPR requires controllers to assess 'the *impact* of the envisaged processing operations on the protection of personal data' where such processing operations are 'likely to *result* in a high risk to the rights and freedoms of natural persons'. In addition, Recital 150 GDPR requires supervisory authorities to take the consequences of a GDPR infringement into consideration when determining any administrative fine to be imposed on a controller.

The second major element of substantive fairness concerns fairness between *the parties* in question.[2193] It recalibrates imbalanced situations and is used in other areas of law, such as employment law.[2194] In the context of data protection law, substantive fairness as suggested here concerns fairness between the *controller* and the *data subject*. This element of substantive fairness aligns with other provisions in the GDPR. The relationship between controller and a data subject is mentioned in Article 6 (4) lit b and Recital 50 GDPR. According to these provisions, the controller needs to take its relationship with the data subject into consideration. The same applies in the context of the Legitimate Interest Assessment. When assessing whether to rely on its legitimate interest for a certain processing

[2189] Stephen A Smith, 'In Defence of Substantive Fairness' (1996) Vol 112 Iss 1 Law Quarterly Review 138-158.
[2190] Pinar Akman, *The Concept of Abuse in EU Competition Law* (Hart Publishing Ltd 2012) 166.
[2191] Article 6 (4) lit d and Recital 50 GDPR.
[2192] See also Recital 60 GDPR.
[2193] Stephen A Smith, 'In Defence of Substantive Fairness' (1996) Vol 112 Iss 1 Law Quarterly Review 138-158.
[2194] Giulia Gentile, 'Two Strings to One Bow? Article 47 of the EU Charter of Fundamental Rights in the EU Competition Case Law: Between Procedural and Substantive Fairness' (2020) Vol 4 No 2 Market and Competition Law Review 169, 173.

activity, the controller needs to take the reasonable expectations of data subjects into account, based on the controller's relationship with the data subjects.[2195]

It is often easier to determine whether a particular outcome is unfair rather than to agree on whether the outcome is fair.[2196] This is indicated in the title of the two major directives in EU consumer protection law which both use the term 'unfair'. Likewise, EU competition law explicitly prohibits certain unfair behaviour as abuse of a dominant position.[2197] Therefore, I suggest focussing on components of fairness that may lead to unfair processing of personal data, rather than to fair processing. Table 6.3 lists the components that must be considered when assessing fairness in the context of processing personal data. The components are divided into the two major elements of substantive fairness, i.e. fairness between the parties and fairness of the outcome.

Components concerning fairness between the parties	
No power inequalities / dominant positions	Is the controller exploiting power inequalities and/or dominant market positions?
Vulnerability	Is the data subject vulnerable?
Good faith	Does the balancing of interests violate the concept of good faith?
Components concerning fairness of the outcome	
Autonomy	Is it likely that the processing will negatively affect the data subject's autonomy and, in particular, decisional privacy?
Non-manipulation	Does the processing create risks regarding the manipulation of the data subject?
No detrimental effects	Does the processing likely lead to detrimental effects for the data subject, e.g. due to the nature of the personal data processed?
Accuracy	Is the processed personal data likely to be inaccurate or is it difficult to determine the accuracy of the processed personal data?
Non-discrimination	Is the outcome of the processing likely to be discriminatory?

Table 6.3 Components of substantive fairness to be considered under the fairness principle in EU data protection law.

The components of substantive fairness listed in Table 6.3 comprehensively protect data subjects from unfair processing because they focus on both the relationship between the data subject and the controller *as well as* on possibly unfair outcomes of processing. These components of substantive fairness specifically address the legal problems identified in this thesis. Obviously, it might be necessary to add additional components in the future as new or additional legal problems arise.

[2195] Recital 47 GDPR.
[2196] Francis Herbert Buckley, 'Three Theories of Substantive Fariness' (1990) Vol 19 Hofstra Law Review 33, 56.
[2197] Article 102 TFEU.

At first sight, it might be surprising that the component '*power inequalities/dominant positions*' should be assessed in the context of fairness under data protection law, as these concepts originate from EU consumer and competition law, which have different legislative aims. Nevertheless, there is often a power inequality between the controller and data subject: It is the controller that determines the purpose of processing, the legal basis for processing, how long data will be stored, whether personal data are accurate, with whom data will be shared and for which purposes personal data will be further processed after collection. Data subjects have enforceable rights, but they cannot influence most of the decisions the controller takes regarding these rights. There is a clear power inequality between the data subjects and the controller, and this power inequality should be considered when assessing fairness in data protection law. In terms of *abusing dominant positions*, which is a concept from EU competition law, competition authorities increasingly take non-compliance into consideration when assessing whether an undertaking abuses its dominant position or engages in other anti-competitive practices.

The Bundeskartellamt, which is Germany's Competition Authority, initiated proceedings due to Google's data processing terms, which allegedly amount to prohibited anticompetitive practices.[2198] AG Rantos argued that competition authorities may take compliance with the rules enshrined in the GDPR into consideration when examining an undertaking's conduct under EU competition law.[2199] The CJEU followed the AG's opinion, provided that the competition authority fulfils its duty of 'loyal cooperation' and consults the competent data protection supervisory authority.[2200] Also, the circumstance in which a controller holds a dominant market position is a relevant factor when assessing whether consent according to Article 4 (11) GDPR is freely given, because a dominant market position affects the freedom of choice of the data subject.[2201] Thus, the CJEU confirms that dominant market position and power imbalance are relevant factors to be considered in the context of data protection law. For this reason, it must be possible to also consider a controller's dominant market position and power imbalances between the controller and the data subject when assessing fairness in EU data protection law.

Vulnerability is mentioned in Recital 75 GDPR in the context of security of processing. The recital states that children must be considered in particular as 'vulnerable natural persons'. However, it is not only children who are potentially vulnerable data subjects. In my view, data subjects are also particularly vulnerable when special categories of personal data relating to them are being processed. Due to the sensitivity of such data, processing is particularly eligible to create harm.[2202] Vulnerability

[2198] See < https://www.bundeskartellamt.de/SharedDocs/Meldung/EN/Pressemitteilungen/2023/11_01_2023_Google_Data_Processing_Terms.html > accessed 8 February 2024.
[2199] Case C-252/21 *Meta Platforms Inc.* [2022] ECR I-704, Opinion of AG Rantos paras 23-33.
[2200] Case C-252/21 *Meta Platforms Inc.* [2023] ECR I-537 paras 56-63.
[2201] Ibid paras 148-149, 154.
[2202] Art 29 Working Party, 'Advice paper on special categories of data ('sensitive data')' (20 April 2011) at 4.

also plays an important role in the processing of emotion data. In fact, revealing emotions makes an individual potentially more vulnerable.[2203] Although not specifically mentioned in any specific provisions, data protection law arguably manifests the idea that data subjects are vulnerable to power imbalances created by digital technologies[2204] simply by regulating the processing of personal data. It therefore seems reasonable to assess the vulnerability[2205] of data subjects in the context of the fairness principle, and not only in the context of other provisions in the GDPR such as provisions relating to consent, DPIAs and ADM.[2206]

Traditional conceptions of *good faith* have their roots in virtue ethics as well as Roman law and essentially refer to the idea of acting in good conscience or not unconscionably, which would prevent taking advantage of another's trust.[2207] The classical notion of bona fides is today enjoying a renaissance and helps modern lawyers to solve current issues.[2208] This applies particularly to virtue ethics. For example, it has been suggested to adopt a virtue ethics approach to privacy regulation.[2209] Virtue ethics focusses on the notion of the good or virtuous person.[2210] Aristotle is seen as the dominant influence on the conceptual profile of virtue.[2211] He conceptualised virtues as character traits[2212] such as such as honesty, courage and patience that promote the performance of right or excellent actions.[2213] In particular, the virtues honesty and trust[2214] seem to relate to the concept of good faith. Good faith is well suited to prevent controllers from taking advantage of their stronger position and should therefore be considered when assessing the fairness of processing. In fact, some have argued to broaden the understanding of the fairness principle in data protection law with the aim to prevent processing contrary to good faith.[2215]

The fairness components *autonomy* and non-*manipulation* are closely related. The essence of autonomy is indicated by the etymology of the term: autos (self) and nomos (rule or law).[2216] The ruling

[2203] Aaron Ben-Ze'Ev, *The Subtlety of Emotions* (MIT Press 2000) 183.
[2204] Ryan Calo, 'Privacy, Vulnerability, and Affordance' (2017) Vol 66 Iss 2 Depaul Law Review 591, 592-593; Gianclaudio Malgieri, Jędrzej Niklas, 'Vulnerable data subjects' (2020) Vol 37 Computer Law & Security Review 2-16.
[2205] For an extensive analysis of vulnerable data subjects, see Gianclaudio Malgieri, Vulnerable People and Data Protection Law (Oxford University Press 2022).
[2206] Gianclaudio Malgieri, Jędrzej Niklas, 'Vulnerable data subjects' (2020) Vol 37 Computer Law & Security Review 2-16.
[2207] Hugh Collins, 'Good Faith in European Contract Law' (1994) Vol 14 No 2 Oxford Journal of Legal Studies 229, 250.
[2208] Martin Schermaier, 'Bona Fides in Roman Contract Law' in Reinhard Zimmermann, Simon Whittaker (eds) *Good Faith in European Contract Law* (Cambridge University Press 2000) 89.
[2209] Bart van der Sloot, *Privacy as Virtue* (Cambridge University Press 2017) 107-143.
[2210] Nathan R Kollar, 'Virtue Ethics' in John K Roth (ed) *Ethics* (Salem Press Inc 2005) 562.
[2211] Shannon Vallor, *Technology and the Virtues: A Philosophical Guide to a Future Worth Wanting* (OUP 2016) 18.
[2212] Bart van der Sloot, *Privacy as Virtue* (Cambridge University Press 2017) 109.
[2213] Shannon Vallor, *Technology and the Virtues: A Philosophical Guide to a Future Worth Wanting* (OUP 2016) 18.
[2214] The virtues honesty and trust are related; see Shannon Vallor, *Technology and the Virtues: A Philosophical Guide to a Future Worth Wanting* (OUP 2016) 121. See also Aimee van Wynsberghe, 'Artificial intelligence: from ethics to policy' (2020) study prepared for European Parliament, 12.
[2215] Milda Mačėnaitė, 'Protecting Children Online: Combining the Rationale and Rules of Personal Data Protection Law and Consumer Protection Law' in Mor Bakhoum et al (eds) *Personal Data in Competition, Consumer Protection and Intellectual Property Law* (Springer Nature 2018) 368.
[2216] Gerald Dworkin, *The Theory and Practice of Autonomy* (Cambridge University Press 1988) 12, 18.

idea of personal autonomy is 'that people should make their own lives',[2217] which means facing freely both existential and everyday choices.[2218] A person is considered to be autonomous when her decisions and actions are her own and thus self-determined,[2219] i.e. a person acts but is not acted upon.[2220] Autonomy is closely related to privacy, partly because privacy seems to be a precondition for autonomy.[2221] It has become one of the core pillars of the fundamental right to privacy under case law adopted by the ECtHR.[2222]

External influences such as manipulation constitute threats to personal autonomy.[2223] The concept of decisional privacy is well suited to address concerns about manipulation.[2224] Decisional privacy refers to being free to make personal decisions and choices.[2225] This erodes when manipulation invades internal thought processes, affects free will or interferes with an individual's self-interest.[2226] As explained in Section 4.3.3, manipulation aims to influence people's choices in ways that circumvent or counter rational decision-making.[2227] It refers to exercising direct influence on an individual's beliefs, desires or emotions to the detriment of individual self-interest[2228] and may involve the act of altering the actual choices available to a person or changing this person's perception of those choices.[2229] Fairness in data protection law should take into account autonomy and non-manipulation because processing of personal data by means of AI generates personal data that might be used in a way that negatively affects the data subject's autonomy. AC generates emotion data that could be used to the detriment of the data subject. Emotions play an important role in the elicitation of autonomous motivated behaviour.[2230] According to research in behavioural science, especially psychology, emotions

[2217] Joseph Raz, *The Morality of Freedom* (Oxford University Press 1986) 369.
[2218] Daniel Susser, Beate Roessler, Helen Nissenbaum 'Technology, autonomy, and manipulation' (2019) Vol 8 Iss 2 Internet Policy Review 1, 8.
[2219] Gerald Dworkin, *The Theory and Practice of Autonomy* (Cambridge University Press 1988) 13.
[2220] See Berlin, which explains the concept of autonomy under the heading positive liberty: 'Isaiah Berlin, *Liberty* (Hendry Hardy ed Oxford University Press 1969) 185; Marijn Sax, *Between Empowerment and Manipulation* (Kluwer Law International B.V. 2021) 131.
[2221] Hildebrandt Mireille, Koops Bert-Jaap, 'The challenges of Ambient Law and Legal Protection in the Profiling Era' (2010) Vol. 73 (3) The Modern Law Review 428, 435.
[2222] Bart van der Sloot, 'Decisional privacy 2.0: the procedural reqirements implicit in Article 8 ECHR and its potential impact on profiling' Vol 7 No 3 International Data Privacy Law 190, 192, *Munjaz v the United Kingdom* App no 2913/06 (17 July 2012) para 80; *NB v Slovakia* App no 29518/10 (12 June 2012); *IG and others v Slovakia* App no 15966/04 (13 November 2012); *VC v Slovakia* App no 18968/07 (8 November 2011).
[2223] Lawrence Haworth, 'Dworkin on Autonomy' (1991) Vol 102 Ethics 129, 136.
[2224] Marjolein Lanzig, 'Strongly Recommended: Revisiting Decisional Privacy to Judge Hypernudging in Self-Tracking Technologies' (2019) Vol 32 Philosophy & Technology 549-568.
[2225] Bart van der Sloot, 'Decisional privacy 2.0: the procedural reqirements implicit in Article 8 ECHR and its potential impact on profiling' Vol 7 No 3 International Data Privacy Law 190, 192.
[2226] Francisco Lupiáñez-Villanueva et al, 'Behavioural study on unfair commercial practices in the digital environment: dark patterns and manipulative personalisation' (2022) Final Report produced by European Innovation Council and SMEs Executive Agency on behalf of the European Commission 92 < https://op.europa.eu/en/publication-detail/-/publication/606365bc-d58b-11ec-a95f-01aa75ed71a1/language-en/format-PDF/source-257599418 > accessed 8 February 2024.
[2227] Allen W Wood, 'Coercion, Manipulation, Exploitation' in Christian Coons, Michael Weber (eds) *Manipulation* (Oxford University Press 2014) 35.
[2228] Anne Barnhill, 'What is Manipulation?' in Christian Coons, Michael Weber (eds) *Manipulation* (Oxford University Press 2014) 52.
[2229] Ruth Faden, Tom Beachamp, Nancy King, *A History and Theory of Informed Consent* (Oxford University Press 1986) 354.
[2230] Leen Vandercammen et al, 'On the Role of Specific Emotions in Autonomous and Controlled Behaviour' (2014) Vol 28 Iss 5 European Journal of Personality 413, 445.

constitute powerful, pervasive and predictable drivers of decision-making.[2231] Emotions can have significant effects on economic transactions and play a powerful role in everyday economic choices.[2232] Likewise, accurate predictions generated by means of ML through the processing of personal data (e.g. purchase history) might be used to manipulate data subjects through tailored recommendations in a way that actions of the data subject are no longer self-determined.

Detrimental effects are at the core of substantive fairness because they directly refer to the *outcome* or *consequences* of a process.[2233] Output generated by AI systems may have detrimental effects for data subjects in many ways. Predictions facilitated by ML approaches, such as negative score values, can prevent the data subject from obtaining a loan to buy a house, a mobile subscription or health insurance coverage. The emotional state of an applicant detected during an automated video assessment can play a role when the hiring manager decides whether the applicant will be invited for the second round of interviews. Such detrimental effects generated by means of AI are generally problematic in terms of substantive fairness. They become even more problematic when the output generated by AI systems is *inaccurate* or *likely* to be inaccurate. Inaccurate personal data may pose significant risks, for example, in the form of economic or reputational harm.[2234] Predictive profiling powered by ML may be used to predict an individual's behaviour, character, risk (e.g. score values) and to treat the individual accordingly.[2235] Predictions can hardly be absolutely certain and are poorly verifiable in the sense that they cannot be verified in advance or sometimes not at all (e.g. the individual is a 'high credit risk' or 'likely to buy a house in two years').[2236] Essentially, ML-based predictions or classifications constitute 'educated guesses based on large amounts of data'.[2237] Inference 'is always an invasion of the unknown, a leap from the known'.[2238] Examples include predictions about a customer's future life such as estimated advancements in career,[2239] credit risk scores, life expectancy scores or future health.[2240] Emotion data generated by means of AC can also be inaccurate.

[2231] Jennifer S Lerner et al, 'Emotion and Decision Making' (2015) Vol 66 Annual Review of Psychology 799, 802.
[2232] Jennifer S Lerner, Deborah A Small, George Loewenstein, 'Heart Strings and Purse Strings' (2004) Vol 15 No 5 American Psychology Society 337-340.
[2233] Stephen A Smith, 'In Defence of Substantive Fairness' (1996) Vol 112 Iss 1 Law Quarterly Review 138-158.
[2234] Danielle Keats Citron, Daniel J Solove, 'Privacy Harms' (2022) Vol 102 Iss 3 Boston University Law Review 793, 817.
[2235] Helena U Vrabec, 'Uncontrollable: Data Subject Rights and the Data-driven Economy' (Dissertation, Leiden University 2019) 220; Hans Lammerant, Paul de Hert, 'Predictive profiling and its legal limits: Effectiveness gone forever' In Bart van der Sloot et al (eds) *Exploring the boundaries of big data* (2016 Amsterdam University Press/WRR) 145-173.
[2236] Sandra Wachter, Brent Mittelstadt, 'A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI' (2019) Issue 2 Columbia Business Law Review 494, 510.
[2237] Teresa Scantaburlo, Andrew Charleswoth, Nello Cristianini, 'Machine Decisions and Human Consequences' in Karen Yeung and Martin Lodge (eds) *Algorithmic Regulation* (OUP 2019) 57; Lee A Bygrave, 'Machine Learning, Cognitive Sovereignty and Data Protection Rights with Respect to Automated Decisions' (2021) University of Oslo Faculty of Law Legal Studies Research Paper Series No. 202-35 at 5 <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3721118> accessed 8 February 2024.
[2238] John Dewey, *The Middle Works of John Dewey, Volume 9, 1899-1924* (Carbondale Southern Illinois University Press 1980) 165.
[2239] Gianclaudio Malgieri, 'Trade Secrets v. Personal Data: A Possible Solution for Balancing Rights' (2016) Vol 6 No 2 International Data Privacy Law 102, 113-114; Sandra Wachter, Brent Mittelstadt 'A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI' (2019) No 2 Columbia Business Law Review 495, 607.
[2240] OECD Working Party on Security and Privacy in the Digital Economy JT03357584 (2014) <https://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=dsti/iccp/reg(2014)3&doclanguage=en> accessed 8 February 2024.

According to an extensive study on affect recognition from facial expressions, it is not possible to confidently infer happiness from a smile, anger from a scowl or sadness from a frown because these emotion categories are more variable in their facial expressions.[2241] Other means to detect emotions, for example, based on speech (see Section 2.2.4.2) and physiological data (see Section 2.2.4.3), have been challenged due to a lack of scientific consensus whether such methods can ensure accurate or even valid results.[2242] It has been argued to broaden the understanding of the fairness principle in data protection law with the aim to prevent processing, which might have detrimental effects for the data subjects concerned.[2243]

Simply putting someone at risk may have a detrimental effect for the data subject, even if that risk never materialises. Harms relating to the processing of inaccurate personal data are highly contextual and depend on how such data are subsequently used. Adverse effects and actual harm depend on various factors such as by which controller the personal data are used, to whom it is disclosed and whether it is shared with other controllers.[2244] In any case, inaccurate personal data inherently causes the risk of possible detrimental effects, regardless of whether this risk materialises. Therefore, the accuracy of personal data also should be considered when assessing fairness in data protection law.

That *discrimination* must be considered in the context of substantive fairness is obvious. There are many examples that processing personal data by means of AI systems may lead to discriminatory outcomes. Due to deficiencies in reasoning capabilities, AI systems may generate discriminatory output. Google's photo app automatically classified images of black people as gorillas.[2245] In New Zealand, a man of Asian descent had his passport application rejected because the software that approves photos claimed his eyes were closed.[2246] Face recognition systems perform poorly in recognising individuals of different ethnicities. For example, face recognition software of Hewlett Packard could not recognise dark-coloured faces as faces.[2247] ADM based on ML could discriminate by means of classes or groups that lead to emergent forms of discrimination based on patterns that have little or

[2241] Lisa Feldman Barrett et al. 'Emotional Expressions Reconsidered' (2019) Vol 20 (1) Psychological Science in the Public Interest 1, 46.
[2242] Kate Crawford et al, 'AI Now Report' (2019) AI Now Institute 12 <https://ainowinstitute.org/publication/ai-now-2019-report-2> accessed 8 February 2024.
[2243] Milda Mačėnaitė, 'Protecting Children Online: Combining the Rationale and Rules of Personal Data Protection Law and Consumer Protection Law' in Mor Bakhoum et al (eds) *Personal Data in Competition, Consumer Protection and Intellectual Property Law* (Springer Nature 2018) 368.
[2244] Danielle Keats Citron, Daniel J Solove, 'Privacy Harms' (2022) Vol 102 Iss 3 Boston University Law Review 793, 817-818.
[2245] Crawford Kate, 'Artificial Intelligence's White Guy Problem' *The New York Times* (New York, 25 June 2016) <https://www.nytimes.com/2016/06/26/opinion/sunday/artificial-intelligences-white-guy-problem.html> accessed 8 February 2024.
[2246] Titcomb James, 'Robot passport checker reject Asian man's photo for having his eyes closed' The Telegraph (London, 7 December 2016) <https://www.telegraph.co.uk/technology/2016/12/07/robot-passport-checker-rejects-asian-mans-photo-having-eyes/> accessed 8 February 2024.
[2247] Frederik Zuiderveen Borgesius, 'Discrimination, artificial intelligence, and algorithmic decision-making'(2019) Report for the Anti-discrimination department of the Council of Europe, 17 <https://rm.coe.int/discrimination-artificial-intelligence-and-algorithmic-decision-making/1680925d73> accessed 8 February 2024.

no intuitive meaning to human practice and thus are socially unrecognisable.[2248] Newly identified classes or groups by means of ML arguably facilitate new forms of social classification with far-reaching socioeconomic consequences,[2249] such as new types of socioeconomic stratification and social hierarchies,[2250] and could consequently lead to new forms of discrimination.[2251] AI may reflect the conscious and unconscious biases of the people who assemble it and thus produce biased outcomes.[2252] This is called encoded bias because the designer's values are 'frozen into the code, effectively institutionalising those values'.[2253] The interests, needs and life experiences of the AI developers will be reflected in the AI they develop,[2254] potentially including stereotyped thinking in terms of traditional gender roles[2255] or racial/ethnic prejudices.

Because humans label much of the training data, human biases and cultural assumptions may be transmitted by classification choices.[2256] Discriminatory attitudes and stereotypes of developers are translated and reflected in the AI system they build.[2257] The developer's prejudices may be reinforced within the ADM system,[2258] and because ML algorithms are applied to every case in which ADM is deployed, they arguably have a bigger potential to discriminate systematically than human decision makers who may discriminate on a case-by-case basis.[2259] This is not only a theoretical concern. Diversity in the ML and AI community is, in fact, an issue. A study that focussed on the 4,000 researchers who published at leading AI and ML conferences disclosed that 88% of the contributions was by from men and only 12% by women.[2260] People that investigate, design and develop AI systems tend

[2248] Monique Mann, Tobias Matzner 'Challenging algorithmic profiling: The limits of data protection and anti-discrimination in responding to emergent discrimination' (2019) Vol 6 Iss 2 Big Data & Society 2, 6 < https://journals.sagepub.com/doi/pdf/10.1177/2053951719895805> accessed 8 February 2024.
[2249] Shoshana Zuboff, *The age of surveillance capitalism* (PublicAffairs 2019).
[2250] Stratification typically focus on income, wealth, occupational structures, social mobility etc see Cecilia L Ridgeway, 'Why Status Matters for Inequality' (2013) Vol 79 Iss 1 American Sociological Review 1, 3.
[2251] Raphaële Xenidis, 'Tuning EU equality law to algorithmic discrimination> Three pathways to resilience' (2020) Vol 27 Iss 6 Maastricht Journal of European and Comparative Law 7636, 752.
[2252] Brent Daniel Mittelstadt et al, 'The ethics of algorithms: Mapping the debate' (2016) Vol 3 Iss 2 Big Data & Society 1, 7.
[2253] Kevin Macnish, 'Unblinking the eyes: the ethics of automating surveillance' (2012) Vol 14 Ethics and Information Technology 151, 158.
[2254] Alex Campolo et al, 'AI Now Report' (2017) 15 <https://ainowinstitute.org/publication/ai-now-2017-report-2> accessed 8 February 2024.
[2255] Janneke Gerards, Raphaële Xenidis, 'Algorithmic discrimination in Europe: Challenges and Opportunities for geneder equality and non-discrimination law' (2021) at 51 study prepared for the European Commission <https://op.europa.eu/en/publication-detail/-/publication/082f1dbc-821d-11eb-9ac9-01aa75ed71a1> accessed 05 May 2021.
[2256] Alex Campolo et al, 'AI Now Report' (2017) 15 <https://ainowinstitute.org/publication/ai-now-2017-report-2> accessed 8 February 2024.
[2257] Janneke Gerards, Raphaële Xenidis, 'Algorithmic discrimination in Europe: Challenges and Opportunities for geneder equality and non-discrimination law' (2021) at 41 study prepared for the European Commission <https://op.europa.eu/en/publication-detail/-/publication/082f1dbc-821d-11eb-9ac9-01aa75ed71a1> accessed 8 February 2024.
[2258] Kevin Macnish, 'Unblinking the eyes: the ethics of automating surveillance' (2012) Vol 14 Ethics and Information Technology 151, 158.
[2259] Indrė Žliobaitė, 'Measuring discrimination in algorithmic decision making' (2017) Vol 31 Data Mining Knowledge Discovery 1060, 1063.
[2260] Mantha Yoan, Hudson Simon, 'Estimating the Gender Ratio of AI researchers Around the World' <https://medium.com/element-ai-research-lab/estimating-the-gender-ratio-of-ai-researchers-around-the-world-81d2b8dbe9c3> accessed 8 February 2024.

to be male, highly educated and very highly paid.[2261] The AI Now Institute found that there is a diversity crisis in the AI sector across gender and race. It found that more than 80% of AI professors are men and in the private sector only 15% of AI research staff at Facebook and 10% at Google are women. When considering diversity in terms of skin colour, the picture looks even worse: only 2.5% of Google's workforce is black, while Facebook and Microsoft are each at 4%.[2262] Therefore, it seems very important to also consider non-discrimination when assessing fairness in the context of data protection law.

6.2.3 Conclusion

In this section, I have outlined that the legal solution to solve the elusiveness problem consists of interpreting the fairness principle in data protection law as both procedural and substantive fairness. The provisions in the GDPR and the corresponding recitals provide clarity with respect to procedural fairness. Substantive fairness, as suggested here, contains two main elements: fairness between the parties and fairness of the outcomes. Table 6.2 contains six components of substantive fairness, distributed over the two main elements of substantive fairness. These components are no power inequalities/dominant positions, vulnerability, good faith, autonomy, non-manipulation, detrimental effects, accuracy and non-discrimination. They indicate *unfairness*. My *solution* to the elusiveness problem is to adopt extensive EDPB guidelines on the principle of fairness and include these components of substantive fairness. In fact, both regulatory guidance[2263] and regulatory enforcement at the EU level[2264] already point to at least three components[2265] of substantive fairness proposed. However, specific regulatory guidance on the principle of fairness does not yet exist, although this principle merits further substantiation in detailed guidelines. To consider the suggested components of substantive fairness is in line with the CJEU's settled case law to give preference to the method of interpretation that ensures and maintains the effectiveness of the provision.[2266] To ultimately 'solve' the elusiveness problem, judicial action is needed. Thus, the CJEU should interpret fairness in EU data protection law as referring to both procedural and substantive fairness.

[2261] Alex Campolo et al, 'AI Now Report' (2017) 5 <https://ainowinstitute.org/publication/ai-now-2017-report-2> accessed 8 February 2024.
[2262] Sarah West, Meredith Whittacker, Kate Crawford 'Discriminating AI Systems: Gender, Race and Power' (2019) AI Now Institute 3 <https://ainowinstitute.org/publication/discriminating-systems-gender-race-and-power-in-ai-2> accessed 8 February 2024.
[2263] European Data Protection Board, 'Guidelines on Article 6(1)(b) GDPR' (Guidelines 2/2019, 8 October 2019), at 6.
[2264] Binding Decision 3/2022 on the dispute submitted by the Irish SA on Meta Platforms Ireland Limited and its Facebook service (Art. 65 GDPR), adopted on 5 December 2022 paras 219-220, 222-223; Binding Decision 4/2022 on the dispute submitted by the Irish SA on Meta Platforms Ireland Limited and its Instagram service (Art. 65 GDPR), adopted on 5 December 2022 paras 223-224, 226-227; Binding Decision 5/2022 on the dispute submitted by the Irish SA on Meta Platforms Ireland Limited and its WhatsApp service (Art. 65 GDPR), adopted on 5 December 2022.
[2265] These are possible (i) adverse consequences of processing which is the same as my suggested component *detrimental effects*, (ii) the data subject's autonomy and (iii) effects of power imbalance which essentially relate to my suggested component of *power inequalities.*
[2266] Case C-31/17 *Cristal Union* [2018] ECR I-168 para 41; Case C-517/07 *Afton Chemical* [2008] ECR I-751 para 43; Case C-152/13 *Holger Forstmann Transporte* [2014] ECR I-2184 para 26.

The advantage of this approach is that controllers can and should consider these components when performing a Data Protection Impact Assessment (DPIA) as required by Article 35 GDPR. According to this provision, controllers must carry out a DPIA if the envisaged processing is likely to result in a high risk to the rights and freedoms of data subjects. This is particularly the case when the controller uses 'new technologies'[2267] for processing, which arguably applies to processing by AI systems. My proposal is also in line with teleological interpretation in EU law, which tasks the CJEU to give concrete expressions to notions that are too general or of which the meaning is unclear.[2268]

6.3 Enhanced protection for 'special data' – the mental data problem

The mental data problem (Type 3)

ML and AC facilitate the processing of mental data, i.e. any data used to infer mental states of individuals including thoughts, beliefs and underlying mechanisms and processes. Mental data are inherently sensitive and form the core of an individual's private sphere. Despite this, mental data are not specifically protected under the GDPR because the approach to enumerate special categories of personal data exhaustively cannot keep up with the developments in AI. This principle creates a significant gap of protection and is not fit for purpose to protect the fundamental right to data protection.

6.3.1 Setting the scene

As outlined in Section 4.8.3, the approach to exhaustively enumerate special data fails. It cannot keep up with technological developments in AI that facilitate unprecedented ways to generate or otherwise process new types or categories of sensitive personal data. Mental data forms the core of an individual's private sphere.[2269] They may contain information concerning unexecuted behaviour, such as unuttered thoughts and intended actions,[2270] information previously inaccessible to others. Therefore, mental data are particularly sensitive and in need of specific protection.

To solve the mental data problem and other legal problems inextricably linked to it (i.e. emotion data, location data and neurodata problems), new or revised legislation is unavoidable. This is due to the wording of Article 9 (1) GDPR, which does not provide any room to broaden the scope of this

[2267] Article 35 (1) GDPR.
[2268] Pierre Pescatore, 'Les objectifs de la Communauté européenne comme principes d'interprétation dans la jurisprudence de la Cour de justice' (1972) vol 2 Miscellanea W.J. Ganshof van der Meersch 328; Koen Lenaerts, José A Gutiérrez-Fons, 'To Say What the Law of the EU Is: Methods of Interpretation and the European Court of Justice' (2013) European University Institute Working Paper AEL 2013/9 at 6 <https://cadmus.eui.eu/bitstream/handle/1814/28339/AEL_2013_09_DL.pdf?sequence=1&isAllowed=y> accessed 8 February 2024.
[2269] Dara Hallinan et al, 'Neurodata and Neuroprivacy: Data Protection Outdated?' (2014) Vol 12 Iss 1 Surveillance and Society 68 <https://ojs.library.queensu.ca/index.php/surveillance-and-society/article/view/neurodata/neurodata4> accessed 8 February 2024.
[2270] Marcello Ienca, Gianclaudio Malgieri, 'Mental data protection and the GDPR' (2022) Vol 9 Iss 1 Journal of Law and the Biosciences 1, 6.

provision by means of different interpretation methods. Literal (textual) interpretation is the prevailing method of interpretation if the provision to be interpreted is clear and precise.[2271] The wording in Article 9 (1) GDPR clearly points to an exhaustive enumeration of special personal data. Typical words from the legal jargon ('for instance', 'such as', 'inter alia' etc.) used to indicate non-exhaustiveness are absent. According to settled case law,[2272] the literal meaning of a provision cannot be called into question by means of contextual or teleological interpretation if provision is clear and precise.[2273] Thus, the re-interpretation of Article 9 (1) GDPR through judicial action performed by the CJEU is not an option. Having established that new or revised legislation is unavoidable, I now elaborate how this could be done. Before doing so, I briefly reflect on the rationale for regulating special data. To avoid confusion, I use the term 'special data' to refer to data that are, in fact, listed and thus currently protected under the GDPR and 'sensitive data' for data that are currently *not specifically protected* under the GDPR (although they arguably should be).

According to the CJEU, the rationale to ensure enhanced protection for special data stems from their particular sensitivity. Processing of special data is likely to constitute a particularly serious interference with the fundamental rights to privacy and data protection.[2274] Recital 51 GDPR stresses the particularly sensitive nature of such data. According to AG Rantos, the object is to prevent significant risks for data subjects arising from the processing of special data, regardless of any subjective element such as the controller's *intention*. Thus, intentions do *not* play a role when determining whether personal data constitutes special data or not.[2275] In the view of SAs, specific protection for special data is needed because misuse may have more severe consequences for data subjects than misuse of 'regular' personal data.[2276] This is underscored by Recital 51 GDPR, which states that 'processing [of sensitive personal data] could create significant risks to the fundamental rights and freedoms'. Nevertheless, the approach to provide specific protection for certain categories of personal data is not undisputed.[2277]

In what I call the 'context objection', Bygrave claims that the sensitivity of personal data is context-dependent.[2278] In the 'use objection', Moerel and Prins argue that the sensitivity of personal data

[2271] Koen Lenaerts, José A Gutiérrez-Fons, 'To Say What the Law of the EU Is: Methods of Interpretation and the European Court of Justice' (2013) European University Institute Working Paper AEL 2013/9 at 6 <https://cadmus.eui.eu/bitstream/handle/1814/28339/AEL_2013_09_DL.pdf?sequence=1&isAllowed=y> accessed 8 February 2024.
[2272] Case C-220/03 *BCE* [2005] ECR I-10595 para 3; Case C-263/06 *Carboni e derivati* [2008] ECR I-1077 para 48; Case C-48/07 *Les Vergers du Vieux Tauves* [2008] ECR I-10627 para 44.
[2273] Koen Lenaerts, José A Gutiérrez-Fons, 'To Say What the Law of the EU Is: Methods of Interpretation and the European Court of Justice' (2013) European University Institute Working Paper AEL 2013/9 at 7 <https://cadmus.eui.eu/bitstream/handle/1814/28339/AEL_2013_09_DL.pdf?sequence=1&isAllowed=y> accessed 8 February 2024.
[2274] Case C-252/21 *Meta Platforms Inc.* [2023] ECR I-537 para 70; Case C-184/20, *OT* [2022] ECR I-601, para 126; Case C-136/17, GC and Others [2019] ECR I-773 para 44; Recital 51 GDPR.
[2275] Case C-252/21 *Meta Platforms Inc.* [2023] ECR I-537 paras 69-70; see also Opinion of AG Rantos para 41.
[2276] Art 29 Working Party, 'Advice paper on special categories of data ('sensitive data')' (20 April 2011) at 4.
[2277] Ludmila Georgieva, Christopher Kuner Commentary of Article 9 in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020) 370.
[2278] Lee A. Bygrave, *Data Privacy Law: An International Perspective* (OUP 2014) 165.

essentially depends on the specific use of personal data.[2279] In their view, the regime for special categories of personal data is no longer meaningful because it is becoming less and less clear which data are sensitive and that the focus should be on the use of data when determining sensitivity of processing.[2280] One of the examples they provide is the case of an email address, which in itself is not sensitive data, but in combination with a password becomes highly sensitive because many individuals use the same email and password to access different websites.[2281] Similarly, regulatory guidance stresses the importance of a more flexible approach to sensitive personal data because the context plays an important role in determining the sensitivity of a certain processing activity.[2282]

The 'context' and 'use' objections are valid, but they are not new. Already traveaux préparatoire relating to the DPD drafted in the 1990s point to the context and use objections.[2283] More importantly, the GDPR explicitly requires one to take the context into account when it comes to the processing of special data. Recital 51 GDPR states that special data merits specific protection because the *context* of their processing may create significant risks for data subjects. The reference to 'context' in this recital was added at an advanced stage of the legislative procedure and was not included in the European Commission's initial proposal.[2284] Thus, the legislator made a deliberate choice to recognise context as a relevant factor when it comes to the processing of special data. This is precisely what the CJEU did when ruling that also personal data which *indirectly* reveal special data are covered by Article 9 GDPR.[2285] In this case, it was possible to derive information with respect to the sex life or sexual orientation of the data subject from 'non-sensitive' personal data published on the Internet, i.e. name-specific data relating to the spouse, cohabitee or partner of that data subject.[2286] This ruling addresses the context and use objections: arguably non-sensitive personal data might become sensitive depending on its specific use and context.

According to US scholar Solove, the current approach with respect to special data is a dead end, and the only viable solution is to focus on use, harm and risk.[2287] According to his 'dead-end' objection,

[2279] Lokke Moerel, Corien Prins, 'Privacy for the Homo Digitalis: Proposal for a New Regulatory Framework for Data Protection in the Light of Big Data and the Internet of Things' (2016) p 11 and 56 <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2784123> accessed 8 February 2024.
[2280] Ibid 11.
[2281] Ibid 56.
[2282] However, note that EU Supervisory Authorities do not seem to be fully aligned in this point; see Art 29 Working Party, 'Advice paper on special categories of data ('sensitive data')' (20 April 2011) at 9-10.
[2283] 'It is generally accepted that the right to privacy is endangered, *not* by the *contents* of personal data, *but* by the *context* in which the processing of personal data takes place.' Commission, 'Proposal for a Council Directive concerning the protection of individuals in relation to the processing of personal data' COM (90) 314 final, explanatory memorandum p 35 <https://resources.law.cam.ac.uk/cipil/travaux/data_protection/COMPLETETRAVEAU(ENGLISH)DPDIRECTIVE.pdf#page=1P> accessed 8 February 2024.
[2284] See Recital 41 at page 24 < https://www.europarl.europa.eu/registre/docs_autres_institutions/commission_europeenne/com/2012/0011/COM_COM(2012)0011_EN.pdf > accessed 8 February 2024.
[2285] Case C-184/20, *OT* [2022] ECR I-601.
[2286] Case C-184/20, *OT* [2022] ECR I-601 paras 117-128.
[2287] Daniel J Solove, 'Data Is What Data Does: Regulating Use, Harm, and Risk Instead of Sensitive Data' (2024) Vol 11 No 4 Northwestern University Law Review 1081, 1083 <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4322198> accessed 8 February 2024.

the special data categories are arbitrary and based on blurry lines. Moreover, in Solove's view, nearly all personal data are special due to the capabilities of powerful ML algorithms. Processing of non-sensitive personal data by means of ML can generate inferences about special data, which means that most controllers are processing vast amounts of special data in violation of the law.[2288]

Solove's dead-end objection completely ignores the rationale of EU law[2289] to specifically protect special data, which involve both prevention of harm *and* risks. According to the CJEU, a heightened standard of protection for special data is needed because this processing is likely to constitute a particularly serious interference with the fundamental rights to privacy and data protection.[2290] Obviously, interferences relate to both harm and risks. According to AG Rantos, the objective is to prevent *significant risks* for data subjects arising from the processing of special data.[2291] The connotation on risks for data subjects is also stressed in Recital 51 GDPR, which states that 'processing [of special data] could create significant risks to the fundamental rights and freedoms'. Thus, Article 9 GDPR proactively prevents harms *and* risks by prohibiting the processing of special data, unless an exception applies. Thus, contrary to what Solove claims in the dead-end objection, the prevention of harm and risks for data subjects *is* covered by the rationale to specifically protect special data.[2292] In addition, substantive fairness as introduced in Section 6.2 provides additional protection against harm and risk, as it focusses on whether the outcome of processing is fair. Therefore, what is left from Solove's 'dead-end' objection is the call to focus on the use, which ultimately boils down to the 'context' and 'use' objections. Moreover, Solove exaggerates when claiming that nearly all personal data is sensitive simply because inferences by means of ML *are possible*. He presumes that almost all controllers engage in such processing and oversimplifies processing performed by means of ML. Arguably, mainly controllers that have the technological know-how and sufficient financial resources engage in such processing, but not 'most organisations' as claimed in Solove's dead-end objection.[2293] Only controllers that *in fact* infer special data by means of ML need to comply with Article 9 GDPR. Solove's dead-end objection mentions powerful ML algorithms several times, but he ignores new types of highly sensitive personal data (e.g. emotion data, mental and neurodata) that can be generated by means of the various AI disciplines discussed in this thesis. Instead, Solove mentions rather trivial

[2288] Daniel J Solove, 'Data Is What Data Does: Regulating Use, Harm, and Risk Instead of Sensitive Data' (2024) Vol 11 No 4 Northwestern University Law Review 1081, 1083, 1084 <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4322198> accessed 8 February 2024.
[2289] Although he is a US scholar, Solove extensively discusses EU law in his contribution relating to the dead-end objection. The GDPR is mentioned 68 times, and Solove admits that the approach to regulating sensitive data stems from the EU. It can, therefore, also be expected that the EU's rationale to regulate sensitive data is acknowledged and discussed.
[2290] Case C-252/21 *Meta Platforms Inc.* [2023] ECR I-537 para 70; Case C-184/20, *OT* [2022] ECR I-601, para 126; Case C-136/17, GC and Others [2019] ECR I-773 para 44; Recital 51 GDPR.
[2291] Case C-252/21, *Meta Platforms Inc.* [2022] ECR I-704, Opinion of AG Rantos para 41.
[2292] Case C-252/21 *Meta Platforms Inc.* [2023] ECR I-537 para 70; Case C-184/20, *OT* [2022] ECR I-601, para 126; Case C-136/17, GC and Others [2019] ECR I-773 para 44; Case C-252/21, *Meta Platforms Inc.* [2022] ECR I-704, Opinion of AG Rantos para 41; Recital 51 GDPR.
[2293] Daniel J Solove, 'Data Is What Data Does: Regulating Use, Harm, and Risk Instead of Sensitive Data' (2024) Vol 11 No 4 Northwestern University Law Review 1081, 1084 <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4322198> accessed 8 February 2024.

examples, for example inferences concerning political beliefs or opinions, sexual orientation, ethnicity, health status and race derived from Facebook likes.[2294]

There are basically two possible approaches for new or revised legislation concerning the processing of special personal data. The first approach is to enumerate specific categories of special personal data ('current approach'). The second approach is to make the sensitivity of a certain processing dependent on the context and specific use of personal data ('contextual approach'). Obviously, both approaches have their (dis)advantages.

The contextual approach has the main advantage that it is quite flexible, as sensitivity depends on the use and context, not on the content of the personal data processed. For example, processing health data by insurance companies for the benefit of data subjects would not be considered sensitive, whereas processing health data to exclude data subjects from insurance coverage would be. In addition, the contextual approach would allow employers to launch initiatives to improve diversity and inclusion within the company. For example, employers could use unsupervised ML to detect correlations and patterns in data relating to the current workforce, which might be helpful to improve their businesses. The current approach makes such initiatives difficult when considering that none of the exceptions to the processing of sensitive data listed in Article 9 (2) GDPR is applicable in this case.
The main advantage of the contextual approach, i.e. flexibility, is simultaneously also a disadvantage. In my view, this approach gives controllers too much flexibility when considering the power imbalance between controllers and data subjects. Ultimately, it is the controller that determines the use of personal data by defining the purpose of processing. Controllers can define purposes with enough specificity and can demonstrate that such purposes are legitimate, meaning any purpose is valid under the GDPR.[2295] Hence, relying on the sensitive use of personal data is not suitable to actually prevent risks and harms for data subjects because controllers determine the use of personal data. They have considerable freedom when doing so and can be creative in defining it as a 'non-sensitive' use. In addition, it is rather difficult to determine precisely which types of use should be regarded as particularly harmful or risky. It is even more difficult to anticipate and foresee all imaginable harmful uses that might emerge in the future. This approach is questionable from the perspective of legal certainty, which notably constitutes one of the GDPR's legislative aims.[2296]

[2294] Daniel J Solove, 'Data Is What Data Does: Regulating Use, Harm, and Risk Instead of Sensitive Data' (2024) Vol 11 No 4 Northwestern University Law Review 1081, 1099-1109 <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4322198> accessed 8 February 2024.
[2295] Asia Biega, Michèle Finck, 'Reviving Purpose Limitation and Data Minimisation in Data-Driven Systems' (2021) Technology and Regulation 44, 49 <https://arxiv.org/abs/2101.06203> accessed 8 February 2024.
[2296] Recitals 7 and 13 GDPR.

The current approach is more convincing from the perspective of legal certainty, because the GDPR lists all types of special data in Article 9, and some are even defined.[2297] Additionally, it is more suitable because it starts from a general prohibition of processing special data. Controllers need to be able to rely on one of the exceptions in Article 9 (2) GDPR. The current approach is based on the rationale that there are specific types of personal data with an inherently sensitive nature as stressed by Recital 51 GDPR.[2298] When considering the inherently sensitive nature of mental data, neurodata and emotion data generated by AI, it should not play a role in which context or for which purpose such data are processed. *Mental data* refers to the processing of information relating to the mental states of individuals. Mental states comprise all conscious and non-conscious mental representations, events, processes and propositional attitudes, including thoughts, beliefs, emotions and moods, as well as the underlying psychological mechanisms (collectively referred to as 'mental states').[2299] Mental data are perceived to form the core of an individual's private sphere[2300] and are therefore of a particularly sensitive nature. *Neurodata* provide unique insights into people[2301] and their behaviour.[2302] Scholars have argued that neurodata are a particularly sensitive class of data due to their direct link with mental processes[2303] and the strong link to the individual's personhood.[2304] Also, *emotion data* have a strong link to personhood. Information regarding emotions is of sensitive and intimate nature[2305] because there is an inherent relationship between emotions and personhood[2306] and privacy is considered fundamental to the maintenance of human dignity and the boundary to one's personhood.[2307] Thus, neurodata, mental data and emotion data are of inherently sensitive nature and merit

[2297] Genetic data in Article 4 (13), biometric data in Article 4 (14) and health data in Article 4 (15) GDPR.
[2298] The following reasoning contained in preparatory documents for the DPD, on which Article 9 GDPR is built, holds still true. 'Certain categories of data which, by virtue of their *contents* – quite *irrespective* of the *context* in which they are *processed* – carry the risk of infringing the data subject's right to privacy' COM (90) 314 final, explanatory memorandum p 35 <https://resources.law.cam.ac.uk/cipil/travaux/data_protection/COMPLETETRAVEAU(ENGLISH)DPDIRECTIVE.pdf#page=1P> accessed 8 February 2024
[2299] Jan-Christoph Bublitz, 'The Nascent Right to Psychological Integrity and Mental Self-Determination' in Andreas von Arnauld, Kerstin von der Decken, Mart Susi (eds) *The Cambridge Handbook of New Human Rights* (Cambridge University Press 2020) 30; Marcello Ienca, Gianclaudio Malgieri, 'Mental data protection and the GDPR' (2022) Vol 9 Iss 1 Journal of Law and the Biosciences 1, 4.
[2300] Dara Hallinan et al, 'Neurodata and Neuroprivacy: Data Protection Outdated?' (2014) Vol 12 Iss 1 Surveillance and Society 68 <https://ojs.library.queensu.ca/index.php/surveillance-and-society/article/view/neurodata/neurodata4> accessed 8 February 2024.
[2301] Neurodata are of highly personalised nature and allows for identification ('brain fingerprinting').
[2302] Brent J. Lance et al, 'Brain-Computer Interface Technologies in the Coming Decades' (2012) Vol 100 Proceedings of the IEE 1587.
[2303] Marcello Ienca, Roberto Andorno, 'Towards New Human Rights in the Age of Neuroscience and Neurotech- nology' (2017) Vol 13 Iss 1 Life Science, Society and Policy 1, 14; Marcello Ienca, Karolina Ignatiadis, 'Artificial Intelligence in Clinical Neuroscience: Methodological and Ethical Challenges' (2020) Vol 11 Iss 2 AJOB Neuroscience 77-87; Rafael Yuste et al, 'Four ethical priorities for neurotechnolgies and AI' (2017) Vol 551 Nature 159-163.
[2304] Marcello Ienca, Roberto Andorno, 'Towards New Human Rights in the Age of Neuroscience and Neurotech- nology' (2017) Vol 13 Iss 1 Life Science, Society and Policy 1, 14.
[2305] Andrew McStay, 'Emotion AI, soft biometrics and the surveillance of emotional life: An unusual consensus on privacy' (2020) Vol 7 Iss 7 Big Data & Society 1, 4.
[2306] Giovanni Stanghellini, René Rosfort, *Emotions and Personhood: Exploring Fragility – Making Sense of Vulnerability* (OUP 2013) 149.
[2307] William S Brown, 'Technology, Workplace Privacy and Personhood' (1996) Vol 15 Journal of Business Ethics 1237, 1243.

specific protection. It is therefore justifiable to maintain a 'sui generis' regime[2308] for such highly sensitive personal data. Letting it entirely up to the controllers to determine whether the envisaged use qualifies as sensitive, as in the contextual approach, is not a suitable solution.

To adjust the level of protection for special data to the harm or risk of harm as suggested in the dead-end objection seems unworkable in practice. Harms and risks are highly subjective, as they depend on the specific data subject concerned by the processing. What may constitute harm for one data subject might be different for another data subject. The same applies to the corresponding risks. Definitions of specific types of harm relating to the processing of special data are arguably too abstract to actually work in practice.[2309] By analogy, proving harm caused by the processing of personal data is inherently difficult. This is underscored by at least nine cases pending at the CJEU[2310] (at the time of writing beginning 2023) which address the compensation of non-material damages caused by GDPR infringements. According to a petition submitted to the Commission, the legislator failed to sufficiently specify when non-material damages exist and to name examples within the GDPR's recitals.[2311] This omission makes it rather difficult for data subjects to claim compensation for non-material damages because they carry the burden of proof. In its response to the petition, the Commission outlined that Recitals 75, 85 and 146 GDPR provide indications for the concept of non-material damages, and that this concept must be further clarified by national courts.[2312] Notably, Recitals 75 and 85 GDPR only mention examples of possible harms relating to *personal data breaches* as defined in Article 4 (12) GDPR. In addition, AG Campos Sánchez-Bordona seems to recognise the difficulty in determining exactly what constitutes harm and what not. He is 'in no doubt that there is a *fine line* between *mere upset* (which is not eligible for compensation) and *genuine* non-material damage (which is eligible for compensation)'. Likewise, he is aware of 'how complicated it is to *delimit*, in the *abstract*, the two categories and apply them to a particular dispute'.[2313] Arguably, it is exactly for these reasons that the legislator omitted to name examples of harm eligible for the compensation of non-material damages. Thus, the approach to adjust the level of protection for special data to the harm or risk of harm as suggested in the dead-end objection is unworkable in practice. Even the author of the dead-end objection admits that regulating use, harm and risk is a difficult road, fraught with complexity.[2314]

[2308] Koops suggests having sui generis regimes for types of data that have certain effects when they are processed see Bert-Jaap Koops, 'The trouble with European data protection law' (2014) Vol 4 No 4 International Data Privacy Law 250, 260.
[2309] Paul Ohm, 'Sensitive Information' (2015) Vol 88 Southern California Law Review 1125, 1147.
[2310] Cases C-340/21 *Natsionalna agentsia za prihodite*; C-300/21 *UI* [2022] ECR I-756*;* C-741/21 *Juris*; C-687/21 *Saturn Electro;* C-667/21 *Krankenversicherung Nordrhein;* C-189/22 *Scalable Capital;* C-182/22 *Scalable Capital* C-456/22 *Gemeinde Ummendorf*; C-590/22 *PS.*
[2311] Petition No 0386/2021 see <https://www.europarl.europa.eu/doceo/document/PETI-CM-699118_EN.pdf> accessed 8 February 2024.
[2312] Ibid.
[2313] Case C-300/21 *UI* [2022] ECR I-756 Opinion AG Manuel Campos Sánchez-Bordona para 116.
[2314] Daniel J Solove, 'Data Is What Data Does: Regulating Use, Harm, and Risk Instead of Sensitive Data' (2023) George Washingto University Law School Draft Research Paper 5 <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4322198> accessed 8 February 2024.

Instead of regulating based on harm, I suggest focussing on the compensation of non-material harm caused by GDPR infringements. Proving this and obtaining compensation according to Article 82 GDPR is extremely difficult for data subjects. This could be overcome by establishing a typology for non-material damages based on the nature of the infringed provision. Article 83 GDPR, which empowers SAs to impose administrative fines on controllers, contains a similar typology. The legislator seems to have weighed GDPR infringements normatively by setting up two different maximum amounts for fines. Infringements of principles and data subject rights can lead to fines of up to twenty million euros or 4% of a controller's annual worldwide turnover, while infringements of other GDPR provisions can lead to fines of up to ten million euros or 2% of a controller's annual worldwide turnover. This distinction indicates that infringements of principles and data subject rights are considered *more serious* than infringements of other provisions.[2315] Thus, the legislator provided an indication concerning the seriousness on an infringement in an abstract sense: the more serious the infringement, the higher the fine.[2316]

The same mechanism might be used to establish a typology for non-material damages. This typology puts a price on the infringement of GDPR provisions. The amount of non-material damages to be awarded for infringements of principles and data subject rights will be higher than for other GDPR infringements. Setting up this typology and embedding it in the GDPR would enable data subjects to *effectively* enforce their right to the compensation of non-material damages.[2317] Arguably, this will also have a deterrent effect on controllers because it facilitates collective actions pursued by bodies representing data subjects in order to obtain the compensation of non-material damages.[2318]

In my view, the current approach is suitable to *prevent* harm and risks arising from the processing of special data. It contains many layers of protection. Processing of such data is prohibited, unless an exception applies. In addition, processing of special data must always be supported by a legal basis[2319] and comply with other provisions[2320] of the GDPR.[2321] The fairness principle and its components listed in Table 6.3 in Section 6.2.2 form a particularly helpful layer of protection. The fairness components vulnerability, autonomy, non-discrimination and detrimental effects protect data subjects from *possible* harm. The controller's obligation to perform a Data Protection Impact Assessment (DPIA) for processing that is likely to result in a high risk for data subjects could be seen as another layer of

[2315] Article 29 Working Party, 'Guidelines on the application of administrative fines for the purposes of Regulation 2016/679' (WP 253, 3 October 2017) 9.
[2316] European Data Protection Board, 'Guidelines on the calculation of administrative fines under the GDPR' (Guidelines 4/2022, 16 May 2022) 16.
[2317] Article 82 (1) GDPR.
[2318] Article 80 (1) GDPR.
[2319] According to Article 6 GDPR; see also European Data Protection Board, 'Guidelines 3/2019 on the processing of personal data through video devices' (29 January 2020) at 17.
[2320] Such as principles for processing and other rules of the GDPR; see Recital 51 GDPR.
[2321] Ludmila Georgieva, Christopher Kuner Commentary of Article 9 in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020) 374, 376.

protection. According to Article 35 (3) GDPR, such a DPIA is mandatory if the controller processes special categories of personal data on a large scale. Actual harm can then be compensated.

However, the current approach also has disadvantages as it may lead to over or under protection of special data.[2322] Overprotection occurs for instance when the processing of special data is not particularly sensitive and is carried out for the benefit of the data subject. This holds true when health data are processed by insurance companies for the benefit of data subjects or when employers process special personal data to improve diversity and inclusion within the company. Typical examples of under protection are mental data, neurodata and emotion data (Section 4.9.3). These highly sensitive types of data are underprotected because they are not included in the exhaustive list of special data according to Article 9 GDPR.

To sum up, the current approach to specifically regulate special personal data with an inherently sensitive nature is at least better than the alternatives suggested in the 'use', 'context' and 'dead-end' objections. However, this approach is far from perfect and has its disadvantages; for instance, it may lead to over-regulation.

6.3.2 Solution: Introducing a dynamic list for special data

Section 4.8.3 concluded that the approach of enumerating special categories of personal data exhaustively is not fit for purpose to address the challenges caused by AI as it cannot keep up with technological developments. To solve this problem, I suggest a revision of Article 9 GDPR, which contains a dynamic list of special personal data. More specifically, I suggest that the European Commission be empowered to adopt delegated acts for the purpose of updating the list of special personal data where necessary due to technological developments. If new information technologies facilitate processing of new types of sensitive personal data, the Commission can proactively add such new categories to the list. Likewise, the Commission is also empowered to remove categories of personal data from that list when the inherently sensitive nature of such data ceases to exist, for example, due to societal changes. When doing so, the Commission should consider the rationale for the increased standard of protection for special data. The rationale is to prevent particularly serious interferences with the fundamental rights to privacy and data protection[2323] as well as corresponding significant risks for data subjects.[2324] In order to prevent over-regulation, it could be considered to also empower the Commission to add exceptions applicable to the processing of special data if corresponding scientific evidence is available.

[2322] Paul Ohm, 'Sensitive Information' (2015) Vol 88 Southern California Law Review 1125, 1146; Helen Nissenbaum, *Privacy in Context: Technology, Policy and the Integrity of Social Life* (Stanford Law Books 2010) 89-102.
[2323] Case C-184/20, *OT* [2022] ECR I-601, para 126; Case C-136/17, GC and Others [2019] ECR I-773 para 44; Recital 51 GDPR.
[2324] Recital 51 GDPR.

EU consumer law follows a similar approach[2325] in the Unfair Commercial Practices Directive (UCPD)[2326] as introduced in Section 6.2.1. In its annex, the UCPD contains a list with commercial practices which are regarded as unfair. However, this list can only be modified by revising the Directive, which makes it less feasible to anticipate quickly-evolving technological change.[2327]

My suggested solution is comparable to the AI Act's compromise text[2328] concerning high-risk systems referred to in Article 6 (2) and Annex III. According to Article 7 (1), the Commission is empowered to add or modify use-cases of high-risk AI systems contained in Annex III.[2329] A similar approach has been adopted in the Digital Services Act ('DSA').[2330] Article 87 DSA empowers the Commission to adopt delegated acts, for example, by laying down the methodology for calculating the number of average monthly active users[2331] or by laying down rules concerning audits to be pursued under the DSA.[2332] In order to proactively counter the argument that the Commission should not be empowered to enact law, I suggest including a similar provision as contained in Article 87 (6) DSA. This provision foresees that delegated acts by the Commission only enter into force if neither the European Parliament nor the Council raise objections.

The proposed solution provides a basic layer of protection for special personal data, i.e. a default prohibition of processing, and is able to address technological developments. In addition, it comes with legal certainty for all the actors involved in the processing of personal data: the controllers, the data subjects, the supervisory authorities and, in litigious cases, the Courts. The components of the fairness principle outlined in Section 6.2.2 constitute the second layer of protection. In particular, the components vulnerability, autonomy, non-discrimination and detrimental effects protect data subjects from possible harm.

I acknowledge that the suggested solution is far from perfect. However, for now, it seems at least *better* than the alternatives suggested in the 'use' and 'dead-end' objections. There are certainly disadvantages, the risk of over-regulation in particular. For example, it can be doubted whether the Commission would be willing to also remove special categories from the list and not only add new

[2325] Although with a different rationale, i.e. consumer law.
[2326] Directive 2005/29/EC of the European Parliament and of the Council of 11 May 2005 concerning unfair business-to-consumer commercial practices in the internal market OJ L 149/22 furtheron 'UCPD'.
[2327] In May 2022, the Commission launched a fitness check on EU consumer law, focussing on digital fairness. This fitness check determines whether additional legislative action is needed to ensure a high level of consumer protection in the digital environment. See <https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/13413-Digital-fairness-fitness-check-on-EU-consumer-law_en> accessed 8 February 2024.
[2328] On 2 February 2024, the Committee of the Permanent Representatives of the Governments of the Member States to the European Union unanimously approved the compromise text of the AI Act resulting from the trilogue negotiations see < https://data.consilium.europa.eu/doc/document/ST-5662-2024-INIT/en/pdf > accessed 8 February 2024.
[2329] Ibid.
[2330] Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October on a Single Market For Digital Services and amending Directive 2000/31/EC [2022] OJ L277/1 'Digital Services Act' (DSA).
[2331] Article 33 (3) DSA.
[2332] Article 37 (7) DSA.

categories. It also remains to be seen how the approach taken in the DSA plays out in practice. Nonetheless, I think that the suggested solution is still better than the available alternatives.

6.3.3 Conclusion

The legal solution to solve the mental data problem[2333] consists of a revision of Article 9 GDPR. This revision should introduce a dynamic list of special personal data. This list overcomes the current problem related to the approach to exhaustively enumerate special categories of personal data. The current approach is not fit for purpose to address the challenges caused by AI as it does not keep up with technological developments. In my suggested solution, the European Commission is empowered to adopt delegated acts to update the list of special personal data where needed in light of technological developments. This solution is flexible enough to address this and comes with legal certainty for all actors involved.

6.4 Confidentiality – the communication surveillance problem

The communication surveillance problem (Type 3)
ML, NLP and AC facilitate the surveillance of both human-machine and interpersonal communication. Major tech companies that offer human-machine communication services, such as virtual assistants, may easily intercept and otherwise process such communication. Providers of these services do not fall under the strict regime of Article 5 (1) ePD, which regulates the confidentiality of communications. This creates a significant gap in legal protection and outlines that the ePD is not fit for purpose to ensure the confidentiality of both interpersonal and human-machine communication.

6.4.1 Setting the scene

AI and people's interactions with it do not fit neatly into paradigms of communication theory that have focussed on human–human communication.[2334] The same can be said about the legal protection with respect to the confidentiality of human-machine communication. The AI discipline natural language processing (NLP) provides powerful means to analyse voice and speech data obtained by means of human-machine communications, in particular when combined with classification techniques adopted in the AI discipline machine learning (ML). With NLP and ML, rather sensitive information can be derived from human speech and other acoustic elements in recorded audio. In addition to the linguistic content of speech, a speaker's voice characteristics and manner of expression may contain a rich array of personal information, including clues with regard to the speaker's biometric identity, personality, physical traits, geographical origin, level of intoxication and sleepiness,

[2333] In addition to other legal problems that are inextricably linked to it (emotion data, location data and neurodata problems).
[2334] Andrea L Guzman, Seth C Lewis, 'Artificial intelligence and communication: A Human-Machine Communication agenda' (2020) Vol 22 Iss 1 New Media & Society 70-86.

age, gender, health condition and even an individual's socioeconomic status.[2335] In addition, speech-based emotion recognition systems powered by the AI discipline affective computing (AC) measure and quantify emotions of a person by observing speech signals of this person.[2336] Research has demonstrated specific associations between emotions such as fear, anger, sadness, joy and features of speech such as pitch, voice level and speech rate.[2337] Amazon's patented technology enabling the virtual assistant Alexa to recognise the user's emotional state derived from the user's voice constitutes a practical example of this (see Section 4.9.3).[2338] Likewise, tech companies may intercept interpersonal communication. For example, a former Apple employee revealed that he had listened to hundreds of Siri recordings every day, including unintended recordings, for the purpose of quality control.[2339] These recordings concerned sensitive interpersonal communications such as discussions between doctors and patients, business deals, seemingly criminal acts and sexual encounters.[2340] This is not an exception, and press coverage points to similar practices at Google[2341] and Amazon[2342] (see Section 4.9.3). In addition, both human-machine and interpersonal communications might be intercepted in the context of virtual assistant services for the purpose of serving targeted ads.[2343]

The protection gap regarding the confidentiality of human-machine communication and interpersonal communication captured in the context of virtual assistant services can only be solved by means of new or revised legislation. The literal interpretation of Article 5 (1) ePD that regulates the confidentiality of communications is clear: The provision does not apply to providers of virtual assistant services such as Amazon, Google and Apple given that these services do not constitute an electronic communication service (ECS) as defined in European Electronic Communications Code (EECC).[2344] The new definition of an ECS covers three types of services: (i) Internet access services, (ii)

[2335] Jacob Leon Kröger, Otto Hans-Martin Lutz, Philip Raschke, 'Privacy Implications of Voice and Speech Analysis – Information Disclosure by Inference' in Michael Friedewald et al (eds) *Privacy and Identity management. Data for Better Living: AI and Privacy* (Springer 2020) 242.
[2336] Chi-Chun Lee et al, 'Speech in Affective Computing' in Rafael Calvo et al (eds), *The Oxford Handbook of Affective Computing* (OUP 2015) 171.
[2337] Christina Sobn and Murray Alpert, 'Emotion in Speech: The Acoustic Attributes of Fear, Anger, Sandess, and Joy' (1999) Vol 28 No 4 Journal of Psycholinguistic Research, 347.
[2338] Huafeng Jin, Shuo Wang 'Voice-Based Determination of Physical and Emotional Characteristics of Users' US Patent Number US 10096319 B1 (Assignee: Amazon Technologies, Inc.) October 2018 <https://patentimages.storage.googleapis.com/f6/a2/36/d99e36720ad953/US10096319.pdf> accessed 8 February 2024.
[2339] Alex Hern, 'Apple contractors regularly hear confidential details on Siri recordings' *The Guardian* (London, 26 July 2019) < https://www.theguardian.com/technology/2019/jul/26/apple-contractors-regularly-hear-confidential-details-on-siri-recordings > accessed 8 February 2024.
[2340] Alex Hern, 'Apple contractors regularly hear confidential details on Siri recordings' *The Guardian* (London, 26 July 2019) < https://www.theguardian.com/technology/2019/jul/26/apple-contractors-regularly-hear-confidential-details-on-siri-recordings > accessed 8 February 2024.
[2341] Tom Simonite, 'Who's Listening When You Talk to Your Google Assistant?' *Wired* (New York, 10 July 2019) <https://www.wired.com/story/whos-listening-talk-google-assistant/> accessed 8 February 2024.
[2342] Alex Hern, 'Amazon staff listen to customers' Alexa recordings, report says' *The Guardian* (London, 11 April 2019) <https://www.theguardian.com/technology/2019/apr/11/amazon-staff-listen-to-customers-alexa-recordings-report-says> accessed 8 February 2024.
[2343] Joseph Cox, 'Marketing Company Claims That It Actually Is Listening to Your Phone and Smart Speakers to Target Ads' *404 Media* (United States, 14 December 2023) <Marketing Company Claims That It Actually Is Listening to Your Phone and Smart Speakers to Target Ads (404media.co)> accessed 8 February 2024.
[2344] Directive (EU) 2018/1972 of the European Parliament establishing the European Electronic Communications Network OJ L 321/36 further on 'EECC'.

interpersonal communications services and (iii) services consisting wholly or mainly in the conveyance of signals.[2345] It also includes over-the-top (OTT) services such as VoIP[2346] solutions, messaging services and web-based email services, which are functionally equivalent to traditional voice telephony and text message services.[2347] With regard to requirement (i), it is clear that virtual assistant services do not constitute Internet access services.

Concerning requirement (ii), an interpersonal communication service is defined as a 'service normally provided for remuneration that enables *direct interpersonal and interactive exchange of information* via *electronic communications networks* between a *finite number of persons*, whereby the persons initiating or participating in the communication determine the recipient(s) and do not include services which enable interpersonal and interactive communication merely as a minor ancillary feature that is intrinsically linked to another service'.[2348] Recital 17 EECC clarifies what is meant with interpersonal communication: communications between *natural persons*. Communications involving legal persons fall within the definition only to a limited extent, for instance if natural persons act on behalf of those legal persons.[2349] Thus, human-machine communications fall outside the scope of interpersonal communication services as defined in Article 2 (5) EECC.

Concerning requirement (iii), all that matters concerning the conveyance of signals is that a service provider is *responsible* vis-à-vis the end-users for *transmission* of the *signal* which ensures that they are supplied with the service to which they have subscribed.[2350] In the case of web-based services, it is the Internet Access Providers (IAPs) and the *operators* of the *various networks* of which the *open web* is based that convey the signals necessary for the functioning of web-based services.[2351] Providers of web-based services can participate in the conveyance of signals, for example, by means of uploading data packets to the Internet or by splitting messages into data packets. According to the CJEU, however, this is not sufficient to be regarded as an ECS consisting 'wholly or mainly in the conveyance of signals on electronic communications networks'.[2352]

Thus, none of the three types of services (i-iii) contained in the definition of an ECS align with human-machine communication services, such as virtual assistants. As outlined in Section 6.3.1, literal (textual) interpretation is the prevailing method of interpretation if the provision to be interpreted is clear

[2345] Article 2 (4) EECC.
[2346] VoIP solutions, for example, enable individuals to call via computer without the call being routed on to a number in the regular telephony numbering plan.
[2347] Recital 15 EECC.
[2348] Article 2 (5) EECC, emphasis added.
[2349] It seems unclear what the phrase 'or are at least involved on one side of the communication' contained in Recital 15 precisely means.
[2350] Case C-475/12, *UPC* [2014] ECR I-285 para 43.
[2351] Case C-193/18, *Google LLC* [2019] ECR I-498 para 36.
[2352] Case C-193/18, *Google LLC* [2019] ECR I-498 para 36.

and precise.[2353] The definition of an ECS according to Article 2 (4) EECC is clear and the three types of services covered by it are defined further in case law,[2354] the EECC[2355] or elsewhere.[2356] According to settled case law,[2357] the literal meaning of a provision cannot be called into question by means of contextual or teleological interpretation if the provision is clear and precise.[2358] Thus, re-interpretation of the notion ECS and the three types of services covered by it through judicial action performed by the CJEU is not an option. Having established that the communication surveillance problem can only be solved by means of new or revised legislation, I now discuss how such legislation might look.

To be clear, and as explained in Section 4.9, providers of human-machine communication services need to adhere to the GDPR when processing personal data. Thus, only because providers of human-machine communication services fall outside the scope of the ePD does not lead to a complete lacuna in legal protection. However, the provisions of the GDPR are less strict than Article 5 (1) ePD. As outlined in Section 4.9.3, human-machine communications deserve the same level of confidentiality as interpersonal communications. This is due to the sensitivity of such communications, as explained in the first paragraph of this section.

6.4.2 Solution: Regulating human-machine communication

The proposed ePrivacy Regulation,[2359] which is still subject to political negotiations, seems well suited to solve this problem. The proposed ePrivacy Regulation sets rules regarding the protection of the fundamental right to privacy and particularly the confidentiality of communications.[2360] Unfortunately, neither the initial proposal nor the subsequent amendments regulate the confidentiality of human-machine communication. The initial proposal clarifies that the ePrivacy Regulation also applies to the transmission of machine-to-machine communications to ensure full protection of the right to privacy and confidentiality of communications.[2361] The proposal completely ignores human-machine communications and therefore, the communication surveillance problem essentially remains in the initial proposal for the ePrivacy Regulation. Instead of providing an analysis of the initial proposal and subsequent amendments, I propose specific provisions that can fill the current protection gap.

[2353] Koen Lenaerts, José A Gutiérrez-Fons, 'To Say What the Law of the EU Is: Methods of Interpretation and the European Court of Justice' (2013) European University Institute Working Paper AEL 2013/9 at 6 <https://cadmus.eui.eu/bitstream/handle/1814/28339/AEL_2013_09_DL.pdf?sequence=1&isAllowed=y> accessed 8 February 2024.
[2354] Case C-193/18, *Google LLC* [2019] ECR I-498 para 36; Case C-475/12, *UPC* [2014] ECR I-285 para 43;
[2355] Interpersonal communications service is defined in Article 2 (5) EECC.
[2356] Internet access service is defined in point (2) of the second paragraph of Article 2 Regulation (EU) 2015/2120.
[2357] Case C-220/03 *BCE* [2005] ECR I-10595 para 3; Case C-263/06 *Carboni e derivati* [2008] ECR I-1077 para 48; Case C-48/07 *Les Vergers du Vieux Tauves* [2008] ECR I-10627 para 44.
[2358] Koen Lenaerts, José A Gutiérrez-Fons, 'To Say What the Law of the EU Is: Methods of Interpretation and the European Court of Justice' (2013) European University Institute Working Paper AEL 2013/9 at 7 <https://cadmus.eui.eu/bitstream/handle/1814/28339/AEL_2013_09_DL.pdf?sequence=1&isAllowed=y> accessed 8 February 2024.
[2359] Commission, 'Proposal for a Regulation of the European Parliament and of the Council concerning the respect for private life and the protection of personal data in electronic communications and repealing Directive 2002/58/EC (Regulation on Privacy and Electronic Communications)' COM (2017) 10 final 'Proposal ePrivacy Regulation'
[2360] Article 1 Proposal ePrivacy Regulation.
[2361] Recital 12 Proposal ePrivacy Regulation.

First, the future ePrivacy Regulation should clarify that the confidentiality of communication also applies to human-machine communication and that processing of this is only allowed in specific circumstances. Therefore, I suggest including the following or similar provision:

Article y Confidentiality of human-machine communications

(1) Human-machine communications shall be confidential. Any interference with human-machine communications, such as listening, tapping, storing, monitoring, scanning, intercepting or other kinds of interception and surveillance that amount to the processing of human-machine communications, by persons other than end-users, shall be prohibited, except on the following grounds:

(a) Processing is strictly necessary for the sole purpose of facilitating human-machine communication explicitly initiated by the end user; or

(b) The end user has explicitly consented to the processing for one or more explicit purposes.

(2) The prohibition enshrined in paragraph 1 also applies to communication between natural persons captured in the context of human-machine communication.

Paragraph 1 of this proposed article sets the general rule that surveillance of human-machine communication and any other kind of processing is prohibited unless specifically permitted in the ePrivacy Regulation. According to my proposal, processing of human-machine communication is first and foremost permitted if this is strictly necessary for the sole purpose of facilitating human-machine communication expressly initiated by the end user. The term 'strictly necessary' is used to limit this processing. A corresponding recital should clarify that purposes such as quality control, advertisement, emotion detection, drawing inferences from captured recordings of human-machine communications are not 'strictly necessary' to facilitate human-machine communication. In my view, such processing should be subject to consent from the end user according to lit b of paragraph 1. To stipulate in a recital that advertisement is not strictly necessary to facilitate human-machine communication might be superfluous at first sight. Nonetheless, I suggest including this purpose as 'not strictly necessary' because companies are rather innovative when interpreting 'necessity'.[2362] In addition, and as explained in Section 5.5.1, the technology for targeted advertisement facilitated by virtual assistant services is readily available, for example, Amazon's US patent 'Keyword Determinations from Voice Data'.[2363] Drawing inferences from recorded human-machine communication by means of ML and NLP may lead to profiling of the end user and reveal a rich array of personal information, including clues with respect to the speaker's biometric identity, personality, physical traits, geographical origin,

[2362] Think, for example, about Meta, which claims that targeted advertisement is strictly necessary for the performance of the contract between Meta and the Facebook user see Case C-446/21.

[2363] Edara Kiran, 'Key Word Determinations From Voice Data' US Patent Number US 8798995B1 (Assignee: Amazon Technologies, Inc.) August 2014 <https://patentimages.storage.googleapis.com/bd/ed/2b/c4c67cc5a9f1ab/US8798995.pdf>, accessed 8 February 2024.

level of intoxication and sleepiness, age, gender, health condition and even an individual's socioeconomic status.[2364]

Likewise, processing human-machine communication for the purpose of emotion detection should require the consent of the end user, mainly due to the sensitive nature of data derived by AC (see Section 4.8.3). As indicated in Section 6.4, emotion detection systems for virtual assistants already exist. For example, Amazon's patented technology enables Alexa to recognise the user's emotional state derived from the user's voice.[2365] Other purposes such as improvement of services and quality control, should also be subject to the consent of the end user because all recordings might contain highly sensitive information. A former Apple employee revealed that he had listened to hundreds of Siri recordings every day for the purpose of quality control. These recordings concerned sensitive interpersonal communications such as discussions between doctors and patients, business deals, seemingly criminal acts and sexual encounters.[2366] This is not an exception, and press coverage points to similar practices at Google[2367] and Amazon (see Section 4.9.3).[2368]

The term 'explicitly initiated' included in lit a) contained in the first paragraph of proposed Article *y* prevents accidental recordings and other kinds of unsolicited processing of human-machine communication. Accidental recordings are common in virtual assistant services[2369] and occur when virtual assistants activate, transmit and/or record audio from their environment when the wake word is *not* spoken.[2370] Such recordings are caused by accidental triggers, i.e. sounds that wrongfully trigger virtual assistants, and they occur within the whole range of virtual assistants available on the market, including Amazon Alexa, Google Assistant and Siri. Researchers conducted a comprehensive analysis of accidental triggers in eleven smart speakers from eight different manufacturers and have found hundreds of such accidental triggers. The researchers automated the process for finding accidental triggers and measured their prevalence using everyday media such as TV shows, news and other kinds

[2364] Jacob Leon Kröger, Otto Hans-Martin Lutz, Philip Raschke, 'Privacy Implications of Voice and Speech Analysis – Information Disclosure by Inference' in Michael Friedewald et al (eds) *Privacy and Identity management. Data for Better Living: AI and Privacy* (Springer 2020) 242.
[2365] Huafeng Jin, Shuo Wang 'Voice-Based Determination of Physical and Emotional Characteristics of Users' US Patent Number US 10096319 B1 (Assignee: Amazon Technologies, Inc.) October 2018 <https://patentimages.storage.googleapis.com/f6/a2/36/d99e36720ad953/US10096319.pdf> accessed 8 February 2024.
[2366] Alex Hern, 'Apple contractors regularly hear confidential details on Siri recordings' *The Guardian* (London, 26 July 2019) < https://www.theguardian.com/technology/2019/jul/26/apple-contractors-regularly-hear-confidential-details-on-siri-recordings > accessed 8 February 2024.
[2367] Tom Simonite, 'Who's Listening When You Talk to Your Google Assistant?' *Wired* (New York, 10 July 2019) <https://www.wired.com/story/whos-listening-talk-google-assistant/> accessed 8 February 2024.
[2368] Alex Hern, 'Amazon staff listen to customers' Alexa recordings, report says' *The Guardian* (London, 11 April 2019) <https://www.theguardian.com/technology/2019/apr/11/amazon-staff-listen-to-customers-alexa-recordings-report-says> accessed 8 February 2024.
[2369] Nathan Malkin et al, 'Privacy Attitudes of Smart Speaker Users' (2019) Iss 4 Proceedings on Privacy Enhancing Technologies 250, 252.
[2370] Daniel J Dubois et al, 'When Speakers Are All Ears: Characterizing Misactivations of IoT Smart Speakers' (2020) Iss 4 Proceedings on Privacy Enhancing Technologies 255-276.

of audio datasets.[2371] Accidental recordings are problematic because conversations and other audio captured are sent over the Internet and subsequently stored on remote servers, [2372] often in the cloud.[2373] Incidents[2374] reveal that accidental recordings potentially include sensitive data and might be shared with third parties.[2375] An Alexa user listened to four years of his Alexa archive and found thousands of fragments of his life, including sensitive conversations such as medication-related family discussions.[2376]

Paragraph 2 of this proposed article is necessary because processing in the context of virtual assistants and similar services captures not only human-machine communications, but also interpersonal communications. Many of the examples mentioned in the previous paragraph in fact relate to recorded communications between natural persons, such as members of the household, visitors etc. When virtual assistant services are used by means of a smartphone app, basically every communication between the end user and any other natural person might be recorded, intentionally or accidentally. These recordings might be sensitive and include conversations between doctors and patients, business partners, criminals and sex partners.[2377] Therefore, communications between natural persons also should be confidential.

For the sake of legal certainty, I also suggest including a (broad) definition of human-machine communication in the ePrivacy Regulation. This definition could be worded as follows:

Article x (00) lit (z)
Human-machine communication means any information, irrespective of its form or content, relating to human-machine interactions facilitated via electronic communications networks.

[2371] Lea Schönherr et al, 'Unacceptable, where is my privacy? Exploring Accidental Triggers of Smart Speakers' (2020) at 1 <https://arxiv.org/pdf/2008.00508.pdf> accessed 8 February 2024.
[2372] Daniel J Dubois et al, 'When Speakers Are All Ears: Characterizing Misactivations of IoT Smart Speakers' (2020) Iss 4 Proceedings on Privacy Enhancing Technologies 255-276.
[2373] Lea Schönherr et al, 'Unacceptable, where is my privacy? Exploring Accidental Triggers of Smart Speakers' (2020) at 2 <https://arxiv.org/pdf/2008.00508.pdf> accessed 8 February 2024.
[2374] Tim Verheyden et al, 'Hey Google, are you listening?' *VRTB* (Brussels 10 July 2019) <https://www.vrt.be/vrtnws/en/2019/07/10/google-employees-are-eavesdropping-even-in-flemish-living-rooms/ > accessed 8 February 2024; Artem Russakovskii, 'Google is permanently nerfing all Home Minis because mine spied on everything I said 24/7 [Update x2]' (2017) <https://www.androidpolice.com/2017/10/10/google-nerfing-home-minis-mine-spied-everything-said-247/> accessed 8 February 2024.
[2375] Daniel J Dubois et al, 'When Speakers Are All Ears: Characterizing Misactivations of IoT Smart Speakers' (2020) Iss 4 Proceedings on Privacy Enhancing Technologies 255.
[2376] Geoffrey A Fowler, 'Alexa has been eavesdropping on you this whole time' *The Washington Post* (Washington, 6 May 2019) <https://www.washingtonpost.com/technology/2019/05/06/alexa-has-been-eavesdropping-you-this-whole-time/> accessed 8 February 2024.
[2377] Alex Hern, 'Apple contractors regularly hear confidential details on Siri recordings' *The Guardian* (London, 26 July 2019) <https://www.theguardian.com/technology/2019/jul/26/apple-contractors-regularly-hear-confidential-details-on-siri-recordings> accessed 8 February 2024; Alex Hern, 'Apple contractors regularly hear confidential details on Siri recordings' *The Guardian* (London, 26 July 2019) < https://www.theguardian.com/technology/2019/jul/26/apple-contractors-regularly-hear-confidential-details-on-siri-recordings > accessed 8 February 2024; Tom Simonite, 'Who's Listening When You Talk to Your Google Assistant?' *Wired* (New York, 10 July 2019) <https://www.wired.com/story/whos-listening-talk-google-assistant/> accessed 8 February 2024; Alex Hern, 'Amazon staff listen to customers' Alexa recordings, report says' *The Guardian* (London, 11 April 2019) <https://www.theguardian.com/technology/2019/apr/11/amazon-staff-listen-to-customers-alexa-recordings-report-says> accessed 8 February 2024.

The proposed definition is intentionally drafted broadly and is suited to cover all kinds of human-machine communication, including virtual assistant services, smart homes services and any possible future means of human-machine communication. Because it covers information regardless of its form or content, it applies to communication in the form of speech, text, video and any other means of current and future communication. In addition, I have refrained from including the requirement of remuneration of services that facilitate human-machine communication. Making the protection of such communication dependent on remuneration, like in the case of information society services,[2378] is the wrong approach, in particular when considering that individuals often tend to use services that are 'free of charge', while in fact 'paying' with their personal data. The apps for virtual assistant services offered by the major actors in the field, namely, Apple, Amazon and Google can all be downloaded for smartphones, free of charge.[2379] Users of these virtual assistant services might need to purchase hardware in case they wish to have dedicated 'smart speakers'[2380] at home, but the virtual assistant service itself remains free of charge. Therefore, the remuneration requirement would prevent legal protection for human-machine communications.

Additionally, and for the sake of legal certainty, the material scope of the initially proposed ePrivacy Regulation[2381] should be extended as follows (underlined text):

This Regulation applies to the processing of electronic communications data carried out in connection with the provision and the use of electronic communications services, <u>human-machine communications</u> and to information related to the terminal equipment of end-users.

The suggested (extended) scope of the ePrivacy Regulation makes clear that this piece of legislation applies to human-machine communications regardless of whether the provider facilitating such communication qualifies as an ECS. This closes the current gap of protection. Notably, within the initial proposal, the same approach has been taken in terms of information relating to the terminal equipment of end-users.[2382]

[2378] Article 1 (1) Directive (EU) 2015/1535 laying down a procedure for the provision of information in the field of technical regulations and of rules on Information Society services ('Information Society Services Directive'); Case C-62/19 *Star Taxi App SRL* [2020] ECR I-980 paras 41-48; Case C-390/18 *X* [2019] ECR I-1112 paras 39-49.
[2379] See <https://smartgeekhome.com/how-much-does-alexa-cost/>; <https://www.makeuseof.com/tag/what-is-google-assistant/>; <https://appstorechronicle.com/what-does-siri-cost> accessed 8 February 2024.
[2380] Parker Hall, 'The Best Smart Speakers With Alexa, Google Assistant, and Siri' *Wired* (New York, 27 September 2022) <https://www.wired.com/story/best-smart-speakers/> accessed 8 February 2024.
[2381] Article 2 Proposal ePrivacy Regulation.
[2382] The material scope stipulated in Article 2 Proposal ePrivacy Regulation explicitly mentions 'information related to the terminal equipment of end-users', which is a novum compared to the current scope defined in Article 1 ePD.

6.4.3 Conclusion

In this section, I have outlined that the legal solution to solve the communication surveillance problem consists of two new provisions in the future ePrivacy Regulation. The first new provision regulates the confidentiality of human-machine communication. According to this provision, the surveillance of human-machine communication is prohibited unless it is specifically permitted, i.e. if processing of human-machine communication is strictly necessary to facilitate such communication or if the user has explicitly provided consent. The second proposed provision defines human-machine communication broadly. For the sake of legal certainty, I also suggest extending the scope of the future ePrivacy Regulation by specifically including human-machine communication. Together, these provisions solve the current gap of protection regarding the confidentiality of human-machine communication.

6.5 Right of access – the trade secrets problem

The trade secrets problem (Type 2)
Trade secret protection under the TSD covers AI itself, as well as output generated by the AI system, including personal data relating to emotional states and life expectancy predictions. When data subjects invoke their right to obtain a copy of personal data undergoing processing according to Article 15 (3) GDPR, controllers are likely to argue that disclosure of the output generated by the AI system infringes their trade secrets and restrict access to such personal data in accordance with Article 15 (4) GDPR. Consequently, data subjects cannot enforce their right to obtain a copy of their personal data.

6.5.1 Setting the scene

The right of access is arguably the most important data subject right. The CJEU repeatedly stressed the relevance of this right as a prerequisite to other data protection rights.[2383] Article 15 (3) GDPR, which forms part[2384] of this highly important data subject right, empowers the data subject to obtain a copy of the personal data undergoing processing. As mentioned in Section 3.3.4.1, the concept of a 'copy' is not defined in the GDPR. The CJEU ruled that a 'copy' refers to the 'faithful reproduction or transcription' of an original. A purely general description of the data undergoing processing or a reference to categories of personal data does not correspond to that definition.[2385] In addition, the right to obtain a copy not only includes personal data collected by the controller, but also information

[2383] Case C-579/21, *Pankki S* [2023] ECR I-501 paras 56-58; Case C-487/21, *F.F.* [2022] ECR I-1000 paras 34-35; Case C-434/16, *Nowak* [2017] ECR I-994 para 57; Joined Cases C-141/12 & C-372/12, *YS, M* [2014] ECR I-2081 para 44; Case C-553/07 *Rijkeboer* [2009] ECR I-03889, para 51.
[2384] Case C-487/21, *F.F.* [2022] ECR I-1000 para 30.
[2385] Case C-487/21, *F.F.* [2022] ECR I-1000 para 21.

resulting from the processing of personal data, for example, a credit score.[2386] Both the CJEU and AG Pitruzella hesitated to clarify what is meant with 'faithful'. Dictionaries describe this notion as 'true and accurate; not changing anything'[2387] and 'true or not changing any of the details, facts, style, etc. of the original'.[2388] The copy must enable the data subject to effectively exercise its right of access in full knowledge of all personal data undergoing processing, including personal data *generated* by the *controller*.[2389] This is only possible if data subjects receive a faithful reproduction in intelligible form of the personal data requested, and *not only* a list with the categories of personal data, as in the case of Article 15 (1) lit b GDPR.

Copies empower data subjects to achieve the aims of the right of access, which includes to 'be aware of, and verify the lawfulness of processing'[2390] and to obtain 'the rectification, erasure or blocking' [2391] of personal data. For example, enforcing the right to rectification necessitates assessing the accuracy of any given piece of personal data. Such an assessment, however, is only possible if the data subject has access to a copy of the actual personal data processed by the controller. Being aware of the mere category of personal data undergoing processing is insufficient for this assessment, because categories are too imprecise. As an example, to assess whether the controller spells the data subject's name correctly requires actual access to the data subject's name and obviously, the mere category 'name' is insufficient. The same applies to personal data generated by means of AI, such as the specific emotional state detected by the AI system or topics of interests ascribed to a data subject inferred by means of ML (pattern detection) or other outcomes of profiling.

Article 15 (4) GDPR states that the right to obtain a copy of the personal data processed should not adversely affect the rights and freedoms of others, which includes personal data generated by AI that fall under within the broad scope of protection under the TSD.[2392] Rights and interests must be balanced against one another. According to the CJEU, a 'fair balance' must be struck between the various fundamental rights protected by the EU legal order and any restriction on those rights must comply with the principle of proportionality.[2393] The trade secrets problem will also not be solved when the controller provides the data subject with redacted documents, as regulatory guidance suggests.[2394] As

[2386] Case C-487/21, *F.F.* [2022] ECR I-1000, para 26.
[2387] See <https://www.oxfordlearnersdictionaries.com/definition/english/faithful?q=faithful> accessed 8 February 2024.
[2388] See < https://dictionary.cambridge.org/dictionary/english/faithful > accessed 8 February 2024.
[2389] Case C-487/21, *F.F.* [2022] ECR I-1000 paras 26, 39; see also the opinion of AG Pitruzella paras 45, 70.
[2390] Recital 63 GDPR.
[2391] Case C-487/21, *F.F.* [2022] ECR I-1000 para 21; see also the opinion of AG Pitruzella paras 45, 70; Joined Cases C-141/12 & C-372/12, *YS, M* [2014] ECR I-2081 para 44.
[2392] Paul B de Laat, 'Algorithmic decision-making employing profiling: will trade secrecy protection render the right to explanation toothless?' (2022) Vol 24 Iss 1 Ethics and Information Technology 1, 16.
[2393] Case C-275/06 *Promusicae* [2008] ECR I-00271 paras 65, 68.
[2394] European Data Protection Board, 'Guidelines 01/2022 on data subject rights – Right of access Version 2.0' (28 March 2023) at 163.

personal data themselves may constitute trade secrets, the controller could redact them, which is not helpful for the data subject and detrimental to the objectives[2395] of Article 15 GDPR.

As outlined in Section 5.6.2, the rule of non-prevalence constitutes the starting point for the balancing exercise. Based on CJEU case law, the outcome of the balancing exercise might essentially favour both the data subject's fundamental right to data protection and commercial interests pursued by the controller. I refer to trade secrets as commercial interests because commercial value constitutes one of the requirements when assessing whether information qualifies as a trade secret under Article 2 (1) TSD. Case law of the CJEU indicates that the protection of IP rights may prevail over the protection of personal data.[2396] The CJEU considered that the obligation to communicate personal data, for the purpose of ensuring effective protection of copyrights, of private persons in civil proceedings is eligible to strike a fair balance between the protection of IP rights and the fundamental right to data protection.[2397] Also, AG Pikamäe stresses that the legislator clearly did not contemplate sacrificing the fundamental right to intellectual property for the benefit of the fundamental right to data protection or the other way around. Rather, the legislator intended a fair balance between these two rights.[2398] However, the CJEU clarified that a fair balance requires particular consideration of the interests of the data subject. In the words of the CJEU, this fair balance 'may however depend, in specific cases, on the nature of the information in question and its sensitivity for the data subject's private life'.[2399] It is thus not excluded that the CJEU favours the data subject's fundamental right to data protection when balancing it with the controller's commercial interest in the form of a trade secret.

According to the CJEU, the balancing of opposing rights and interests, i.e. IP rights/trade secrets versus the fundamental right to data protection, depends on the *specific circumstances of the case*.[2400] Obviously, this conclusion is not satisfactory, nor does it provide legal certainty. I think it is questionable whether 'fair balancing' is the proper solution here. When considering the highly important role of the right to obtain a copy of the personal data processed and the consequences arising from the restriction of this right, in particular for other data subject rights, the trade secrets problem must be solved differently. I now discuss what this solution could look like.

[2395] Case C-487/21, *F.F.* [2022] ECR I-1000 paras 33-35.
[2396] Case C-597/19 *Telenet BVBA* [2021] ECR I-492 para 132; Case C-580/13 *Stadtsparkasse Magdeburg* [2015] ECR I-485 paras 28-41; C-461/10 *Bonnier Audio AB* [2012].
[2397] See Case C-264/19 *YouTube LLC* [2020] ECR I-542 paras 37-38; C-461/10 *Bonnier Audio AB* [2012] paras 57-60;
[2398] Case C-634/21, *SCHUFA Holding AG* [2023] ECR I-220, Opinion AG Pikamäe para 55.
[2399] Case C-131/12, *Google Spain* [2014] ECR I-317 para 81.
[2400] Case C-597/19 *Telenet BVBA* [2021] ECR I-492 para 111; Case C-13/16 *Rīgas* [2017] ECR I-336 para 31.

6.5.2 Solution: Introducing a new exception in the TSD

Many concerns have been raised with respect to the clash of trade secrets and the right of access in the context of AI.[2401] This is mainly due to the breadth of trade secrets: Any detail of algorithmic processing may be declared as a trade secret by the controller, including personal data generated by AI.[2402] Recital 2 TSD acknowledges that personal data might fall within the scope of information covered as trade secrets by mentioning 'information on customers'. In this very specific case of obtaining a copy of personal data under the right of access, I suggest eliminating the balancing exercise described in Section 6.5.1 and partially restrict trade secret protection. Perhaps the term 'restricting' is not completely accurate. Rather, my approach is to avoid that controllers exploit trade secret protection when data subjects exercise their right to obtain a copy of personal data undergoing processing. I use the term 'exploit' because, in my view, providing data subjects with a copy of their personal data is unlikely to harm the interests of the controller and the ability to compete.

As outlined in Section 5.6, three cumulative criteria must be met to trigger trade secret protection under the TSD. To qualify as trade secret according to Article 2 TSD, the information must be *secret*, have *commercial value* due to its secrecy and shall be subject to *reasonable steps to keep it secret*. It has been suggested to interpret the notion of commercial value as simply referring to the trade secret holder's ability to compete.[2403] However, I deem this interpretation too narrow when consulting the recitals of the TSD as the trade secret holder's ability to compete is simply one of the various ways how interests may be harmed. Protected information or knowledge has commercial value in the sense of the TSD, for example, when its unlawful acquisition, use or disclosure is likely to *harm the interests of the person lawfully controlling it*, in that it undermines that person's business or financial interests, strategic position or ability to compete.[2404] Misappropriation of trade secrets could also lead to costs for internal investigations, increased costs for protective measures and costs for prosecuting and litigating.[2405]

The acquisition, use and disclosure of trade secrets can either be lawful or unlawful under the TSD. I doubt that it is possible to speak of an unlawful disclosure of a trade secret in the context of a data subject's access request to receive a copy of the personal data undergoing processing. Article 3 (2)

[2401] Sandra Wachter, Brent Mittelstadt 'A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI' (2019) No 2 Columbia Business Law Review 494, 608; Gianclaudio Malgieri, 'Trade Secrets v. Personal Data: A Possible Solution for Balancing Rights' (2016) Vol 6 No 2 International Data Privacy Law 102, 113-114; Paul B de Laat, 'Algorithmic decision-making employing profiling: will trade secrecy protection render the right to explanation toothless?' (2022) Vol 24 Iss 1 Ethics and Information Technology 1, 9.
[2402] Paul B de Laat, 'Algorithmic decision-making employing profiling: will trade secrecy protection render the right to explanation toothless?' (2022) Vol 24 Iss 1 Ethics and Information Technology 1, 9.
[2403] Teresa Trallero Ocaña, *The Notion of Secrecy* (Nomos 2021) 415.
[2404] Jens Schovsbo, 'The Directive on trade secrets and its background' in Jens Schovsbo, Timo Minssen and Thomas Riis (eds) *The Harmonization and Protection of Trade Secrets in the EU* (Edward Elgar Publishing 2020) 14.
[2405] Baker McKenzie, 'Study on Trade Secrets and Confidential Business Information in the Internal Market' (MARKT/2011/128/D) (2013), 129 <https://single-market-economy.ec.europa.eu/publications/study-trade-secrets-and-confidential-business-information-internal-market_en> accessed 8 February 2024.

TSD outlines that acquisition, use and disclosure of trade secrets is lawful if 'required or allowed by Union or national law'. In my view, Article 15 (3) GDPR should be considered as a provision which requires the trade secret holder (controller) to lawfully disclose a copy of personal data undergoing processing. This interpretation however is not explicitly affirmed by the corresponding recital. Recital 18 TSD states, in a general manner, that 'the acquisition, use or disclosure of trade secrets, *whenever imposed* or *permitted* by law, should be treated as lawful for the purposes of this Directive'. Examples mentioned in Recital 18 do not refer to rights of data subjects, but focus on the rights of workers, their representatives and acquisitions or disclosures of trade secrets taking place in the context of statutory audits performed in accordance with Union or national law. However, the word 'particularly' hints to a non-exhaustive interpretation. Therefore, it seems reasonable to interpret that Article 3 (2) TSD also applies to the controller's obligation to disclose a trade secret (in the form of a copy of personal data), as required by Article 15 (3) GDPR. Consequently, this disclosure is lawful. From a systematic point of view, this also excludes ex-ante liability for misappropriation of the trade secret.[2406] Controllers might argue that such disclosure harms its interest protected by the TSD and refer to Article 15 (4) GDPR. Hence, the ultimate question is whether disclosing a copy of personal data undergoing processing to the data subject is likely to undermine the controller's business or financial interests, strategic position or ability to compete.[2407] In my view, this is not the case for four reasons.

First, the right to obtain a copy of personal data undergoing processing is an individual, non-transferable right. Only the data subject or a third party on the data subject's behalf can invoke it. In addition, the controller must identify the data subject when responding to a request and confirm the identity of the data subject in case of doubt[2408] to minimise the risk of unlawful disclosure. This limits the possible harm for the controller as personal data will be disclosed solely to the data subject (or its representative) making the request.

Second, after having obtained a copy of the personal data undergoing processing, it seems unlikely that the data subject will use this information in a way that undermines the controller's business or financial interests, strategic position or ability to compete. More specifically, data subjects will hardly make their copies of personal data available to the public or to other controllers, for example, to competitors because of privacy considerations. Thus, the risk of subsequent disclosure of personal data in ways that harm the interests of the controllers, in particular their position to compete, seems to be small.[2409] In cases in which data subjects use their right to obtain a copy of personal data in an

[2406] Teresa Trallero Ocaña, *The Notion of Secrecy* (Nomos 2021) 250.
[2407] Recital 14 TSD.
[2408] Article 12 (6) GDPR.
[2409] Teresa Trallero Ocaña, *The Notion of Secrecy* (Nomos 2021) 313.

abusive manner, controllers may regard such requests as manifestly unfounded. Controllers may refuse to comply with such requests or charge a reasonable fee.[2410]

Third, personal data does not have commercial value per se and does not automatically undermine a controller's business or financial interests when disclosed to the data subject. One single piece of personal data may qualify as a trade secret, but will hardly have a commercial value. It is mostly the composition of various pieces of personal data, in particular in the form of profiles, that constitute commercial value.[2411] There is no established approach to measuring the economic value of data, arguably because this very much depends on the content and the context and because it is difficult to quantify the benefits of data.[2412] Nevertheless, there are three common approaches to measure the monetary value of personal data from a firm's perspective, considering (i) the stock value of the firm, (ii) the revenues of the firm or (iii) the price of personal data records on the market.[2413] The conceptual challenges linked to each approach (every approach has its drawbacks)[2414] also come with various practical challenges. For example, markets for data and datasets are underdeveloped, and there is also no universal standard for categorising data into 'types' for statistical purposes.[2415] Hence, due to the challenges for measuring the value of personal data, it is difficult for controllers to substantiate that the disclosure of personal data copies to the data subject indeed harms their business and financial interests. In addition, the disclosure of individual personal data, even if generated by AI, arguably does not affect the trade secret holder's ability to compete. Likewise, it does not involve a disclosure to competitors. In addition, the relative value of individuals' data is typically rather low.[2416]

Fourth, providing data subjects with a copy of their personal data does not facilitate reverse engineering that may unlock trade secrets and consequently harm the controller's interests. Reverse engineering originates from mechanical engineering but is now increasingly used in the context of digital technologies.[2417] It is a technique whereby a product is being analysed in order to understand how it was designed and how it operates.[2418] In the context of IT systems, reverse engineering may simply

[2410] Case C-307/22, *FT* [2023] ECR I-315, Opinion AG Emiliou paras 32-35; European Data Protection Board, 'Guidelines 01/2022 on data subject rights – Right of access Version 2.0' (28 March 2023) at 188-191.
[2411] Marc van Lieshout, 'The value of personal data' in Jan Camenisch et al (eds) *Privacy and Identity 2014 IFIP AICT vol. 457* (Springer 2015) 29; Teresa Trallero Ocaña, *The Notion of Secrecy* (Nomos 2021) 313.
[2412] John Mitchell et al, 'Going Digital Toolkit Note: Measuring the economic value of data' OECD Document DSTI/CDEP/GD(2021)2/FINAL at 8, 10, 22 <https://one.oecd.org/document/DSTI/CDEP/GD(2021)2/FINAL/en/pdf> accessed 8 February 2024.
[2413] Marc van Lieshout, 'The value of personal data' in Jan Camenisch et al (eds) *Privacy and Identity 2014 IFIP AICT vol. 457* (Springer 2015) 29.
[2414] Gianclaudio Malgieri, Bart Custers, 'Pricing privacy – the right to know the value of your personal data' (2017) Vol 34 Iss 2 Computer Law & Security Review 289-303.
[2415] John Mitchell et al, 'Going Digital Toolkit Note: Measuring the economic value of data' OECD Document DSTI/CDEP/GD(2021)2/FINAL at 15 <https://one.oecd.org/document/DSTI/CDEP/GD(2021)2/FINAL/en/pdf> accessed 8 February 2024.
[2416] Teresa Trallero Ocaña, *The Notion of Secrecy* (Nomos 2021) 313.
[2417] Frank Apunkt Schneider, Günther Friesinger, 'Technology v Technocracy' in Günther Friesinger and Jana Herwig (eds) *The Art of Reverse Engineering* (transcript Verlag 2014) 10.
[2418] Noam Shemtov, Beyond the Code: Protection of Non-Textual Features of Software (Oxford University Press 2017) 71.

be described as 'the process of analysing a system to create representations of the system at a higher level of abstraction'.[2419] Therefore, reverse engineering starts with the final product and analyses backwards in order to determine the methods, components and logic used to generate the final product.[2420] A simple copy of personal data however prevents reverse engineering as it does not facilitate any access to software artefacts. The goal of reverse engineering is to derive information from available software artefacts and to translate it into abstract representations. Software artefacts are requirements, design, code, test case, manual pages etc.[2421] Providing a copy of personal data does not facilitate access to the system that generated the personal data nor does it facilitate access to the system's internal components expressed in source code format[2422] or other system artefacts. In addition, the TSD indicates that reverse engineering requires access to the *product* or *object* in which the trade secret is embodied.[2423] However, this is impossible when simply a copy of personal data is disclosed.

The risks related to reverse engineering are different, however, when a part of the algorithm would need to be disclosed to the data subject for complying with Article 15 (1) lit h GDPR (meaningful information about the logic involved in ADM). In a case pending at the CJEU, the technical expert appointed by the referring court suggested that at least a part of the algorithm needs to be disclosed to comprehend the logic involved in ADM[2424] (see Section 5.6.2). Although it seems unlikely that the CJEU follows the expert's opinion, such information is more likely to indeed harm the controller's business or financial interests, strategic position or ability to compete. Disclosing a part of the algorithm, together with additional information,[2425] allows one to analyse the system used to understand how it was designed and how it operates[2426] which ultimately unlocks the trade secret of the controller. If successful, reverse engineering facilitates the generation of a new program which is functionally equivalent to or even better than the program which was subject to reverse engineering.[2427] Obviously, this undermines the controller's business or financial interests, strategic position or ability to compete. However, the outcome is different when only a copy of personal data is provided.

[2419] Gerardo Canfora, Massimiliano Di Penta, 'New Frontiers of Reverse Engineering' (2007) Future of Software Engineering (FOSE '07) 326-341.
[2420] Noam Shemtov, Beyond the Code: Protection of Non-Textual Features of Software (Oxford University Press 2017) 71.
[2421] Gerardo Canfora, Massimiliano Di Penta, 'New Frontiers of Reverse Engineering' (2007) Future of Software Engineering (FOSE '07) 326, 327.
[2422] Noam Shemtov, Beyond the Code: Protection of Non-Textual Features of Software (Oxford University Press 2017) 71.
[2423] Article 3 (1) lit b TSD; Teresa Trallero Ocaña, *The Notion of Secrecy* (Nomos 2021) 537.
[2424] Case C-203/22 *Dun & Bradstreet Austria* see page 12 <https://www.ris.bka.gv.at/Dokumente/Lvwg/LVWGT_WI_20220211_VGW_101_042_791_2020_44_00/LVWGT_WI_20220211_VGW_101_042_791_2020_44_00.pdf> accessed 8 February 2024.
[2425] E.g. information such as the concrete factors and mathematical formula used, the concrete value assigned to the data subject, the disclosure of the intervals within which different data on the same factor are assigned to the same value; see Case C-202/22 *Dun & Bradstreet Austria*.
[2426] Noam Shemtov, Beyond the Code: Protection of Non-Textual Features of Software (Oxford University Press 2017) 71.
[2427] Andrew Johnson-Laird, 'Software Reverse Engineering in the Real World' (1994) Vol 19 Iss 3 University of Dayton Law Review 843, 846.

Based on these arguments, providing data subjects with a copy of their personal data seems unlikely to harm the controller's business or financial interests, strategic position or ability to compete. It could harm the interests of the controller, but it does not harm the rights or interests specifically protected by the TSD. Therefore, there is no need for a balancing exercise as outlined in Section 6.5.1. Instead, a solution is needed which allows data subjects to effectively enforce their right to obtain a copy of their personal data. Currently, controllers can buttress their (arguable) trade secrets protection.[2428] Already in 2011, Facebook denied a data subject access to his personal data because such disclosures 'would adversely affect trade secrets'.[2429] As I have outlined in this section, these claims are unjustified regarding obtaining a copy of personal data undergoing processing. Empowering data subjects to effectively enforce their right to obtain a copy of their personal data must entail the elimination of the power imbalance between the data subject and the controller. In the current situation, it is the controller who decides whether to provide a copy, and the data subject can only influence the controller's decision by means of costly, lengthy and burdensome litigation. My suggested solution aims to overcome the current issues by extending the exceptions to trade secrets protection currently enshrined in Article 5 TSD as follows:

New exception in Article 5 TSD:

e) for exercising the right to obtain a copy of the personal data undergoing processing as set out in Article 15 (3) of Regulation (EU) 2016/679

The proposed solution solves the trade secrets problem by clarifying that trade secrets protection under the TSD does not apply when data subjects enforce their right to obtain a copy of their personal data undergoing processing enshrined in Article 15 (3) GDPR. This solution is needed because the right of access is a precondition for the enforcement of other data subject rights.[2430] It allows data subjects to verify the lawfulness[2431] of processing and empowers them to request controllers to rectify, erase or block their personal data.[2432] As outlined in Section 6.5, an actual copy of the personal data is the only way for data subjects to obtain rectification of inaccurate personal data. The right to rectification will become more important in the future considering the developments in AI. These developments facilitate the generation of vast amounts of personal data in the form of predictions, profiles, emotion data and any other types of inferred personal data. As outlined in Sections 4.3.1, 4.7.1 and 5.7.2, such personal data are likely to be sometimes inaccurate. This can only be rectified when data

[2428] Paul B de Laat, 'Algorithmic decision-making employing profiling: will trade secrecy protection render the right to explanation toothless?' (2022) Vol 24 Iss 1 Ethics and Information Technology 1, 14.
[2429] See <http://www.europe-v-facebook.org/FB_E-Mails_28_9_11.pdf> accessed 8 February 2024.
[2430] Case C-434/16, *Nowak* [2017] ECR I-994 para 57; Joined Cases C-141/12 & C-372/12, *YS, M* [2014] ECR I-2081 para 44; Case C-553/07 *Rijkeboer* [2009] ECR I-03889, para 51.
[2431] Recital 63 GDPR.
[2432] Joined Cases C-141/12 & C-372/12, *YS, M* [2014] ECR I-2081 para 44.

subjects obtain a copy of the personal data processed, for example, the exact emotional state detected by the AI system or the precise topics of interests ascribed to a data subject. By extending the exceptions in Article 5 TSD, five legislative aims of the GDPR will be achieved, namely, ensuring a high level of protection in the EU,[2433] providing data subjects with control concerning the processing of their personal data,[2434] enhancing legal certainty,[2435] strengthening the data subject's rights and the effective protection of personal data.[2436] Simultaneously, it does not necessarily negatively affect the controller's commercial interests protected by the TSD, nor does it hinder the free flow of personal data between the Member States, which is another legislative goal of the GDPR.[2437]

6.5.3 Conclusion

In this section, I have outlined that the legal solution to solve the trade secrets problem consists of introducing a new provision in Article 5 TSD. This new provision, in the form of an exception, clarifies that trade secrets protection under the TSD does not apply when data subjects enforce their right of access according to Article 15 (3) GDPR. This exception strengthens the position of data subjects. It enables subjects to enforce their data subject rights regarding personal data generated by means of AI. Such an exception is justified because providing data subjects with a copy of their own personal data seems unlikely to harm the controller's business or financial interests, strategic position or ability to compete.

6.6 Right to rectification – the verifiability standard problem

The verifiability standard problem (Type 3)

Data subjects need to meet the objective verifiability standard to have output generated by ML and AC powered systems rectified. Output generated by means of ML may constitute unverifiable personal data. Emotion data are by nature highly subjective. Therefore, data subjects cannot provide evidence that meets the objective verifiability standard. Thus, the right to rectification is not fit for purpose to protect the fundamental right to data protection, as this standard hinders data subjects from exercising their right.

6.6.1 Setting the scene

The right to rectification enables the data subject to request the controller to rectify inaccurate personal data and to have incomplete personal data completed.[2438] As the name of the right indicates,

[2433] Recitals 6 and 10 GDPR; Case C-534/20, *Leistritz AG* [2022] ECR I-594 para 26; Case C-645/19 [2021] *Facebook Ireland* ECR I-483 para 45; Case C-511/18, *La Quadrature du Net* [2020] ECR I-791 para 106; Case C-131/12, *Google Spain* [2014] ECR I-317 para 66; Case C-154/21 *Österreichische Post AG* [2023] ECR I-3 para 44; Case C-132/21 *Nemzeti* [2023] ECR I-2 para 42.
[2434] Recitals 7 GDPR.
[2435] Recitals 7 and 13 GDPR.
[2436] Recital 11 GDPR.
[2437] Recitals 3 and 6 GDPR.
[2438] Article 16 GDPR.

rectification implicitly relies upon the notion of verification in the sense that something may demonstrably be shown to be inaccurate or incomplete.[2439] The CJEU seems to put the emphasis on *factual* evidence, ruling that facts in particular are susceptible to provable evidence.[2440] This task is straightforward when personal data are verifiable (such as a name, date of birth, email address or the weight of an individual).[2441] Nonetheless, predictions produced by ML, such as life expectancy, score value ratings and career perspectives, are essentially educated guesses based on large amounts of data.[2442] Such data are neither factual nor counter-factual data. Predictions may prove to be wrong or true, but in essence they are simply probabilistic and not objectively verifiable,[2443] mainly because they relate to the future and lack 'truth' as a baseline for comparison.[2444] Also, other types of personal data generated by AI such as emotion data are not objectively verifiable due to the subjective perception of emotion. Emotions are subjectively verifiable: emotion data can uniquely be verified by the individual experiencing the emotional state.[2445] Thus, due to the unverifiable or subjective nature of personal data generated by means of AI, it is impossible for data subjects to provide factual data meeting the objective verifiability standard. Consequently, they cannot enforce their right to rectification for personal data which is likely to be inaccurate (Sections 4.3.1, 4.7.1 and 5.7.2).

The right to rectification according to Article 16 GDPR is an underexplored provision in both academia and regulatory guidance. The same can be said about case law on this from the CJEU. There are only three rulings[2446] on the matter which explicitly deal with the right (under the DPD). Only one case relating to the right to rectification is pending at the CJEU.[2447] Nevertheless, I reckon that the right to rectification will have a more prominent role in the future due to developments in AI and the nature of the personal data generated by it.

Let me start with the scope of the right to rectification in the context of personal data generated by means of AI. There are no cases yet at the CJEU which specifically relate to the verifiability standard problem. Regulatory guidance suggests interpreting the scope of the right to rectification broadly,

[2439] Sandra Wachter, Brent Mittelstadt 'A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI' (2019) No 2 Columbia Business Law Review 494, 548.
[2440] Case C-460/20, *TU* [2022] ECR I-962 para 66.
[2441] Sandra Wachter, Brent Mittelstadt 'A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI' (2019) No 2 Columbia Business Law Review 494, 548.
[2442] Teresa Scantaburlo, Andrew Charleswoth, Nello Cristianini, 'Machine Decisions and Human Consequences' in Karen Yeung and Martin Lodge (eds) *Algorithmic Regulation* (OUP 2019) 57; Lee A Bygrave, 'Machine Learning, Cognitive Sovereignty and Data Protection Rights with Respect to Automated Decisions' (2021) University of Oslo Faculty of Law Legal Studies Research Paper Series No. 202-35 at 5 <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3721118> accessed 8 February 2024.
[2443] Jef Ausloos, Michael Veale, René Mahieu, 'Getting Data Subject Rights Right' (2019) Vol 10 Iss 3 JIPITEC 283, 302.
[2444] Diana Dimitrova, 'The rise of the personal data quality principle: is it legal and does it have an impact on the right to rectification?' (2021) Vol 12 No 3 European Journal of Law and Technology 21.
[2445] Jennifer Healey, 'Physiological Sensing of Emotion' in Rafael Calvo et al (eds), *The Oxford Handbook of Affective Computing* (OUP 2015) 213, 214.
[2446] Case C-434/16, *Nowak* [2017] ECR I-994; Joined Cases C-141/12 & C-372/12, *YS, M and S v Minister voor Immigratie, Integratie en Asiel* [2014] ECR I-2081; Case C-553/07 *Rijkeboer* [2009] ECR I-03889.
[2447] Case C-247/23, *VP*.

including both derived and inferred personal data.[2448] According to EU supervisory authorities, the right to rectification not only applies to 'input data' but also to 'output data'.[2449] In this context, input data means the personal data used by the AI system to generate the output, for example, bank statements, income, zip-code of an individual or the facial expressions of an individual recorded during an automated video assessment. The output data are the prediction with respect to the individual (e.g. non-reliable borrower) or the individual's emotional state detected by the AI system (e.g. anger). Both types of output constitute personal data as they concern information relating to an identified or identifiable natural person. It is therefore clear that the right to rectification applies to both types of output generated by AI.

There are views which suggest limiting the right to rectification to factual data. AG Sharpston takes the view that 'only information relating to *facts* about an individual can be personal data.'[2450] Such facts may be expressed in different forms, for example a person's weight might be expressed objectively in kilos or in subjective terms such as 'underweight' or 'obese'.[2451] Guidelines of the EDPS bluntly state that the right to rectification 'only applies to *objective and factual data*, not to subjective statements (which, by definition, cannot be factually wrong).'[2452] By referring to CJEU case law, legal scholars Wachter and Mittelstad suggest that inferred personal data are being excluded from the scope of the right to rectification.[2453] Implicitly, AG Pikamäe also seems to take this view concerning the automated establishment of a credit score performed by a credit rating agency. In his view, data subjects may enforce their right to rectification 'if the *personal data used to carry out the scoring* should prove to be inaccurate'.[2454] This limits the right to rectification to the input, i.e. to the personal data used to established the credit score. Simultaneously, it excludes the output in the form of the established credit score (inferred personal data).

When these views are applied to predictions generated by ML or emotion data generated by means of AC, none of them could be rectified. To be considered a non-reliable borrower is simply a probabilistic prediction which cannot be verified currently as it relates to the future. Thus, it does not constitute factual data. Likewise, the emotional state detected by the AI system is simply subjective and thus cannot constitute factual data. Obviously, this outcome is undesirable and, in my view, simply wrong, because the text of Article 16 GDPR does not at all suggest such a limitation. Article 16 GDPR applies to the 'rectification of inaccurate personal data' and it does not play a role whether such

[2448] Art 29 Working Party 'Guidelines on Automated Individual Decision-Making and Profiling for the Purposes of Regulation 2016/679', (WP251rev.01, 6 February 2018) at 8-9.
[2449] Ibid at 17-18.
[2450] Joined Cases C-141/12 & C-372/12, *YS, M* [2014] ECR I-2081, Opinion of AG Sharpston para 56.
[2451] Ibid para 57.
[2452] European Data Protection Supervisor, 'Guidelines on the Rights of Individuals with regard to the Processing of Personal Data' (25 February 2014) at 18.
[2453] Sandra Wachter, Brent Mittelstadt 'A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI' (2019) No 2 Columbia Business Law Review 494, 550.
[2454] Case C-634/21, *SCHUFA Holding AG* [2023] ECR I-220, Opinion AG Pikamäe para 50, emphasis added.

personal data constitutes factual data, inferred data, input data or output data *as long as is personal data*, i.e. information relating to an identified or identifiable natural person. In addition, there is nothing in the preparatory documents of the GDPR, which indicates the legislator's intention to limit this right to factual data. In addition, such a limitation would be contradictory to the CJEU's contextual and teleological approach to interpret data subject rights.[2455] As a result, both the prediction as a 'non-reliable borrower' and the emotional state detected by the AI system do fall within the scope of the right to rectification.

It could be argued that inferred personal data by means of AI such as the classification as a non-reliable borrower and detected emotional states constitute opinions (i.e. judgements, thoughts or beliefs about someone[2456]) cannot be rectified. In fact, similar claims about opinions have been made with respect to the accuracy principle. According to Herbst and Dienst, since opinions are not directly related to an objectively provable or disprovable reality about individuals, they cannot be labelled as accurate or inaccurate and thus lie beyond the scope of the accuracy principle.[2457] According to their view, personal data in the form of opinions are simply not the type of information to which the accuracy principle de facto can apply.[2458] When transposing this view to the right to rectification, personal data in the form of opinions cannot be rectified if the personal data does not constitute an objectively provable or disprovable reality about the data subject (a fact)[2459]. Arguably, this applies to the non-reliable borrower prediction and emotional states detected by the AI system. Due to their unverifiable and/or subjective nature, this output in the form of opinions does not constitute an objectively provable or disprovable reality (i.e. a fact) about the data subjects concerned. Consequently, it cannot be rectified.

Wachter and Mittelstad, by referring to CJEU case law, argue that inferred personal data cannot be rectified under data protection law as it constitutes *opinions* and/or *assessments*.[2460] This view is based on a non-contextual reading of the CJEU's case law and assumes that opinions and/or assessments are not rectifiable under Article 16 GDPR. This assumption is wrong. Opinions and/or assessments relating to a particular data subject constitute personal data according to the CJEU. In the words of the CJEU, the concept of personal data 'encompasses all kinds of information, not only *objective* but also *subjective*, in the form of *opinions* and *assessments*, provided that it "relates" to the data

[2455] Case C-434/16, *Nowak* [2017] ECR I-994 paras 53, 54.
[2456] See <https://dictionary.cambridge.org/dictionary/english/opinion> accessed 8 February 2024.
[2457] Tobias Herbst, 'Art. 5 Grundsäze für die Verarbeitung personenbezogener Daten' in Jürgen Kühling and Benedikt Buchner (eds) DatenschutzGrundverordnung/BDSG (2nd edn Beck 2018) 229, para 60; Sebastian Dienst, 'Lawful Processing of Personal Data in Companies under the GDPR' in Daniel Rücker and Tobias Kugler (eds) New European General Data Protection Regulation: A Practitioner's Guide (Beck/Hart/Nomos 2018) 68, para 326.
[2458] See also Dara Hallinan, Frederik Zuiderveen Borgesius 'Opinions can be incorrect (in our opinion)! On data protection law's accuracy principle' (2020) Vol 10 No 1 IDPL 1, 5.
[2459] Case C-460/20, *TU* [2022] ECR I-962 para 68.
[2460] Sandra Wachter, Brent Mittelstadt 'A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI' (2019) No 2 Columbia Business Law Review 494, 550.

subject'.[2461] This condition is satisfied if the information, by reason of its content, purpose or effect, is linked to a particular person.[2462] According to the CJEU, personal data in the form of assessments or opinions fall under the scope of the right to rectification. The data subject to whom the assessment or opinion relates has, at least in principle, a right to rectification because opinions and assessment qualify as personal data.[2463]

The right to rectification is not absolute and not intended to enable data subjects to object and change unfavourable opinions and assessments relating to them. Obviously, the right to rectification should not result in situations in which a candidate in a professional examination may correct his answers in an exam retroactively.[2464] Neither should a person involved in an immigration case be able to rectify the content of a legal analysis.[2465] This contextual and normative limitation is justified and necessary in order to avoid an interpretation of the right to rectification that is excessively broad or 'over-inclusive'.[2466] To add another example, if a controller's employee classifies a data subject as a complete idiot, the data subject cannot use Article 16 GDPR to change this opinion. This would by contrary to the freedom of expression and information according to Article 11 EUCFR. This statement arguably amounts to a value judgement which is not susceptible to proof according to the CJEU.[2467] In common language usage, value judgements are 'a personal opinion about whether something is good or bad' based on 'on personal opinion rather than facts'.[2468] However, the data subject could correct the incorrect representation of this opinion and point out why the subject is not an idiot, for example, by adding a supplementary statement as foreseen by the second sentence of Article 16 GDPR.

Thus, opinions and assessments regarding a specific data subject do fall under Article 16 GDPR. This conclusion also holds true when personal data inferred by means of AI are seen as opinions and assessments. It seems likely that the CJEU will rely on a specific type of teleological interpretation, i.e. functional interpretation 'effet utile'.[2469] If personal data in the form of opinions or assessments established by humans are subject to the right to rectification, the same must apply to opinions and assessments established by machines. Nonetheless, qualifying personal data generated by AI as opinions or assessments might be premature or simply wrong. As outlined in Section 4.7.1, inferences generated by machines are *not* based on human reasoning. Whereas humans have been conditioned to look for

[2461] Case C-434/16, *Nowak* [2017] ECR I-994 para 34. Emphasis added.
[2462] Ibid para 35.
[2463] Ibid para 46.
[2464] Case C-434/16, *Nowak* [2017] ECR I-994, para 54.
[2465] Joined Cases C-141/12 & C-372/12, *YS, M and S v Minister voor Immigratie, Integratie en Asiel* [2014] ECR I-2081, para 45.
[2466] Koen Lenaerts, José A Gutiérrez-Fons, 'To Say What the Law of the EU Is: Methods of Interpretation and the European Court of Justice' (2013) European University Institute Working Paper AEL 2013/9 at 27 <https://cadmus.eui.eu/bitstream/handle/1814/28339/AEL_2013_09_DL.pdf?sequence=1&isAllowed=y> accessed 8 February 2024.
[2467] Case C-460/20, *TU* [2022] ECR I-962 para 66.
[2468] See < https://dictionary.cambridge.org/dictionary/english/value-judgment > accessed 8 February 2024.
[2469] Koen Lenaerts, José A Gutiérrez-Fons, 'To Say What the Law of the EU Is: Methods of Interpretation and the European Court of Justice' (2013) European University Institute Working Paper AEL 2013/9 at 25 <https://cadmus.eui.eu/bitstream/handle/1814/28339/AEL_2013_09_DL.pdf?sequence=1&isAllowed=y> accessed 8 February 2024.

causes (why), AI focusses on correlations and probabilities (what).[2470] As indicated in Section 4.3.1, current AI systems have been called to be clueless[2471] to understand cause and effect and to be devoid of common sense.[2472] It seems that humans are much better at this than machines.[2473] Common sense reasoning still constitutes a challenge in AI applications.[2474] AI is unable to think in a manner on par with human thinking[2475] which is underscored by the shortcomings in the AI discipline of automated reasoning (Section 2.2.5). Personal data generated by AI systems cannot qualify as opinions and/or assessments when considering that such systems do not adopt human reasoning and lack common sense capabilities. The correct qualification for personal data generated by AI systems is 'personal data inferred by automated means'.

It is crucial for data subjects that personal data generated by AI systems fall under the right to rectification, in particular when considering that such data are highly scalable and riskier than personal data derived by humans. Actions taken based on probabilistic predictions and correlations may have real impact on human interests[2476] (e.g., to receive a loan or to be employed). This holds particularly true when such predictions or correlations are essentially considered as *facts*, although such personal data generated by ML are simply probabilistic and relate to future conduct that has not yet happened. As outlined in Sections 4.3.1, 4.7.1 and 5.7.2, output generated by AI can be problematic in terms of accuracy. Personal data inferred by AI are not based on human reasoning, and AI is currently subject to severe reasoning deficiencies, in particular regarding common sense reasoning (see also Sections 2.2.5, 4.3.1, 4.4.1 and 4.7.1). Personal data generated by AI can be shared with third parties on a large scale (e.g. advertisers and other service providers).

After having discussed these views that interpret the scope of the right to rectification too narrowly, I also want to mention a view that interprets the right to rectification too broadly. Dimirova suggests that the right to rectification should be seen as a tool 'having the potential to rectify algorithm model issues', meaning that this right can also be invoked to correct the quality of the data processing

[2470] Viktor Mayer-Schönberger; Kenneth Cukier, *Big Data: A Revolution That Will Transform How We Live, Work and Think* (Houghton Mifflin Harcourt 2013) 14, 18.
[2471] Brian Bergstein, 'What AI still can't do' MIT Technology Review (Cambridge 31 January 2020) <https://www.technologyreview.com/2020/01/31/304844/ai-common-sense-reads-human-language-ai2/> accessed 8 February 2024.
[2472] Cade Metz, 'Paul Allen Wants to Teach Machines Common Sense' *The New York Times* (New York, 28 February 2018) < https://www.nytimes.com/2018/02/28/technology/paul-allen-ai-common-sense.html > accessed 09 November 2019.
[2473] Davide Castelvecchi, 'AI pioneer: The dangers of abuse are very real' *Nature* (London, 4 April 2019) < https://www.nature.com/articles/d41586-019-00505-2> accessed 8 February 2024.
[2474] Shoham Yoav et al, 'The AI Index 2018 Annual Report' (AI Index Steering Committee Stanford University 2018) 64 <https://hai.stanford.edu/sites/default/files/2020-10/AI_Index_2018_Annual_Report.pdf > accessed 8 February 2024.
[2475] Lance Eliot, 'AI Ethics And The Quagmire Of Whether You Have A Legal Right To Know Of AI Inferences About You, Including Those Via AI-Based Self-Driving Cars' *Forbes* (New York, 25 May 2022) < https://www-forbes-com.cdn.ampproject.org/c/s/www.forbes.com/sites/lanceeliot/2022/05/25/ai-ethics-and-the-quagmire-of-whether-you-have-a-legal-right-to-know-of-ai-inferences-about-you-including-those-via-ai-based-self-driving-cars/amp/> accessed 8 February 2024.
[2476] Brent Daniel Mittelstadt et al, 'The ethics of algorithms: Mapping the debate' (2016) Vol 3 Iss 2 Big Data & Society 1, 5; Solon Barocas, 'Data Mining and the Discourse on Discrimination' (2014) <https://dataethics.github.io/proceedings/DataMiningandtheDiscourseOnDiscrimination.pdf> accessed 8 February 2024.

model.[2477] Obviously, when assessing the accuracy of personal data generated by AI, the model upon which the personal data are based also must be considered in order to ensure a comprehensive assessment. This is because the quality of the information, i.e. the personal data generated by AI, is affected by the quality of the AI system used.[2478] In my view, the right to rectification should not be interpreted so broadly as to empower data subjects to request the rectification of models deployed by an AI system. In itself, AI models do not constitute personal data. They process (and are trained with) personal data. Models cannot be 'information relating to an identified or identifiable natural person' simply because they operate and are trained with personal data from many data subjects. Thus, the right to rectification should be limited to input and output data. Extending this right to the rectification of models deployed by AI systems is not needed from a conceptual point of view. It is the accuracy principle, together with the accountability principle further substantiated in Article 24 (1) GDPR, that obliges controllers to ensure that the AI system generates accurate output. Controllers must 'implement appropriate and effective measures to ensure and demonstrate' that processing of personal data occurs in accordance with the rules laid down in the GDPR.[2479]

After having established the proper scope of the right to rectification in the context of AI, the question remains how data subjects may enforce their right to rectification concerning inferred personal data that by nature is either unverifiable or subjective. I now discuss possible solutions.

6.6.2 Solution: Amending the right to rectification

The problems surrounding the rectification of personal data generated by means of AI have not gone unnoticed. The scholars Wachter and Mittelstadt have claimed that inferences increasingly determine how data subjects are being viewed and evaluated, and that the GDPR attributes only limited rights regarding inferences to data subjects.[2480] They suggest closing this gap and proposing the 'right to reasonable inferences'. This right should apply to 'high-risk' inferences that cause damage to privacy or reputation or have low verifiability in the sense of being predictive or opinion-based while being used for 'important decisions'.[2481] The suggested right has an ex-ante and ex-post component. This right obliges controllers, ex-ante, to establish whether an inference is reasonable, by disclosing to the data subject (i) why certain data are normatively acceptable bases to draw inferences, (ii) why these inferences are normatively acceptable and relevant for the chosen processing purpose or type of automated decision and (iii) whether the data and methods used to draw the inferences are accurate and

[2477] Diana Dimitrova, 'The rise of the personal data quality principle: is it legal and does it have an impact on the right to rectification?' (2021) Vol 12 No 3 European Journal of Law and Technology 28.
[2478] See Lee A Bygrave, who discusses information quality in the context of information systems 'Ensuring Right Information on the Right Person(s)' (1996) University of Oslo, Institute for Private Law <https://www.jus.uio.no/ifp/om/organisasjon/afin/forskning/notatserien/1996/4_96.html> accessed 8 February 2024.
[2479] Art. 24 (1), Recital 74 GDPR.
[2480] Sandra Wachter, Brent Mittelstadt 'A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI' (2019) No 2 Columbia Business Law Review 494, 611 and 613.
[2481] Ibid 611, 613.

statistically reliable. Then, an ex-post component allows data subjects to challenge unreasonable inferences which could support the right to contest ADM as enshrined in Article 22 (3) GDPR.[2482] The ex-post component relates to the verifiability problem discussed here. It allows data subjects to raise objections on the ground that the inference or its source data is irrelevant, unreliable or non-verifiable and, concerning unverifiable and subjective inferences, to provide supplementary information to convince the controller to change its assessment.[2483] According to Wachter and Mittelstadt, the right to reasonable inferences 'would embed an answer to the verifiability question in law' and thus strengthen data protection rights, including the right to rectification which arguably already offers 'a remedy for non-verifiable and subjective inferences and opinions'.[2484] I assume that these statements refer to the ex-ante component of the right which obliges controllers to inform data subjects whether the data and methods used to draw the inferences are accurate and statistically reliable. If the controller cannot demonstrate this, data subjects can enforce their right to rectification because they can establish that the inference is not accurate.

The proposed right to reasonable inferences is an important contribution to the field and contains several valid points and suggestions. However, it is beyond the scope of this thesis to analyse this broad right in depth. I therefore restrict myself to assess whether the right to reasonable inferences solves the verifiability standard problem. In essence, it does not solve the problem because controllers are likely to claim that the methods used to draw the inferences are accurate and statistically reliable. If not, controllers would incriminate themselves and indicate non-compliance with the accuracy principle which could lead to both regulatory and private enforcement. In addition, controllers need results from reliable practices. To state not using accurate and statistically reliable methods would be of no use for controllers. Consequently, data subjects may not receive information that empowers them to effectively enforce their right to rectification concerning unverifiable or subjective personal data generated by AI. It will arguably become even more difficult for data subjects to enforce this right because controllers, when confronted with a rectification request, can simply claim that the methods used to draw the inferences are accurate and statistically reliable and refer to the information already disclosed in the context of the right to reasonable inferences. The suggested scope of the right contains several ambiguous terms, such as 'high-risk' inferences causing 'damage to privacy or reputation' and 'important decisions'. I opine that this right, when implemented as suggested, would lead to similar problems as those occurring in the context to the right not to be subject to ADM (see Section 5.11). In addition, data subjects should be able to enforce their right to rectification irrespective whether the personal data are used for 'important decisions'. This holds particularly true when considering the extensive data sharing which takes place in the context of IoT solutions which leverage

[2482] Sandra Wachter, Brent Mittelstadt 'A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI' (2019) No 2 Columbia Business Law Review 494, 613.
[2483] Ibid 494, 619.
[2484] Ibid.

data captured using Internet of Things devices. IoT is defined as the cyber-physical ecosystem of interconnected physical and potentially virtual sensors and actuators.[2485] If shared with other controllers, inaccurate personal data may cause harm to data subjects because it is disclosed and subsequently used by third parties.

Another solution for the verifiability standard problem is proposed by Ausloos, Veale and Mahieu. They suggest construing the right to rectification as an addendum rather than a replacement of data. In contentious cases, neither the data subject nor the controller should act as 'the arbiter of truth'. Rather, when the controller has 'good reasons' to disagree with the data subject with respect to a requested rectification, the best solution is to ensure that both views co-exist in the data processing system and to oblige the controller to consider both the suggested rectification and the original data.[2486] The data subject has a right to provide 'a supplementary statement' as enshrined in the second sentence of Article 16 GDPR. However, it is unclear what specific obligations such a supplementary statement imposes on the controller,[2487] also when consulting regulatory guidance.[2488] Thus, the right to have incomplete personal data completed does not prove to be particularly helpful in the context of AI because it does not solve the problem of inaccurate data. Furthermore, the proposed solution does not effectively protect the data subject. The data subject has no means to control how the controller shares the 'original data' of which the accuracy the data subject contests. Third, the controller's 'good reasons' to disagree with the requested rectification seem to be too vague and gives the controller significant leeway. Conclusively, the suggested solution does not really solve the problem, as potentially inaccurate personal data will be further processed by the controller, including the risk of subsequent sharing with third parties.

The solution I have in mind is more straightforward. In essence, I suggest slightly broadening the right to rectification concerning the processing of personal data generated by automated means and empower data subjects to easily contest the accuracy of such personal data. When the data subject contests the accuracy of such personal data, the controller shall either cease processing or rectify the personal data as requested by the data subject, unless it can demonstrate that the controller's interest prevail. I therefore suggest adding a second paragraph to Article 16 GDPR, worded as follows:

[2485] European Union Agency for Network and Information Security, 'Good Practices for Security of Internet of Things' (2018) 45 <https://www.enisa.europa.eu/publications/good-practices-for-security-of-iot/@@download/fullReport> accessed 8 February 2024.
[2486] Jef Ausloos, Michael Veale, René Mahieu, 'Getting Data Subject Rights Right' (2019) Vol 10 Iss 3 JIPITEC 283, 302.
[2487] Diana Dimitrova, 'The rise of the personal data quality principle: is it legal and does it have an impact on the right to rectification?' (2021) Vol 12 No 3 European Journal of Law and Technology 27.
[2488] Which simply states that Article 16 GDPR contains a right for the data subject to complement the personal data with additional information see Art 29 Working Party 'Guidelines on Automated Individual Decision-Making and Profiling for the Purposes of Regulation 2016/679', (WP251rev.01, 6 February 2018) at 18.

> *(2) The data subject shall have the right to contest the accuracy of personal data generated by automated means, including to obtain the rectification of such personal data. The controller shall cease the processing and, if requested by the data subject, rectify the personal data, unless the controller demonstrates that its interest to process the personal data in the contested form and for the specified purpose override the interests, rights and freedoms of the data subject.*

First, I propose to use the term 'generated by automated means' to overcome discussions whether personal data are inferred or observed, as is the case concerning the right to data portability (see Section 5.9). Furthermore, the term 'automated' means is widely used in the GDPR[2489] and is broad enough to capture any kind of processing facilitated by means of AI. At the same time, the term 'automated' means limits the extended scope of the right to rectification by excluding personal data inferred or generated by humans such as opinions and conclusions with respect to the data subject. This avoids creating regulatory overreach and limits the right for data subjects to (i) exercise influence (control) over personal data generated by means of AI and other automated means, (ii) concerns related to the accuracy of such personal data and (iii) possible harm for data subjects caused by the automated processing of personal data, like the rationale concerning Article 22 GDPR.[2490]

The right of data subjects to contest the accuracy of personal data generated by automated means allows them to exercise effective control over the processing of such data. Data subjects may request the rectification of such personal data without having to provide evidence that meets the objective verifiability standard. As pointed out in Sections 6.6.1 and 5.7.3, this might be impossible due to the unverifiable and subjective nature of the personal data generated by AI. Reversing the burden of proof and demanding the controller to provide evidence that the personal data meets the objective verifiability standard 'does the trick'. The proposed solution imposes the duty on the controller to demonstrate why its interest to process personal in the contested form prevails over the interests, rights and freedoms of the data subject. Thus, the controller bears the burden of proof to demonstrate that its interests prevail when the controller intends to process the personal data contested by the data subject. The proposed solution intentionally excludes specific requirements to which data subjects must adhere when exercising this right. This allows data subjects to effectively enforce this right, which is needed when considering that personal data generated by AI may be unverifiable or subjective. It protects data subjects from harms arising due to the processing of personal data of which the accuracy cannot be verified due to the lack of truth as a baseline for comparison, as is the case with predictions. When data subjects contest the accuracy of predictions, controllers need to cease processing and, if

[2489] Articles 2 (1), 4 (2), 20 (1) lit b, 21 (5) and Recitals 15, 68 GDPR.
[2490] Isak Mendoza, Lee A Bygrave, 'The Right Not to be Subject to Automated Decisions Based on Profiling' in Tatiana-Eleni Synodinou et al (eds) *EU Internet Law: Regulation and Enforcement* (Springer International 2017) 83-84; Lee A. Bygrave, Commentary of Article 22 in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020) 526.

requested by the data subject, rectify the prediction. A controller may only continue with processing the prediction if it can demonstrate that its interests prevail. This might be quite challenging and requires the controller to carefully assess the interests at hand. The proposed solution resembles the concept of the right to object according to Article 21 (1) GDPR in which the controller has to prove that it has compelling legitimate grounds to processing.[2491] If the nature of personal data generated by AI is highly subjective, as is the case with detected emotional states, the data subject may easily contest the accuracy and ask the controller to rectify the detected emotional state as perceived by the data subject.

If a controller cannot demonstrate that its interests to process the personal data for the specified purpose prevail, it must ultimately erase such personal data in accordance with Article 17 (1) lit a GDPR. In this case, processing the personal data is no longer necessary for the specified purpose when the controller cannot demonstrate prevailing interests. This provides effective protection[2492] for the data subject because personal data of which the nature is unverifiable or subjective may only be processed if the controller's interests indeed prevail, and in all other cases such personal data must be either rectified or erased after the data subject has contested the accuracy.

It might be argued that the proposed solution is overly broad and reinforces the data subjects' interests too strongly. However, I think this is not the case. In my view, if controllers engage in speculative processing of personal data of which the nature is unverifiable or subjective, data subjects need a powerful counterweight to contest to such processing. This solution does not prohibit such processing from the outset, as data subjects need to enforce their right to create an impact on the controller. In addition, this solution does not intervene with the controller's fundamental right to have a business or the controller's freedom of contract. It simply obliges controllers to assess their own interests and the data subject's fundamental rights, freedoms and interests when engaging in arguably speculative processing that relates to unverifiable or subjective personal data. If the controller's interests do not prevail, it can no longer process such data. The decision of whom to hire or accept as a client remains in full discretion of the controller and there is no impact on the freedom of contract. The latter is covered by the freedom to conduct a business according to Article 16 EUCFR (as confirmed by the CJEU)[2493] and grants the controller legal freedom to enter a contract and decide on its content.[2494]

[2491] Gabriela Zanfir-Fortuna, Commentary of Article 21 in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020) 517.
[2492] Case C-645/19 [2021] *Facebook Ireland* ECR I-483 para 45; Case C-319/20, *Meta Platforms Ireland Limited* [2022] ECR I-322 para 73.
[2493] Case C-426/11, *Alemo-Herron* [2013] ECR I-521 para 32; Case C-283/11, *Sky Österreich* [2013] ECR-28 paras 42, 43.
[2494] Olha O Cherednychenko, 'Fundamental Freedoms, Fundamental Rights, and the Many Faces of Freedom of Contract in the EU' in Mads Andenas, Tarjei Bekkedal, Luca Pantaleo (eds) *The Reach of Free Movement* (Springer 2017) 273, 276.

The proposed solution does not negatively affect the accomplishment of an economic union and economic progress, which is one of the legislative goals of the GDPR.[2495] It restricts the processing of personal data generated by automated means when data subjects enforce their right to contest the accuracy of such data or to have it rectified. If the proposed solution has an economic impact at all, it seems likely to be minimal when considering that the majority of data subjects do not invoke their rights granted by the GDPR. According to empirical research conducted in the Netherlands, 83% of the participants reported to not have taken any action to enforce their data subject rights.[2496] Unfortunately, the study does not specifically outline the practical use of the right to rectification. When referring to the practical use of other data subject rights (object 8%, access 5%, erasure 4%), one can expect similarly low figures for the right to rectification.[2497] If there is economic impact for the controllers and the economic union, it will be minimal. The low practical usage of data subject rights does not imply that these rights are superfluous. They empower data subjects to effectively influence the processing of personal data. To couple the justification of such rights with practical usage is ill-founded and would make many enforceable rights, for example, those enshrined in consumer law, superfluous.

The proposed solution is well aligned with a couple of legislative aims envisaged by the GDPR. It ensures a consistent and high level of protection of natural persons,[2498] and strengthens the data subject right's effectiveness.[2499] Likewise, the solution provides the same level of legally enforceable data subject rights[2500] by avoiding difficulties concerning procedural autonomy as discussed in Section 5.7.1. The rectification of unverifiable or subjective personal data generated by automated means depends not on objectively verifiable evidence but on the balancing of the interests at hand.

6.6.3 Conclusion

In this section, I have outlined that the legal solution to solve the verifiability standard problem consists of amending the right to rectification. I suggest adding an additional paragraph in Article 16 GDPR. This paragraph broadens the right to rectification regarding the processing of personal data generated by automated means and empowers data subjects to easily contest the accuracy of such personal data. When data subjects contest the accuracy, the controller shall either cease processing or rectify the personal data as requested by the data subject, unless it can demonstrate that the controller's

[2495] Recital 2 GDPR.
[2496] Joanna Strycharz, Jef Ausloos, Natali Helberger, 'Data Protection or Data Frustration? Individual Perceptions and Attitudes towards the GDPR' (2020) Vol 6 Iss 3 European Data Protection Law Review 407, 414-415.
[2497] See Table 4: Joanna Strycharz, Jef Ausloos, Natali Helberger, 'Data Protection or Data Frustration? Individual Perceptions and Attitudes towards the GDPR' (2020) Vol 6 Iss 3 European Data Protection Law Review 407, 417.
[2498] Recital 10 GDPR; Case C-534/20, *Leistritz AG* [2022] ECR I-594 para 26; Case C-645/19 [2021] *Facebook Ireland* ECR I-483 para 45; Case C-511/18, *La Quadrature du Net* [2020] ECR I-791 para 106; Case C-131/12, *Google Spain* [2014] ECR I-317 para 66; Case C-154/21 *Österreichische Post AG* [2023] ECR I-3 para 44.
[2499] Recital 11 GDPR.
[2500] Recital 13 GDPR.

interest prevail. This new paragraph solves the verifiability standard because data subjects are not required to provide objectively verifiable evidence when they intend to rectify unverifiable and subjective personal data generated by AI.

6.7 Automated decision-making – cumulativeness problem

The cumulativeness problem (Type 3)

The cumulative and vague requirements in Article 22 GDPR render it inapplicable to many decisions enabled, taken by or generated with the support of AI. Therefore, Article 22 GDPR is not fit for purpose to effectively protect data subjects from the particular risks associated with the automated processing of personal data, which is the main rationale of this provision according to the CJEU.

6.7.1 Setting the scene

As outlined in Section 3.3.4.6, Article 22 (1) GDPR rests on three cumulative conditions: (i) a decision is made that is (ii) based solely on automated processing or profiling and (iii) has either legal effects or similarly significant effects for the data subject concerned.[2501] Most output generated by AI, i.e. ML predictions such as future behaviour, potential interests or characteristics of data subjects, do not necessarily constitute decisions in the sense of requirement (i). The same can be said about output produced by an AI system that intends to detect the emotional state of an individual, combining ML with other AI disciplines (AC, CV and NLP). Requirement (ii) excludes AI systems that 'only' provide decisional support for decision-making from the scope of Article 22 GDPR.[2502] In fact, a limited degree of human involvement is sufficient to render Article 22 GDPR inapplicable.[2503] For example, the Amsterdam Court of Appeal considers a personal conversation as sufficient to satisfy the requirement of actual human intervention.[2504] Also, requirement (iii) seems difficult to satisfy considering that AI systems used for ADM utilise relatively obscure logic and come with covert consequences.[2505] Thus, due to the cumulative requirements which must be met simultaneously, this right is often not applicable. It therefore protects data subjects ineffectively from decisions enabled, generated or supported by AI. This starkly contrasts with the rationale of the provision as identified by the CJEU, which is *effective protection* against the risks associated with the *automated* processing of personal

[2501] Case C-634/21, *SCHUFA Holding AG* [2023] ECR I-957 para 43; Lee A. Bygrave, Commentary of Article 22 in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020) 532.
[2502] Lee A Bygrave, 'Minding the machine v2.0: The EU General Data Protection Regulation and Autmated Decision Making' in Karen Yeung and Martin Lodge (eds) *Algorithmic Regulation* (OUP 2019) 253.
[2503] Sebastião Barros Vale, Gabriela Zanfir-Fortuna, 'Automated Decision-Making Under the GDPR: Practical Cases from Courts and Data Protection Authorities' (2022) 8 <https://fpf.org/wp-content/uploads/2022/05/FPF-ADM-Report-R2-singles.pdf> accessed 8 February 2024.
[2504] Amsterdam Court of Appeal 4 April 2023, ECLI:NL:GHAMS:2023:793 para 3.25.
[2505] Lee A. Bygrave, Commentary of Article 22 in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary – 2021 Update* (OUP 2021) 100.

data, including profiling.[2506] Despite CJEU's broad interpretation[2507] of the notion of a decision, the cumulativeness problem is not solved. The other two cumulative conditions (ii) and (iii) must still be met *simultaneously*. Often, processing is not 'solely automated' as required by condition (ii), and the required effects foreseen by condition (iii) remain vague.

Article 22 GDPR is heavily debated in academia, mostly focussing on the question whether the GDPR contains a right to explanation[2508] of ADM as indicated in Sections 4.4.1 and 5.6.2.[2509] Binns and Veale[2510] discuss particular challenges with respect to conditions (i) to (iii) that arise when human intervention and/or a decision's significance is layered by stages or by particular decision outcomes. These challenges include, for example, the difficulty to locate the decision itself and whether the significance should be interpreted in terms of potential or realised effects.[2511] Brkan compares Article 22 GDPR with a Swiss cheese with giant holes in it due to the limitations and exceptions enshrined in this provision.[2512] Bygrave uses a different metaphor for pointing to the issues of Article 22 (1) GDPR. If one of the three requirements is not met, the house of cards collapses and the provision does not apply in its entirety.[2513] This metaphor underscores the essence of the cumulativeness problem. I now discuss how this problem could be solved.

6.7.2 Solution: Redrafting the right not to be subject to ADM

In essence, there are three approaches to solve the cumulativeness problem. The first is to consider Article 22 GDPR a regulatory failure and focus on other means enshrined in the GDPR to counter the challenges and risks of ADM. The fairness and accountability principle, data protection by design and default, data protection impact assessments and certifications could be suitable instruments for this. In particular, data protection impact assessments ('DPIAs') according to Article 35 GDPR could be helpful because they demand controllers to consider the rights, freedoms and interests of data subjects rather than focussing on the degree of automation involved in ADM.[2514] However, to leave

[2506] Case C-634/21, *SCHUFA Holding AG* [2023] ECR I-957 paras 57, 60.
[2507] Case C-634/21, *SCHUFA Holding AG* [2023] ECR I-957 paras 44-46; Opinion AG Pikamäe paras 37, 38, 42, 43.
[2508] Thomas Wischmeyer, 'Artificial Intelligence and Transparency: Opening the Black Box' in Thomas Wischmeyer and Timo Rademacher (eds) *Regulating Artificial Intelligence* (Springer Nature 2020) 75-101; Sandra Wachter, Brent Mittelstadt, Luciano Floridi, 'Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation' (2017) Vol 7 Iss 2 IDPL 76-99; Giancludio Malgieri, Giovanni Comandé, 'Why a Right to Legibility of Automated Decision-Makig Exists in the General Data Protection Regulation' (2017) Vol 7 Iss 4 IDPL 243-265.
[2509] For an overview, see Maja Brkan, 'Do algorithms rule the world? Algorithmic decision-making and data protection in the framework of the GDPR and beyond' (2019) Vol 27 Iss 2 International Journal of Law and Information Technology 91, 110-119.
[2510] Reuben Binns, Michael Veale, 'Is that your final decision? Multi-stage profiling, selective effects, and Article 22 of the GDPR' (2021) Vol 11 No 4 International Data Privacy Law 319, 332.
[2511] Reuben Binns, Michael Veale, 'Is that your final decision? Multi-stage profiling, selective effects, and Article 22 of the GDPR' (2021) Vol 11 No 4 International Data Privacy Law 319, 332.
[2512] Maja Brkan, 'Do algorithms rule the world? Algorithmic decision-making and data protection in the framework of the GDPR and beyond' (2019) Vol 27 Iss 2 International Journal of Law and Information Technology 91, 97.
[2513] Lee A Bygrave, 'Minding the machine v2.0: The EU General Data Protection Regulation and Autmated Decision Making' in Karen Yeung and Martin Lodge (eds) *Algorithmic Regulation* (OUP 2019) 253.
[2514] Reuben Binns, Michael Veale, Is that your final decision? Multi-stage profiling, selective effects, and Article 22 of the GDPR' (2021) Vol 11 No 4 International Data Privacy Law 319, 331.

the task to mitigate possible risks for data subjects related to ADM to controllers is insufficient. Apart from formalistic bureaucratic overkill and a lack of substantive change,[2515] it is fairly predictable that controllers will seize the opportunity to claim that AI systems and ADM generated by it are not really 'risky'.[2516] Ultimately, controllers are responsible for processing of personal data and need to perform risk assessments, such as DPIAs. Hence, the first approach is not suitable to actually solve the cumulativeness problem.

The second approach is to find the solution beyond data protection law, such as EU consumer law. In May 2022, the European Commission launched a fitness check on EU consumer law focussing on digital fairness. This fitness check determines whether additional legislative action is needed to ensure a high level of consumer protection in the digital environment.[2517] The Commission stressed the risks for consumers associated with the digital transformation, specifically difficulties for consumers to make informed choices and safeguard their interests.[2518] More specifically, the Commission points to commercial practices that distort consumers decision-making processes and abuse their behavioural biases by means of personalisation and profiling. It specifically links these practices with the processing of personal data: 'underlying data collection and processing combined with analysis of consumers behaviour and their cognitive biases can be used to influence consumers to take decisions that are detrimental to their best interests'.[2519]

In the digital economy, personal data constitute an integral part of products, services and transactions. In this context, personal data may be seen as an economic asset (e.g., use of a service in exchange for personal data), part of the service (e.g. virtual assistants and IoT services), means to determine the conditions of the service (e.g. personalisation) or as a means to influence consumer's decision-making process (e.g. exploiting consumer behavioural biases).[2520] EU consumer law and policy aims to ensure a high level of consumer protection, in particular with regard to the health, safety and *economic interests of consumers*.[2521] An important aspect of this is to avoid possible exploitations of the consumer as the economically weaker party.[2522] Thus, the scope and objectives of EU consumer and data protection law are different. Nonetheless, these two areas of law might complement each other.[2523]

[2515] Lilian Edwards, Michael Veale, 'Slave to the Algorithm: Why a 'Right to Explanation' is Probably not the Remedy You are Looking for' (2017) Vol 16 Iss 1 Duke Law & Technology Review 19, 77-80.
[2516] Reuben Binns, Michael Veale, 'Is that your final decision? Multi-stage profiling, selective effects, and Article 22 of the GDPR' (2021) Vol 11 No 4 International Data Privacy Law 319, 331.
[2517] See < https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/13413-Digital-fairness-fitness-check-on-EU-consumer-law_en > accessed 8 February 2024.
[2518] Commission, 'New Consumer Agenda' COM (2020) 696 final at 10 < https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020DC0696&from=EN > accessed 8 February 2024.
[2519] Ibid.
[2520] Natali Helberger et al, 'The perfect match? a closer look at the relationship between eu consumer law and data protection law' Vol 54 Iss 5 Common Market Law Review 1427, 1430-1431.
[2521] Article 169 TFEU.
[2522] Stephen Weatherill, *EU Consumer Law and Policy* (2nd edn Elgar Publishers 2013) 310.
[2523] Natali Helberger et al, 'The perfect match? a closer look at the relationship between eu consumer law and data protection law' Vol 54 Iss 5 Common Market Law Review 1427, 1464.

EU data protection law governs the processing of personal data by means of AI and EU consumer law protects the economic interests of consumers. Using personal data generated by AI (e.g., emotion data) to distort a consumer's decision-making capacity may be prohibited under EU consumer law. Consider a trader that exploits a consumer's emotional state by manipulating the consumer into the conclusion of a contract that is detrimental to its economic interest. This specific use of personal data potentially constitutes a prohibited unfair commercial practice under the current and future EU consumer law framework. EU consumer law protects the economic interests of data subjects acting in the capacity of a consumer by prohibiting unfair commercial practices that rely on the use of personal data generated through AI. However, this is a complementary protection to the protection provided by Article 22 GDPR, which does not primarily protect the data subject's economic interests. Rather, Article 22 GDPR aims to effectively protect individuals against the particular risks associated with the automated processing of personal data, including profiling.[2524] It also envisages to let data subjects exercise influence over ADM, to reduce concerns over the quality of ADM,[2525] and to uphold human dignity by ensuring that humans maintain the primary role in constituting themselves.[2526] This is emphasised by Recital 4 GDPR, which states that 'the processing of personal data should be designed to serve mankind'. Hence, the cumulativeness problem cannot be simply solved by current or future EU consumer law.

Another relevant area of law to address the cumulativeness problem is the AI Act. In 2021, the EU Commission proposed[2527] the AI Act. After multiple amendments and trilogue negotiations, the AI Act's compromise text[2528] was published in February 2024. The latter tries to achieve the ambitious aim to be a far-reaching regulation envisaging a high level of protection for Union values, fundamental rights and principles. At the same time, it focusses on new rules relating to placing on the market, putting into service and use of AI systems, promotes innovation and aims to improve the functioning of the internal market.[2529] Thus, it is regulation covering aspects of product safety and fundamental rights. Due to its scope, [2530] the AI Act's compromise text does not specifically regulate risks for data subjects arising from the processing of personal data in the context of ADM. This does not mean that the AI Act is not beneficial for individuals and the society, but it simply does not address the specific

[2524] Case C-634/21, *SCHUFA Holding AG* [2023] ECR I-957 paras 57, 60.
[2525] Recital 71 GDPR.
[2526] Isak Mendoza, Lee A Bygrave, 'The Right Not to be Subject to Automated Decisions Based on Profiling' in Tatiana-Eleni Synodinou et al (eds) *EU Internet Law: Regulation and Enforcement* (Springer International 2017) 83-84; Lee A Bygrave, 'Minding the machine v2.0: The EU General Data Protection Regulation and Automated Decision Making' in Karen Yeung and Martin Lodge (eds) *Algorithmic Regulation* (OUP 2019) 249.
[2527] AI Act proposal adopted by the Commission COM (2021) 206 final <https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52021PC0206&from=EN> accessed 8 February 2024.
[2528] AI Act compromise text resulting from the trilogue negotiations see <https://data.consilium.europa.eu/doc/document/ST-5662-2024-INIT/en/pdf> accessed 8 February 2024.
[2529] Article 1 and Recitals 1, 5, 28 AI Act compromise text <https://data.consilium.europa.eu/doc/document/ST-5662-2024-INIT/en/pdf> accessed 8 February 2024.
[2530] Ibid, Article 2 (5a) states that the AI Act shall not affect the GDPR nor the ePD.

risks relating to the processing of personal data. For this, secondary law on the fundamental right to data protection remains the proper regulatory instrument.

Interestingly, Article 68 c of the AI Act's compromise text introduces a 'right to explanation of individual decision-making.'[2531] Reading this provision leads to a deja vu: the wording is very similar to Article 22 GDPR, with some variations. Article 68 c (1) compromise text reads as follows: '*Any affected person subject to a decision which is taken by the deployer on the basis of the output from a high-risk AI system listed in Annex III, with the exception of systems listed under point 2, and which produces legal effects or similarly significantly affects him or her in a way that they consider to adversely impact their health, safety and fundamental rights shall have the right to request from the deployer clear and meaningful explanations on the role of the AI system in the decision-making procedure and the main elements of the decision taken.*' This points clearly to the academic discussions on the existence of a right to explanation for ADM under the GDPR.[2532] In the AI Act, the emphasis lies on meaningful *explanation*, as opposed to meaningful *information* under the GDPR. The notion of 'main elements' of the decision seems to be a new concept. Article 68 c (3) of the AI Act's compromise text states that this right '*shall only apply to the extent that the right referred to in paragraph 1 is not already provided for under Union legislation*.' Undoubtedly, this paragraph refers to Article 22 GDPR and will lead to tricky demarcation issues, blended with legal uncertainty. What seems clear, however, is that the AI Act aims to provide *complementary* protection from ADM. Hence, the second approach to finding a solution beyond data protection law is unsuitable to solve the cumulativeness problem.

The third approach is to redraft Article 22 GDPR. In my view, this is the most suitable solution. In fact, some scholars already suggested to 'radically' redraft Article 22 or 'let it die'.[2533] These scholars suggested to redraft paragraph 1 of Article 22 GDPR as follows:

The data subject shall ~~have the right not to~~ **not** be subject to a decision based ~~solely~~ on automated processing **without meaningful human intervention**, including profiling, which produces ~~legal effects concerning him or her or similarly significantly affects~~ **a significant effect on** him or her.

This suggestion is a good starting point, but in my view not suited to solve the cumulativeness problem. Whereas paragraph 1 gets rid of requirement (ii) 'based solely on automated processing or profiling', it introduces a new requirement, i.e. 'without meaningful human intervention'. Debates will

[2531] Article 68 c AI Act compromise text <https://data.consilium.europa.eu/doc/document/ST-5662-2024-INIT/en/pdf> accessed 8 February 2024.
[2532] See references contained in the last paragraph of Section 3.3.4.6 for an overview.
[2533] Paul De Hert, Guillermo Lazcoz, 'Radical rewriting of Article 22 GDPR on machine decisions in the AI era' *European Law Blog* (13 October 2021) <https://europeanlawblog.eu/2021/10/13/radical-rewriting-of-article-22-gdpr-on-machine-decisions-in-the-ai-era/> accessed 8 February 2024.

arise which requirements must be met to qualify as meaningful human intervention, similar to the discussions in Sections 5.11.2 and 5.11.3. Also, this new requirement comes with some ambiguity which is likely to be buttressed by controllers. In addition, it is not entirely clear whether the requirement of a 'significant effect' for the data subject must materialise or also includes potentially significant effects. Whether the reference 'including profiling' should be understood as 'involving profiling' or rather as an alternative baseline criteria for application (either ADM or profiling)[2534] remains unclear. In sum, the ambiguities with respect to the cumulative requirements that must be met to render Article 22 GDPR applicable remain to a large extent.

I suggest redrafting Article 22 GDPR as follows:

Harmful profiling and automated inferences

1. *The data subject shall not be subject to profiling or automated inferences which potentially harm its interests, rights and freedoms. Controllers must assume harm if profiling or automated inferences is intended to be used for decision-making regarding that data subject.*
2. *Paragraph 1 shall not apply if such profiling or automated inferences:*
   a) *is necessary for entering into, or performance of, a contract between the data subject and a data controller;*
   b) *is authorised by Union or Member State law to which the controller is subject and which also lays down suitable measures to safeguard the data subject's rights and freedoms and legitimate interests; or*
   c) *is based on the data subject's explicit consent.*
3. *The data subject shall have the right to obtain the controller's assessment which is required to comply with paragraph 1.*
4. *In the cases referred to in points a) and c) of paragraph 2, the data controller shall implement suitable measures to safeguard the data subject's rights and freedoms and legitimate interests, at least the right to obtain human intervention on the part of the controller, to express his or her point of view and to contest the decision.*
5. *Profiling and automated inferences referred to in paragraph 2 shall not be based on special categories of personal data referred to in Article 9(1), unless point (a) or (g) of Article 9(2) applies and suitable measures to safeguard the data subject's rights and freedoms and legitimate interests are in place.*

Paragraph 1 entails two cumulative requirements: (i) profiling *or* automated inferences and (ii) possible harm to the data subject's interests, rights and freedoms. As indicated in Section 3.3.4.6 and confirmed by the CJEU, [2535] Article 22 GDPR constitutes a prohibition which is subject to the

[2534] Lee A Bygrave, 'Minding the machine v2.0: The EU General Data Protection Regulation and Automated Decision Making' in Karen Yeung and Martin Lodge (eds) *Algorithmic Regulation* (OUP 2019) 252.
[2535] Case C-634/21, *SCHUFA Holding AG* [2023] ECR I-957 paras 52, 64; Opinion AG Pikamäe para 31.

exceptions listed in paragraph 2. The nature of this provision should be clarified by a corresponding recital to avoid another discussion in academia.

The term profiling in requirement (i) is defined in Article 4 (4) GDPR. The core element of this definition is 'to evaluate certain personal aspects' relating to the data subject. Evaluation includes efforts to 'analyse' or 'predict' aspects with respect to data subjects, for example, their economic situation, personal preferences, interests, reliability and behaviour. In addition, profiling refers to any form of automated processing of personal data to evaluate data subjects. The wording 'in particular' is typically used to indicate non-exhaustiveness. Thus, the examples of specific personal aspects mentioned in the definition are not exhaustive. The definition of profiling is broad enough to capture personal data generated by AI systems, for example, to establish probabilistic predictions (ML) or to detect the data subject's emotional state (AC) based on behaviour (e.g. facial expressions). Profiling also covers any kind of score attributed to a data subject. Think about an insurance company that ascribes a risk score to a data subject as a 'risky driver'. A dating app which attributes an 'attractiveness' score to the data subject to suggest a match with individuals having a similar score is another example.

I have added automated inferences as an additional requirement triggering this provision. In everyday use, inferences are defined as 'a *guess* that you make or an opinion that you form based on the information you have'[2536] or 'something that you *can find* out indirectly from what you already know'.[2537] Both definitions point to the predictive nature of inferences. Although profiling arguably covers most types of automated inferences, some AI systems may be beyond the scope of profiling. Think about speech-based emotion recognition systems as introduced in Section 2.2.4.2 and the real-world examples mentioned in Sections 4.7.1 and 4.9.3. Amazon's patented technology enables Alexa to detect the user's emotional state derived from the user's voice.[2538] Spotify's patented voice assistant[2539] recognises when a user sounds sad and then offers encouragement by 'cheering' the user.[2540] A bank used a speech-based emotion recognition system to predict the emotional states of customers calling the bank's customer support.[2541] In these examples, emotional states are inferred from speech recorded or

[2536] See < https://dictionary.cambridge.org/dictionary/english/inference?q=inferences > accessed 8 February 2024.
[2537] See < https://www.oxfordlearnersdictionaries.com/definition/english/inference?q=inference > accessed 8 February 2024.
[2538] Huafeng Jin, Shuo Wang 'Voice-Based Determination of Physical and Emotional Characteristics of Users' US Patent Number US 10096319 B1 (Assignee: Amazon Technologies, Inc.) October 2018 <https://patentimages.storage.googleapis.com/f6/a2/36/d99e36720ad953/US10096319.pdf> accessed 8 February 2024.
[2539] Daniel Bromand et al, 'Systems and Methods for Enhancing Responsiveness to Utterances Having Detectable Emotion' US Patent Number US 10566010 B2 (Assignee: Spotify AB) February 2020 11 < https://patentimages.storage.googleapis.com/2a/9d/2d/926b58a2bd956f/US10566010.pdf >, accessed 8 February 2024.
[2540] Josh Mandell, 'Spotify Patents A Voice Assistant That Can Read Your Emotions' *Forbes* (New York, 12 March 2020) <https://www.forbes.com/sites/joshmandell/2020/03/12/spotify-patents-a-voice-assistant--that-can-read-your-emotions/> accessed 8 February 2024.
[2541] Sebastião Barros Vale, Gabriela Zanfir-Fortuna, 'Automated Decision-Making Under the GDPR: Practical Cases from Courts and Data Protection Authorities' (2022) 48 <https://fpf.org/wp-content/uploads/2022/05/FPF-ADM-Report-R2-singles.pdf> accessed 8 February 2024.

streamed in daily life environments. Arguably, uttered speech and the emotional state derived from this is not necessarily behaviour or a 'personal aspect' as referred to in the definition of profiling. The detected emotional state constitutes an automated inference, as it is a guess based on information (recorded speech) the controller already has. It seems possible that new forms of automated inferences will arise in the future that do not fit the definition of profiling.

Requirement (ii) is intentionally phrased broadly to address some of the problems caused by AI. The term 'potentially' makes clear that not only realised harm is covered by Article 22 but also potential harm. This is needed due to the probability & inaccuracy (Section 4.3.1), common sense and rebuttal (Section 4.7.1), as well as the verification (Section 4.6.2) problems. These problems show that personal data generated by AI may harm the data subject's interest, rights or freedoms. Personal data generated by AI that is inaccurate, contradictory to common sense or cannot be verified due to its probabilistic nature is likely to harm the data subject's interest, rights or freedoms.

For example, ML generates uncertain knowledge such as predictions and correlations that are probabilistic. This may lead to inaccurate evaluations and representations of data subjects because ML often generalises. The use of probabilistic information in the context of a controller's decision-making process can have adverse and detrimental effects for data subjects. Predictions facilitated by ML, such as negative score values, may prevent the data subject from obtaining a loan for buying a house or a mobile subscription. This occurred in a case pending at the CJEU. Due to a poor score value ascribed to the data subject, the mobile network operator denied to prolong a mobile contract subscription with a rather low monthly fee of 10 €.[2542] The AC-powered HireVue software analyses the emotions a candidate portrays during the video assessment[2543] and automatically assigns the candidate with an average rating (score) and recommendation whether the candidate should be employed. It clearly harms the data subject's interest to find employment if the recruiter relies on inaccurate emotion data. Notably, AC technology is also used in sectors other than human resources, including marketing, customer service, healthcare, insurance, retail, autonomous driving, education and gaming.[2544]

Harm may be less obvious for output created by the AI system that merely constitutes the product of probability-based analytic processes (and thus inferred data as outlined in Section 4.4.1 and 4.4.3). Think about ML models that apply dimensionality reduction according to Section 2.2.1.2 on easily

[2542] Case C-203/22 *Dun & Bradstreet Austria* p 2 <https://www.ris.bka.gv.at/Dokumente/Lvwg/LVWGT_WI_20220211_VGW_101_042_791_2020_44_00/LVWGT_WI_20220211_VGW_101_042_791_2020_44_00.pdf> accessed 8 February 2024.
[2543] Nathan Mondragon, Clemens Aicholzer, Kiki Leutner, 'The Next Generation of Assessments' (HireVue 2019) <http://hrlens.org/wp-content/uploads/2019/11/The-Next-Generation-of-Assessments-HireVue-White-Paper.pdf> accessed 8 February 2024.
[2544] Cem Dilmegani, 'Top 24 Affective Computing (Emotion AI) Use Cases in 2023' <https://research.aimultiple.com/affective-computing-applications/> accessed 8 February 2024; Deepanshu Gahlaut, 'Top Emotion AI Companies to Watch out for in 2023' <https://deepanshugahlaut.medium.com/top-emotion-ai-companies-to-watch-out-for-in-2023-db925868fd9f> accessed 8 February 2024.

accessible digital records of behaviour, for example Facebook likes. These models predict the data subject's personality traits[2545] and could be used by a provider of a dating app. When implemented in the dating app, these personality traits could influence 'potential matches' and thus limit the data subject's freedom to choose between possible dating partners.

For these reasons, paragraph 1 of my proposal introduces a *rebuttable presumption* that profiling or automated inferences intended to be used for decision-making harm the data subject's interests, rights and freedoms. If controllers intend to engage in AI-powered processing, they may rebut this assumption and document the corresponding assessment mentioned in paragraph 3 accordingly. The rebuttable presumption of harm contained in paragraph 1 of my proposal is inspired by the EU Commission's proposal for an Artificial Intelligence Liability Directive.[2546] This proposal contains rebuttable presumptions, that are seen as the least interventionist tools because they balance the interests of claimants and defendants. Rebuttable presumptions are common in national liability systems of EU Member States.[2547]

When read together with paragraph 3, requirement (ii) enshrines a two-part human-in-the-loop approach for two reasons. First, it places human involvement at the very start of the processing chain according to the principle of data protection by design and default.[2548] It reinforces this principle which obliges controllers to assess the risks for the data subject's rights and freedoms posed by the envisaged processing and implement the data protection principles enshrined in Article 5 GDPR. In particular, the fairness (as suggested in Section 6.2.2) and accuracy principle will play an important role in this context. Second, a context-driven assessment which takes the interests, rights and freedoms of a particular data subject concerned into consideration is required. For example, profiling or automated inferences in the context of targeted advertisement are less likely to be harmful than profiling or automated inferences that influence the decision-making pursued in a recruitment context. Potential harm is subjective and will always depend on the context and the data subject concerned. A human assessment is needed due to the reasoning and common sense deficiencies in the AI discipline AR. The balancing problem explained in Section 4.2.1 shows that autonomous AI systems cannot balance the fundamental rights and freedoms of the parties involved due to the reasoning and cognitive deficiencies in the AI discipline AR.

[2545] Michal Kosinski, David Stillwell, Thore Graepel, 'Private traits and attributes are predictable from digital records of human behaviour' (2013) Vol 110 No 15 PNAS, 5802.
[2546] Commission, 'Proposal for a Directive of the European Parliament and of the Council on adapting non-contractual civil liability rules to Artificial Intelligence (AI Liability Directive) COM (2022) 496 final <https://commission.europa.eu/system/files/2022-09/1_1_197605_prop_dir_ai_en.pdf> accessed 8 February 2024.
[2547] AI Liability Directive Proposal at 6 <https://commission.europa.eu/system/files/2022-09/1_1_197605_prop_dir_ai_en.pdf> accessed 8 February 2024.
[2548] Article 25 GDPR.

Paragraph 3 enables the data subject to obtain the assessment performed by the controller as required by paragraph 1. This assessment outlines why the controller reached the conclusion that profiling or automated inferences are unlikely to harm the data subject's interests, rights and freedoms. When the data subject is not convinced by this assessment, it can exert real influence concerning such processing. Based on the information contained in this assessment, the data subject can enforce its rights provided by the GDPR, namely:

- Lodging a complaint with the competent supervisory authority (Article 77 GDPR)
- Enforcing the right to an effective judicial remedy against the controller (Article 79 GDPR)
- Mandating a representative to exercise its rights (Article 80 GDPR)

The suggested redrafting of Article 22 GDPR arguably achieves what is currently envisaged by this provision. It aims to *effectively protect data subjects* against the particular risks associated with the automated processing of personal data, including profiling.[2549] It also supports data subjects to exercise influence over profiling and decision-making, to reduce concerns over its quality[2550] and to uphold human dignity by ensuring that humans keep the primary role in constituting themselves.[2551] The latter is emphasised by Recital 4 GDPR, which states that 'the processing of personal data should be designed to serve mankind' and requirement (ii) reflects this aim.

The redrafted version significantly broadens this right. By removing the requirement that decision-making involving profiling must be fully automated, it also applies to decisions which are *influenced* by AI. This addresses the problem that personal data generated by AI may create harm for the data subject, in particular when it is subsequently shared with and used by other parties. For example, a poor score value generated by a credit rating agency may prevent data subjects from obtaining a mobile subscription. A low attractivity score in a dating app might suggest potential dating partners that do not match the data subject's expectations and thus limit the data subject's freedom to choose between possible dating partners. The suggested redrafting renders Article 22 GDPR applicable regardless of whether the decisions taken regarding the data subjects are fully automated. The example of the score value used by the mobile network operator for the decision whether or not to prolong a mobile subscription would thus fall under the prohibition of Article 22 GDPR. The revised text of Article 22 GDPR also clarifies that potential harm is sufficient to trigger the protection granted by this right (i.e. the prohibition). Instead of providing data subjects with a procedural safeguard such as the current right to contest to ADM (see Section 5.11.3), it empowers the data subject to obtain the assessment performed by the controller as required by paragraph 1. This information enables the data

[2549] Case C-634/21, *SCHUFA Holding AG* [2023] ECR I-957 paras 57, 60.
[2550] Recital 71 GDPR.
[2551] Isak Mendoza, Lee A Bygrave, 'The Right Not to be Subject to Automated Decisions Based on Profiling' in Tatiana-Eleni Synodinou et al (eds) *EU Internet Law: Regulation and Enforcement* (Springer International 2017) 83-84; Lee A Bygrave, 'Minding the machine v2.0: The EU General Data Protection Regulation and Automated Decision Making' in Karen Yeung and Martin Lodge (eds) *Algorithmic Regulation* (OUP 2019) 249.

subject to exert real influence over such processing and facilitates the enforcement of the data subject's rights enshrined in the GDPR.

6.7.3 Conclusion

In this section, I have argued that the legal solution to solve the cumulativeness problem consists of amending the right not to be subject to ADM. The proposed wording focusses on profiling and automated inferences that potentially harm the data subject's rights, interests or freedoms rather than on 'automated decision-making'. The proposed wording covers decisions which are influenced by profiling generated by AI. It also requires controllers to assess whether profiling, automated inferences and the intended decision-making potentially harm the data subject. Data subjects can obtain this assessment, which allows them to enforce their rights provided by the GDPR, lodging a complaint with an SA or initiating legal proceedings in particular.

6.8 Conclusions

This chapter aimed to answer Subquestion 5, i.e. how the incompatibilities of the current legal framework identified in Subquestions 3 and 4 should be addressed. Based on the selection criteria effectiveness, urgency and novelty, I have addressed six legal problems: the elusiveness, mental data, communication surveillance, trade secrets, verifiability standards and cumulativeness problems.

This chapter has focussed on legal solutions, although technological solutions[2552] should also be explored and developed. I argued that the incompatibilities of the current legal framework can be addressed by the following legal solutions: (i) new interpretations of existing provisions, (ii) amendments of existing provisions or (iii) the introduction of entirely new provisions. Table 6.4 provides an overview of which legal problem should be addressed by which type of legal solution.

Problem (type)	**AI Disciplines**	**Suggested Legal Solution (i, ii or iii)**
Elusiveness (2, 3)	ML, NLP, CV, AC, AR	New interpretation as substantive fairness (i)
Mental data (3)	ML, AC	Introducing dynamic list for special data (iii)
Comm. surveillance (3)	ML, NLP, AC	Regulating human-machine communication (iii)
Trade secrets (2,3)	ML, NLP, CV, AC, AR	Adding a new exception in the TSD (iii)
Verifiability standard (3)	ML, AC	Amending right to rectification (ii)
Cumulativeness (3)	ML, NLP, CV, AC, AR	Redrafting right not to be subject to ADM (ii)

Table 6.4 Outlining legal problems (type), AI disciplines concerned and suggested legal solutions.

[2552] As mentioned in Section 6.1 e.g. randomisation techniques, secure multiparty computation, homomorphic encryption, differential privacy, knowledge-infused learning.

The *elusiveness problem* should be addressed by a new interpretation of the fairness principle. The legal solution consists of interpreting the fairness principle as both procedural and *substantive* fairness. The provisions in the GDPR and the corresponding recitals already provide clarity with respect to procedural fairness. Substantive fairness as suggested here contains two major elements: fairness between the parties and fairness of the outcomes. Several components of substantive fairness should be considered, distributed among the two major elements of substantive fairness. These components are power inequalities/dominant positions, vulnerability, good faith, autonomy, non-manipulation, detrimental effects, accuracy and non-discrimination. To ultimately 'solve' the elusiveness problem, judicial action is needed. The CJEU should interpret fairness in EU data protection law as including both procedural and substantive fairness.

The proposed legal solution for the *mental data problem* consists of the introduction of a new dynamic list for special data. This solution overcomes the current problem that the approach to enumerate special data exhaustively is not fit for purpose to address the challenges caused by AI as it does not keep up with technological developments. In my suggested solution, the European Commission is empowered to adopt new delegated acts for the purpose to update the list of special data where needed due to technological developments. This solution is flexible and comes with legal certainty for all actors involved.

The proposed legal solution for the *communication surveillance problem* consists of two new provisions to be included in the future ePrivacy Regulation. The first new provision specifically regulates the confidentiality of human-machine communication. According to this provision, the surveillance of human-machine communication is prohibited unless it is specifically permitted, i.e. if processing of human-machine communication is strictly necessary to facilitate such communication or if the user has explicitly provided consent. The second new provision defines human-machine communication broadly. For the sake of legal certainty, the scope of the future ePrivacy Regulation should be extended by specifically including human-machine communication. Taken together, these provisions solve the current gap of protection regarding the confidentiality of human-machine communication.

The proposed legal solution for the *trade secrets problem* consists of a new exception to be included in Article 5 TSD. This new provision clarifies that trade secret protection under the TSD does not apply if data subjects enforce their right of access according to Article 15 (3) GDPR. This strengthens the position of data subjects. It enables them to enforce their data subject rights with regard to personal data generated by AI. This exception is justified because the right of access constitutes a conditio sine qua non for all other data subject rights. In addition, providing data subjects with a copy of their own personal data seems unlikely to harm the controller's interests specifically protected by the TSD. This protects a company's business and financial interests, strategic position and ability to compete.

The *verifiability standard problem* should be addressed by amending the right to rectification. I suggest adding an additional paragraph in Article 16 GDPR. This paragraph broadens the right to rectification regarding the processing of personal data generated by automated means and empowers data subjects to easily contest the accuracy of such personal data. When data subjects contest the accuracy, the controller shall either cease processing or rectify the personal data as requested by the data subject, unless it can demonstrate that its own interests prevail. This new paragraph solves the verifiability standard because data subjects are not required to provide objectively verifiable evidence when asking for the rectification of unverifiable and subjective personal data generated by AI.

The proposed legal solution for the *cumulativeness problem* consists of the redrafting of the right not to be subject to ADM. The proposed wording focusses on profiling and automated inferences instead of 'automated decision-making'. It requires controllers to perform an assessment of whether the envisaged profiling or automated inferences potentially harm the data subject's interests, rights and freedoms. The redrafted provision assumes harm if profiling or automated inferences is intended to be used for decision-making on the data subject concerned. Data subjects can obtain the assessment performed by the controller, which allows them to enforce their rights enshrined in the GDPR, in particular lodging a complaint with an SA or initiate legal proceedings. My proposed solution gets rid of the cumulativeness problem and enables data subjects to exercise real influence regarding profiling and automated inferences enabled by AI.

7 Conclusion

This chapter draws the conclusions for this thesis. Section 7.1 answers the main research question. Section 7.2 provides recommendations for future legislation and Section 7.3 presents some ideas for future research.

7.1 Answer to the research question

Before providing an answer to the research question, I quickly recap the *AI disciplines*, the *current EU legal framework* and the *three types of legal problems* discussed in this thesis.

AI refers to adaptive machines that can autonomously execute activities and tasks that require capabilities usually associated with humans. In this thesis, I have focussed on five *AI disciplines*: machine learning (ML), natural language processing (NLP), computer vision (CV), affective computing (AC) and automated reasoning (AR). *ML* is a set of computational methods using experience to improve its performance and to make accurate predictions. Three methods are used for ML, i.e. supervised, unsupervised and reinforcement learning. Deep learning (DL) is a particular kind of ML that uses many layers. Approaches in DL feed a large set of input data into an artificial neural network (ANN) that produces successive transformations of the input data. Each hidden layer combines the values in the preceding layer. *NLP* aims to give computers the ability to process human language. It includes both the generation and understanding of natural language. *CV* is a discipline of AI devoted to perceive objects, described as the science and technology of machines that 'see'. *AC*, sometimes called 'emotion AI', is computing that relates to emotions and aims to develop machines with emotional capabilities. *AR* is the discipline that aims to develop computers that can use stored information to answer questions and to draw new conclusions. Research in AR focusses on logical reasoning, probabilistic reasoning and common sense reasoning.

The fundamental right to privacy according to Article 7 EUCFR protects everyone's 'right to respect for his private and family life, his home and communications'. The fundamental right to data protection as enshrined in Article 8 EUCFR grants everyone 'the right to the protection of personal data concerning him or her'. These fundamental rights are closely linked, but they are not identical,[2553] as they differ in terms of material and personal scope.[2554] Both fundamental rights are further substantiated in EU secondary law. The most relevant legislation in EU secondary law is the GDPR and the

[2553] Juliane Kokott, Christoph Sobotta 'The distinction between privacy and data protection in the jurisprudence of the CJEU and ECtHR' (2013) Vol 3 No 4 International Data Privacy Law 222, 223, 228; Herke Kranenborg, Commentary of Article 8 in Steve Peers et al (eds), *The EU Charter of Fundamental Rights* (Hart/Beck 2014) 229.
[2554] The material scope of the fundamental right to data protection seems to be broader whereas it is more narrow in terms of personal scope as it exludes legal persons; see Juliane Kokott, Christoph Sobotta 'The distinction between privacy and data protection in the jurisprudence of the CJEU and ECtHR' (2013) Vol 3 No 4 International Data Privacy Law 222, 225; Joined Cases C–92/09 and C–93/09 *Volker und Markus Schecke and Eifert* [2010] ECR I-11063, paras 52, 53 and 87.

ePrivacy Directive. Article 7 & 8 EUCFR as well as the GDPR and the ePrivacy Directive form the '*current legal framework*'. I have focussed on principles[2555] and enforceable rights[2556] contained in the current EU legal framework.

This thesis distinguishes between *three types of legal problems*: *(1)* legal provisions that are violated, *(2)* legal provisions that cannot be enforced and *(3)* legal provisions that are not fit for purpose to protect the fundamental right at stake. These three types of legal problems arise or may arise when principles and enforceable rights contained in the current EU legal framework are applied to the five AI disciplines.

The main research question of this thesis is:

To what extent do the developments in AI require a new legal framework for the fundamental rights to privacy and the protection of personal data?

My answer to that question is as follows. It is *not* needed to establish a *new* legal framework for the fundamental rights to privacy and the protection of personal data. Rather, the current legal framework must be adjusted to ***some extent***. The extent to which this adjustment is needed largely depends on (i) the type of legal problem and (ii) the AI discipline.

(i) Type of legal problem

	Type 1	Type 2	Type 3
Principles	12	4	16
Enforceable rights	11	8	9
Total	**23**	**12**	**25**

Table 7.1 Number of legal problems (per type) distributed among principles and enforceable rights contained in the current legal framework.

A total of sixty[2557] Type 1, 2 and 3 legal problems were identified in this thesis when the principles and enforceable rights enshrined in the current legal framework are applied to the AI disciplines. This is shown in Table 7.1. Type 1 and 3 legal problems arise almost just as often and roughly occur twice

[2555] Proportionality, Lawfulness, Fairness, Transparency, Accuracy, Purpose limitation, Data minimisation, Confidentiality, Exhaustive enumeration, Accountability.
[2556] Informational privacy, bodily privacy, mental privacy, communicational privacy, right of access, right to rectification, right to erasure, right to data portability, right to object, right not to be subject to ADM.
[2557] This total must not be confused with the 55 legal problems identified in Chapters 4 and 5 of this thesis. The difference between the two totals is caused by the fact that the elusiveness, interpretability, precision, trade secrets and training data problems each lead to two *types* of legal problems.

as much as Type 2 legal problems. To what extent the current legal framework should be adjusted is highly influenced by the *type* of legal problem.

For *Type 1* legal problems, the current legal framework suffices. These types of legal problems do not necessarily require adjustments of the current legal framework. The solution for these problems is obvious. Violations of provisions contained in the current legal framework need to be enforced through data subjects and/or representative bodies ('private enforcement') as well as through supervisory authorities ('regulatory enforcement'). Thus, regarding Type 1 legal problems, the current legal framework is fit for purpose, provided that violations are in fact enforced.

Conversely, the current legal framework *does not suffice* for Type *2* and *3* legal problems. Unenforceable and 'unfit' provisions are simply not appropriate to protect individuals. These two types of legal problems require either *adjustments* of current provisions or *new interpretations*. In the latter case, judicial action instead of legislative action is needed. Take, for example, the elusiveness problem (Section 4.3.2). The elusive role and meaning of the fairness principle reduces legal certainty and makes it difficult for data subjects to challenge the fairness of processing enabled by AI systems and enforce this principle (Type 2). When interpreted by the CJEU as both procedural and substantive fairness, this principle would prevent potential harm for data subjects resulting from the processing of personal data by AI systems (see Section 6.2.2). In other cases, legislative action is unavoidable. The legislator needs to adjust the provisions in the current legal framework. This applies, for example, to the communication surveillance problem discussed in Section 4.9.3. Article 5 (1) ePD regulates the confidentiality of communication, but excludes human-machine communication services facilitated by AI (e.g. virtual assistants) from its scope. This creates a significant gap of protection (Type 3). The legislator could include new provisions in the future ePrivacy Regulation and specifically regulate the confidentiality of human-machine communication.

In light of the types of legal problems, the *extent* of adjustments to the current legal framework is also influenced by the distinction between *principles* and *enforceable rights*. As shown in Table 7.1, principles cause the *majority* of Type *3* legal problems. Conversely, Type *2* legal problems occur more often with *enforceable rights* than with principles.

(ii) AI disciplines

AI Discipline	Type 1		Type 2		Type 3	
	Principles	**Rights**	**Principles**	**Rights**	**Principles**	**Rights**
Machine Learning	8	9	4	7	15	8
Total	**17**		**11**		**23**	
Natural Language Processing	4	7	3	3	9	7
Total	**11**		**6**		**16**	
Computer Vision	4	3	3	2	7	7
Total	**7**		**5**		**14**	
Affective Computing	7	7	2	5	13	7
Total	**14**		**7**		**21**	
Automated Reasoning	7	3	2	2	7	7
Total	**10**		**4**		**14**	

Table 7.2 Overview of each discipline of AI causing different types of legal problems when applied to the principles and enforceable rights in the current legal framework.

Table 7.2 shows which AI discipline causes which type of legal problem when applied to the principles and enforceable rights in the current legal framework. As apparent from Table 7.2, the current legal framework does *not* suffice regarding all *AI disciplines* discussed in this thesis. Each discipline causes Type 1, 2 and 3 legal problems. The extent of adjustments varies per discipline of AI. I use the number of Type 2 and 3 legal problems as indicators for the extent of adjustments. *ML* and *AC* clearly stand out in terms of number of legal problems. There is a clear need for adjustments of the legal framework with regard to these two AI disciplines. ML leads to thirty-five legal problems, of which eleven are Type 2 and twenty-four are Type 3. *AC* is within the same range and leads to twenty-nine legal problems, of which seven are Type 2 and twenty-one are Type 3. To a *lesser extent*, the AI disciplines NLP, CV and AR also necessitate adjustments of the current legal framework. *NLP* causes twenty-three legal problems, six of which are Type 2 and seventeen are Type 3. *CV* causes slightly fewer legal problems than NLP, i.e. twenty problems, of which five are Type 2 and fifteen are Type 3. *AR* leads to eighteen legal problems, of which four are Type 2 and fourteen are Type 3.

When considered together, the five AI disciplines discussed in this thesis lead to thirty-three Type 2 legal problems and eighty-eight Type 3 legal problems. The Type 2 legal problems expose a clear enforcement problem. Many principles and enforceable rights in the current legal framework cannot

be enforced when applied to AI. The considerably higher total of Type 3 legal problems points to an obvious mismatch between the current legal framework and the developments in AI. Altogether, Type 2 and 3 legal problems unveil a clear need for adjustments of the current legal framework. They also disclose a difference between law in the books and law in action.

With so many legal problems, the question arises what to focus on and how. In my view, the legal problems that are most urgent and have the biggest impact on individuals should be prioritised. The elusiveness, mental data, communication surveillance, trade secrets, verifiability standard and cumulativeness problems discussed in Chapter 6 meet the requirements of urgency and impact. In terms of the how to address these problems, I suggest relying on three types of possible legal solutions: (i) new interpretations of existing provisions through guidelines and courts (ii) amending existing provisions or (iii) introducing new provisions.

Notably, the best solution is *not always* to be found within the current legal framework. The legislator could consider other areas of law, ensuring that these interact properly with the current legal framework for privacy and data protection. EU consumer law, competition law and product safety law are crucial to holistically protect individuals from actual and potential harm caused by AI. In February 2024, the AI Act's compromise text[2558] was published. It remains to be seen whether the legislator is diligent enough to ensure that the current legal framework and the AI Act genuinely complement each other rather than creating confusion about their interplay. When looking at the 'right not to be subject to automated individual decision-making' (Article 22 GDPR) and the 'right to explanation of individual decision-making' (Article 68 c AI Act compromise text),[2559] confusion seems more likely.

7.2 Recommendations for future legislation

Type 2 and 3 legal problems are the most problematic because they relate to situations in which data subjects cannot enforce their rights and to provisions which are not fit for purpose to protect the fundamental rights to privacy and the protection of personal data. The law appears to protect individuals, but in reality this protection is flawed. Future legislation should focus on these two types of legal problems. Admittedly, there is no silver bullet to solve Type 2 and 3 legal problems. Nonetheless, future legislation should put the emphasis on legal provisions that are *effective*. By effective, I mean provisions that are enforceable and fit for purpose to actually protect individuals.

To enact effective legal provisions, I recommend the legislator to use two particular instruments more often: *rebuttable presumptions* and *reversal of proof*.

[2558] AI Act compromise text resulting from the trilogue negotiations <https://data.consilium.europa.eu/doc/document/ST-5662-2024-INIT/en/pdf> accessed 8 February 2024.
[2559] Ibid.

The first instrument is rebuttable presumptions. Rebuttable presumptions assume something to be true until proven otherwise. It is an evidentiary instrument and shifts the evidential burden on the party to prove the contrary.[2560] The party to whose detriment the presumption is devised must adduce evidence to demonstrate that it is incorrect. This constitutes the rebuttal. Presumptions are used in law to improve the effectiveness of enforcement or to strengthen the claimant's position.[2561] Thus, presumptions are a particularly suitable instrument to improve enforcement and strengthen the position of individuals as the holders of fundamental rights. Rebuttable presumptions are seen as the least interventionist tool and are common in national liability systems of EU Member States. Also, the EU Commission's proposal for an Artificial Intelligence Liability Directive contains several rebuttable presumptions.[2562] *Rebuttable presumptions of harm* would make provisions contained in the legal framework more effective. My suggested legal solution for the *cumulativeness problem* contains a presumption of harm (see Section 6.7.2) The proposed legal solution consists of redrafting the right not to be subject to ADM and *assumes harm* if profiling or automated inferences is intended to be used for making decisions about the data subject. It requires controllers to perform an assessment of whether the envisaged profiling or automated inferences potentially harm the data subject's interests, rights and freedoms. Data subjects can obtain this assessment from the controller, which allows them to enforce their rights enshrined in the GDPR (e.g. lodging a complaint with an SA or initiate legal proceedings). A rebuttable presumption of harm might also be helpful with respect to the compensation of *non-material damages* caused by infringements of provisions contained in the current legal framework.

The second instrument is reversal of the burden of proof. To enact provisions that are more effective, the legislator should consider *reversal of proof* to favour the rights and interests of natural persons. The burden of proof facilitates courts to arrive at a decision in a legal dispute in favour of one of the parties involved in the case.[2563] Usually, the party that asserts a certain claim must prove it.[2564] With the reversal of the burden of proof, this burden shifts to the other party who must demonstrate that the claim put forward does not stand. Rules on the burden of proof have proven to be successful instruments in EU non-discrimination law.[2565] This instrument translates legal provisions in the 'books' to effective rights that protect individuals. Reversal of the burden of proof may ease problems

[2560] David Bailey, 'Presumptions in EU competition law' (2010) Vol 34 Iss 11 European Competition Law Review 362, 363.
[2561] Cyrill Ritter, 'Presumptions in EU competitionlaw' (2018) Vol 6 Iss 2 Journal of Antitrust Enforcement 189, 206.
[2562] Commission, 'Proposal for a Directive of the European Parliament and of the Council on adapting non-contractual civil liability rules to Artificial Intelligence (AI Liability Directive) COM (2022) 496 final at 6, 11, 13 and particularly Articles 3 and 4 <https://commission.europa.eu/system/files/2022-09/1_1_197605_prop_dir_ai_en.pdf> accessed 8 February 2024.
[2563] Douglas Walton, *Burden of Proof, Presumption and Argumentation* (Cambridge University Press 2014) 1.
[2564] Christopher Roberts, 'Reversing the burden of proof before human rights bodies' (2021) Vol 25 Iss 10 The International Journal of Human Rights 1682, 1684.
[2565] Lilla Farkas, Orlagh O'Farrell, 'Reversing the burden of proof: Practical dilemmas at the European and national level' (2015) document prepared for the European Commission at 9 <https://op.europa.eu/en/publication-detail/-/publication/a763ee82-b93c-4df9-ab8c-626a660c9da8/language-en> accessed 8 February 2024.

of data subjects regarding the enforcement of their rights. My proposed legal solution for the verifiability standard problem makes use of this instrument. I suggest adding an additional paragraph in Article 16 GDPR that broadens the right to rectification regarding the processing of personal data generated by automated means. This empowers data subjects to easily contest the accuracy of such personal data. When data subjects do so, the controller shall either cease processing or rectify the personal data as requested by the data subject, unless it can demonstrate that its own interests to process the personal data in the form as contested by the data subject prevail. Thus, it is the controller that bears the burden of proof. The reversal of the burden of proof makes the right to rectification more effective regarding personal data generated by AI systems.

7.3 Future research

The plethora of legal problems identified in this thesis indicates a clear need for future research. In my view, future research should be interdisciplinary, connecting different disciplines like law, technology, sociology, philosophy, economics and behavioural sciences.

In a world full of probabilistic predictions, scores and other inferences generated by means of AI, the accuracy principle is more important than ever. I call for interdisciplinary research in the fields of computer science and law to better substantiate the accuracy principle. Such research should develop specific standards of accuracy for personal data processed in the context of AI. Information quality, accuracy and completeness in computer science as well as validation accuracy in ML are relevant for this.

There is a clear need for interdisciplinary research with respect to the AI Act, for example, regarding manipulation enabled by AI systems. The AI Act's compromise text bans AI systems that deploy 'subliminal' or 'purposefully manipulative' or 'deceptive' techniques.[2566] However, the effect of 'subliminal techniques' appears to be statistically insignificant.[2567] Interdisciplinary research should further investigate how AI systems could manipulate individuals, how this affects personal autonomy and creates other ethical issues, which techniques are most harmful and effective, how individuals react to manipulation attempts and its economic consequences and how the law should address manipulation. Interdisciplinary research involving the disciplines law, technology, sociology, philosophy, economics and behavioural sciences is needed for this. The results of such research allows the legislator to adopt effective legal provisions which actually protect individuals.

[2566] Article 5 (1) lit a AI Act compromise text <https://data.consilium.europa.eu/doc/document/ST-5662-2024-INIT/en/pdf> accessed 8 February 2024.
[2567] Matija Franklin et al, 'The EU's AI Act needs to address critical manipulation methods' *The OECD.AI Policy Observatory* (Paris, 21 March 2023) <https://oecd.ai/en/wonk/ai-act-manipulation-methods?utm_source=substack&utm_medium=email> accessed 8 February 2024; Randolph J Trappey, Arch G Woodside, *Brand Choice* (Palgrave Macmillan London 2005).

In the context of the AI Act, the transparency of AC systems is another topic that requires interdisciplinary research. For example, Article 3 (1) point 34 of the AI Act's compromise text directly relates to AC systems. It defines an emotion recognition system ('ERS') as 'an AI system for the purpose of identifying or inferring emotions or intentions of natural persons on the basis of their biometric data'.[2568] Article 52 (2) compromise text requires deployers of ERS to inform individuals concerned about *the operation of the system*. Accompanying Recital 70 explains that natural persons should be notified when exposed to systems that can identify or infer their emotions or intentions. Thus, deployers of AI systems are not obliged to inform individuals about which *specific emotion* the system detected. This contradicts what Picard, the pioneer in AC, propagated: individuals should be able to know which emotion the machine recognised.[2569] Thus, the AI Act's compromise text does not fill the current loophole in EU data protection law. Interdisciplinary research is needed to explore possible solutions for closing this loophole in a legally and technologically sound manner. Scientists in the fields of computer science, psychology, philosophy and law will need to work together to achieve this goal.

Future research should also explore purely technological solutions.[2570] The problem of common sense discussed in Section 4.7.1 discloses reasoning deficiencies in the AI discipline of automated reasoning. This legal problem certainly meets the priorisation criteria of urgency and impact. But the solution to this problem is not a legal one. Since a long time, scientists had tried to understand and formalise how humans reason and whether reasoning methods may be automatised.[2571] The lack of progress in developing general automated common sense reasoning capabilities underscores that this is a very difficult problem in the field of AI.[2572] Common sense reasoning appears not only to be the hardest problem for AI, but also the most important one.[2573] The solution to this problem is technological, and future research in AI should prioritise it. For example, approaches such as qualitative spatial representation and reasoning[2574] should be further explored.

[2568] Article 3 (1) point 34 AI Act compromise text <https://data.consilium.europa.eu/doc/document/ST-5662-2024-INIT/en/pdf> accessed 8 February 2024.
[2569] Rosalind W Picard, *Affective Computing* (MIT Press 1997) 122.
[2570] E.g. randomisation techniques, secure multiparty computation, homomorphic encryption, differential privacy, synthetic data or knowledge-infused learning.
[2571] Marco Gavanelli, Toni Mancini, 'Automated Reasoning' (2013) Vol. 7 No. 2 Intelligenza Artificiale 113.
[2572] Brandon Bennet, Anthony G Cohn, 'Automated Common-sense Spatial Reasoning: Still a Hughe Challenge' in Stephen Muggleton, Nicholas Chater (eds) *Human-Like Machine Intelligence* (Oxford University Press 2021) 405.
[2573] Gary Marcus, Ernest Davis, *Rebooting AI: Buidling Artificial Intelligence we can trust* (Pantheon Books 2019).
[2574] Brandon Bennet, Anthony G Cohn, 'Automated Common-sense Spatial Reasoning: Still a Hughe Challenge' in Stephen Muggleton, Nicholas Chater (eds) *Human-Like Machine Intelligence* (Oxford University Press 2021) 410, 423.

Within this thesis, I have focussed on the potential and actual legal problems for individuals caused by developments in the field of AI. However, I am also fully aware of all the potential and actual benefits for individuals. In essence, everything depends on the *actual use of AI*. Decades ago, Kranzberg put it so accurately: 'Technology is neither good nor bad; nor is it neutral'.[2575] Clearly, such a maxim applies to the use of AI as well.

[2575] Melvin Kranzberg, 'Technology and History: "Kranzberg's Laws"' (1995) Vol 15 Iss 1 Bulletin of Science, Technology & Society 5-13.

Samenvatting proefschrift 'EU-wetgeving inzake privacy en gegevensbescherming toegepast op AI: een analyse van de juridische problemen voor individuen'

Het is algemeen bekend dat elke baanbrekende technologie risico's en complexe beleidsuitdagingen met zich meebrengt. Dit geldt in het bijzonder voor AI en de fundamentele rechten op privacy en de bescherming van persoonsgegevens. Dit proefschrift analyseert deze uitdagingen en heeft als doel antwoord te geven op de vraag in hoeverre de ontwikkelingen op het gebied van AI een nieuw juridisch kader voor deze fundamentele rechten vereisen.

Hoofdstuk 1 introduceert de context en maatschappelijke relevantie van dit proefschrift, evenals de onderzoeksvraag, de gebruikte methodologie en de reikwijdte van het proefschrift.

Hoofdstuk 2 onderzoekt wat AI is. Allereerst worden de bestaande definities van AI besproken. Vervolgens wordt ingegaan op de AI-disciplines die in de context van fundamentele rechten op privacy en gegevensbescherming het meest problematisch zijn. Deze disciplines zijn onder meer machinaal leren, natuurlijke taalverwerking, computervisie, affectief computergebruik en geautomatiseerd redeneren. *Machinaal leren* (ML) is een belangrijke discipline van AI, en richt zich op computers die zichzelf programmeren op basis van ervaring. ML kan worden toegepast door middel van verschillende methoden, variërend van gesuperviseerd tot niet gesuperviseerd leren, tot versterkend leren. *Diep leren* (DL) is een zeer krachtige vorm van machinaal leren. De resultaten op dit gebied zijn bereikt met *kunstmatige neurale netwerken* (KNN), die, in vergelijking met de neurale netwerken in het menselijk brein, uit een verbazingwekkend klein aantal neuronen bestaan. Door middel van *natuurlijke taalverwerking* (NLP) zijn machines in staat menselijke taal te verwerken. Dit omvat zowel het genereren als het begrijpen van natuurlijke taal. NLP draagt aanzienlijk bij aan het verbeteren van interacties tussen machines en mensen. *Computervisie* (CV) bevordert geautomatiseerd begrip van visuele beelden, waardoor machines in staat zijn om "te zien". Gezichtsherkenning, een van de toepassingen van computervisie, zorgt ervoor dat machines de identiteit van mensen die zichtbaar zijn in afbeeldingen of video's kunnen identificeren of verifiëren op basis van biometrische gegevens. Omdat emoties een belangrijk element van menselijke intelligentie zijn en een grote rol spelen in het dagelijks leven, is *affectief computergebruik* (AC) erop gericht om machines emotionele capaciteiten te geven. AC-methoden waarmee emoties afgeleid worden uit gezichtsuitdrukkingen en spraak, kunnen eenvoudig worden toegepast en op grote schaal worden gebruikt. Werkwijzen op het gebied van *geautomatiseerd redeneren* (GR) zijn gericht op het automatisch uitvoeren van individuele redeneringen.

Hoofdstuk 3 introduceert het huidige EU-rechtskader met betrekking tot de fundamentele rechten op privacy en de bescherming van persoonsgegevens. Ook gaat het hoofdstuk in op relevante secundaire EU-wetgeving. In dit kader worden de Algemene Verordening Gegevensbescherming (AVG), het

belangrijkste secundaire EU-recht op het gebied van gegevensbescherming, en de ePrivacy-richtlijn (ePR) besproken. Bijzondere nadruk wordt gelegd op de beginselen en afdwingbare rechten in de AVG.

De hoofdstukken 4 en 5 gaan in op de juridische problemen die zich voordoen, of kunnen voordoen, wanneer de *beginselen* en *afdwingbare rechten* uit het huidige rechtskader worden toegepast op de in hoofdstuk 2 geïntroduceerde AI-disciplines. Drie categorieën juridische problemen worden besproken: (1) wettelijke bepalingen worden geschonden, (2) wettelijke bepalingen zijn niet handhaafbaar en (3) wettelijke bepalingen zijn ongeschikt om het fundamentele recht in kwestie te beschermen. Deze juridische problemen worden onderzocht vanuit het perspectief van natuurlijke personen (individuen).

Door de tekortkomingen in geautomatiseerd redeneren (GR) zijn AI-systemen niet in staat om de logica van systemen die werken met geautomatiseerde besluitvorming (GB) weer te geven. De redeneringen of criteria die ten grondslag liggen aan een geautomatiseerd besluit, zijn zodoende onduidelijk. AI-systemen die gebruikmaken van DL- en KNN-benaderingen van machinaal leren, produceren waarschijnlijk niet-interpreteerbare resultaten. Wanneer ze worden gebruikt in de context van GB, kunnen de verwerkingsverantwoordelijken geen zinvolle informatie over de logica achter de GB aan de betrokkenen verstrekken. Hierdoor schenden ze het transparantiebeginsel (Type 1).

AI-systemen kunnen persoonsgegevens verwerken op een manier die doorgaans als oneerlijk wordt beschouwd, bijvoorbeeld wanneer door machinaal leren gegenereerde waarschijnlijkheidsvoorspellingen als feiten worden beschouwd. De onduidelijke rol en betekenis van het behoorlijkheidsbeginsel verminderen de rechtszekerheid en maken het moeilijk voor betrokkenen om de eerlijkheid van een verwerking aan te vechten. Het behoorlijkheidsbeginsel is hierdoor lastig handhaafbaar (Type 2).

AI-systemen vergemakkelijken de geautomatiseerde verwerking van nieuwe soorten gevoelige gegevens, zoals emotiegegevens en mentale gegevens. Ondanks hun zeer gevoelige aard worden dergelijke gegevens in de AVG niet specifiek beschermd als bijzondere gegevens. Dit komt doordat ervoor is gekozen om alle bijzondere gegevens uitputtend op te sommen. Aangezien de ontwikkelingen op het gebied van AI niet zijn bij te houden, loopt de wetgever hierdoor achter de feiten aan. Als gevolg ontstaat een hiaat in de bescherming, waardoor de AVG niet geschikt is om het fundamentele recht op bescherming van persoonsgegevens te waarborgen (Type 3).

Machinaal leren, natuurlijke taalverwerking en affectief computergebruik faciliteren toezicht op de communicatie tussen mens en machine. Grote techbedrijven die mens-machine communicatiediensten (zoals virtuele assistenten) aanbieden, kunnen dergelijke communicatie gemakkelijk onderscheppen en op een andere manier verwerken. Met natuurlijke taalverwerking en machinaal leren kan

gevoelige informatie worden afgeleid uit menselijke spraak en andere akoestische elementen in opgenomen audio. Naast de inhoud van spraak, kunnen de stemkarakteristieken en uitdrukkingswijze van een spreker een breed scala aan persoonlijke informatie bevatten. Dit omvat aanwijzingen over de biometrische identiteit, persoonlijkheid, fysieke kenmerken, geografische herkomst, het niveau van dronkenschap/slaperigheid, leeftijd, geslacht, gezondheidstoestand en zelfs de sociaaleconomische status van de spreker. Aanbieders van mens-machine communicatiediensten vallen niet onder het strikte regime van Artikel 5 (1) ePR, dat de vertrouwelijkheid van communicatie regelt. Deze leemte in de wet geeft aan dat de ePR niet geschikt is om de vertrouwelijkheid van mens-machine communicatie te waarborgen (Type 3).

Betrokkenen moeten voldoen aan de objectieve controleerbaarheidsnorm om gegevens te laten rectificeren die gegenereerd zijn door systemen op het gebied van machinaal leren en affectief computergebruik. Persoonsgegevens die worden gegenereerd door middel van machinaal leren kunnen oncontroleerbaar zijn. Gegevens over emoties zijn van nature zeer subjectief. Betrokkenen kunnen hierdoor geen bewijs leveren dat voldoet aan de objectieve controleerbaarheidsnorm. Het recht op rectificatie is dus niet geschikt om het fundamentele recht op bescherming van persoonsgegevens te beschermen, aangezien de norm betrokkenen belemmert in de uitoefening van hun recht (Type 3).

Hoofdstuk 6 heeft als doel om antwoord te geven op de vraag hoe de tekortkomingen van het huidige wettelijke kader die in hoofdstuk 4 en 5 zijn geïdentificeerd, moeten worden aangepakt. Op basis van de selectiecriteria effectiviteit, urgentie en nieuwheid bespreek ik zes juridische problemen: onduidelijkheid, mentale data, communicatiesurveillance, bedrijfsgeheimen, controleerbaarheid en cumulatieproblemen. Hoofdstuk 6 onderzoekt geschikte juridische oplossingen voor deze juridische problemen. Juridische oplossingen zijn (i) nieuwe interpretaties van bestaande bepalingen, (ii) het wijzigen van bestaande bepalingen of (iii) het introduceren van nieuwe bepalingen als antwoord op de betreffende juridische problemen. Wat dit laatste betreft, worden twee specifieke instrumenten onderzocht: weerlegbare aannames en omkering van bewijs.

Hoofdstuk 7 bespreekt de conclusies van dit proefschrift en geeft antwoord op de vraag in hoeverre de ontwikkelingen in AI een nieuw juridisch kader voor de fundamentele rechten vereisen. De mate waarin aanpassingen nodig zijn hangt grotendeels af van (i) het soort juridisch probleem en (ii) de AI-discipline. Dit laatste wordt uitgedrukt door het totaal aantal Type 2 en Type 3 problemen per AI discipline.

Voor juridische problemen van Type 1 volstaat het huidige rechtskader. De oplossing voor deze problemen ligt voor de hand. Schendingen van bepalingen binnen het huidige rechtskader moeten gesanctioneerd worden door betrokkenen en/of vertegenwoordigende organen ("particuliere handhaving") en door toezichthoudende autoriteiten ("regelgevende handhaving"). Het huidige rechtskader

volstaat daarentegen niet voor juridische problemen van Type 2 en 3. Niet-handhaafbare en "ongeschikte" bepalingen zijn simpelweg niet toereikend om privacy en persoonsgegevens te beschermen. Deze twee soorten juridische problemen vereisen *aanpassingen in wetgeving* of nieuwe interpretaties van de huidige bepalingen. In het laatste geval zijn gerechtelijke maatregelen in plaats van wetgevende maatregelen nodig. Neem bijvoorbeeld het probleem onduidelijkheid. De onduidelijke rol en betekenis van het behoorlijkheidsbeginsel vermindert de rechtszekerheid en maakt het moeilijk voor betrokkenen om de eerlijkheid van verwerkingen door AI-systemen aan te vechten en het beginsel af te dwingen (Type 2). Wanneer dit beginsel door het HvJ-EU wordt geïnterpreteerd als zowel procedurele als materiële eerlijkheid, zou het potentiële schade voor betrokkenen als gevolg van de verwerking van persoonsgegevens door AI-systemen voorkomen. In andere gevallen zijn wetgevende maatregelen onvermijdelijk. De wetgever moet bepalingen binnen het huidige rechtskader aanpassen. Dit geldt bijvoorbeeld voor het probleem met communicatiesurveillance. Artikel 5 (1) ePR regelt de vertrouwelijkheid van communicatie, maar sluit mens-machine communicatiediensten gefaciliteerd door AI (bijv. virtuele assistenten) uit van het toepassingsgebied. Hierdoor ontstaat een aanzienlijk hiaat in de bescherming (Type 3). De wetgever zou nieuwe bepalingen in de toekomstige ePrivacy-verordening kunnen opnemen, en specifiek de vertrouwelijkheid van communicatie tussen mens en machine kunnen regelen.

Bibliography

Ahuja S, Kumar J, 'Conceptualizations of user autonomy within the normative evaluation of dark patterns' (2022) Vol 24 Iss 4 Ethics and Information Technology

Akman P, *The concept of abuse in EU competition law* (Hart Publishing Ltd 2012)

Alonso E, 'Actions and agents' in Frankish Keith and Ramsey William M. (eds) The *Cambridge Handbook of Artificial Intelligence* (2014)

Alpaydin E, *Machine Learning: The New AI* (3rd edn MIT Press 2016)
-- -- *Introduction to Machine Learning* (4th edn MIT Press 2020)

Amit Y, Felzenszwalb P, 'Object Detection' in Katsushi Ikeuchi (ed) *Computer Vision – A Reference Guide* (Springer 2014)

Ariely D, Norton M, 'How actions create - not just reveal – preferences' (2007) Vol 12 Iss 1 Trends in Cognitive Sciences

Ausloos J, Veale M, Mahieu R, 'Getting Data Subject Rights Right' (2019) Vol 10 Iss 3 JIPITEC

Awasthi A, Mandal M, 'Facial Expressions of Emotions: Research Perspectives' in Manas K. Mandal, Avinash Awasthi (eds) *Understanding Facial Expressions in Communication* (Springer 2015)

Ayata D, Yaslan Y, Kamasak M, 'Emotion Recognition from Multimodal Physiological Signals for Emotion Aware Healthcare Systems' (2020) Vol 40

Bagger Tranberg C, 'Proportionality and data protection in the case law of the European Court of Justice' (2011) Vol 1 No 4 International Data Privacy Law

Bailey D, 'Presumptions in EU competition law' (2010) Vol 34 Iss 11 European Competition Law Review

Banakar R, Travers M, 'Introduction' in Reza Banakar, Max Travers (eds) *Theory and Method in Socio-Legal Research* (Hart Publishing 2005)

Barnhill A, 'What is Manipulation?' in Christian Coons, Michael Weber (eds) *Manipulation* (OUP 2014)

Barnhizer D, 'Inequality of bargaining power' (2005) Vol 76 Iss 1 University of Colorado Law Review

Barocas S, 'Data Mining and the Discourse on Discrimination' (2014)
-- -- and Nissenbaum H, 'Big Data's End Run around Anonymity and Consent' in *Privacy, Big Data, and the Public Good: Frameworks for Engagement* (CUP 2014)
-- -- and Selbst A, 'Big Data's disparate impact' (2016) Vol. 104 California Law Review

Barrett L et al. 'Emotional Expressions Reconsidered' (2019) Vol 20 (1) Psychological Science in the Public Interest

Bartlett M et al, 'Toward Automatic Recognition of Spontaneous Facial Actions' in Paul Ekman, Erika L. Rosenberg, *What the Face Reveals* (2nd edn OUP 2005)

Batini C, Palmonari M, Viscusi G, 'Opening the Closed World: A Survey of Information Quality Research in the Wild' in Luciano Floridi, Phyllis Illari (eds) *The Philosophy of Information Quality* (Springer Nature 2014)
-- -- and Scannapieco M, *Data Quality* (Springer 2006)

Baumann H, Dörig S, 'Emotion-Oriented Systems and the Autonomy of Persons' in Paolo Petta, Catherine Pelachaud, Roddie Cowie (eds) *Emotion-Oriented Systems* (Springer 2011)

Bellman R, *An Introduction to Artificial Intelligence: Can computers think*? (Boyd & Faser 1978)

Ben-Ze'Ev A, *The Subtlety of Emotions* (MIT Press 2000)

Bennet B, Cohn A, ‘Automated Common-sense Spatial Reasoning: Still a Huge Challenge’ in Stephen Muggleton, Nicholas Chater (eds) *Human-Like Machine Intelligence* (OUP 2021)

Berlin I, ‘Counter Enlightenment’ in Dictionary of the History of Ideas (1973)
-- -- *Liberty* (Hendry Hardy ed OUP 1969)

Bernhardt C, *Turing’s Vision: The Birth of Computer Science* (MIT Press 2016)

Betancourt M, ‘A Unified Treatment of Predictive Model Comparison’ (2015)

Biega A, Finck M, ‘Reviving Purpose Limitation and Data Minimisation in Data-Driven Systems’ (2021) Technology and Regulation 44-61

Biermann J, Horton J, Walter J, ‘Algorithmic Advice as a Credence Good’ (2022) Centre for European Economic Research Discussion Paper No 22-071

Binns R, Veale M, Is that your final decision? Multi-stage profiling, selective effects, and Article 22 of the GDPR’ (2021) Vol 11 No 4 International Data Privacy Law

Björkegren D, Grissen D, ‘Behavior Revealed in Mobile Phone Usage Predicts Credit Repayment’ (2020) Vol 34 Iss 3 The World Bank Economic Review

Black J, ‘Forms and paradoxes of principles-based regulation’ (2008) Vol 3 No 4 Capital Markets Law Journal

Blok P, ‘Het recht op privacy: een onderzoek naar de betekenis van het begrip 'privacy' in het Nederlandse en Amerikaanse recht’ (DPhil thesis, Tilburg University 2002)

Blumenthal-Barby J, ‘A Framework for Assessing the Moral Status of Manipulation’ in Christian Coons, Michael Weber (eds) *Manipulation* (OUP 2014)

Bolton T et al, ‘On the Security and Privacy Challenges of Virtual Assistants’ (2021) Vol 21 Iss 7 Sensors

Bourtoule L et al, ‘Machine Unlearning’ (IEEE Symposium on Security and Privacy, San Jose, May 2021)

Brkan M, ‘Do algorithms rule the world? Algorithmic decision-making and data protection in the framework of the GDPR and beyond’ (2019) Vol 27 Iss 2 International Journal of Law and Information Technology

-- -- and Bonnet G, ‘Legal and Technical Feasibility of the GDPR’s Quest for explanation of Algorithmic Decisions: of Black Boxes, White Boxes and Fata Morganas (2020) Vol 11 Iss 1 European Journal of Risk Regulation

Brown S, ‘European regulation of consumer credit: enhancing consumer confidence and protection from a UK perspective?’ in James Devenney et al (eds) *Consumer credit, debt and investment in Europe* (CUP 2012)

Brown W, ‘Technology, Workplace Privacy and Personhood’ (1996) Vol 15 Journal of Business Ethics

Bruno G, ‘The Importance of the European Convention on Human Rights for the Interpretation of the Charter of Fundamental Rights of the European Union’ in Giuseppe Palmisano (ed) *Making the Charter of Fundamental Rights a Living Instrument* (Brill Publishing 2014)

Bublitz J, ‘The Nascent Right to Psychological Integrity and Mental Self-Determination’ in Andreas von Arnauld, Kerstin von der Decken, Mart Susi (eds) *The Cambridge Handbook of New Human Rights* (CUP 2020)
-- -- and Merkel R, ‘Crimes against Minds: On Mental Manipulations, Harms and a Human Right to Mental Self-Determination’ (2014) Vol 8 Iss 1 Criminal Law and Philosophy

Buckley F, ‘Three Theories of Substantive Fariness’ (1990) Vol 19 Hofstra Law Review

Burr C, Cristianini N, ‘Can machines read our mind?’ (2019) Vol 29 Iss 3 Minds and Machines
-- -- and Cristianini N, Lydmann J, ‘An Analysis of the Interaction Between Intelligent Software Agents and Human Users’ (2018) Vol 28 Minds and Machines

Burrel J, 'How the machine 'thinks': understanding opacity in machine learning algorithms' (2016) Vol 3 Iss 1 Big Data Society

Bommasani R et al, 'On the Opportunities and Risks of Foundation Models' (2022) Center for Research on Foundation Models Standford University < https://arxiv.org/pdf/2108.07258.pdf > accessed 8 February 2024.

Bygrave L, 'Ensuring Right Information on the Right Person(s)' (1996) University of Oslo Institute for Private Law
-- -- 'Automated Profiling, minding the machine: Article 15 of the EC Data Protection Directive and Automated Profiling' (2001) Vol 17 No. 1 Computer & Law Security Report 1
-- -- *Data Privacy Law: An International Perspective* (OUP 2014)
-- -- 'Minding the machine v2.0: The EU General Data Protection Regulation and Autmated Decision Making' in Karen Yeung and Martin Lodge (eds) *Algorithmic Regulation* (OUP 2019)
-- -- Commentary of Articles 22 and 25 in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020)
-- -- 'Machine Learning, Cognitive Sovereignty and Data Protection Rights with Respect to Automated Decisions' (2021) University of Oslo Faculty of Law Legal Studies Research Paper Series No. 202-35
-- -- and Tosoni L, Commentary of Articles 4 (1) and 4 (7) in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020)

Calders T, Custers B, 'What is Data Mining and How Does it Work?' in Bart Custers et al. (eds) *Discrimination and Privacy in the Information Society* (Springer 2013)

Calders T, Žliobaitė I, 'Why Unbiased Computational Processes Can Lead to Discriminative Decision Procedures' in Bart Custers et al (eds) Discrimination and Privacy in the Information Society (Springer 2013)

Calix R, Javadpour L, Gerald M. Knapp, 'Detection of Affective States From Text and Speech For Real-Time Human-Computer Interaction' (2012) Vol 54 No 4 Human Factors and Ergonomics Society

Calo R, 'Privacy, Vulnerability, and Affordance' (2017) Vol 66 Iss 2 Depaul Law Review

Calvo R et al, 'Introduction to Affective Computing' in Rafael Calvo et al (eds), *The Oxford Handbook of Affective Computing* (OUP 2015)

Campolo A et al, 'AI Now Report' (2018)

Cand S, Alpert M, 'Emotion in Speech: The Acoustic Attributes of Fear, Anger, Sadness, and Joy' (1999) Vol 28 No 4 Journal of Psycholinguistic Research

Canfora G, Di Penta M, 'New Frontiers of Reverse Engineering' (2007) Future of Software Engineering (FOSE '07)

Cao Y, Yang J, 'Towards Making Systems Forget with Machine Unlearning' (IEEE Symposium on Security and Privacy, San Jose, May 2015)

Carlini N et al, 'The Secret Sharer: Evaluating and Testing Unintended Memorization in Neural Networks' (USENIX Security Symposium, Santa Clara, August 2019)

Caplier A, 'Visual Emotion Recognition: Status and Key Issues' in Catherine Pelachaud (ed) Emotion-oriented Systems (Wiley-ISTE 2012)

Carr N, *The Big Switch: Rewiring the World, from Edison to Google* (W. W. Norton & Company 2009)

Chang S et al, 'Turn-Taking Prediction for Natural Conversational Speech' (Interspeech Conference Incheon, September 2022)

Chang Y et al, 'A Survey on Evaluation of Large Language Models' (2023)

Chee W et al, 'Brain Structure in Young and Old East Asians and Westerners: Comparison of Structural Volume and Cortical Thickness' (2011) Vol 23 Iss 5 Journal of Cognitive Neuroscience

Cherednychenko O, 'Fundamental Freedoms, Fundamental Rights, and the Many Faces of Freedom of Contract in the EU' in Mads Andenas, Tarjei Bekkedal, Luca Pantaleo (eds) *The Reach of Free Movement* (Springer 2017)

Cho S, 'Exploiting machine learning techniques for location recognition and prediction with smartphone logs' (2016) Vol 176 Neurocomputing

Chow T, Siu-Yeung C, *Neural Networks and Computing: Learning Algorithms and Applications* (Imperial College Press 2007)

Chuang Z, Wu C, 'Multi-Modal Emotion Recognition from Speech and Text' (2004) Vol. 9 No. 2 Computational Linguistics and Chinese Language Processing

Chung H et al, 'Alexa, Can I Trust You?' (2017) Vol 50 Iss 9 Computer
-- -- and Park J, Lee S, 'Digital forensic approaches for Amazon Alexa ecosystem' (2017) Vol 22 Digital Investigation

Citron Keats D, Pasquale F, 'The scored society: Due process for automated predictions' (2014) Vol 89 Iss 1 Washington Law Review

Clifford D, 'Citizen-consumers in a personalised Galaxy: Emotion influenced decision-making, a true path to the dark side?' (2017) CiTiP Working Paper 31/2017
-- -- and Ausloos J, 'Data Protection and the Role of Fairness' (2018) Vol 37 No 1 Yearbook of European Law

Cohen J, 'Affording Fundamental Rights' (2017) Volume 4 Iss 1 Critical Analysis of Law

Cohn J, De La Torre F, 'Automated Face Analysis for Affective Computing' in Rafael Calvo et al (eds), *The Oxford Handbook of Affective Computing* (OUP 2015)

Collins H, 'Good Faith in European Contract Law' (1994) Vol 14 No 2 Oxford Journal of Legal Studies

Cowie R, 'Ethical Issues in Affective Computing' in Rafael Calvo et al (eds), *The Oxford Handbook of Affective Computing* (OUP 2015)

Cramer R, Damgård I, Nielsen J, *Secure Multiparty Computation and Secret Sharing* (CUP 2015)

Crawford K et al, 'AI Now Report' AI Now Institute (2018)

Creutzfeldt N et al, 'Socio-legal theory and methods: introduction' in Naomi Creutzfeld, Marc Mason, Kirsten McConnachie (eds) *Routledge Handbook of Socio-Legal Theory and Methods* (Routledge 2020)

Culnan M, 'Protecting Privacy Online: Is Self-Regulation Working?' (2000) Vol 19 Iss 1 Journal of Public Policy & Marketing

Custers B, *The Power of Knowledge* (Wolf Legal Publishers 2004)
-- -- 'Data Dilemmas in the Information Society' in Bart Custers et al (eds), *Discrimination and Privacy in the Information Society* (Springer 2013)
-- -- 'Profiling as inferred data. Amplifier effects and positive feedback loops' in Emre Bayamlıoğlu, Irina Baraluic, Liisa Janssens and Mireille Hildebrandt (eds), *Being Profiled: Cogitas Ergo Sum* (Amsterdam University Press 2018)
-- -- and Heijne A, 'The right of access in automated decision-making: The scope of article 15(1)(h) GDPR in theory and practice' (2022) Vol 46 Computer Law & Security Review
-- -- and Ursic H, 'Big data and data reuse: a taxonomy of data reuse for balancing big data benefits and personal data protection' (2016) Vol 6 Iss 1 International Data Privacy Law

Dahl J, Rudinow Sætnan A, 'It all happended so slowly – On controlling function creep in forensic DNA databases' (2009) Vol 37 International Journal of Law, Crime and Justice

Daly I et al, 'Affective brain–computer music interfacing' (2016) Vol 13 No 4 Journal of Neural Engineering

Das S et al., 'Applications of Artificial Intelligence in Machine Learning: Review and Prospect' (2015), Vol. 115, No. 9 International Journal of Computer Applications

Davis E, Morgenstern L, 'Introduction: Progress in formal common sense reasoning' (2004) Vol 153 Artificial Intelligence

de Barcelos Silva A et al, 'Intelligent personal assistants: A systematic literature review' (2020) Vol 147 Expert Systems With Applications

De Laat P, 'Algorithmic Decision-Making based on Machine Learning from Big Data: Can Transparency restore Accountability' (2017) Vol. 31 Issue 4 Philosophy & Technology

de Laat P, 'Algorithmic decision-making employing profiling: will trade secrecy protection render the right to explanation toothless?' (2022) Vol 24 Iss 1 Ethics and Information Technology

de Terwangne C, Commentary of Article 5 and 16 in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020)

de Vries K, 'Right to Respect for Private and Family Life' in Pieter van Dijk et al (eds) *Theory and Practice of the European Convention on Human Rights* (Intersentia 2018)

Dewey J, *The Middle Works of John Dewey, Volume 9, 1899-1924* (Carbondale Southern Illinois University Press 1980)

Dezfouli A, Nock R, Dayan P, 'Adversarial vulnerabilities of human decision-making' (2020) Vol 117 Iss 46 PNAS

Dienst S, 'Lawful Processing of Personal Data in Companies under the GDPR' in Daniel Rücker and Tobias Kugler (eds) *New European General Data Protection Regulation: A Practitioner's Guide* (Beck/Hart/Nomos 2018)

Dimitrova D, 'The rise of the personal data quality principle: is it legal and does it have an impact on the right to rectification?' (2021) Vol 12 No 3 European Journal of Law and Technology

Docksey C, 'Four fundamental rights: finding the balance' (2016) Vol 6 No 3 International Data Privacy Law

Dodge Y, 'Algorithm' in: *The Concise Encyclopedia of Statistics* (Springer New York 2006)

Domínguez-Jiménez J, 'A machine learning model for emotion recognition from physiological signals' (2020) Vol 55 Biomedical Signal Processing and Control

Doshi-Velez F et al, 'Accountability of AI Under the Law: The Role of Explanation' (2017) Berkman Klein Center Working Group on Explanation and the Law Working Paper

Douglas T, Forsberg L, 'Three Rationales for a Legal Right to Mental Integrity' in: Sjors Ligthart et al (eds) *Neurolaw Palgrave Studies in Law, Neuroscience, and Human Behavior* (Palgrave Macmillan 2021)

Dreyfusss R, van Eechoud M 'Choice of law in EU trade secrecy cases' in Jens Schovsbo, Timo Minssen and Thomas Riis (eds) *The Harmonization and Protection of Trade Secrets in the EU* (Edward Elgar Publishing 2020)

Dubois D et al, 'When Speakers Are All Ears: Characterizing Misactivations of IoT Smart Speakers' (2020) Iss 4 Proceedings on Privacy Enhancing Technologies

Dupré D et al, 'A performance comparison of eight commercially available automatic classifiers for facial affect recognition' (2020) 15 (4) PLoS ONE

Dwork C, 'Differential Privacy' in Michele Bugliesi et al (eds) *Automata, Languages and Programming* (Springer 2006)
-- -- and Roth A, *The Algorithmic Foundations of Differential Privacy* (Now Publishers Inc 2014)

Dworkin G, *The Theory and Practice of Autonomy* (CUP 1988)

Edwards L, Veale M, 'Slave to the Algorithm: Why a "Right to Explanation" is Probably not the Remedy You are Looking for' (2017) Vol 16 Iss 1 Duke Law & Technology Review

Eke D et al, 'Pseudonymisation of neuroimages and data protection: Increasing access to data while retaining scientific utility' (2021) Vol 1 Iss 4 Neuroimage

Ekman P, Friesen W, 'Constants across cultures in the face and emotion' (1971) Vol 17 (2) Journal of Personality and Social Psychology
-- -- 'Facial Action Coding System: A Technique for the Measurement of Facial Movements' (1978) Consulting Psychologists Press

El Emam K, Mosquera L, Hoptroff R, *Practical Synthetic Data Generation* (O'Reilly Media Inc 2020)

Elfering S, *Unlocking the Right to Data Portability* (Nomos 2019)

Engelbrecht A, *Computional Intelligence – An Introduction* (2 edn John Wiley & Sons 2007)

Eskens S, 'Profiling the European Citizen in the Internet of Things: How Will the General Data Protection Regulation Apply to this Form of Personal Data Processing, and How Should It? (2016) Master thesis University of Amsterdam

Etteldorf C, 'EDPB on the Interplay between the ePrivacy Directive and the GDPR' (2019) Iss 5 No 2 European Data Protection Law Review

Etzioni A, Etzioni, O 'Keeping AI Legal' (2016) 19 Vand. J. Ent. & Tech. L

Eyal A, 'Reasoning and decision making' in Frankish Keith and Ramsey William M. (eds) *The Cambridge Handbook of Artificial Intelligence* (2014)

Faden R, Beachamp T, King N, *A History and Theory of Informed Consent* (OUP 1986)

Fang C et al, 'The Overfitting Iceberg' (Machine Learning Carnegie Mellon University 31 August 2020)

Farkas L, O'Farrell O, 'Reversing the burden of proof: Practical dilemmas at the European and national level' (2015) document prepared for the European Commission

Fayek H, Lech M, Cavedon L, 'Evaluating deep learning architectures for Speech Emotion Recognition' (2017) Vol 92 Neural Networks

Fayyad M, 'Measures of the Principle of Good Faith in European Consumer Protection and Islamic Law, a Comparative Analysis' (2014) Vol 28 Arab Law Quarterly

Finck M, Pallas F, 'They who must not be identified- distinguishing personal from non-personal data under the GDPR' (2020) Vol 10 No 1 International Data Privacy Law

Finn R, Wright D, Friedewald M, 'Seven Types of Privacy' in S. Gutwirth et al. (eds.), *European Data Protection: Coming of Age* (Springer 2013)

Floridi L, Illari P, 'Information Quality, Data and Philosophy' in Luciano Floridi, Phyllis Illari (eds) *The Philosophy of Information Quality* (Springer Nature 2014)

Fogg B, *Persuasive Technology*: *Using Computers to Change What We Think and Do* (EBSCO Publishing 2003)

Fosch Villaronga E, Kieseberg P, Li T, 'Humans forget, machines remember: Artificial intelligence and the Right to Be Forgotten' (2018) Vol 34 Iss 2 Computer Law & Security Review

Fox R, 'Someone to watch over us: Back to the panopticon?' (2001) Vol 1 Iss 3 Criminal Justice

François-Lavet V et al, 'An Introduction to Deep Reinforcement Learning' (2018), Vol. 11, No. 3-4 Foundations and Trends in Machine Learning

Franklin M et al, 'Recognising the importance of preference change: A call for a coordinated multidisciplinary research effort in the age of AI' (2022)

Franklin S, 'History, motivations, and core themes' in Frankish Keith and Ramsey William (eds) *The Cambridge Handbook of Artificial Intelligence* (2014)

Freudiger J, Shokri R, Hubaux J, 'Evaluating the Privacy Risk of Location-Based Services' in Danezis Georg (ed) *Financial Cryptography and Data Security* (Springer 2012)

Gaur M et al, 'Knowledge-Infused Learning: A Sweet Spot in Neuro-Symbolic AI' (2022) Vol 26 Iss 4 IEE Internet Computing

Gavanelli M, Mancini T, 'Automated Reasoning' (2013) Vol. 7 No. 2 Intelligenza Artificiale

Gavrilescu M, Vizireanu N, 'Predictiong Depression, Anxiety, and Stress Levels from Videos Using the Facial Action Coding System' (2019) Vol 19 No 17

Generosi A, Ceccacci S, Mengoni M 'A deep learning-based system to track and analyse customer behaviour in retail store' (IEEE 8th International Conference on Consumer Electronics, Berlin 2018)

Gentile G, 'Two Strings to One Bow? Article 47 of the EU Charter of Fundamental Rights in the EU Competition Case Law: Between Procedural and Substantive Fairness' (2020) Vol 4 No 2 Market and Competition Law Review

Georgieva L, Christopher Kuner Commentary of Article 9 in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020)

Gerard D, 'Fairness in EU Competition Policy: Significance and Implications' (2018) Vol 9 No 4 Journal of European Competition Law & Practice

Ginart A et al, 'Making AI Forget You: Data Deletion in Machine Learning', Advances in Neural Information Processing Systems (2019)

Goffey A, 'Algorithm' in Matthew Fuller (ed) *Software Studies: A Lexicon* (MIT Press 2008)

Goldberg Y, *Neural Network Methods in Natural Language Processing* (Morgan & Claypool Publishers 2017)

Gonçalves dos Santos C, Papa J 'Avoiding Overfitting: A Survey on Regularization Methods for Convolutional Neural Networks' (2022) Vol 54 No Iss 10s ACM Computing Surveys

González Fuster G, *The Emergence of Personal Data Protection as a Fundamental Right of the EU* (Springer 2014)
-- -- 'Study on the essence of the fundamental rights to privacy and to protection of personal data' (2022)
-- -- and Peeters M, 'Person identification, human rights and ethical principles. Rethinking biometrics in the era of artificial intelligence' (2021)

Goodfellow I, Bengio Y, Courville A, *Deep Learning* (MIT Press 2016)

Goodman B, Flaxman B, 'European Union regulations on algorithmic decision-making and a right to explanation' (2017) Vol 38 No 3 AI Magazine

Gorin M, 'Towards a Theory of Interpersonal Manipulation' in Christian Coons, Michael Weber (eds) *Manipulation* (OUP 2014)

Graef I, 'Blurring Boundaries of Consumer Welfare' in Mor Bakhoum et al (eds), *Personal data in competition, consumer protection and intellectual property law* (Springer 2018)
-- -- and Husovec M, Purtova N, 'Data Portability and Data Control: Lessons for an Emerging Concept in EU Law' Vol 19 No 06 German Law Journal
-- -- and Clifford D, Valcke P, 'Fairness and enforcement: bridging competition, data protection, and consumer law' (2018) Vol 8 No 3 International Data Privacy Law

Guinchard A, 'Taking proportioanality seriously: The use of contextual integrity for a more informed and transparent analysis in EU data protection law' (2018) Vol 24 Iss 6 European Law Journal

Guo L, 'Randomization Based Privacy Preserving Categorical Data Analysis' (DPhil thesis, University of North Carolina 2010)

Gutwirth S, *Privacy and the information age* (Lanham:Rowman & Littlefield Publishers 2002)
-- -- and Hildebrandt M, 'Some Caveats on Profiling' in Serge Gutwirth et al (eds), *Data Protection in a Profiled World* (Springer Nature 2010)

Guzman A, Lewis S, 'Artificial intelligence and communication: A Human-Machine Communication agenda' (2020) Vol 22 Iss 1 New Media & Society

Hacker P, 'Teaching fairness to artificial intelligence: Existing and novel strategies against algorithmic discrimination under EU law' (2018) Vol 55 Iss 4 Common Market Law Review

Hallinan D et al, 'Neurodata and Neuroprivacy: Data Protection Outdated?' (2014) Vol 12 Iss 1 Surveillance and Society
-- -- and Zuiderveen Borgesius F, 'Opinions can be incorrect (in our opinion)! On data protection law's accuracy principle' (2020) Vol 10 No 1 International Data Privacy Law

Hammer L, *The international human right to freedom of conscience: some suggestions for its development and application* (Ashgate 2001)

Harris D et al, *Law of the European Convention on Human Rights* (4th edn OUP 2018)

Harrison J, *Handbook of Practical Logic and Automated Reasoning* (CUP 2009)

Hastie T, Tibshirani R, Friedman J, *The Elements of Statistical Learning* (2nd edn 2008)

Häuselmann A, 'Profiling and the GDPR: Harmonised Confusion' (2018) Jusletter 13

Haworth L, 'Dworkin on Autonomy' (1991) Vol 102 Ethics

Hazelrigg L, 'Inference' in Melissa Hardy, Alan Bryman (eds) *Handbook of Data Analysis* (Sage Publications 2004)

Healey J, 'Physiological Sensing of Emotion' in Rafael Calvo et al (eds), *The Oxford Handbook of Affective Computing* (OUP 2015)

Heigold G et al, 'End-to-End Text-Dependent Speaker Verification' (2015)

Heinrichs J, 'The Sensitivity of Neuroimaging Data' (2012) Vol 5 Iss 2 Neuroethics

Helberger N et al, 'The perfect match? a closer look at the relationship between eu consumer law and data protection law' Vol 54 Iss 5 Common Market Law Review

Herbst T, 'Art. 5 Grundsätze für die Verarbeitung personenbezogener Daten' in Jürgen Kühling and Benedikt Buchner (eds) DatenschutzGrundverordnung/BDSG (2nd edn Beck 2018)

Hijmans H, 'Commentary of Article 1' in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds) *The EU General Data Protection Regulation: A Commentary* (OUP 2020)

Hildebrandt M, 'Primitives of legal protection in the era of data-driven platforms' (2018)
-- -- and Koops B, 'The challenges of Ambient Law and Legal Protection in the Profiling Era' (2010) Vol. 73 (3) The Modern Law Review

Hoeren T, 'The EU Directive on the Protection of Trade Secrets and its Relation to Current Provisions in Germany' (2018) Vol 9 Iss 2 JIPITEC

Holzinger A, 'From Machine Learning to Explainable AI' (IEEE DISA Conference, Kosice, August 2018)

Hourri S, Kharroubi J, 'A deep learning approach for speaker recognition' (2020) Vol. 23 Iss. 1 International Journal of Speech and Technology

Hoy M, 'Alexa, Siri, Cortana, and More: An Introduction to Voice Assistants' (2018) Vol 37 No 1 Medical Reference Servics Quarterly

Hsu Y et al, 'Automatic ECG-Based Emotion Recognition in Music Listening' (2020) Vol 11 No 1 IEEE Transactions on Affective Computing

Hwand H, David Matsumoto, 'Emotional Expression' in Catharine Abell, Joel Smith (eds) *The Expression of Emotion* (CUP 2016)

Ienca M, Andorno R, 'Towards New Human Rights in the Age of Neuroscience and Neurotech- nology' (2017) Vol 13 Iss 1 Life Science, Society and Policy
-- -- and Haselager P, Emanuel E, 'Brain Leaks and Consumer Technology' (2018) Vol 36 Iss 9 Nature Biotechnology
-- -- and Ignatiadis K, 'Artificial Intelligence in Clinical Neuroscience: Methodological and Ethical Challenges' (2020) Vol 11 Iss 2 AJOB Neuroscience
-- -- and Malgieri G, 'Mental data protection and the GDPR' (2022) Vol 9 Iss 1 Journal of Law and the Biosciences

Jampani V, 'Learning Inference Models for Computer Vision' (Dissertation, Universität Tübingen 2016)

Jansen K et al, *Approximation, Randomization, and Combinatorial Optimization* (Springer 2004)

Javed Y, Sethi S, Jadoun A, 'Alexa's Voice Recording Behavior: A Survey of User Understanding and Awareness' (ARES '19, Canterbury 26-29 August 2019)

Jebelean T et al, 'Automated Reasoning' in Buchberger Bruno et al (eds) *Hagenberg Research* (Springer 2009)

Jeunet C, N'Kaoua B, Lotte F, 'Chapter 1 - Advances in user-training for mental-imagery-based BCI control: Psychological and cognitive factors and their neural correlates' in Damien Coyle (ed) *Progress in Brain-Computer Inerfaces: Lab Experiments to Real-World Applications* (Elsevier 2016)

Jianxin Z et al., 'Privacy-preserving Machine Learning Based Data Analytics on Edge Devices' (AIES Conference, New Orleans, January 2018)

Johnson-Laird A, 'Software Reverse Engineering in the Real World' (1994) Vol 19 Iss 3 University of Dayton Law Review

Joseph Raz, *The Morality of Freedom* (OUP 1986) 369.

Joshi Jet al, 'Multimodal assistive technologies for depression diagnosis and monitoring' (2013) Vol 7 Journal on Multimodal User Interfaces

Jurafsky D, James M, *Speech and Language Processing* (2 edn, Pearson Education Limited 2014)
Kanjo E et al, 'Emotions in context: examining pervasive affective sensing systems, applications, and analyses' (2015) Vol 19 Personal and Ubiquitous Computing

Kapitein M, 'Personalized Persuasion in Ambient Intelligence' (Doctoral Thesis, TU/e Eindhoven 2012)

Keats Citron D, Solove D, 'Privacy Harms' (2022) Vol 102 Iss 3 Boston University Law Review

Kellermayr P, 'Big Neurodata: On the Responsible Use of Neurodata from Clinical and Consumer-Directed Neurotechnological Devices' (2018) Vol 14 Neuroethics

Keltner D et al. 'Emotional Expression: Advances in Basic Emotion Theory' (2019) Vol 43 Iss 2 Journal of Nonverbal Behaviour

Klimas T, Vaitiukait J, 'The Law of Recitals in European Community Legislation' (2008) Vol 15 No 1 ILSA Journal of International & Comparative Law

Kline R, 'Cybernetics, Automata Studies, and the Dartmouth Conference on Artificial Intelligence' (2011) 4, EEE Computer Society

Kokott J, Sobotta C, 'The distinction between privacy and data protection in the jurisprudence of the CJEU and ECtHR' (2013) Vol 3 No 4 International Data Privacy Law

Kollar N, 'Virtue Ethics' in John K Roth (ed) *Ethics* (Salem Press Inc 2005*)*

Kollmer T, Eckhardt A, 'Dark Patterns' (2022) Vol 64 Iss 6 Business & Information Systems Engineering

Koning M, 'The purpose and limitations of purpose limitation' (Doctoral thesis, Radboud University Nijmegen 2020)

Koops B, 'The trouble with European data protection law' (2014) Vol 4 No 4 International Data Privacy Law
-- -- et al 'A Typology of Privacy' (2017) Vol. 38 Iss. 2 University of Pennsylvania Journal of International Law

Kosinski M, Stillwell D, Graepel T, 'Private traits and attributes are predictable from digital records of human behaviour' (2013) Vol 110 No 15 PNAS

Kosta E, Dumortier J, 'ePrivacy Directive: Assessment of transposition, effectiveness and compatibility within the proposed Data Protections Regulaton' (2015)

Kostas D et al, 'Please Forget Where I Was Last Summer: The Privacy Risks of Public Location (Meta) Data' (2019)

Kotschy W, Commentary of Article 6 in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020)

Kotu V, Deshpande B, *Data Science* (2nd edn Elsevier 2019)

Kovač J, Štruc V, Peer P 'Frame-based classification for cross-speed gait recognition' (2019) Vol 78 Multimedia Tools and Applications

Kranenborg H, Commentary of Articles 8 and 17 in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020)

Krans B, Nylund A, 'Aspects of Procedural Autonomy' in Bart Krans, Anna Nylund (eds) *Procedural Autonomy Across Europe* (Intersentia 2020)

Kranzberg M, 'Technology and History: "Kranzberg's Laws" ' (1995) Vol 15 Iss 1 Bulletin of Science, Technology & Society

Kröger J et al, 'Privacy Implications of Voice and Speech Analysis – Information Disclosure by Inference' in Michael Friedewald et al (eds) *Privacy and Identity management. Data for Better Living: AI and Privacy* (Springer 2020)

Kröger J, Raschke P, Is My Phone Listening in? On the Feasibility and Detectability of Mobile Eavesdropping' in: Simon N Foley (eds) *Data and Applications security and Privacy XXXIII* (Springer 2019)

Krysa J, Sedek G, 'Source Code' in Matthew Fuller (ed) *Software Studies: A Lexicon* (MIT Press 2008)

Kumar K et al, 'Multi-modal brain fingerprinting: A manifold approximation based framework' (2018) Vol 183 Neuro-Image

Kuner C et al, 'Expanding the artificial intelligence-data protection debate' (2018) Vol 8 No 4 International Data Privacy Law

Kursuncu U, Gaur M, Sheth A, 'Knowledge Infused Learning (K-IL): Towards Deep Incorporation of Knowledge in Deep Learning' (2020)

La Diega G, ‘Against the Dehumanisation of Decision-Making: Algorithmic Decisions at the Crossroads of Intellectual Property, Data Protection, and Freedom of Information’ (2018) Iss 1 Vol 9 JIPITEC

Lance B et al, ‘Brain-Computer Interface Technologies in the Coming Decades’ (2012) Vol 100 Proceedings of the IEE

Lanzig M, ‘Strongly Recommended: Revisiting Decisional Privacy to Judge Hypernudging in Self-Tracking Technologies’ (2019) Vol 32 Philosophy & Technology

Laud P, Kamm L, *Applications of Secure Multiparty Computing* (IOS Press BV 2015)

Lavazza A, ‘Freedom of Thought and Mental Integrity: The Moral Requirements for Any Neural Prosthesis’ (2018) Vol 12 Frontiers in Neuroscience

Lazaro C, Le Métayer D, ‘The Control over Personal Data: True Remedy or Fairy Tale?’ (2015) Vol 12 Iss 1 SCRIPT-ed

Lazarus R, *Emotion and Adaption* (OUP 1991)

Lech M et al, ‘Real-Time Speech Emotion Recognition Using a Pre-trained Image Classification Network: Effects of Bandwidth Reduction and Computing’ (2020) Vol 2 Frontiers in Computer Science

Lammerant H, de Hert P, ‘Predictive profiling and its legal limits: Effectiveness gone forever’ In Bart van der Sloot et al (eds) *Exploring the boundaries of big data* (2016 Amsterdam University Press/WRR)

Lee C et al, ‘Speech in Affective Computing’ in Rafael Calvo et al (eds), *The Oxford Handbook of Affective Computing* (OUP 2015)

Lee Y et al, ‘AIMQ: a methodology for information quality assessment’ (2002) Vol 40 Iss 2 Information & Management

Leenes R, De Conca S, ‘Artificial intelligence and privacy – AI enters the house through the Cloud’ in Woodrow Barfield, Ugo Pagallo (eds) *Research handbook on the law of artificial intelligence* (Edward Elgar Publishing Inc. 2018)

Lenaerts K, ‘National Remedies for Private Parties in the Light of the EU Law Principles of Equivalence and Effectiveness’ (2011) Vol 46 Irish Jurist
-- -- and Gutiérrez-Fons J, ‘To Say What the Law of the EU Is: Methods of Interpretation and the European Court of Justice’ (2013) European University Institute Working Paper AEL 2013/9

Lench H, Koebel Capenter Z, 'What Do Emotions Do for Us?' in Heather C Lench (ed) *The Function of Emotions* (Springer 2018)

Lerner J et al, ‘Emotion and Decision Making’ (2015) Vol 66 Annual Review of Psychology
-- -- and Small A, Loewenstein G, ‘Heart Strings and Purse Strings’ (2004) Vol 15 No 5 American Psychology Society

Leuner J, ‘A Replication Study: Machine Learning Models Are Capable of Predicting Sexual Orientation From Facial Images’ (2018)

Li D, Yang L, ‘A Joint Introduction to Natural Language Processing and Deep Learning’ and ‘Epilogue: Frontiers of NLP in the Deep Learning Era’ in Deng Li and Liu Yang (eds) *Deep learning in natural language processing* (Springer 2018)

Li S, Jain A, ‘Introduction’ in Li Stan, Jain Anil (eds) *Handbook of Face Recognition* (2nd edn, Springer 2011)

Liao L, ‘Location-Based Activity Recognition’ Dissertation University of Washington 2006

Liu Y et al, ‘Summary of ChatGPT-Related research and perspective towards the future of large language models’ (2023) Vol 1 Meta-Radiology 1 – 14

Ligthart S, 'Freedom of thought in Europe: do advances in 'brain-reading' technology call for revision?' (2020) Vol 7 Iss 1 Journal of law and the biosciences
-- -- et al, 'Forensic Brain-Reading and Mental Privacy in European Human Rights Law: Foundations and Challenges' (2021) Vol 14 Neuroethics

Lipton Z, 'The Mythos of Model Interpretability' (2018) Vol 16 Iss 3 ACMQueue

Lühmann T, Schumacher P, Stegemann L, 'Gegenwart und Zukunft kollektiver Rechtsdurchsetzung im Datenschutzrecht' (2023) Volume 3 Zeitschrift für Datenschutzrecht

Lupiáñez-Villanueva F et al, 'Behavioural study on unfair commercial practices in the digital environment: dark patterns and manipulative personalisation' (2022)

Lynskey O, Commentary of Article 20 in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020)

Mačėnaitė M, 'Protecting Children Online: Combining the Rationale and Rules of Personal Data Protection Law and Consumer Protection Law' in Mor Bakhoum et al (eds) *Personal Data in Competition, Consumer Protection and Intellectual Property Law* (Springer Nature 2018)

Macnish K, 'Unblinking the eyes: the ethics of automating surveillance' (2012) Vol 14 Ethics and Information Technology

Makin J, Moses D, Chang E, 'Machine translation of cordial activity to text with an encoder-decoder framework' (2020) Vol 23 Nature Neuroscience

Malgieri G, 'Trade Secrets v. Personal Data: A Possible Solution for Balancing Rights' (2016) Vol 6 No 2 International Data Privacy Law
-- -- The concept of Fairness in the GDPR' (FAT* '20: Conference on Fairness, Accountability, and Transparency, Barcelona, January 2020)
-- -- *Vulnerable People and Data Protection Law* (OUP 2022)
-- -- and Comandé G, 'Why a Right to Legibility of Automated Decision-Making Exists in the General Data Protection Regulation' (2017) Vol 7 Iss 4 International Data Privacy Law
-- -- and Custers B, 'Pricing privacy – the right to know the value of your personal data' (2017) Vol 34 Iss 2 Computer Law & Security Review
-- -- and Niklas J, 'Vulnerable data subjects' (2020) Vol 37 Computer Law & Security Review

Malkin N et al, 'Privacy Attitudes of Smart Speaker Users' (2019) Iss 4 Proceedings on Privacy Enhancing Technologies

Mann M, Matzner T, 'Challenging algorithmic profiling: The limits of data protection and anti-discrimination in responding to emergent discrimination' (2019) Vol 6 Iss 2 Big Data & Society

Mantello P, Manh-Tung H, 'Why we need to be weary of emotional AI' (2022) AI & Society

Marcus G, Davis E, *Rebooting AI: Building Artificial Intelligence we can trust* (Pantheon Books 2019)

Marechal C et al, 'Survey on AI-Based Multimodal Methods for Emotion Detection' in Joanna Kołodziej, Horacio González-Vélez (eds) *High-Performance Modelling and Simulation for Big Data Applications* (Springer 2019)

Marengo F, *Privacy and AI: Protecting Individuals' Rights in the Age of AI* (2023)

Markou C, Deakin S, 'Ex Machina Lex: Exploring the Limits of Legal Computability' in Simon Deakin, Christopher Markou (eds) *Is Law Computable?: Critical Perspectives on Law and Artificial Intelligence* (Hart Publishing 2020)

Marshall J, *Personal Freedom through Human Rights Law? Autonomy, Identity and Integrity under the European Convention on Human Rights* (Martinus Nijhoff Publishers 2009)

Marsland S, *Machine Learning: An Algorithmic Perspective* (2nd edn Chapman & Hall 2015)

Mary L, *Extraction of Prosody for Automatic Speaker, Language, Emotion and Speech Recognition* (2nd edn Springer 2019)

Mathur A, Mayer J, Kshirsagar M, 'What Makes a Dark Pattern... Dark?' (CHI Conference on Human Factors in Computing Systems, Yokohama, May 2021)

Matz S et al, 'Privacy in the age of psychological targeting' (2020) Vol 31 Current Opinion in Psychology
-- -- 'Psychological targeting as an effective approach to digital mass persuasion' (2017) Vol 114 No 48 PNAS

Mavroudis V, Veale M 'Eavesdropping Whilst You're Shopping: Balancing Personalisation and Privacy in Connected Retail Spaces' (Living in the Internet of Things Conference, London, March 2018)

Maxwell W, 'Principle-based regulation of personal data: the case of 'fair processing' (2015) Vol 5 No 3 International Data Privacy Law

Mayer-Schönberger V; Cukier K, *Big Data: A Revolution That Will Transform How We Live, Work and Think* (Houghton Mifflin Harcourt 2013)
-- -- and Padova Y, 'Regime change? Enabling Big Data through Europe's new Data Protection Regulation' (2016) Vol 17 No 2 Science and Technology Law Review

McStay A, 'Emotion AI, soft biometrics and the surveillance of emotional life: An unusual consensus on privacy' (2020) Vol 7 Iss 7 Big Data & Society
-- -- and Urquhart L, 'This time with feeling? Assessing EU data governance implications of out of home appraisal based emotional AI' (2019) Vol 24 No 10 First Monday

Mendoza I, Bygrave L, 'The Right Not to be Subject to Automated Decisions Based on Profiling' in Tatiana-Eleni Synodinou et al (eds) *EU Internet Law: Regulation and Enforcement* (Springer International 2017)

Menkel-Meadow C, 'Uses and Abuses of Socio-Legal Studies' in Naomi Creutzfeld, Marc Mason, Kirsten McConnachie (eds) *Routledge Handbook of Socio-Legal Theory and Methods* (Routledge 2020)

Michie S, van Stralen M, West R, 'The behaviour change wheel: A new method for characterising and designing behaviour change interventions' (2011) Vol 6 Implementation Science

Minsky M, *Semantic Information Processing* (MIT Press 1968)

Mitchell T, 'The discipline of Machine Learning' Carnegie Mellon University Paper CMU-ML-06-108 (2006)

Mittelstadt B et al, 'The ethics of algorithms: Mapping the debate' (2016) Vol 3 Iss 2 Big Data & Society

Mitchell M, Krakauer C, 'The debate over understanding in AI's large language models' (2023) Vol 120 Iss 3 PNAS 1-5

Moerel L, Priens C, 'Privacy for the Homo Digitalis: Proposal for a New Regulatory Framework for Data Protection in the Light of Big Data and the Internet of Things' (2016)

Mohri M, Rostamizadeh A, Talwalkar A, *Foundations of Machine Learning* (MIT Press 2012)

Molyneux C, Oyarzabal R 'What Is a Robot (Under EU Law)?' (2018) Vol 1 RAIL: The Journal of Robotics, AI & Law

Moore D, Commentary of Article 23 in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020)

Moses D et al, 'Real-time decoding of question-and-answer speech dialogue using human cortical activity' (2019) 10 Nature Communication

Mowbray A, 'The Creativity of the European Court of Human Rights' (2005) Vol 5 Iss 1 Human Rights Law Review

Mulligan K, Scherer K, 'Toward a Working Definition of Emotion' (2012) Vol. 4 No. 4 Emotion Review

Munakata T, *Fundamentals of the New Artificial Intelligence* (2nd edn, Springer 2008)

Munk T, 'Does Online Privacy Exist in the GDPR Era? The Google Voice Assistant Case' in Tatiana-Eleni Synodiou et al (eds) *EU Internet Law in the Digital Single Market* (Springer Nature 2021)

Murphy K, *Machine Learning: A Probabilistic Perspective* (MIT Press 2012)

Narayanan A et al, 'A Precautionary Approach to Big Data Privacy' in Serge Gutwirth et al (eds) *Data Protection on the Move: Current Developments in ICT and Privacy/Data Protection* (Springer Netherlands 2014)

Narayanan A, Shmatikov V, 'Myths and Fallacies of Personally Identifiable Information' (2010) 53 Communications of the ACM

Nautsch A et al, 'Preserving privacy in speaker and speech characterisation' (2019) Vol 58 Computer Speech & Language

Newell S, Marabelli M, 'The Crowd and Sensors Era: Opportunities and Challenges for Individuals, Organizations, Society, and Researchers' (ICIS, Auckland, December 2014)
-- -- 'Strategic opportunities (and challenges) of algorithmic decision-making: A call for action on the long-term societal effects of datafication' (2015) Vol. 24 Iss. 1 The Journal of Strategic Information Systems

Nielsen M, 'Why are deep neural networks hard to train' in: *Neural Networks and Deep Learning* (Determination Press 2015)

Nilsson N, *The Quest for Artificial Intelligence: A History of Ideas and Achievements* (CUP 2010)

Niolenko S, 'Synthetic Data for Deep Learning' (2019)

Nissenbaum H, *Privacy in Context: Technology, Policy and the Integrity of Social Life* (Stanford Law Books 2010)

Nixon M, Aguado A, *Feature Extraction & Image Processing for Computer Vision* (3rd edn Elsevier 2012)

Nordberg A, 'Trade secrets, big data and artificial intelligence innovation: a legal oxymoron?' in Jens Schovsbo, Timo Minssen and Thomas Riis (eds) *The Harmonization and Protection of Trade Secrets in the EU* (Edward Elgar Publishing 2020)

Ohm P, 'Sensitive Information' (2015) Vol 88 Southern California Law Review
Paszkiel S, *Analysis and Classification of EEG Signals for Brain–Computer Interfaces* (Springer Nature 2020)

Pedreschi D et al, 'Open the Black Box: Data-Driven Explanation of Black Box Decision Systems' (2018)

Pescatore P, 'Les objectifs de la Communauté européenne comme principes d'interprétation dans la jurisprudence de la Cour de justice' (1972) Vol 2 Miscellanea W.J. Ganshof van der Meersch

Picard R, 'Affective Computing' (1995) MIT Media Laboratory Perceptual Computing Section Technical Report No 321
-- -- *Affective Computing* (MIT Press 1997)

Pieraccini R, *AI Assistants* (MIT Press 2021)

Pipino L et al, 'Developing Measurement Scales for Data Quality Dimensions' in Richard Y Wang et al (eds) *Information Quality* (Routledge 2005 1 edn)

Pound R, 'Law in Books and Law in Action' (1910) Vol 44 Iss 1 Amercian Law Review

Prasad D, Reddy A, Vasumathi D, 'Privacy-Preserving Naive Bayesian Classifier for Continuous Data and Discrete Data' in Raju Surampudi Bapi et al (eds) *First International Conference on Artificial Intelligence and Cognitive Computing* (Springer Nature 2019)

Priyanka A, Gawali B, Suresh M, *Introduction to EEG- and speech-based emotion recognition* (Elsevier Inc 2016)

Purcell R, Rommelfanger K, ‘Internet-Based Brain Training Games, Citizen Scientists, and Big Data: Ethical Issues in Unprecedented Virtual Territories’ (2015)

Purtova N, ‘Do property rights in personal data make sense after the Big Data turn?’ (2017) Vol 10 No 2 Journal of Law & Economic Regulation
-- -- ‘The law of everything. Broad concept of personal data and future of EU data protection law' (2018) Vol 10 Iss 1 Law, Innovation and Technology

Quentin A et al, ‘Consumer Choice and Autonomy in the Age of Artificial Intelligence and Big Data’ (2018) Vol 5 Customer Needs and Solutions

Quintana D et al, ‘Heart rate variability is associated with emotion recognition: Direct evidence for a relationship between the automatic nervous system and social cognition’ (2012) Vol 86 No 2 International Journal of Psychophysiology

Rainey S et al, ‘Is the European Data Protection Regulation sufficient to deal with emerging data concerns relating to neurotechnology?’ (2020) Journal of Law and the Biosciences

Rashid M et al, ‘The classification of EEG Signal Using Different Machine Learning Techniques for BCI Application’ in J.-H. Kim et al (Eds) *Robot Intelligence Technology and Applications* (Springer 2018)

Raz J, *The Morality of Freedom* (OUP 1986)

Redman T, ‘Measuring Data Accuracy: A Framework and Review’ in Richard Y Wang et al (eds) *Information Quality* (Routledge 2005 1 edn)

Reece A, Danforth C, ‘Instagram photos reveal predictive markers of depression’ (2017) Vol. 6 No. 15 EPJ Data Science

Ridgeway C, ‘Why Status Matters for Inequality’ (2013) Vol 79 Iss 1 American Sociological Review

Ritter C, ‘Presumptions in EU competition law’ (2018) Vol 6 Iss 2 Journal of Antitrust Enforcement
Roberto Pieraccini, *AI Assistants* (MIT Press 2021)

Roberts C, ‘Reversing the burden of proof before human rights bodies’ (2021) Vol 25 Iss 10 The International Journal of Human Rights

Rocher L, Hendrickx J, Yves-Alexandre de Montjoye, ‘Estimating the success of re-identifications in incomplete datasets using generative models’ (2019) Vol 10 Nature Communications

Rodotà S, ‘Data Protection as a Fundamental Right’ in Serge Gutwirth et al (eds), *Reinventing Data Protection?* (Springer 2009)

Rommelfanger K et al, ‘Mind the Gap: Lessons Learned from Neurorights’ AAAS Center for Sience Diplomacy (2022)

Rosenberg E, ‘Introduction: The Study of Spontaneous Facial Expressions in Psychology’ in Paul Ekman, Erika L. Rosenberg, *What the Face Reveals* (2nd edn OUP 2005)

Rouvroy A, Poullet Y, ‘The Right to Informational Self-determination and the Value of Self-development: Reassessing the Importance of Privacy for Democracy’ in Serge Gutwirth et al (eds), *Reinventing Data Protection?* (Springer 2009)

Rumelhart D, Hinton G, Williams R ‘Learning representations by backpropagating errors’ (1986) Vol. 323 Nature

Russel S, Norvig P, *Artificial Intelligence, A Modern Approach* (3rd edn, Pearson Education 2016)

Saeed W, Omlin C, ‘Explainable AI (XAI): A systematic meta-survey of current challenges and future opportunities’ (2023) Vol 263 Knowledge-Based Systems

Sanders N, ‘A Balanced Perspective on Prediction and Inference for Data Science in Industry’ (2019) Iss 1.1 Harvard Data Science Review

Sartor G, Lagioia F, ‘The impact of the General Data Protection Regulation (GDPR) on artificial intelligence’ (2020) Study for the European Parliament’s Panel for the Future of Science and Technology

Savin A, *EU Telecommunications Law* (Elgar 2018)

Sax M, *Between Empowerment and Manipulation* (Kluwer Law International B.V. 2021)

Scantaburlo T, Charleswoth A, Cristianini N, ‘Machine Decisions and Human Consequences’ in Karen Yeung and Martin Lodge (eds) *Algorithmic Regulation* (OUP 2019)

Schabas W, *The European Convention on Human Rights: A Commentary* (OUP 2015)

Schäfer B, ‘Information Quality and Evidence Law: A New Role for Social Media, Digital Publishing and Copyright Law?’ in Luciano Floridi, Phyllis Illari (eds) *The Philosophy of Information Quality* (Springer Nature 2014)

Schauer F, Profiles, Probabilities, and Stereotypes (Harvard University Press 2006)

Schelter S, ‘Amnesia – A Selection of Machine Learning Models That Can Forget User Data Very Fast’ (Conference on Innovative Data Systems, Amsterdam, January 2020)

Schermaier M, ‘Bona Fides in Roman Contract Law’ in Reinhard Zimmermann, Simon Whittaker (eds) *Good Faith in European Contract Law* (CUP 2000)

Schneider F, Friesinger G, ‘Technology v Technocracy’ in Günther Friesinger and Jana Herwig (eds) *The Art of Reverse Engineering* (transcript Verlag 2014)

Schönherr L et al, ‘Unacceptable, where is my privacy? Exploring Accidental Triggers of Smart Speakers’ (2020)

Schovsbo J, ‘The Directive on trade secrets and its background’ in Jens Schovsbo, Timo Minssen and Thomas Riis (eds) *The Harmonization and Protection of Trade Secrets in the EU* (Edward Elgar Publishing 2020)

Selbst A, Powles J, ‘Meaningful information and the right to explanation’ (2017) Vol 7 Iss 4 International Data Privacy Law

Serban I et al. ‘A Deep Reinforcement Learning Chatbot’ Montreal Institute for Learning Algorithms (2017)

Shan H, ‘Towards High Performance and Efficient Brain Computer Interface Character Speller: Convolutional Neural Network based Methods’ Dissertation Universiteit Leiden 2020

Shemtov N, Beyond the Code: Protection of Non-Textual Features of Software (OUP 2017)

Shi Z, *Advanced Artificial Intelligence* (World Scientific 2011)

Shu L et al, ‘A Review of Emotion Recognition Using Physiological Signals’ (2018) Vol 18 Iss 7

Smith S, ‘In Defence of Substantive Fairness’ (1996) Vol 112 Iss 1 Law Quarterly Review

Sobn C, Alpert M, ‘Emotion in Speech: The Acoustic Attributes of Fear, Anger, Sandess, and Joy’ (1999) Vol 28 No 4 Journal of Psycholinguistic Research

Sokolova A, Konushin A ‘Methods of Gait Recognition in Video’ (2019) Vol 45 No 4 Programming and Computer Software

Solove D, ‘Data Is What Data Does: Regulating Use, Harm, and Risk Instead of Sensitive Data’ (2024) Vol 11 No 4 Northwestern University Law Review 1081 - 1138

Song C, Ristenpart T, Shmatikov V, 'Machine Learning Models that Remember Too Much' (2017) in Bhavani Thuraisingham et al (eds) Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security, CCS 2017 Dallas

Stanghellini G, Rosfort R, *Emotions and Personhood: Exploring Fragility – Making Sense of Vulnerability* (OUP 2013)

Steinert S, Friedrich O, 'Wired Emotions: Ethical Issues of Affective Brain–Computer Interfaces' (2020) Vol 26 Science and Engineering Ethics

Strauß S, 'From Big Data to Deep Learning: A Leap Towards Strong AI or Intelligentia Obscura' (2018) 2 (3), Big Data and Cognitive Computing

Strycharz J, Ausloos J, Helberger N, 'Data Protection or Data Frustration? Individual Perceptions and Attitudes towards the GDPR' (2020) Vol 6 Iss 3 European Data Protection Law Review

Sturari M et al, 'Robust and affordable retail customer profiling by vision and radio beacon sensor fusion' (2016) Vol. 81 Pattern Recognition Letters

Sunstein C, 'The Ethics of Nudging' (2015) Vol 32 Yale Journal of Regulation

Surblytė-Namavičienė G, *Competition and Regulation in the Data Economy* (Edward Elgar Publishing 2020)

Susser D, Roessler B, Nissenbaum H, 'Technology, autonomy, and manipulation' (2019) Vol 8 Iss 2 Internet Policy Review

Szeliski R, *Computer Vision: Algorithms and Applications* (Griets David, Schneider Fred Springer eds 2011)

Talesh S, Mertz E, Klug H, 'Introduction to the Research Handbook on Modern Legal Realism' in Shauhin Talesh, Elizabeth Mertz and Heinz Klug (eds) *Research Handbook on Modern Legal Realism* (Edward Elgar Publishing Limited 2021)

Tandon N, Varde A, de Melo G, 'Commonsense Knowledge in Machine Intelligence' (2017) Vol 46 No 4 SIGMOD Record

Taylor J, Practical Autonomy and Bioethics (Routledge 2009)

Taylor P, *Freedom of Religion UN and European Human Rights Law and Practice* (2005 CUP)

Tene O, Polonetsky J, 'Big Data for All: Privacy and User Control in the Age of Analytics' (2013) 11 Northwestern Journal of Technology and Intellectual Property

Toch E et al, 'Analyzing large-scale human mobility data: a survey of machine learning methods and applications (2019) Vol 58 Knowledge and Information Systems

Tomba K et al, 'Stress Detection Through Speech Analysis' (2018) Vol 1 ICETE 2018

Tome P et al., 'Facial soft biometric features for forensic face recognition' (2015) Vol 257 Forensic Science International

Tosoni L, Commentary of Article 4 (6) in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020)
-- -- 'The right to object to automated individual decisions: resolving the ambiguity of Article 22(1) of the General Data Protection Regulation' (2021) Vol 11 Iss 2 International Data Privacy Law

Trallero Ocaña T, *The Notion of Secrecy* (Nomos 2021)

Trappey R, Woodside A, *Brand Choice* (Palgrave Macmillan London 2005)
Treleaven P et al, 'The Future of Cybercrime: AI and Emerging Technologies Are Creating a Cybercrime Tsunami' (2023)

Trigueros D, Meng L, Hartnett M, 'Face recognition: From Traditional to Deep Learning Methods' (2018)

Tur G et al, 'Deep Learning in Conversational Language Understanding' in Deng Li and Liu Yang (eds) *Deep learning in natural language processing* (Springer 2018)

Turing A, 'Computing Machinery and Intelligence' (1950) Vol LIX Iss 236 Mind

Tzirakis P et al, 'End-to-End Multimodal Emotion Recognition using Deep Neural Networks' (2015) Vol. 14 No. 8 Journal of Latex Class Files

Usuelli M, *R machine learning essentials* (Packt Publishing 2014)

Uuk R, 'Manipulation and the AI Act' (2022)

Vallor S, *Technology and the Virtues: A Philosophical Guide to a Future Worth Wanting* (OUP 2016)

Valstar M, 'Automatic Facial Expression Analysis' in Manas K. Mandal, Avinash Awasthi (eds) *Understanding Facial Expressions in Communication* (Springer 2015)

van Canneyt T et al, 'Data Protection: CJEU case law review – 1995-2020' (2021) Vol 56 Computerrecht

van der Sloot B, 'Decisional privacy 2.0: the procedural requirements implicit in Article 8 ECHR and its potential impact on profiling' (2017) Vol 7 No 3 International Data Privacy Law
-- -- *Privacy as Virtue* (CUP 2017)

Van Hoecke M, *Methodologies of Legal Research: What Kind of Method for What Kind of Discipline?* (Hart Publishing 2011)

van Lieshout M, 'The value of personal data' in Jan Camenisch et al (eds) *Privacy and Identity 2014 IFIP AICT vol. 457* (Springer 2015)

Vandercammen L et al, 'On the Role of Specific Emotions in Autonomous and Controlled Behaviour' (2014) Vol 28 Iss 5 European Journal of Personality

Veale M et al, 'Algorithms that Remember: Model Inversion Attacks and Data Protection Law' (2018) A 376 Philosophical Transactions of the Royal Society A 376

Verhagen T, van Dolen W, 'The influence of online store beliefs on consumer online impulse buying: A model and empirical application' (2011) Vol. 48 Information & Management

Voigt P, von dem Bussche A, *The EU General Data Protection Regulation (GDPR): A Practical Guide* (Springer 2017)

von Grafenstein M, *The Principle of Purpose Limitation in Data Protection Laws* (Nomos 2017)

Vorras A, Mitrou L, 'Unboxing the Black Box of Artificial Intelligence: Algorithmic Transparency and/or a Right to Functional Explainability' in Ttiana-Eleni Synodinou et al (eds) *EU Inernet Law in the Digital Single Market* (Springer Nature 2021)

Vrabec H, 'Uncontrollable: Data Subject Rights and the Data-driven Economy' (Dissertation, Leiden University 2019)
-- --, *Data Subject Rights under the GDPR* (OUP 2021)

Wachter S, Mittelstadt B, Floridi L, 'Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation' (2017) Vol 7 Iss 2 International Data Privacy Law
-- -- and Mittelstadt S, 'A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI' (2019) No 2 Columbia Business Law Review

Wajnerman Paz A, 'Is Mental Privacy a Component of Personal Identity?' (2021) Vol 15 Frontiers in Human Neuroscience
-- -- 'Is Your Neural Data Part of Your Mind? Exploring the Conceptual Basis of Mental Privacy' (2022) Vol 32 Minds and Machines

Wallerman A, ‘Towards an EU law doctrine on the exercise of discretion in national courts? The Member States’ self-imposed limits on national procedural autonomy’ (2016) Vol 53 Iss 2 Common Market Law Review

Walton D, *Burden of Proof, Presumption and Argumentation* (CUP 2014)

Wang R, Strong D, ‘Beyond Accuracy: What Data Quality Means to Data Consumers’ (1996) Vol 12 No 4 Journal of Management Information Systems

Wang Y et al, ‘A systematic review on affective computing: emotion models, databases, and recent advances’ (2022) Volumes 83-84 Information Fusion

Warwick K, *Artificial Intelligence: The basics* (Routledge 2012)

Weatherill S, *EU Consumer Law and Policy* (2nd edn Elgar Publishers 2013)

Welinder Y, Palmer A, ‘Face Recognition, Real-Time Identification, and Beyond’ in Selinger Evan, Polonetsky Jules, Tene Omer (eds) *The Cambridge Handbook of Consumer Privacy* (CUP 2018)

Wen H et al, ‘Neural Encoding and Decoding with Deep Learning for Dynamic Natural Vision’ (2018) Vol 28 Iss 12 Cerebral Cortex

West S, Whittacker M, Crawford K, ‘Discriminating AI Systems: Gender, Race and Power’ (2019) AI Now Institute

Williams B, Brooks F, Shmargad Y, ‘How Algorithms Discriminate Based on Data They Lack: Challenges, Solutions, and Policy Implications’ (2018) Vol 8 Journal of Information Policy

Wischmeyer T, ‘Artificial Intelligence and Transparency: Opening the Black Box’ in Thomas Wischmeyer and Timo Rademacher (eds) *Regulating Artificial Intelligence* (Springer Nature 2020)

Wisman T, ‘Privacy, Data Protection and E-Commerce’ in Arno L. Lodder, Andrew D. Murray (eds) *EU regulation of e-commerce. A commentary* (Elgar 2017)

Wood A, ‘Coercion, Manipulation, Exploitation’ in Christian Coons, Michael Weber (eds) *Manipulation* (OUP 2014)

Xenidis R, ‘Tuning EU equality law to algorithmic discrimination - Three pathways to resilience’ (2020) Vol 27 Iss 6 Maastricht Journal of European and Comparative Law

Yang M et al, ‘Speech Reconstruction from Human Auditory Cortex with Deep Neural Networks’ (Interspeech Conference, Dresden, September 2015)

Yang Z, Zhong S, Wright R, ‘Privacy-Preserving Classification of Customer Data without Loss of Accuracy’ (2005)

Yoav S et al, ‘The AI Index 2018 Annual Report’ (AI Index Steering Committee Stanford University 2018)

Yoshida S, *Computer Vision* (Nova Science Publisher 2011)

Yuste R et al, ‘Four ethical priorities for neurotechnolgies and AI’ (2017) Vol 551 Nature

Zanfir-Fortuna G, Commentary of Articles 13-15 and 21 in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), *The EU General Data Protection Regulation: A Commentary* (OUP 2020)

Zarsky T, ‘Mine your own business!’ (2003) 5 Yale Journal of Law and Technology
-- -- ‘Incompatible: The GDPR in the Age of Big Data’ (2017) Vol 47 Iss 4 Seton Hall Law Review

Zbancioc M, Feraru S, ‘A study about the automatic recognition of the anxiety emotional state using Emo-DB’ (E-Health and Bioengineering Conference, Iasi, 2015)

Zhan J, ‘Using Homomorphic Encryption for Privacy-Preserving Collaborative Decision Tree Classification’ (IEEE Symposium on Computational Intelligence and Data Mining, Honolulu 2007)

-- -- et al, ‘Emotion recognition using multi-modal data and machine learning techniques: A tutorial and review’ (2020) Vol 59 Information Fusion

Zhao D et al, ‘Learning joint space–time–frequency features for EEG decoding on small labeled data’ (2019) Vol 114 Neural Networks

Zheng S, ‘User Perceptions of Smart Home IoT Privacy’ (2018) Vol 2 Proceedings of the ACM on Human-Computer Interaction

Zhi-Hua Z, Ji F, ‘Deep Forest: Towards an Alternative to Deep Neural Networks’ (IJCAI Conference, Melbourne, August 2017)

Zhu B, Shin U, Shoaran M, ‘Closed-Loop Neural Prostheses with On-Chip Intelligence: A Review and A Low-Latency Machine Learning Model for Brain State Detection’ (2021)

Žliobaitė I, ‘Measuring discrimination in algorithmic decision making’ (2017) Vol 31 Data Mining Knowledge Discovery

Zuboff S, *The age of surveillance capitalism* (PublicAffairs 2019)

Zuiderveen Borgesius F, ‘Improving Privacy Protection in the area of Behavioural Targeting’ (Doctoral thesis, Universiteit van Amsterdam 2015)
-- -- ‘Discrimination, artificial intelligence, and algorithmic decision-making’ Report for the Anti-discrimination department of the Council of Europe (2019)

Acknowledgments

It has been an exciting, absorbing, and, at times, frustrating ride to delve into the depth of artificial intelligence as well as EU privacy and data protection law. Some individuals have supported and helped me tremendously along this 'PhD ride'. I want to express my gratitude to the following data subjects by singling them out accordingly.

I am grateful to my supervisor, Bart Custers, for his excellent guidance. Also, I am thankful for the input from Gerrit-Jan Zwenne in his role as supervisor. I also thank Bart Schermer, Gianclaudio Malgieri, Nadya Purtova, Aline Klingenberg and Michèle Finck for their commitments as members of the doctorate committee.

To the eLaw community: thank you for creating such a welcoming and inspiring environment for my research. Although I was an external PhD candidate, you have always made me feel like a truly 'internal' part of eLaw. Lex, it was a pleasure to experience the PhD track together, with all the ups and downs. Particular thanks also to Roy de Kleijn from Leiden University for reviewing the chapter that describes AI from a technological perspective.

Writing a PhD thesis next to applying the law in practice is challenging and requires flexibility from the employer's side. Susanne Hofmann (PwC Legal), Joke Bodewits (Hogan Lovells), Geert Potjewijd and Axel Arnbak (De Brauw): thank you for always supporting my research endeavours and for granting me the freedom of mind that is required for writing a PhD thesis. Also, eternal thanks to the team members for picking up the work when needed.

Following the external PhD track can be a lonely journey. Susanna, thank you for always being so kind and letting me stay at your home in my beloved Züri. Tettje, Louis and Malouke, I want to express my gratitude for warmly welcoming me into your family and for all your support. I am also impressed with how patiently you endured my scientific monologues at the dinner table. Santiago, thanks for inspiring me back then in Oslo and supporting me in my research. To my friends Marco, Yannick and Woflis, thank you for keeping our friendship alive despite the distance. To my brother Lukas, thanks for encouraging me to go on a journey abroad together occasionally - it always helped me relax. Vanessa, many thanks for suffering with me during the last phase of my PhD endeavour. To my brother Johannes and to my Opi, thanks for keeping me in your prayers. Dad, your dedication to education has always inspired me. Siebe and Shubhanyu – thank you for being my paranymphs; I feel truly supported. To Thom, the creative mastermind: thanks for creating a marvellous cover (together with MidJourney and inspirations from Hinke).

Finally, Winnie, I am deeply grateful for your constant encouragement and support during these years, including your meticulous attention to missing full stops in the footnotes. I love you and look forward to the time ahead of us.

Curriculum Vitae

Andreas Nicolas Häuselmann works for De Brauw Blackstone Westbroek N.V. as a Senior Legal Advisor for its Privacy, Data & Cybersecurity practice.

Before joining De Brauw, Andreas worked for the Privacy and Cybersecurity practice of Hogan Lovells International LLP based in Amsterdam. Prior to that, Andreas was a member of the Privacy & ICT law practice at PwC Legal Switzerland located in Zurich.

Andreas is an external PhD candidate at eLaw, Center for Law and Digital Technologies at Leiden Law School of Leiden University, since October 2018. He regularly publishes in leading peer-reviewed journals on topics related to AI and EU data protection law. Andreas also gives guest lectures and talks at international conferences, for instance at the MIT.

Andreas holds a Master of Laws (LL.M.) in Information and Communication Technology Law from the University of Oslo and a Master of Laws (LL.M.) in Information Technology and Intellectual Property Law from the Leibniz University Hannover.