Universiteit
Leiden
The Netherlands

Big data, big risks, big power shifts: Evaluating the General Data Protection Regulation as an instrument of risk control and power redistribution in the context of big data

Rhoen, M.H.C.

Citation
Rhoen, M. H. C. (2019, September 12). *Big data, big risks, big power shifts: Evaluating the General Data Protection Regulation as an instrument of risk control and power redistribution in the context of big data*. Retrieved from https://hdl.handle.net/1887/77748

Version:	Accepted Manuscript
License:	Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden
Downloaded from:	https://hdl.handle.net/1887/77748

Note: To cite this publication please use the final published version (if applicable).

Cover Page

Universiteit Leiden

The handle http://hdl.handle.net/1887/77748 holds various files of this Leiden University dissertation.

Author: Rhoen, M.H.C.
Title: Big data, big risks, big power shifts: Evaluating the General Data Protection Regulation as an instrument of risk control and power redistribution in the context of big data
Issue Date: 2019-09-12

Big data, big risks, big power shifts

Evaluating the General Data Protection Regulation as an instrument of risk control and power redistribution in the context of big data

ISBN 978 94 6375 465 1

Copyright information
Chapters 1 and 6 © 2019 Michiel Rhoen, Creative Commons CC BY-SA 4.0;
Chapters 2, 3, 4 and 5: see the prelude of each chapter for copyright information;
All other content © 2019, Michiel Rhoen, all rights reserved.

Cover image credits
Cover design by Michiel Rhoen
Cover photograph by Alejandro Piñero Amerio (vjgalaxy) via pixabay.com
Additional cover illustrations by wanicon, Linector and Smartlne via www.flaticon.com

Print: Ridderprint | www.ridderprint.nl

Big Data, Big Risks, Big Power Shifts

Evaluating the General Data Protection Regulation
as an instrument of risk control and power redistribution
in the context of big data

PROEFSCHRIFT

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van Rector Magnificus prof. mr. C.J.J.M. Stolker,
volgens besluit van het College voor Promoties
te verdedigen op 12 september 2019
klokke 11:15 uur

door

Michiel Hendrik Christiaan Rhoen

geboren te Sittard

in 1969

Promotors:	Prof. dr. G.J. Zwenne
	Prof. dr. A.H.J. Schmidt
Co-promotor:	Dr. M. van der Linden-Smith (Vrije Universiteit Amsterdam)
Promotiecommissie:	Prof. dr. P.H. Blok (Utrecht University)
	Prof. dr. G.P. van Duijvenvoorde
	Dr. Q.Y. Feng (Utrecht University)
	Dr. R. Polčák (Masaryk University, Brno, Czech Republic)

Preface

This book presents the results of my research that was conducted from 2014 up to and including 2018. I want to acknowledge the following people who have offered their time and effort to help me improve the results:

Prof. Dr. Andreas Wiebe LLM. Doc. JUDr of the Georg-August Universität in Göttingen, Germany, presided the Research Forum Göttingen on 10 and 11 November 2016. A close-to-final draft of chapter 4 was presented at this forum. I wish to thank prof. Wiebe for extending the invitation to present at the Forum.

Doc. JUDr. Radim Polčák, Ph.D. of Masaryk University in Brno, Czech Republic, acted as commentator during the aforementioned Forum; he offered some insightful comments and corresponded with me after the Forum to further clarify these comments. I wish to thank dr. Polčák for his insights. His remarks have been worked into chapter 4.

Chapter 5 was co-authored by dr. Qing Yi Feng of Utrecht University, the Netherlands. I am grateful to dr. Feng for her willingness to collaborate with me, someone from another field without a track record in complexity science. Complexity is a nuanced subject where mathematics is often as important as language for making valid statements. Comprehensibly stating the points made in chapter 5 would have been a lot more difficult and time-consuming – and perhaps even impossible – without her help.

This work is dedicated to the loving memory of my father Wim Rhoen (1939–2018).

Houten, June 2019

Michiel Rhoen
michiel.rhoen@gmail.com

Table of Contents

1 Introduction

This research explores how "big data" leads to shifts in the distribution of power and risk between natural persons and data controllers, and how the General Data Protection Regulation (GDPR) addresses these shifts.

Big data currently is the subject of lively scholarly discourse. In their well-known book, Mayer-Schönberger and Cukier call it a "revolution".[1] However, the underlying technological developments have been shaping our society for a long time and the dilemmas facing law and society have been dealt with before. This chapter presents the central question that my research project aims to answer, the delineation of the research and the methodology. It will also position this research within the large body of legal scholarly work that has already been laid down – and still is being written – in this field. To provide context to the research question, this chapter starts with an outline of the technical origins of the term "big data", its applications in the consumer market and the emergence of platforms (sections 1.1 and 1.2). Subsequently, it will briefly explore some of the ways that power and risk shift as a result of the deployment of big data and the history European data protection law from 1968 to 2018.

1.1 Big data as technology: definition and origins

This research is not intended to conclude – or even take part in – any discussion on the meaning of "data", "information" or other concepts, but a working definition of both terms may still be useful to avoid unnecessary confusion when discussing big data. Therefore, in this research, "information" is interpreted as in the General Definition of Information (GDI): information – or semantic content – consists of "meaningful, well-formed data"; "data" is plural of "datum" which the GDI defines as "a lack of uniformity", for example between two physical states or the symbols describing those states. Using those definitions, "Big data" is a large collection of symbols discerning non-uniform states, which serves to derive or generate

[1] Viktor Mayer-Schönberger and Kenneth Cukier, *Big Data: A Revolution That Will Transform How We Live, Work, and Think* (Houghton Mifflin Harcourt 2013).

information.[2] The term "big data" purportedly originated in the mid-1990s in commercial computer graphics. By the year 2000, the term had appeared in academic papers in the field of computer science and statistics/econometrics.[3] In 2008, the popular science magazine Wired published "Visualizing big data: Bar charts for words" as part of an issue dedicated to the "Petabyte Age" (a Petabyte equals 1000 Terabytes).[4] According to the New York Times, 2012 was the "crossover year" for big data, due to the use of the term in a mainstream photo book, the Davos world economic forum and a Dilbert comic, among others.[5] Two months after publication of Mayer-Schönberger and Cukier's book with the same title, the term "big data" entered the Oxford Dictionary of English, with the following definition:[6]

> "extremely large data sets that may be analysed computationally to reveal patterns, trends, and associations, especially relating to human behaviour and interactions."

Apart from its large volume, Zikopoulos and Eaton attribute two additional technical characteristics to big data to help distinguish it from traditional or "small" data in the context of computational analysis: variety and velocity. Variety indicates that big data consists of both structured (or relational) data and unstructured data. Structured data is data that originates from, or could be added to, traditional database systems; unstructured data is a catch-all name for data from "web pages, web log files (...), search indexes, social media forums, e-mail, documents, sensor data from active and passive systems".[7] Velocity means not only that data is generated faster than before – which would essentially be similar to volume per unit of time – but also that analysis

---

2 Luciano Floridi, 'Philosophical Conceptions of Information' in Giovanni Sommaruga (ed), *Formal theories of information: from Shannon to semantic information theory and general concepts of information* (Springer 2009) 16, 18.

3 Francis X Diebold, 'A Personal Perspective on the Origin(s) and Development of "Big Data": The Phenomenon, the Term, and the Discipline' (2012) <http://www.ssc.upenn.edu/~fdiebold/papers/paper112/Diebold_Big_Data.pdf> accessed 20 March 2019.

4 Mark Horowitz, 'Visualizing Big Data: Bar Charts for Words' (2008) 16 Wired Magazine <https://www.wired.com/2008/06/pb-visualizing/>.

5 Steve Lohr, 'How Big Data Became So Big - Unboxed' *The New York Times* (11 August 2012) <https://www.nytimes.com/2012/08/12/business/how-big-data-became-so-big-unboxed.html> accessed 20 March 2019.

6 Oxford Dictionaries, 'Tweet Geekery and Epic Crowdsourcing: An Oxford English Dictionary Update' (*OxfordWords blog*, 13 June 2013) <https://blog.oxforddictionaries.com/2013/06/13/oed-june-2013-update/> accessed 20 March 2019.

on the data is performed sooner, while the data "is *still in motion*, not just after it is *at rest*" (emphasis in original).[8]

Big data was not purposefully designed as a separate technology. Instead, it is the result of continuing technological progress. The following developments appear especially relevant for the advent of big data: datafication, the digital transformation, telecommunications and the exponential capacity growth of computers, digital storage and telecommunications.

Datafication: Many processes have been automated over the years: from doing laundry and transferring money between bank accounts to delivering news and entertainment to mobile devices. Due to the possibility of errors or malfunctions, automation requires some form of monitoring. In simple processes, such as performed in washing machines, a direct readout of the cycle status on the device itself may suffice. But automation of more complex processes, or of tasks that are performed remotely, often requires that events in these processes are somehow recorded, not only for the short term (to make automated branching decisions), but also for a longer period to enable review and monitoring. Recordable events, such as the dialling of a phone number in a telephone network, generate data which is laid down in log files. Analysis of log files enables billing, the detection and correction of malfunctions, and the discovery of hacking and crime. As more processes become automated, more events are recorded, which leads to ever larger data sets as indicated by the moniker "petabyte age".

But datafication did not start in the petabyte age: the automated generation and processing of log files has been done for over half a century. An early example is the data from automated switches in the telephone network, which was already collected and analysed at larger scales in 1949.[9] ERMA, the first automated bookkeeping system for retail banking, had similar capabilities; it became operational in 1959.[10] Both developments constituted the datafication of their respective processes and both were essential in the development of new services: automated billing for phone services and linking of bank accounts and credit cards, respectively. They could have qualified

7 Paul Zikopoulos and Chris Eaton, *Understanding Big Data: Analytics for Enterprise Class Hadoop and Streaming Data* (McGraw-Hill Osborne Media 2011) 7.

8 ibid 9.

9 Godfrey Hammond, 'Your Phone Dial Computes Your Bill' (1949) 154 *Popular Science* 135.

10 AW Fisher and JL McKenney, 'The Development of the ERMA Banking System: Lessons from History' (1993) 15 IEEE Annals of the History of Computing 44.

as "big data" applications in their time. Due to the progress of technology, the number of recordable events in telecommunications has dramatically increased. Telecommunication networks have shifted away from recording "numbers dialled" towards technologies that record smaller events at a higher frequency.[11] Furthermore, many more phenomena are now converted to data and the number of processes generating data has grown with the number of sensors. This is not limited to events relating to accounting and computing: today, a typical smartphone contains an accelerometer, a gyroscope, a GPS sensor, an ambient light sensor and a thermometer, plus the capacity to connect with other systems that can register other variables pertaining to the environment or the human body; all these sensors provide data that can be made available for further processing.

Digital transformation: This term has no strict definition, but can be understood to be "the change associated with the application of digital technology in all aspects of human society."[12] An important effect of this transformation is the increased availability of information in digital form: information on pricing and availability of goods and services, text, images, audio and video are commonly digitally available. But digital technology has also become the basis of economical and social interaction. This has increased the number of human activities that employ automation at the end user level. As a result, an increasing number of activities generate data. This includes the buying and selling of goods and services, the retrieval of information, and keeping up with friends next door or at the other side of the world. This has led to increased logging of everyones' everyday activities. Where a person reading information in a physical book does not generate data, the same person reading the same information online involves sending a request identifying the user's device and, consequently, the user; this request is logged at several nodes of internet infrastructure. Similarly, a financial transaction using cash tends not to generate identifying information, whereas the same transaction paid through a check or a bank card generates a log entry recording both the payer and the payee. As a result of the digital transformation, the rate of production of digital data has increased significantly: in 2016, IBM asserted that "90 percent of the data in the world today has been created in the last two years alone".[13]

---

11 JS Turner, 'New Directions in Communications (or Which Way to the Information Age?)' (2002) 40 IEEE Communications Magazine 50, 50–51.

12 Digital transformation, 'Digital Transformation', *Wikipedia* (2018) <https://en.wikipedia.org/wiki/Digital_transformation> accessed 19 March 2019.

Electronic Telecommunications: Collecting single points of data into big data is not possible without some form of centralisation: data from all relevant locations needs to be transferred to a central point to enable logging for a large number of processes or sensors. In the big data context, this centralisation is typically achieved through telecommunication. The layout of the telephone network has included switching centres at several levels of centralization for a long time. This made telephony especially suitable for early datafication.[14] Centralisation across national borders became much easier after 1988, when the International Telecommunications Union (ITU, a United Nations specialised agency) decided on a treaty containing a new set of telecommunication regulations. These regulations empowered private entities to establish telecommunication links through "special arrangements", i.e., outside of the scope of the public switched telephone network (PSTN, the network accessible through the ITU numbering plan).[15] Specifically, article 9 of the new Regulations established that "for the first time, private operators were explicitly allowed to use leased lines to provide services, including data services."[16]

This could have resulted in a panoply of non-interoperable networks, if a standard for network interoperability had not already been available. In October 1982, the United States Department of Defense adopted the Internet Protocol Suite enabling the establishment of an interconnected network of computer networks, effectively creating the internet per the 1st of January, 1983.[17] The availability of a suitable protocol and the legal possibility of establishing network arrangements outside the PSTN enabled the emergence of the internet as the global open "network of networks" we know today: Hill asserts that "the Internet would not exist without article 9".[18] The

[13] Watson Marketing, '10 Key Marketing Trends for 2017 and Ideas for Exceeding Customer Expectations' (IBM Marketing Cloud 2017) WRL12345USEN 3 <https://www-01.ibm.com/common/ssi/cgi-bin/ssialias?htmlfid=WRL12345USEN> accessed 21 March 2019.

[14] Hammond (n 9); For banking, centralisation was originally achieved by mail: Fisher and McKenney (n 10).

[15] International Telecommunication Union, *International Telecommunication Regulations. Final Acts of the World Administrative Telegraphe and Telephone Conference, Melbourne, 1988, (WATTC-88).* (ITU 1989) 11; ITU-T, 'The International Public Telecommunication Numbering Plan - Recommendation ITU-T E.164' (Telecommunication Standardization Sector of ITU 2010) Recommendation E 36438.

[16] Richard Hill, *The New International Telecommunication Regulations and the Internet* (Springer Berlin Heidelberg 2014) 8.

[17] J Postel, 'NCP/TCP Transition Plan' (1981) Request for Comments RFC 801 4–5 <https://tools.ietf.org/html/rfc801> accessed 20 March 2019.

[18] Hill (n 16) 44.

internet has undercut the monopolies of national telephone carriers, reducing the price of telecommunications. This, in turn, has reduced the costs associated with providing services remotely. For example, a US-based company like Uber can offer a digital platform for taxi services in several European cities simultaneously without establishing a physical presence there. This platform also allows for the gathering of status and location information of drivers and customers in several countries simultaneously, and users can book a taxi in different countries through a single interface – features unavailable in traditional taxi services that are centralised at the city or district level.

Exponential growth: The capacity and performance of computers, storage and telecommunications has increased exponentially over the past decades. These observations are often referred to as Moore's law (for microprocessors), Keck's law (for data cables) and Kryder's law (for storage).[19] These "laws" have pushed the envelope of big data ever since its first applications. Because the amount of data that can be efficiently generated, stored, analysed and transferred has exponentially increased, the scope of big data has expanded to include an ever wider range of applications.

Two important use cases of big data are the use of analytics for the creation of knowledge (a term that is used loosely here, indicating the ability to assign attributes to objects or persons of interest) and the automation of decision-making.[20] This has not only assisted scientists in their search for subatomic particles and remote celestial objects:[21] it has also helped commercial parties to record and analyse an ever larger number of human activities and leverage the created knowledge in economic activities through data-generating platforms.[22]

---

19 Gordon E Moore, 'Cramming More Components onto Integrated Circuits' (1965) 38 Electronics Magazine 114 ff <https://ieeexplore.ieee.org/abstract/document/4785860> accessed 20 March 2019; Jeff Hecht, 'Is Keck's Law Coming to an End?' (2016) 2016 IEEE Spectrum 11 <https://spectrum.ieee.org/semiconductors/optoelectronics/is-kecks-law-coming-to-an-end>; Chip Walter, 'Kryder's Law' (2005) 293 Scientific American 32.

20 OECD, 'Data-Driven Innovation for Growth and Well-Being: Interim Synthesis Report' (OECD 2014) 30–33; OECD (ed), *Data-Driven Innovation: Big Data for Growth and Well-Being* (OECD 2015) 150.

21 Clifford Lynch, 'How Do Your Data Grow?' (2008) 455 Nature 28 <http://dx.doi.org/10.1038/455028a>.

22 OECD, *Data-Driven Innovation* (n 20) 90.

1.2 Personalisation, two-sided markets and the dominance of platforms

Big data has given rise to a number of opportunities for increased economic efficiency. This research focuses on the application of one particular type of data: personal data. The EU General Data Protection Regulation (GDPR) defines personal data as:

> "any information relating to an identified or identifiable natural person ('data subject'); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person".[23]

Initial use of big data in the consumer market involved *knowledge creation*. An early example is the datafication of retail transactions. From 1996, "integrated customer-facing front-ends" enabled retailers to record the link between transaction details and individual consumers.[24] A well-known example is a customer loyalty program extending benefits to customers when they present a personalised card at the time of purchase. Effective personalisation can reduce costs, for example if it reduces spending on ineffective marketing for the merchant; it has potential value for the consumer through the extension of attractive offers and the reduction of irrelevant advertising. Additionally, it offers insights into characteristics like brand loyalty and price sensitivity of individuals and groups, which is useful for market segmentation. This form of knowledge creation expands on loyalty programs based simply on "amount of money spent" like airlines' frequent flyer programs or retail trading stamp campaigns. Similar forms of knowledge creation result from individual credit scoring, which can provide more detailed insights in financial risk than earlier forms of knowledge creation based on smaller amounts of personal data, like bonus-malus systems and actuarial tables in the insurance market.

---

[23] 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, art 4(1); a data subject is a natural person identifiable by personal data (art. 4(1), GDPR).

[24] Vineet Kumar and Werner Reinartz, *Customer Relationship Management: Concept, Strategy, and Tools* (Springer Science & Business Media 2012) 17.

Apart from the creation of knowledge, datafication has also brought about the *automation of decision-making*. For example, big data enables algorithms to autonomously make data-driven pricing decisions. These algorithms use personal data (like the location of a consumer or their browser history), fluctuations in demand and competitor's prices as inputs – a technique known as dynamic pricing.[25] A Dutch insurance company now offers to adjust car insurance premiums based on real-time monitoring of location and driving behaviour in an attempt to "improve traffic safety".[26] Credit reporting firms, themselves controllers of large data sets, offer services that provide real-time risk assessment for merchants to use in their pricing algorithms and in their decisions to extend credit.[27]

Automated decision-making plays an even more important part in two-sided markets, defined here as "markets in which one or several platforms enable interactions between end-users, and try to get the two (or multiple) sides 'on board' by appropriately charging each side."[28] Two-sided markets existed long before datafication – newspapers charging fees to both subscribers and advertisers are a well-known traditional example.[29] However, datafication has enabled new two-sided markets, while automation has made them more efficient. Today's largest processors of personal data, also known as the "Tech's Frightful Five" (Amazon, Apple, Facebook, Google and Microsoft)[30] all provide platforms where consumers and merchants can

25 PK Kannan and Praveen K Kopalle, 'Dynamic Pricing on the Internet: Importance and Implications for Consumer Behavior' (2001) 5 International Journal of Electronic Commerce 63 <https://doi.org/10.1080/10864415.2001.11044211> accessed 20 March 2019; R Preston McAfee and Vera L Te Velde, 'Dynamic Pricing in the Airline Industry' in Terrence Hendershott (ed), *Economics and Information Systems* (1st edition, Elsevier 2006) 551–552.

26 Consumentenbond, 'Review: ANWB Veilig Rijden' (*Consumentenbond*, 21 July 2016) <https://www.consumentenbond.nl/autoverzekering/anwb-veilig-rijden> accessed 19 March 2019.

27 See, for example, Experian, 'Determine the Best Offer: Make Credit Decisions That Yield the Best Results' (2018) <http://www.experian.com/business-services/customer-leads.html> accessed 19 March 2019.

28 Jean-Charles Rochet and Jean Tirole, 'Two-Sided Markets: An Overview' (Institut d'Economie Industrielle working paper 2004) <https://pdfs.semanticscholar.org/1181/ee3b92b2d6c1107a5c899bd94575b0099c32.pdf> accessed 20 March 2019.

29 Jean-Charles Rochet and Jean Tirole, 'Platform Competition in Two-Sided Markets' (2003) 1 Journal of the European Economic Association 990, 992 <http://onlinelibrary.wiley.com/doi/10.1162/154247603322493212/abstract> accessed 20 March 2019.

30 Farhad Manjoo, 'Tech's Frightful Five: They've Got Us' *The New York Times* (10 May 2017) <https://www.nytimes.com/2017/05/10/technology/techs-frightful-five-theyve-got-

meet. The popularity of these platforms is based on commercial success in the sales of books and consumer products, smartphones, a social network, web search and operating systems, respectively. One reason that these companies deserve the "frightful" qualifier is their large number of users, indicating possible market dominance. For example: Facebook reported over 2 billion monthly active users in the second quarter of 2017, Google reported one billion active Gmail users in February 2016 and Apple reported 800 million iTunes accounts in 2014 although it did not estimate the number of active users.[31]

Platform providers use several business models. Revenue streams typically consist of charging consumers and sellers for access (e.g., through subscriptions or tying to hardware purchases), charging sales commissions to sellers and charging sellers for personalisation options for advertising and pricing decisions. Platform providers have a strong incentive to both maximise the number of consumers and to improve the accuracy of consumer profiling: both increase the value of the platform to sellers.[32] Platforms can attract additional consumers both by reducing the costs for consumers and by increasing the perceived value of their services. As a result, many platform services use one or more forms of community building, and many services are offered to consumers at no charge or tied to another purchase. Increased accuracy of profiling is achieved through the analysis of big data. Personalised advertising services offer market segmentation based on user profiles. The necessary data is generated by using a persistent method for consumer identification and the logging of events such as web search, maintaining lists of contacts, sending e-mail messages, mobile device use,

---

us.html> accessed 20 March 2019.

31 Nigam Arora, 'Seeds Of Apple's New Growth In Mobile Payments, 800 Million ITune Accounts' (*Forbes*, 24 April 2014) <https://www.forbes.com/sites/nigamarora/2014/04/24/seeds-of-apples-new-growth-in-mobile-payments-800-million-itune-accounts/> accessed 19 March 2019; statista.com, 'Facebook Users Worldwide 2018' (*Statista*, 2018) <https://www.statista.com/statistics/264810/number-of-monthly-active-facebook-users-worldwide/> accessed 20 March 2019; statista.com, 'Gmail: Global Active Users Worldwide 2016' (*Statista*, 2017) <https://www.statista.com/statistics/432390/active-gmail-users/> accessed 20 March 2019.

32 A phenomenon known as "network externalities". Carl Shapiro and Hal R Varian, *Information Rules: A Strategic Guide to the Network Economy* (Harvard Business Press 1998) 194; for networks, this is called "Metcalfe's law": B Briscoe, A Odlyzko and B Tilly, 'Metcalfe's Law Is Wrong - Communications Networks Increase in Value as They Add Members-but by How Much?' (2006) 43 Spectrum, IEEE 34; For a brief introduction to network externalities, see Paul Krugman and Robin Wells, *Economics* (Third Edition, Worth Publishers 2012) 469–472.

interaction with content, social interaction, buying behaviour, the logging of geographical location and the output of additional sensors.

The growth of the "Frightful Five's" two-sided markets has economic significance: since 2007, the five tech firms have replaced energy firms and financial institutions in dominating the rankings of "largest market capitalisation in the world".[33] A number of high-profile EU competition law cases against these companies underlines this significance: the European Commission fined Google €2.42 billion in 2017; Microsoft was fined €561 million in 2013, both for matters relating to article 102 of the Treaty on the Functioning of the European Union (TFEU) regarding abuse of a dominant position.[34] The German Bundeskartellamt (German national competition authority) initiated proceedings based on a similar complaint against Facebook in 2016.[35] Political effects are also visible: in 2017, the European Commission deemed a 1991 Republic of Ireland decision for a low tax rate for Apple to be illegal state aid and announced to take the Republic to court over its failure to reclaim a possible €13 billion in "illegal benefits" from the firm.[36]

For the individual consumer, a platform provider may appear to be "just another contract partner". But without the platforms of the Frightful Five dominating large sectors of the economy, the use of data analytics could very well have taken more

33 List of public corporations by market capitalization, 'List of Public Corporations by Market Capitalization', *Wikipedia* (2018) <https://en.wikipedia.org/wiki/List_of_public_corporations_by_market_capitalization> accessed 20 March 2019.

34 European Commission, 'Antitrust: Commission Fines Google €2.42 Billion for Abusing Dominance as Search Engine by Giving Illegal Advantage to Own Comparison Shopping Service' (27 June 2017) <http://europa.eu/rapid/press-release_IP-17-1784_en.htm> accessed 19 March 2019; European Commission, 'Antitrust: Commission Fines Microsoft for Non-Compliance with Browser Choice Commitments' (6 March 2013) <http://europa.eu/rapid/press-release_IP-13-196_en.htm> accessed 19 March 2019.

35 Bundeskartellamt, 'Bundeskartellamt Eröffnet Verfahren Gegen Facebook Wegen Verdachts Auf Marktmachtmissbrauch Durch Datenschutzverstöße' (*Meldung*, 3 March 2016) <https://www.bundeskartellamt.de/SharedDocs/Meldung/DE/Pressemitteilungen/2016/02_03_2016_Facebook.html> accessed 19 March 2019.

36 European Commission, 'State Aid: Commission Refers Ireland to Court for Failure to Recover Illegal Tax Benefits from Apple Worth up to €13 Billion' (4 October 2017) <http://europa.eu/rapid/press-release_IP-17-3702_en.htm> accessed 19 March 2019; Vanessa Houlder, Alex Barker and Arthur Beesley, 'Apple's EU Tax Dispute Explained' *Financial Times* (London, 30 August 2016) <https://www.ft.com/content/3e0172a0-6e1b-11e6-9ac1-1055824ca907> accessed 20 March 2019.

time, or could have been less pervasive. After all, their platforms have gathered data from a wide range of markets and a large consumer base, and made this data available to a large number of economic actors. Therefore, platforms have probably increased the use and the capabilities of data analytics.

1.3 Big data as a source of risk for individuals

Apart from the effects indicated above, datafication has introduced new risks. A small number of controllers now have access to large amounts of personal data about millions or billions of consumers. This causes a pronounced information asymmetry between controllers and data subjects and between controllers and governments. This information asymmetry is often qualified as a threat to individual privacy. It has introduced risks for both the individual and for society.

Theoretically, privacy risks are associated with the loss of individual autonomy – the opportunity to have one's own identity and independently make individual choices – and a diminishing separation between self and society that threatens the opportunities for dissent and critique.[37] Such a loss of autonomy can have far-reaching legal effects, because important legal concepts – *eg*, the right to vote, individual liability and freedom of contract – are based on the assumption that this autonomy is protected. Stated in terms of fundamental rights, big data can threaten the right to respect for private and family life and the prohibition of discrimination (articles 8 and 14, European Convention on Human Rights).

The threat to the right to private life can take many forms. A few examples:

- Information asymmetry can result in privacy losses by exposing information regarding contexts where data subjects have a "reasonable expectation of privacy".[38] This reduces the private sphere for data subjects, especially if the data is collected during the time that a data subject otherwise has a reasonable expectation of privacy or if data is combined from different contexts.

[37] Julie E Cohen, 'Turning Privacy Inside Out' (2019) 20 Theoretical Inquiries in Law (forthcoming), 3 <https://papers.ssrn.com/abstract=3162178> accessed 19 March 2019.

[38] ECtHR *Halford v. the United Kingdom*, 25 June 1997, 1997-III, para 45.

- Being "observed in all matters" puts data subjects under constant "threat of correction, judgement and criticism"[39] with possibly far-reaching psychological effects.
- A sufficiently large amount of personal data could enable a controller to digitally emulate a person, exposing data subjects to the threat of "plagiarism of their own uniqueness",[40] a well-known form of which is identity theft.
- Datafication extending across many contexts of social interaction can reduce the opportunities for anonymous expression, increasing personal inhibitions on personal expression due to the pressures associated with the "tyranny of the majority".[41]
- Storage of large amounts of data increases the adverse effects of data loss and breach of confidentiality.[42]
- The free flow of personal data can reduce informational self-determination as recognised by, for example, German law.[43] It also increases the likelihood that unlawful use of personal data can be hidden from view.

1.4 Big data and power relations: risks for society

Several authors claim that big data will bring about change at the societal level. Some of these changes are seen as risks. Mayer-Schönberger and Cukier mention the risks of endangered privacy, penalties based on propensities (instead of evidence) and the misuse of big data as a means for oppression.[44] According to Constantiou and Kallinikos, big data "reawakens the ghost of abstract or generic descriptions that may carry dubious social relevance".[45] Zuboff sees big data as a new "logic of accumulation": a new form of wealth inequality that she has dubbed "surveillance capitalism".[46] This leads to "substantial asymmetries of knowledge and power": the users of Google know less about themselves than Google does, they know little of

39 Bruce Schneier, 'The Eternal Value of Privacy' (*WIRED*, mei 2006) <http://archive.wired.com/politics/security/commentary/securitymatters/2006/05/70886> accessed 20 March 2019.

40 ibid.

41 Alexis de Tocqueville, *Democracy in America* (JP Mayer and George Lawrence eds, Harper Perennial Modern Classics 2006) ch 15.

42 Hal Berghel, 'Equifax and the Latest Round of Identity Theft Roulette' (2017) 50 IEEE Computer 72.

43 Bundesverfassungsgericht: *Volkszählungsurteil* [1983] BVerfGE 65,1 para C II 1 a.

44 Mayer-Schönberger and Cukier (n 1) ch 8.

45 Ioanna D Constantiou and Jannis Kallinikos, 'New Games, New Rules: Big Data and the Changing Context of Strategy' (2015) 30 Journal of Information Technology 44.

Google's operations, and this difference is insurmountable because the data-gathering happens through "undetectable functions of a global infrastructure that is also [...] essential for basic social participation."[47] This infrastructure is ominously called "big other".[48]

Indeed, datafication and the associated information asymmetries can cause power to shift from data subjects towards controllers. Firstly, collecting large amounts of data on a large number of natural persons is comparable to mass surveillance,[49] and surveillance is a well-known means of exerting power over individuals or groups of people. It is the central idea behind Bentham's Panopticon;[50] philosophical analysis was offered by Foucault.[51] But the risks of surveillance have been known for much longer. The two notions, that relations governing the availability of information are also power relations, and that society benefits if individuals are somehow protected against information asymmetry, are much older than data protection law – perhaps even dating back to biblical times.[52] Because of the resulting power differences, the notion of a surveillance state is alarming to many people, and the power that private controllers of large datasets can accumulate can have comparable effects.

Depending on the desired purpose, controllers can either practice *overt* or *covert* surveillance. Overt surveillance is often used to enforce conformity, both by governments and private actors: people who feel that they are being watched, tend to

46 Shoshana Zuboff, 'Big Other: Surveillance Capitalism and the Prospects of an Information Civilization' (2015) 30 Journal of Information Technology 75, 77.

47 ibid 83.

48 ibid 85.

49 Bruce Schneier, 'Metadata = Surveillance' (2014) 12 IEEE Security & Privacy 84, 84 <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6798571> accessed 20 March 2019.

50 Jeremy Bentham, *The Panopticon Writings* (Miran Božovic ed, Verso 1995) 31.

51 Michel Foucault, *Discipline and Punish: The Birth of the Prison* (Second Vintage Books edition, Vintage 1991) 201–202.

52 Kai Von Lewinski, 'Zur Geschichte von Privatsphäre Und Datenschutz-Eine Rechtshistorische Perspektive' [2012] Datenschutz: Grundlagen, Entwicklungen und Kontroversen, Bundeszentrale für politische Bildung, Bonn 23, 23; see also; Omer Tene, 'Vint Cerf Is Wrong. Privacy Is Not An Anomaly' (*Center for Internet and Society at Stanford Law School - Other writing*, 22 November 2013) <https://cyberlaw.stanford.edu/publications/vint-cerf-wrong-privacy-not-anomaly> accessed 21 March 2019 in response to; Gregory Ferenstein, 'Google's Cerf Says "Privacy May Be An Anomaly". Historically, He's Right.' (*TechCrunch*, 20 November 2013) <http://social.techcrunch.com/2013/11/20/googles-cerf-says-privacy-may-be-an-anomaly-historically-hes-right/> accessed 19 March 2019.

modify their behaviour both to comply with applicable norms and to not stand out too much from the crowd. The Chinese government is known to use overt surveillance to feed its "social credit" system, a system that enforces social norms by combining ubiquitous face scanning technology, internet traffic interception and data mining to control access to credit and public transport.[53] Overt surveillance by private parties can be seen on many internet discussion boards (including social networks), where content moderators can enforce standards by editing or deleting comments and by visibly removing offenders from the community.

On the other hand, covert surveillance is often used to discover individual features that could otherwise remain hidden. If individuals are permanently aware that they are being watched, the resulting increased conformity in their behaviour can make it more difficult to discover meaningful differences in traits of interest to a party using the data. Many governments permit their secret services to use covert mass surveillance in the interest of national security. Similarly, private entities that use personal data as part of their business model – especially, the aforementioned "Frightful Five" – tend to include data collection as a part of some other function of the platform while disclosing their data collection activities only in their terms and conditions or privacy statements.

Regardless of whether it is employed by governments or private parties, the power resulting from permanent mass surveillance can have far-reaching effects that can easily be qualified as risks:

- Governments can intend to use mass surveillance for the benefit of society, but Schneier asserts that it is "poor civic hygiene to install technologies that could someday facilitate a police state."[54] Apart from that, permanent surveillance of a society has been linked to negative effects on social capital and economic performance.[55]

---

53 Rene Chun, 'Big In… China: Machines That Scan Your Face' [2018] *The Atlantic* <https://www.theatlantic.com/magazine/archive/2018/04/big-in-china-machines-that-scan-your-face/554075/> accessed 19 March 2019; Adam Greenfield, 'China's Dystopian Tech Could Be Contagious' [2018] *The Atlantic* <https://www.theatlantic.com/technology/archive/2018/02/chinas-dangerous-dream-of-urban-control/553097/> accessed 20 March 2019.

54 Bruce Schneier, *Secrets and Lies: Digital Security in a Networked World* (1 edition, Wiley 2004) 53.

55 Andreas Lichter, Max Loeffler and Sebastian Siegloch, 'The Economic Costs of Mass Surveillance: Insights from Stasi Spying in East Germany' (IZA Discussion Papers 2015) s

- Enhanced possibilities for the identification of group membership and group attributes can exacerbate discriminatory trends in society and reduce solidarity, especially to the detriment of underprivileged groups.
- Similarly, if private controllers achieve a dominant position in their markets, this can have adverse effects at the societal level. A dominant position could be abused for rent-seeking,[56] to drive existing competitors out of the market or to raise barriers to entry for new competitors, thereby reducing economic vitality.
- If a dominant platform provider can control the flow of news and other information, this platform can facilitate reduced social, political or cultural pluralism and solidarity, for example by creating "filter bubbles." A platform provider can subsequently offer the creation of filter bubbles as a service to providers of other services. This can have far-reaching effects for an economy if it distorts the marketplace of goods and services, for example by a platforms' rent-seeking behaviour; it can have social and political effects if it fragments the marketplace of ideas[57] in a society.[58] Distortions in the marketplace of ideas can be disproportionally effective where electoral margins are thin and electoral systems can enable single-party dominance: activity on Facebook has been linked to meddlesome activity in the 2016 United States presidential elections.[59]

One of the aims of data protection law is to manage the risks associated with the processing of personal data and the resulting power differentials. The need for

5.3 and 6.

56 Richard A Posner, 'The Social Costs of Monopoly and Regulation' (1975) 83 Journal of Political Economy 807, 809–812 <https://www.jstor.org/stable/1830401> accessed 20 March 2019.

57 "(...) that the ultimate good desired is better reached by free trade in ideas – that the best test of truth is the power of the thought to get itself accepted in the competition of the market (...)". *Jacob Abrams, et al. v. United States* [1919] 250 U.S. 616, p. 630.

58 Frederik J Zuiderveen Borgesius and others, 'Online Political Microtargeting: Promises and Threats for Democracy' (2018) 14 Utrecht Law Review 82, 89 <https://www.utrechtlawreview.org/article/10.18352/ulr.420/> accessed 21 March 2019.

59 Eli Pariser, *The Filter Bubble: What the Internet Is Hiding from You* (Penguin Press 2011); Hannes Grassegger and Mikael Krogerus, 'Ich Habe Nur Gezeigt, Dass Es Die Bombe Gibt' [2016] *Das Magazin* <https://www.dasmagazin.ch/2016/12/03/ich-habe-nur-gezeigt-dass-es-die-bombe-gibt/> accessed 20 March 2019; Emma Graham-Harrison and Carole Cadwalladr, 'Revealed: 50 Million Facebook Profiles Harvested for Cambridge Analytica in Major Data Breach' (*the Guardian*, 17 March 2018) <http://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election> accessed 20 March 2019.

regulating automated processing arose almost as soon as its large-scale deployment by governments in post-war Europe.

1.5 Development of European data protection law, 1968-2018

In the late 1960's, political bodies began responding to the perceived threat of data processing technology for human rights and freedoms.[60] In 1968, the Council of Europe Parliamentary assembly recommended that the Committee of Ministers study whether "the national legislation in the member States adequately protects the right to privacy against violations which may be committed by the use of modern scientific and technical methods."[61] The first known European data protection law entered into force in 1970 in Hesse, Germany.[62] It was aimed at data processing by government bodies. Since this "Datenschutzgesetz" did not mention personal data specifically, it covered processing of both personal data and other data. Its enactment followed a recent Hessian innovation: the establishment of a government body for data processing and five associated computer centres processing a wide range of personal and non-personal data. The state of Hesse was well aware of possible power shifts associated with the processing of large data sets: the law provided for a *data protection supervisor* charged with observing the effects of processing, and with preventing shifts in the balance of powers between government bodies.[63] According to its Prime Minister, the law was enacted "to prevent the Orwellian vision of the all-knowing State seeking out the utmost intimate corners of the human sphere from becoming reality."

Legal protections were seen as necessary because citizens greeted the introduction of the use of computers in government administration with considerable skepticism. Apart from Germany, this was also registered in the Netherlands, where the 1971 census met resistance due to privacy concerns. These concerns intensified once the bureau of statistics stressed that the data would be processed largely by computers. The assertion that humans would hardly see the data did not provide the reassurance that officials had expected. Instead, activists considered the use of computers as

60 Sian Rudgard, 'Origins and Historical Context of Data Protection Law' in Eduardo Ustaran and others (eds), *European Privacy: Law and Practice for Data Protection Professionals* (International Association of Privacy Professionals 2012) 6.

61 Parliamentary Assembly, 'Human Rights and Modern Scientific and Technological Developments' (Council of Europe 1968) Recommendation 509 (1968).

62 Datenschutzgesetz vom 7. Oktober 1970, GVBl. II 300-10, Gesetz- und Verordnungsblatt für das Land Hessen nr. 41 (Teil I), 12 October 1970, p. 625 ("Datenschutzgesetz 1970").

63 Datenschutzgesetz 1970, §10(2).

increasing the threat of government intrusion into private lives. The census was nevertheless concluded, although some 250.000 people declined to participate and the results were severely delayed, due to both accidental errors in the question forms and deliberate errors in the answers. The Dutch government established a Commission to investigate the possibility of the introduction of privacy legislation in 1972.[64]

The increasing use of computers soon gave rise to legislative efforts in the field of data protection on both the European and the national levels. In 1972, "Resolution No. 3 on the protection of privacy in view of the increasing compilation of personal data into computers" was adopted by the seventh Conference of European Ministers of Justice in Basel. Resolutions 73 (22) and (74) 29 by the Committee of Ministers of the Council of Europe (COE) can be seen as the first steps towards "establish[ing] a framework of specific principles and norms to prevent unfair collection and processing of personal data". Many of the principles in these resolutions are still relevant: article 2 of the Annex to Resolution (74) 29 states that the information stored in data banks in the public sector should be "obtained by lawful and fair means, accurate and kept up to date, and appropriate and relevant to the purpose for which it has been stored" – principles that are carried over to the GDPR's articles 5(1)(a, c-d), and 6(2).[65]

In 1977 the Member States of the COE started negotiations for a treaty on data protection.[66] This initiative was at least partly attributable to events in France: in 1974, the newspaper *Le Monde* unveiled the French national government's plans to link all personal administrative data of the French citizenry in a computer system called "Système automatisé pour les fichiers administratifs et le répertoire des individus" (SAFARI) under the headline "SAFARI or the hunt for the French". The wide-ranging concern resulting from this report eventually resulted in the *Loi n° 78-17 du 6 janvier*

64 Jan Holvast, 'Op weg naar een risicoloze maatschappij? De vrijheid van de mens in de informatie-samenleving' (Leiden University 1986) ch 5; quoted in: Maurice Blessing, 'Het Verzet Tegen de Volkstelling van 1971' (2005) 15 *Historisch Nieuwsblad* <https://www.historischnieuwsblad.nl/nl/artikel/6697/het-verzet-tegen-de-volkstelling-van-1971.html> accessed 19 March 2019.

65 Council of Europe Committee of Ministers, *Resolution (74) 29 on the protection of the privacy of individuals vis-a-vis electronic data banks in the public sector 1974*; Council of Europe Committee of Ministers, *Resolution (73) 22 on the protection of the privacy of individuals vis-a-vis electronic data banks in the private sector* 1973.

66 Council of Europe, 'Convention 108 and Protocol: Background' (*Data Protection*) <https://www.coe.int/en/web/data-protection/convention108/background> accessed 19 March 2019.

1978 relative à l'informatique, aux fichiers et aux libertés.[67] This law is notable for being the first data protection law where special categories of personal data where recognised as a separate concern and worthy of specific protections.

Several other European countries enacted data protection legislation in the same period, in some cases based on a new and specific constitutional foundation.[68] Supranational efforts soon followed: in 1980, the Council of the OECD published a "Recommendation concerning guidelines for the processing of personal data and cross-border data flows".[69] In 1981, the "Convention nr. 108 for the protection of individuals with regard to automatic processing of personal data" ("Convention 108" or "Strasbourg Convention") was concluded.[70] It has since been ratified by all the Member States of the COE.[71] This Convention, like the OECD guidelines before it, covers personal data processed by both public bodies and private entities. The accompanying Explanatory Report all but recognises Moore's, Kryder's and Keck's laws when it states:

> "There is a need for such legal rules [strengthening data protection] in view of the increasing use made of computers for administrative purposes. Compared with manual files, automated files have a vastly superior storage capability and offer possibilities for a much wider variety of transactions, which they can perform at high speed. Further growth of automatic data processing in the administrative field is expected in the coming years inter alia as a result of the lowering of data processing costs, the availability of "intelligent" data processing devices

---

67 Loi n° 78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés 1978 (JORF [Journal Officiel de la Republique Française]) 227; Philippe Boucher, 'Safari Ou La Chasse Aux Français' *Le Monde* (Paris, 21 March 1974) <http://rewriting.net/2008/02/11/safari-ou-la-chasse-aux-francais/>; Molly Guiness, 'France Maintains Long Tradition of Data Protection' (*DW.COM*, 26 January 2011) <http://www.dw.com/en/france-maintains-long-tradition-of-data-protection/a-14797711> accessed 20 March 2019.

68 Rudgard (n 60) 15–17.

69 Council of the OECD, 'Recommendation of the Council Concerning Guidelines Governing the Protection of Privacy and Transborder Flows of Personal Data - C(80)58/FINAL' <http://www.oecd.org/internet/ieconomy/oecdguidelinesontheprotectionofprivacyandtransborderflowsofpersonaldata.htm> accessed 19 March 2019.

70 Council of Europe, Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data, done at Strasbourg 28 January 1981 (ETS 108).

71 https://www.coe.int/en/web/conventions/full-list/-/conventions/treaty/108/signatures.

> and the establishment of new telecommunication facilities for data transmission."[72]

The Netherlands and Germany were two countries where protests against the national census persisted. The 1981 Dutch census was never performed due to continuing protests based on privacy concerns. The Dutch national government has since decided to end the practice of door-to-door census and has performed only partial and virtual censuses.[73] In Germany, the Constitutional Court of the Federal Republic of Germany granted an injunction against the April 1981 census, followed by its final verdict against the Census law on 15 December. This verdict declared the law underlying the 1981 census incompatible with the German constitution and introduced the right to "informational self-determination" into German jurisprudence.[74]

In the years following 1981, European states implemented national data protection laws based on the OECD Guidelines and Convention no. 108. These national laws could differ in scope and in the extent of the protection. Since the OECD guidelines and Convention no. 108 both required that cross-border data flows were to be encouraged only if the receiving state had a similar level of data protection enshrined in law,[75] differences in national laws could stand in the way of the free flow of personal data between Member States. When the Member States of the (then) European Community signed the Single European Act with the purpose to establish a single European market in 1986, these differences in national law were seen as a possible hindrance in development of this market.[76] Therefore, in 1990, the European Commission published its first proposal for a data protection directive.[77] In 1992, a

---

72 Council of Europe, 'Explanatory Report to the Convention for the Protection of Individuals with Regard to Automatic Processing of Personal Data' (Council of Europe 1981) Explanatory Report 108 1.

73 E Schulte Nordholt and others, *Dutch Census 2011: Analysis and Methodology.* (Statistics Netherlands 2014).

74 Bundesverfassungsgericht, *Volkszählungsurteil* [1983] BVerfGE 65,1

75 Council of the OECD (n 69) para 17; Convention 108, art. 12(3).

76 Article 13, Single European Act [1987] OJ L 169/1, p. 7.

77 European Commission, 'Proposal for a Council Directive on the Protection of Individuals with Regard to the Processing of Personal Data' (European Commission 1990) COM(90) 314 final.

revised proposal was published.[78] Eventually, the finalised Data Protection Directive (DPD) was published in the Official Journal in 1995.[79]

Article 39 of the Treaty on the European Union and article 16 of the Treaty on the Functioning of the European Union provided a new legal basis for EU legislation in the area of data protection.[80] Preparations for data protection reform commenced in 2009 by means of two public consultations; a first draft for the GDPR was proposed in 2012.[81] The final version was published in the Official Journal of the European Union on 4 May 2016 and became applicable 25 May 2018 (art. 99(2)).

1.6 Interaction between science, policy and law

Considering that the processing of personal data forms a source of risks for data subjects and society, the GDPR can be seen as a policy response to these risks.[82] Presumably, this response is based on an assessment of the threat and an appropriate solution that is testable to a reasonable degree, to allow for meaningful evaluation of the legislation and to promote coherence in judicial decisions. However, the extent of

78 European Commission, 'Amended Proposal for a Council Directive on the Protection of Individuals with Regard to the Processing of Personal Data and on the Free Movement of Such Data' (European Commission 1992) COM (92) 422 final.

79 European Parliament and Council Directive 95/46/EC 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, [1995] OJ L 281/31, p. 31–50 (Data Protection Directive).

80 Treaty of Lisbon amending the Treaty on European Union and the Treaty establishing the European Community, signed at Lisbon, 13 December 2007, OJ C 306, p. 1–271

81 European Commission, 'Safeguarding Privacy in a Connected World. A European Data Protection Framework for the 21st Century' (European Commission 2012) Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions COM(2012) 9 final 3 <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2012:0009:FIN:EN:PDF> accessed 20 March 2019; European Commission, 'Proposal for a Regulation of the European Parliament and of the Council on the Protection of Individuals with Regard to the Processing of Personal Data and on the Free Movement of Such Data (General Data Protection Regulation) COM(2012) 11 (FINAL)' (European Commission 2012) <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52012PC0011> accessed 19 March 2019.

82 'Regulation can be seen as being inherently about the control of risks, whether these relate to illnesses caused by the exposure to carcinogens, inadequate utility services, or losses caused by incompetent financial advice.' Robert Baldwin, Martin Cave and Martin Lodge, *Understanding Regulation: Theory, Strategy, and Practice* (2nd edition, Oxford University Press 2013) 83; TNS Opinion and Social, 'Special Eurobarometer 359: Attitudes on Data Protection and Electronic Identity in the European Union' (European Commission 2011) Survey s 1.4.1.

the risks of big data is not yet fully clear.[83] For-profit surveillance of a large part of the populace has no precedent in modern history. Government surveillance at the scale of entire populations used to be expensive and labor-intensive, and was therefore practiced only by the most totalitarian or authoritarian of regimes. But it is now becoming a viable option for almost any government, especially if governments dominate large areas of a society's economic and social life, or if private companies can be convinced or coerced to cooperate in surveillance efforts.[84]

Even though the development of big data applications is relatively recent, societies have some experience dealing with power differentials and unknown risks of new technologies through legislation. The interplay between risk perception, power relations, fairness and legislation has been described and modelled, mainly in economics and the social sciences. Competition law and consumer protection law have the preservation of fairness and the moderation of the effects of power differentials as their focus. Similarly, questions surrounding the regulation of technological risks have also raised matters of fairness and power differentials, and models have been developed to better understand the interplay between relevant actors. These models have also been used in legislation, e.g. in environmental protection law. This provides a number of points of reference to compare data protection legislation with legislative efforts in other areas.

The GDPR aims to regulate several types of risks. A number of examples from the recitals:

- risks against the "rights and freedoms" of natural persons (Recitals 3 and 9), sometimes focused on sensitive data (recital 51);

---

83 Nadezhda Purtova, 'Who Decides on the Future of Data Protection? Role of Law Firms in Shaping European Data Protection Regime' (2014) 28 International Review of Law, Computers & Technology 204, 209 <http://dx.doi.org/10.1080/13600869.2013.801591> accessed 20 March 2019.

84 Maya Wang, 'China's Chilling "Social Credit" Blacklist' *Wall Street Journal* (11 December 2017) <https://www.wsj.com/articles/chinas-chilling-social-credit-blacklist-1513036054> accessed 21 May 2019; Sharon Weinberger, 'Son of TIA: Pentagon Surveillance System Is Reborn in Asia' (*WIRED*, 22 March 2007) <https://www.wired.com/2007/03/son-of-tia-pentagon-surveillance-system-is-reborn-in-asia/> accessed 21 March 2019; See also the now-defunct Directive 2006/24/EC of the European Parliament and of the Council of 15 March 2006 on the retention of data generated or processed in connection with the provision of publicly available electronic communications services or of public communications networks and amending Directive 2002/58/EC, [2006] OJ L 105/54 (Data Retention Directive).

- the risk that children are not fully aware of the risks involved by the processing of their data at the moment they consent to processing (recital 65);
- more generally, "risks to the interests and rights of the data subject" or "risks inherent in the processing" of personal data, including the risk of discriminatory effects (recitals 71, 83, 122).

This focus on risk management justifies an exploration into the degree to which the GDPR employs current theories on risk identification, evaluation and management. Such an exploration seems especially justified when considering, as will become clear in subsequent chapters, that several other fields of EU legislation have indeed incorporated testable models developed and verified in a scientific context.

In this research, the GDPR is evaluated using a limited number or models regarding distribution of power and technological risk. These models have originated in the social and the exact sciences. They are briefly mentioned here; their relevance and application in this book will be discussed in section 1.8 below (Methodology):

- Neil Komesar's method of *comparative institutional analysis* from the field of law and economics is used to evaluate or model the results of choosing a large-scale decision-making process to which a class of decisions is (to be) assigned. In this research, this method is applied to compare several options of decision-making where the processing of personal data is part of a consumer contract;
- Michael Barnett and Raymond Duvall's theory of *power in social relations* from the social sciences is used to compare the GDPR with EU consumer protection law to assess the GDPR's protection against unfair contract terms and unfair commercial practices where the processing of personal data is part of a consumer contract;
- Ulrich Beck's theory of the *risk society,* Charles Perrow's theory of *normal accidents,* and Andreas Klinke and Ortwin Renn's *approach to risk evaluation and management,* also stemming from the social sciences but partly based in the exact sciences, are used to compare how the GDPR and various EU legal instruments of environmental protection law acknowledge and deal with technological risks;
- The science of *complex systems* is used to evaluate the expected effectiveness of two articles relating to the processing of sensitive personal data as defined in article 9(1) of the GDPR.

European Data protection law has shown periods of relative stability punctuated by moments of substantial change. The development of new iterations of regulation can take over a decade and is likely to involve finding acceptable compromises between conflicting interests and viewpoints. The GDPR, for example, replaces a directive that came into force 23 years earlier; the directive from 1995 succeeded a Council of Europe treaty from 1981. The European Commission hopes that the GDPR will be future proof for decades to come.[85]

But long periods of legislative standstill increase the risk that data protection law becomes less effective due to technological progress. The years between subsequent iterations could therefore be used to increase our understanding of the effects of both innovation and legislation on risks and power relations, and to build a body of jurisprudence where the assumptions of legislators are tested against the outcomes of real-life disputes before the courts. The aim of gaining these insights is to systematically improve the efficacy of the law. Still, we must recognise, as Coase did, that both the presence *and* the absence of regulation will rarely result in any sort of optimal solution.[86]

1.7 Introducing the research question

In the case of the GDPR, improving our understanding of the interaction between law and technology stands a good chance of being useful because a number of experts seem to have doubts about its expected effectiveness. Criticism emerged already in the period leading up to the GDPR's passing into law. Three examples:

- Moerel has opined that the GDPR needs to be made *future proof*. Technological developments will negate the effects of the informed consent requirement, the profiling prohibition and overly specific documentation requirements; she dismisses the purpose limitation principle as "at odds with the reality of big data".[87]

---

85 European Commission, 'Proposal for a General Data Protection Regulation' (n 81) 104.

86 'It is obvious that if you are comparing the performance of an industry under regulation with what it would be without regulation, there is no reason to assume (indeed there is good reason not to assume) that either of these situations will correspond to anything an economist would call optimal. (...) Until we realize that we are choosing between social arrangements which are all more or less failures, we are not likely to make much headway.' Ernest W Williams and Ronald H Coase, 'Discussion' (1964) 54 The American Economic Review 192, 194–195 <http://www.jstor.org/stable/1818503> accessed 21 March 2019.

- Koops has put forth that there is a number of *fallacies* underlying the GDPR: it focuses too much on the concept of informational self-determination, it puts too much faith in controllers to perform certain actions, and it attempts to regulate developments like behavioural advertising and profiling that require their own kinds of regulation.[88]
- Zarsky claims that the GDPR is *incompatible* with "the data environment that the availability of big data generates", which could either lead to the Regulation's irrelevance or to making big data analysis "suboptimal and inefficient."[89]

Considering the possible impact of big data on individuals and societies discussed in sections 1.3–1.4, data protection law should be future proof, free from obvious fallacies and compatible with both its social and technological contexts. The above criticisms therefore give rise to the following question:

To what extent does the GDPR reflect or employ theories of power relations and risk management presented by Komesar, Barnett and Duvall, Beck, Perrow, Klinke and Renn, and complex systems science?

The question is approached through the following sub-questions:

- How do the decision-making mechanisms in the GDPR itself, and in the EU lawmaking process that produced the GDPR, compare to other available decision-making mechanisms with regards to opportunities for effective participation by data subjects?
- How do the GDPR's protections for data subjects giving consent or entering into a contract compare to the protections in EU consumer protection law?
- To what extent were existing insights from the social sciences and environmental law applied in the GDPR insofar as it deals with the identification of risks of big data or with the addressing of new or unknown risks?

---

87 Lokke Moerel, *Big Data Protection: How to Make the Draft EU Regulation on Data Protection Future Proof. Oratie 14 Februari 2014* (Tilburg University 2014) 51–54.

88 Bert-Jaap Koops, 'The Trouble with European Data Protection Law' (2014) 4 International Data Privacy Law 250, ss II–IV <https://academic.oup.com/idpl/article-abstract/4/4/250/2569063/The-trouble-with-European-data-protection-law> accessed 20 March 2019.

89 Tal Zarsky, 'Incompatible: The GDPR in the Age of Big Data' (2017) 47 Seton Hall Law Review 995, 996 <http://scholarship.shu.edu/shlr/vol47/iss4/2>.

- Is the GDPR's protection of sensitive personal data adequate in the context of big data and relevant insights in the field of Complex Systems Science?

1.7.1 Delineation

Geographically, this research deals primarily with the European Union. The treaties underlying the institutions and the workings of the Union, secondary EU law, jurisprudence of the Court of Justice, but also the European Convention on Human Rights and the case law of the European Court of Human Rights, are the foundations of the EU legal order and therefore count as primary sources. Additionally, other treaties and Member States' domestic law and jurisprudence will be referenced where appropriate. However, the subject matter of the question implies that developments outside of the EU can be of significance: they will be included where relevant.

The primary focus is on the processing of personal data based on the necessity for the performance of a contract and on consent. Observations are mostly limited to the private and consumer context and the provisions of Chapters I to III of the GDPR (General provisions principles and rights of the data subject). The processing of personal data (including profiling) based on the need to comply with a legal obligation, the vital interest of the data subject or the legitimate interest of the controller will not be covered: this mostly excludes use cases from the administrative law and criminal law contexts from the scope of this work. The specific processing situations of chapter IX (*e.g.*, freedom of expression, employment and archiving) are not covered as they have only limited relevance to the consumer context. This research also excludes the provisions specifically regarding the consent of minors and the specific national provisions on the capabilities of minors to enter into contracts.

Provisions pertaining to the obligations of controllers and processors towards each other and towards supervisory authorities, as well as the provisions regarding transfers of personal data to third countries and the authority of supervisory authorities and their cooperation and consistency are not covered in depth for the same reason, although they can be mentioned in passing.

This research does only occasionally identify differences between platform providers and non-consumer end users of a platform. Even though platform providers in two-sided markets play an essential role in the development of datafication and the effects of big data, the GDPR does not distinguish platform providers from other types of controllers. Both a platform provider and the non-consumer end users of the platform

tend to count as controllers in the sense of article 4(7) of the GDPR, especially if they have separate contracts with the consumer. Also, consumer contracts are held to the same legal standards, regardless of whether the other party is a platform provider or not.

1.8 Methodology

This section accounts for the relevance of the proposed models and their application in this thesis, and describes how they will be used in the analysis of the GDPR in the following chapters. A more comprehensive overview of the relevant elements of these models is presented in the relevant chapters.

Because a large part of the GDPR is outside the scope of the research, a complete overview of GDPR provisions is omitted. Where necessary, reference is made to the relevant handbooks published by the European Agency for Fundamental Rights and the Council of Europe.[90]

1.8.1 Komesar's theory of Comparative Institutional Analysis

The effects of power differentials between the individual and the government and, more generally, between the "haves and the have nots"[91] have been moderated to various extents extent in the political systems and the economies of modern nations. In the social democracies typical for the European Union, application of the principles of the *Rechtsstaat* has led to the emergence and regulation of large-scale decision making processes, specifically the legislative process, the market and the courts. These processes – or *institutions* – can redistribute power through general principles (like "one man, one vote" or "equality before the law") as well as through more focused instruments like consumer protection law, competition law or forum

[90] European Union Agency for Fundamental Rights, European Court of Human Rights and Council of Europe, *Handbook on European Data Protection Law* (Publications Office of the European Union 2014); European Union Agency for Fundamental Rights and European Court of Human Rights, *Handbook on European non-discrimination law* (Publications Office of the European Union 2011) <http://fra.europa.eu/sites/default/files/fra_uploads/1510-fra-case-law-handbook_en.pdf> accessed 19 March 2019.

[91] Marc Galanter, 'Why the Haves Come out Ahead: Speculations on the Limits of Legal Change' (1974) 9 Law & Society Review 95 <https://www.jstor.org/stable/3053023> accessed 13 February 2019.

choice rules.[92] The goal of these processes can be manifold, but achieving a more level playing field for participants wielding comparatively little power when compared to governments or large corporations is often one of them.

The creation or recognition of institutions then gives rise to the question of *institutional choice*: which decision-making process should handle a certain class of decisions by default? In this book, Komesar's method of *comparative institutional analysis*[93] will be used to compare the effect of different institutional choices on the opportunities for effective consumer participation in decision-making regarding contracts that involve the processing of personal data. After all, where opportunities for effective participation are reduced for one class of participants, power differentials can be expected to increase in favour of the party whose opportunities for effective participation are greater.

Komesar proposes a comparative analysis by considering and comparing the *dynamics of participation*, or the costs and benefits of participating in each decision-making process. The analysis offers an opportunity to estimate which decision-making processes are more prone to favour the special interests of the few, and which processes are more suitable to favour the interests of the many. The analysis accounts for variables such as the costs of information, the costs of organisation and the height and the distribution of the stakes that participants have in the outcome of the decision-making process for each individual decision.

The relevance of this model for analysis of decision-making processes lies in article 6 of the GDPR, enumerating the grounds for lawfulness of the processing of personal data. Article 6(1)(a–b) states that processing of personal data can be lawful if a data

92 Note that the term 'institutions' can refer to the 'forms, outcomes and dynamics of economic organisation', the 'rules of the game in a society' and 'legal systems, political systems' and other form of organisation. Morgan, Glenn and others, 'Introduction' in Glenn Morgan and others (eds), *The Oxford Handbook of Comparative Institutional Analysis* (Oxford University Press 2010) 2; Masahiko Aoki, *Toward a Comparative Institutional Analysis* (1st edition, The MIT Press 2001) 1; Douglass C North, *Institutions, Institutional Change and Economic Performance* (Cambridge University Press 1990) 4 includes among institutions "any form of constraint that human beings devise to shape human interaction".

93 Neil K Komesar, 'Governance, Economics and the Dynamics of Participation' in Neil Komesar and others (eds), *Understanding global governance : institutional choice and the dynamics of participation* (European University Institute 2014); Neil K Komesar, *Law's Limits: The Rule of Law and the Supply and Demand of Rights* (Cambridge University Press 2001).

subject gives consent or if the processing is "necessary for the performance of a contract"; article 6(1)(c) and 6(1)(e) state that processing is lawful if it is "necessary for compliance with a legal obligation" and if it is "necessary for the performance of a task carried out in the public interest or in the exercise of official authority vested in the controller." In other words, often a choice exists between providing lawfulness through decisions made in the market or in the legislative process, and in cases where no choice is made, courts will eventually decide. In this context, comparative institutional analysis and considering the dynamics of participation in the market, the political process and the courts can offer insights in the effects of developments like the emergence of big data and the widespread use of personal devices that use telecommunications to send personal data of large groups of consumers to controllers.

1.8.2 Barnett and Duvall's theory of power in social relations

According to Barnett and Duvall, "power is the production, in and through social relations, of effects that shape the capacities of actors to determine their circumstances and fate."[94] As was mentioned earlier, many controllers of personal data can effectively keep data subjects under surveillance, which increases their power over these data subjects. This is observed in its most visible form in social network services, but many other types of commercial contracts can offer similar insights in a data subjects' personal life and use the results of knowledge creation in commercial enterprise.

A power differential opens an avenue for *unfair treatment* of the entity that holds less power. This risk of unfairness is a rationale for EU consumer protection law.[95] For example, if a commercial contract meets the criteria for an unfair commercial practice as defined in article 5(2), Unfair Commercial Practices Directive and falls within its scope, it is prohibited according to article 5(1) of that Directive.[96] However, in jurisdictions where the *lex specialis* doctrine is prevalent, an argument could be made

94 Michael Barnett and Raymond Duvall, 'Power in International Politics' (2005) 59 International Organization 39, 42 <http://www.jstor.org/stable/3877878> accessed 13 February 2019 (paraphrasing John Scott's 2001 work 'Power').

95 Stephen Weatherill, *EU Consumer Law and Policy* (Edward Elgar Publishing 2013) 93.

96 European Parliament and Council Directive 2005/29/EC of 11 May 2005 concerning unfair business-to-consumer commercial practices in the internal market and amending Council Directive 84/450/EEC, European Parliament and Council Directives 97/7/EC, 98/27/EC and 2002/65/EC and European Parliament and Council Regulation (EC) No 2006/2004, [2005] OJ L 149/22 (Unfair Commercial Practices Directive).

that unfairness concerning the processing of personal data should not be judged against consumer protection criteria, but exclusively against the data protection provisions that have their own context-specific definition of fairness.[97] The same case could be made if the processing falls outside the scope of the Unfair Commercial Practices Directive.

Even though Barnett and Duvall wrote their article with international politics in mind, their framework has been applied to power relations between citizens and firms before.[98] In this research, Barnett and Duvall's theory provides a framework for evaluating and qualifying the power differentials resulting from processing, and a comparison between the expected results of fairness provisions in two directives containing EU consumer protection law and the GDPR.

1.8.3 Beck's theory of the risk society

In 1986, Ulrich Beck proposed that the nature of risk had changed in modern times due to technological progress. He asserted that technology had transformed from a way to reduce hazards into a source of possible disaster *e.g.*, in the form of nuclear technology and persistent poisons. Where during the "first modernity", societies were mainly concerned with the distribution of wealth, in second modernity it was also concerned with the distribution of risk. At the same time, individualisation of social inequality reduced the efficacy of collective decision-making.[99]

Beck's theory of risk has been subject to valid criticisms, mainly because it views risks in terms of worst-case scenarios and because it does not consider individual choice and risk acceptance.[100] In the context of this research, these criticisms are considered to be of limited relevance. First of all, the processing of personal data is not seen as a harbinger of impending catastrophe. Instead, it is assumed that big data be seen as a net gain to society, but having possible adverse side-effects that merit consideration

97 European Union Agency for Fundamental Rights, European Court of Human Rights and Council of Europe (n 90) 76–78.

98 Andreas Dür and Dirk De Bièvre, 'The Question of Interest Group Influence' (2007) 27 Journal of Public Policy 1, 6 <https://www.cambridge.org/core/product/identifier/S0143814X07000591/type/journal_article> accessed 17 May 2019.

99 A more complete overview is presented in chapter 4 below. Ulrich Beck, *Risikogesellschaft. Auf Dem Weg in Eine Andere Moderne* (1st ed., Suhrkamp Verlag 1986).

100 Gabe Mythen, *Ulrich Beck: A Critical Introduction to the Risk Society* (Pluto Press 2004) 180–182.

and prevention measures. Secondly, this research focuses on the possibilities for data subjects to exercise their own agency. Thirdly, these criticisms, valid as they are, do not diminish the relevance of risk society theory in the context of consumer contracts and consent. As it turns out, the development of EU environmental protection legislation shows some signs of the adoption of some of Beck's core notions. Together with Perrow's normal accident theory, they provide a useful frame of reference for evaluating legislation that deals with technological risks.

1.8.4 Perrow's theory of normal accidents

In 1984, Charles Perrow proposed that systems that are difficult to intuitively comprehend and that are highly time-sensitive to escalation in the event of failure could suffer *system accidents*. A system accident occurs when timely recovery from a partial failure is so time-sensitive that it can lead to catastrophic breakdown of an entire system. In his opinion, system accidents could occur in high-risk systems such as nuclear and chemical plants. They could be so inherent to these systems that they could be called *normal accidents*.[101] Perrow proposes a system of *social and cultural rationality*, where societies decide through political discourse whether technologies should be abandoned, restricted, or tolerated and improved.

Perrow's theory of normal accidents has faced considerable criticism regarding its usefulness in understanding and preventing large disasters.[102] In this regard, it has been pitted against High Reliability Theory, among others. In response, Perrow expanded his theory to include power and social relations. In a further development of the theory, he added the idea of *fantasy documents*, being used in decision-making processes leading to the acceptance of high-risk systems. Such documents will claim that system accidents will be virtually impossible; they serve to make the many bear the risk of these systems for the benefit of the few.[103]

---

[101] Charles Perrow, *Normal Accidents: Living with High-Risk Technologies* (Basic Books 1984).

[102] For an overview, see Andrew Hopkins, 'Discussion: The Limits of Normal Accident Theory' (1999) 32 Safety Science 93 <https://linkinghub.elsevier.com/retrieve/pii/S0925753599000156> accessed 16 May 2019.

[103] Charles Perrow, 'The Limits of Safety: The Enhancement of a Theory of Accidents' (1994) 2 Journal of Contingencies and Crisis Management 212 <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1468-5973.1994.tb00046.x> accessed 16 May 2019; Charles Perrow, 'Accidents, Normal', *International Encyclopedia of the Social & Behavioral Sciences* (Elsevier 2001) <http://linkinghub.elsevier.com/retrieve/pii/B0080430767045095> accessed 13 February 2019.

The usefulness of Perrow's theory, as expanded, lies not so much in its ability to predict and avert disasters as in its framework for looking at who decides that risks are acceptable and why. Indeed, EU environmental law instruments as well as national environmental protection law in Member States have implemented several aspects of Perrow's model of risk evaluation and control.

The relevance of the models proposed by Beck and Perrow can be seen as follows: like the risks addressed by environmental protection law, datafication has its origin in technological progress. Furthermore, datafication potentially affects not only individuals but also societies, and if societies want to address these risks, the qualities of decision-making processes and the properties of control mechanisms become relevant. In this research, Beck's and Perrow's theories are the starting point for a two-part comparative analysis between the GDPR and EU environmental protection legislation.

1.8.5 Klinke and Renn's approach to risk evaluation and management

Klinke and Renn have proposed a classification of technological risks based on whether the risk potential of technology is known, whether the damage potential is known, whether the disaster potential is high and whether social mobilisation on the technology is high. Based on the result of this classification, a risk management strategy could then be based on quantitative risk analysis, guided by an attitude of precaution, or negotiated in discourse.

Alternatives to Klinke and Renn's model have been proposed, but these seem to have left the basic premise of their model intact. Kristensen et al, for example, have proposed a more refined method for arriving at a risk management strategy but the list of characteristics that of risky technologies appears to span a similar space as Klinke and Renn's. The same counts for the management strategies that they offer.[104] The validity of the approach by Klinke and Renn therefore appears to be hitherto essentially undisputed.

Klinke and Renn's model is used in to determine what is the nature of the risk that the drafters of the GDPR aimed to address, and whether the risk management strategy matches the implicit risk assessment. This analysis is performed separately for the

[104] V Kristensen, T Aven and D Ford, 'A New Perspective on Renn and Klinke's Approach to Risk Evaluation and Management' (2006) 91 Reliability Engineering & System Safety 421, 422, 428 <https://linkinghub.elsevier.com/retrieve/pii/S0951832005000785> accessed 16 May 2019.

general case of consumer contracts, and for the special case of the processing of sensitive personal data as defined in article 9(1) GDPR.

1.8.6 Complex systems science

Complex systems science (or *complexity theory*) considers the similarities between seemingly unrelated observable phenomena. It has been observed that some types of events, such as nonlinear response to relatively small inputs, and the emergence of spontaneous order, can occur in many types of systems like individual organisms, cells, rainforests and weather systems. The shared property of these systems is called *complexity*, and the aforementioned classes of events can be described or predicted using similar techniques across disciplines.

Due to the widespread use of individual electronic devices with telecommunications capabilities, occurrences like spontaneous order in groups of people can now be distilled from a distance by centralised analysis of personal data emitted by these devices. This opens new possibilities for knowledge creation, for example through covert or semi-covert surveillance. If this knowledge pertains to sensitive traits as defined in article 9(1) of the GDPR, it can trigger prohibitions on the processing of sensitive personal data and automatic decision-making including profiling based on sensitive data (article 22(4)).

Complexity theory is used to evaluate the expected efficacy of articles 9 and 22 of the GDPR, regarding the processing of sensitive personal data and automated decision-making (including profiling) based on the processing of sensitive data.

1.9 Structure

Chapters 2, 3, 4 and 5 consist of four original research articles that were accepted by peer-reviewed journals between November 2014 and April 2018. These four chapters each deal with one of the sub-questions, in order, and contain their own conclusions as presented at the time of publication.

These articles are presented essentially as they were published, and in the same chronological order. The edits to the chapters have been kept to a minimum. The article appearing as chapter 2 was published before the GDPR appeared in the Official Journal of the European Union. Therefore, the references to GDPR articles in this chapter have been updated for consistency. Other edits are limited to the correction of misspellings and grammatical errors, the application of a uniform scheme for

references in footnotes and the bibliography, the re-numbering of the footnotes, updating the uniform resource locators (URLs) in the footnotes – identifiable by an "accessed" date more recent than the original publication date for almost all URLs) – and the combination of the published bibliographies into a single bibliography at the end of this book. References to articles that appear as chapters in this book have been converted to internal references.

Chapters 4 and 5 were published in journals that prohibited the referencing of the authors' own works, and each article was aimed to be self-contained. The conclusions of each chapter in this book do therefore not always refer to earlier chapters. The chapters' preludes illustrate the timeline and the context in which each article originated: they are not part of the research.

Chapter 5 was co-authored with dr. Qing Yi Feng from Utrecht University. Dr. Feng kindly provided section 5.2 (on emergence); Dr. Feng and I co-wrote sections 5.7 and 5.8; I provided the first drafts for these sections.

The final chapter answers the research question. Additionally it merges the conclusions of the four preceding chapters to identify underlying issues, propose remedies and suggest further research. Due to the fact that it was written after the completion of the preceding chapters, it occasionally contains pointers to works not previously cited. The final chapter also contains some additional insights gained during the research project that did not find their way into any of the articles due to word limits, scope restrictions or timing.

2 Big data and consumer participation in privacy contracts

Deciding who decides on privacy

Prelude

A short remark in Guibault and others, *Digital Consumers and the Law. Towards a Cohesive European Framework* (Kluwer Law International 2012) inspired me to write the text that is now included in this book as Chapter 2. Page 144 of Digital Consumers and the Law reads: "Given the relatively small (and costly) scope of the judicial process, further legislative action may serve to strengthen the position of consumers of digital content." The accompanying footnote referred to Komesar's method of Comparative Institutional Analysis and his 2001 book *Law's Limits*.

The remark made me curious: I could readily believe that the position of digital consumers were better strengthened through law than trough court cases. But how did the footnote support this? It turns out that the way that Komesar looks at decision-making processes can help point out strengths and weaknesses in the GDPR.

Timeline and citation

The original article was published in the Privacy special issue of the Utrecht Journal of International and European Law on 27 February 2015 under a "Creative Commons Attribution 3.0 Unported" open-access license. The call for papers was published on 16 July 2014 and is archived at https://www.utrechtjournal.org/announcement/.

The original version was submitted on 16 November 2014. The Utrecht Journal conditionally accepted the first version on 21 December. The revised version was submitted 10 January 2015 and was accepted 11 February. The Utrecht Journal recommends the following citation:

> Rhoen, M., (2015). Big Data and Consumer Participation in Privacy Contracts: Deciding who Decides on Privacy. Utrecht Journal of International and European Law 31(80), pp. 51–71. DOI: http://doi.org/10.5334/ujiel.cu

2.1 Introduction

In late spring of 2011, Dutch Parliament debated the transposition of the revised EU Telecoms Package into national law.[105] This debate became the centre of public interest after telecom provider KPN proudly explained to its investors that they were ready to start using Deep Packet Inspection (DPI) to see which applications generated data traffic over their wireless network. Marco Visser stated: 'We will not block services but [...] we will price them'.[106] Packets in this context are units of data at the network layer level of telecommunications. Each packet consists of control information (containing, among others, the origin and destination of the packet) and user data (the actual data being sent). DPI involves analysing data packets for both their control information and their user data.

Responding to declining Short Message Service (SMS) revenues, KPN announced the company had the intention to use this DPI technology to charge for the use of instant messaging (IM) applications on smartphones (apps). Use of IM apps (like WhatsApp) substituted consumers' use of individually priced SMS messages. This dramatically reduced the profits of KPN (and later other telecom providers all over the world).[107] DPI would allow KPN to reverse this trend, as it enabled the company to distinguish IM traffic from other traffic, and charge a higher price for the IM services.

105 European Parliament and Council Directive 2009/140/EC of 25 November 2009 amending Directives 2002/21/EC on a common regulatory framework for electronic communications networks and services, 2002/19/EC on access to, and interconnection of, electronic communications networks and associated facilities, and 2002/20/EC on the authorisation of electronic communications networks and services, [2009] OJ L 337/, p. 37–69 (the Framework Directive); European Parliament and Council Directive 2009/136/EC of 25 November 2009 amending Directive 2002/22/EC on universal service and users' rights relating to electronic communications networks and services, Directive 2002/58/EC concerning the processing of personal data and the protection of privacy in the electronic communications sector and Regulation (EC) No 2006/2004 on cooperation between national authorities responsible for the enforcement of consumer protection laws, [2009] OJ L 337/, p. 11–36 (the Rights Directive); and European Parliament and Council Regulation (EC) No 1211/2009 of 25 November 2009 establishing the Body of European Regulators for Electronic Communications (BEREC) and the Office, [2009] OJ L 337/1, p. 1–10.

106 Marco Visser, 'KPN Investor Day: Consumer Wireless. Strengthen - Simplify - Grow' (KPN Investor Day, London, 10 May 2011) 6 <https://ir.kpn.com/download/companies/koninkpnnv/Presentations/KPN_Investor_Day_-_Selective_topics.pdf> accessed 21 March 2019 (video no longer available).

107 Dawinderpal Sahota, 'Global SMS Revenue Declines for First Time' (*Telecoms.com*, 14 January 2014) <http://telecoms.com/212062/global-sms-revenue-declines-for-first-time/> accessed 20 March 2019.

Consequently, KPN would no longer provide a "neutral" network, understood here as a network that treats all types of traffic equally, but would be able to distinguish between different origins of traffic (i.e. different apps) and use different policies for the corresponding services. For this purpose, KPN would analyse the user data within packets sent across its network, as well as their control information – hence the name *Deep* Packet Inspection. This is, to some extent, analogous to the postman reading a letter to see whether it is priority mail, instead of looking at the indication and stamps on the envelope. Although imperfect in many ways, this analogy indicates the privacy implications of DPI.

2.1.1 Parliament steps in

KPN's announcement brought the net neutrality debate to Dutch Parliament. Essentially, net neutrality is about control: are network operators allowed to 'block [...] or prioritise [...] certain network traffic or traffic from particular sources'?[108] Should KPN be allowed to charge its subscribers a premium for their use of specific applications?

Several non-governing minority parties proposed an amendment to the Bill implementing the revised Telecoms Package, demanding network neutrality from all telecommunications providers and specifically prohibiting the analysis of traffic by content other than for technical reasons (e.g. ensuring network integrity or security). The proposed amendment was meant to secure the possible benefits of a neutral network,[109] but also to secure consumers' privacy. DPI, these parties argued, gave telecommunications providers an unhealthy degree of insight into consumers' private communications, since it must include analysis of their content.[110]

Initially, the cabinet minister responsible for the Bill opposed the amendment. In his view, telecoms law already provided safeguards against DPI. For example, it prohibited telecom providers from secretly 'limiting access to and/or use of services' as well as 'procedures to measure and shape traffic': providers could use DPI for price discrimination only if they told consumers beforehand.[111] If consumers objected to a

[108] Paul Ganley and Ben Allgrove, 'Net Neutrality: A User's Guide' (2006) 22 Computer Law & Security Review 454, 457 <http://www.sciencedirect.com/science/article/pii/S0267364906000902> accessed 19 March 2019.

[109] ibid 461.

[110] KST II 2010-2011, 24095 nr. 285, p. 8, 28 (Dutch Parliamentary documents).

[111] Rights Directive, article 1(14).

provider's use of DPI, they could choose another provider. Imposing net neutrality, he contended, could make the Dutch market less attractive to investors because it would close an avenue of revenue maximisation, left open in the Telecoms Package. He also suggested that BEREC, the European body of regulators,[112] was better suited to regulate net neutrality than the national legislator. A specific requirement within the Netherlands might undo the EU efforts at harmonising the internal market for telecom services. In the end however, after some debate, the minister accepted the proposed amendment and Parliament adopted the amendment, thereby including the net neutrality obligation in the Dutch Telecommunications Act.[113]

2.1.2 Consumer participation options for privacy contracts

The network neutrality debate hints at a wider privacy issue: nowadays, consumer contracts for everyday services allow private parties to collect and use large quantities of data. This data identifies individuals, for example by making use of cookies, e-mail addresses, shipping addresses, device identifiers (and other hardware properties), subscriber information, account numbers or unique tokens like loyalty cards. The term "data" can describe the contents of communications but also traffic data or metadata – "data about data", e.g. timestamps and location identifiers. All this data can reveal many aspects of individuals. If, and to the extent that, such data is about identified or identifiable individuals, it qualifies as personal data as per article 2(a) of the Data Protection Directive and article 4(1) of the GDPR, which implies that the processing of the data has to comply with national law based on said instruments and, since the GDPR has become applicable, with the GDPR itself.[114]

Usually, these consumer contracts offer benefits to consumers that are unrelated to their personal data. Sometimes, these contracts are unavoidable for the consumer: everybody needs banking, telecommunications and public transport. Sometimes consumers enter into these contracts to obtain a side benefit to another transaction (e.g. using a loyalty program to obtain a rebate from a retailer). Sometimes the

112 BEREC is established by Regulation no. 1211/2009, art 1(1).

113 Article 7.4a Telecommunicatiewet, as amended by Act of 10 May 2012, Stb 2012, nr. 235 (Dutch National Journal); KST II 2010-2011, 24095 nr. 285, p. 29 (Dutch Parliamentary documents); Handelingen II, 8 June 2011, nr. 90, item 3, p. 90-3-36 (Dutch Parliamentary proceedings).

114 European Parliament and Council Directive 95/46/EC 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, [1995] OJ L 281/, p. 31–50 (Data Protection Directive).

exchange of personal data for services is an essential performance of the contract (e.g. when using a social network site) and sometimes the data only has value for the party collecting it, for example when a web site uses tracking cookies to follow the behaviour of every website visitor.

The ubiquity of these contracts is the result of the increased datafication[115] of daily life and the increasing sophistication of algorithms for analysis. This has made such data valuable to an increasing number of businesses. Often the aim is to retain data indefinitely in order to create an enduring profile of the individuals, for example to analyse (online) shopping habits to determine whether a consumer is pregnant.[116] These and similar developments have spawned the phrase "big data" to indicating the volume, variety and velocity of the data streams.[117] Obviously big data can have privacy implications.[118]

In this research, the term 'privacy contracts' is used for all varieties of such contracts.[119] Privacy contracts can take the form of written two-party agreements, but they are usually "agreed upon" by means of non-negotiable terms and conditions or unilateral privacy statements on websites. These contracts often govern services that consumers cannot easily do without like telecommunications, banking or grocery shopping, or that are increasingly a part of modern life like being part of a social network or using household appliances like smart TV's.[120]

Consumers have reduced opportunities for participation in determining the contents of privacy contracts: they cannot voice their opinion on the contents, the contents are not influenced by their opinions, and not entering into the contract is often not an option. KPN's announcement to start pricing IM apps by using DPI increased awareness among consumers and legislators that contracts about the use of

115 "the ability to render into data many aspects of the world that have never been quantified before": Viktor Mayer-Schönberger and Kenneth Cukier, 'The Rise of Big Data: How It's Changing the Way We Think about the World' (2013) 92 Foreign Affairs 28, 29.

116 Charles Duhigg, *The Power of Habit: Why We Do What We Do in Life and Business* (Random House Trade Paperback Edition, Random House Trade Paperbacks 2014) 194–195; 209–210.

117 Zikopoulos and Eaton (n 7) 5–8; Mayer-Schönberger and Cukier (n 1).

118 Neil M Richards and Jonathan H King, 'Three Paradoxes of Big Data' (2013) 66 (2013) Stanford Law Review Online (passim) and referenced literature.

119 Eric W Verhelst, *Recht Doen Aan Privacyverklaringen: Een Juridische Analyse van Privacyverklaringen Op Internet* (Kluwer 2012) ch 3.

120 Walter Peissl, 'Information Privacy in Europe from a TA Perspective' in Serge Gutwirth, Yves Poullet and Paul De Hert (eds), *Data protection in a profiled world* (Springer 2010) 251.

telecommunication services are indeed privacy contracts, i.e. contracts pertaining to the use of personal data and its privacy implications.

The Dutch Parliamentary net neutrality debate compared two large-scale decision-making processes that could decide on the terms of privacy contracts: the market will and the political process. Komesar calls these large-scale processes institutions; he distinguishes the market, the political process and the courts.[121] Seen from this perspective, the case of DPI is an example of institutional choice. This can be illustrated by the subsequent (and partly hypothetical) stages in the net neutrality discussion:

- First stage: after intense lobbying by interest groups including the telecommunications industry,[122] the aforementioned revised Telecoms package explicitly left net neutrality decisions to the market. If too many consumers refuse to enter into a contract with KPN, another provider will offer a better deal or KPN will change their offering.
- Second stage: In the Netherlands, a number of interest groups and Parliament felt that users could be forced or misled to agree to DPI contracts. For them, DPI was equal to permanent unwarranted eavesdropping on private communications.[123] They claimed these decisions should be made in the political process. The permissible terms of all privacy contracts between subscribers and telecoms providers were changed as a result.[124]
- Third (hypothetical) stage: Had both the market and the political process failed to sufficiently protect consumer privacy, a consumer would be entitled to involve the national courts. Article 8 of the European Convention of Human Rights (ECHR) guarantees the right to respect for one's 'private and family life, his home

121 Komesar, *Law's Limits* (n 93) 31.

122 Yana Breindl, 'Promoting Openness by "Patching" European Directives: Internet-Based Campaigning during the EU Telecoms Package Reform' (2011) 8 Journal of Information Technology & Politics 346, 354 <http://dx.doi.org/10.1080/19331681.2011.595326> accessed 13 February 2019.

123 Handelingen II, 8 June 2011, nr. 90, item 3, p. 90-3-36 (Dutch Parliamentary proceedings); Daphne van der Kroft, 'Persbericht: Bits of Freedom Roept KPN-Abonnees Op Om Aangifte Te Doen Tegen Aftappen' (12 May 2011) <https://www.bof.nl/2011/05/12/persbericht-bits-of-freedom-roept-kpn-abonnees-op-om-aangifte-te-doen-tegen-aftappen/> accessed 13 February 2019.

124 Similarly: Lucie MCR Guibault and others, *Digital Consumers and the Law. Towards a Cohesive European Framework* (Kluwer Law International 2012) 144.

and (...) correspondence'. Article 13 of the Convention requires an effective remedy before a national court for violations of this right. However, a verdict from a court would typically affect only the contracts to which the litigants are a party.

Different institutions might offer different outcomes. Evidently, the market, the political process and the courts all offer different opportunities for effective consumer participation: the market allows for negotiations on particular contract terms; consumers can elect legislators and authorise them to impose rules to govern all contracts; courts settle disputes between consumers and their contract partners (such as telecoms providers).

Participation opportunities for consumers are of interest to privacy contracts because participation is one of three traditional legitimacy requirements, together with transparency and accountability, for an act that affects a fundamental right.[125] A society, such as a State or the European Union, can deliberately choose which institution decides on privacy contracts. A clear choice of institution promotes efficient decision-making by increasing legal certainty and pre-arranging channels for dispute resolution and business development.

2.1.3 Consumer participation as a question of Institutional Choice

Consumer participation takes different forms in different institutions: consumers can act as citizens, voters, litigants and participants in consumer interest groups. They are the holders of specific protections in the Charter of Fundamental Rights of the European Union (the Charter).[126] The term "consumers" is used throughout this chapter, to easily distinguish them from their contract partners, indicated as "producers". Producers are providers of goods and services and the controllers and

[125] These three elements of due process are generally believed to be necessary to provide legitimacy to any legal act that affects a fundamental right. For examples, see Danielle Keats Citron, 'Technological Due Process' (2007) 85 Wash. UL Rev. 1249, 1256–1257; note the similarity between these safeguards and the concept of due process in United States law: Serge Gutwirth and Paul de Hert, 'Een Theoretische Onderbouw Voor Een Legitiem Strafproces. Reflecties over Procesculturen, de Doelstellingen van de Straf, de Plaats van Het Strafrecht En de Rol van Slachtoffers'' (2001) 31 Delikt & delinkwent 1048, paras 12–13 <http://www.vub.ac.be/LSTS/pub/Gutwirth/006.pdf> accessed 13 February 2019 ("Theoretical underpinnings for legitimate criminal procedure. Reflections on process cultures, the aims of punishment, the place of criminal law and the role of victims").

[126] Art. 38, Charter of Fundamental Rights of the European Union, [2010] OJ C 83/02.

processors of personal data within the market,[127] whilst in other institutions they can operate as lobbyists or litigants.

Certain participation possibilities may be more desirable than others, depending on the policy objectives a society wants to achieve, and the level of privacy protection it wishes to offer. In this context, the following question becomes relevant:

When deciding on privacy contracts in the age of big data, how does institutional choice affect consumer participation opportunities and how does it affect the feasibility of policy objectives in the European multilevel jurisdiction?

The question will be addressed by answering the following sub-questions:

1. Why does institutional choice matter for privacy protection?
2. How do the possibilities of participation for consumers and producers qualitatively compare between institutions, if decisions on consumer privacy were to be left to the market, the political process or the courts, respectively?
3. What does the analysis imply for different policy objectives concerning the impact of big data on society?

The scope of this chapter is restricted to privacy contracts in which one party qualifies as a consumer.[128] This implies that criminal or national security investigations, employment relationships and torts are not covered. Further, since the comparison of the effectiveness of institutions is not directly dependent on substantive law, aspects of substantive law are not addressed. Moreover, as data protection and privacy protection are not always easily distinguished and sometimes used interchangeably, this chapter distinguishes between both concepts only when this is required for the subject at hand.[129]

---

127 Art. 2(d-e), Data Protection Directive.

128 As defined by article 2(b) of Council Directive 93/13/EEC of 5 April 1993 on unfair terms in consumer contracts, [1993] OJ L95/29 (Unfair Terms Directive), art 2b:, "any natural person who, in contracts covered by this Directive, is acting for purposes which are outside his trade, business or profession".

129 DLA Piper, 'Part 4: The Future of Online Privacy and Data Protection' (European Union 2009) SMART 2007/0037 s 1.1.3 <http://ec.europa.eu/newsroom/dae/document.cfm?doc_id=833> accessed 13 February 2019; Gerrit-Jan Zwenne, *Diluted Privacy Law* (Leiden University 2013) 12 <http://papers.ssrn.com/abstract=2488486> accessed 13 February 2019.

2.1.4 Comparative Institutional Analysis – Methodological notes

This chapter answers the questions set out above by performing a comparative institutional analysis between the political process, the market and the courts at both national and European Union levels. The comparison focuses on each institution's possibilities for providing effective participation opportunities to consumers.

Comparison between institutions is possible if the same individuals participate in all compared institutions. In the case of big data and privacy contracts, we assume that buyers and sellers, or consumers and producers, or litigants before the courts and voters and lobbyists are indeed members of the same mass of people. Comparison is useful, even though no single institution can be expected to perform perfectly. Institutional performance deteriorates when the number and complexity of the required decisions increase. In these cases, 'institutions tend to move together'.[130] Even the best available institutional option may leave much to be desired.

In the model offered by Komesar, the essence of institutional comparison lies in comparing the incentives that drive the actions of the mass of participants in these institutions (consumers, producers, litigants, voters, lobbyists). He calls this the dynamics of participation. These dynamics are determined by a simple comparison of costs and benefits.

- The benefits of participation are dependent on the distribution of stakes at play for the participants. This distribution is determined by the average per capita stakes within the population and by the extent to which the stakes vary within the population.
- The costs of participation are the costs of information and the costs of organising collective action. Depending on the institution, participation costs are known as transaction costs, litigation costs or political participation costs. In the model of regulatory capture offered by Levine and Forrence, they can also include monitoring costs.[131]

[130] Komesar, *Law's Limits* (n 93) 23, 28.

[131] Michael E Levine and Jennifer L Forrence, 'Regulatory Capture, Public Interest, and the Public Agenda: Toward a Synthesis' (1990) 6 Journal of Law, Economics, and Organization 167, 171 <http://jleo.oxfordjournals.org/cgi/doi/10.1093/jleo/6.special_issue.167> accessed 13 February 2019.

2.1.5 Structure of this chapter

The relevance and the specifics of the application of comparative institutional analysis to the research question are addressed in sections 2.2-2.4. Sections 2.5-2.9 contain the actual analysis. Sections 2.9-2.10 then apply the outcome of this analysis to a number of possible policy objectives. The final section contains some concluding remarks.

2.2 How institutions matter for consumer privacy

To effectively make use of a legal right, it must be reasonably achievable. If two parties' interests are not fully aligned and a decision is required, a rule of substantive law usually cannot fully provide its' intended protection because decisions always have a cost. Coase's theory of transaction costs suggests that some parties will not seek a decision if the cost of getting that decision outweighs its benefits.

The level of legal protection offered by substantive law can therefore be expected to be lower if transaction costs are higher.[132] Transaction costs vary within institutions, and they may vary from State to State, causing different levels of legal protection. To prevent these differences from becoming excessive, both the European Union and the European Convention on Human Rights (ECHR)[133] set minimum standards for the effectiveness of legal decision-making processes, at least before the courts.[134]

Transaction costs will rise in every institution if numbers and complexity increase. Therefore, it is of particular relevance that, as a result of datafication, both the number and the complexity of decisions on privacy contracts have increased substantially. The number of decisions is related to the rising number of privacy contracts and the number of transactions generating personal data, as indicated earlier. The complexity of privacy contracts is also increasing, particularly in terms of the technology used for executing them, and the number of parties involved.

[132] 'I'll let you write the substance... you let me write the procedure, and I'll screw you every time.' Regulatory Reform Act: Hearing on H.R. 2327. House Comm. on the Judiciary, 98th Cong. 312 (1983) (United States Parliamentary proceedings, statement of Rep. John Dingell).

[133] European Convention for the Protection of Human Rights and Fundamental Freedoms, Sept. 3, 1953, ETS 5, 213 UNTS 221.

[134] Two examples: Art. 13, ECHR of the Convention (right to an *effective* remedy); CJEU 19 June 1990, Case 213/89 *Q. v Secretary of State for Transport, ex parte: Factortame Ltd and others* (Factortame I), [1990] ECR I-02433, paras. 21-23 (availability of interim relief as a condition for effectiveness).

Technology has increased the complexity of privacy contracts. The fact that technological aspects need not be disclosed to consumers[135] makes it difficult to determine whether a particular use of personal data remains within the boundaries set in the contract, or even to determine those boundaries themselves. The increasing number of parties involved increases complexity, because data will be collected from multiple sources and shared with multiple users. This may result in multilateral exchanges of personal data based on bilateral contracts.

In the market, increasing numbers and complexity can lead to the use of form contracts, eliminating party bargaining options. In the political process, increasing numbers and complexity make any law less likely to be well suited to the majority of transactions. The courts have limited capacity to efficiently provide every market party with the decisions they need. Institutional choice when deciding on privacy contracts is a matter of choosing among 'imperfect alternatives'.[136]

2.3 Everything has a price: privacy analysis by cost and benefit

Coase's theory of transaction costs dictates that if an institution decides on privacy contracts, the value of these contracts plays an important part in the dynamics of participation. This presents a problem for two reasons. Firstly, privacy – which, in the European legal tradition, qualifies as a fundamental right and an aspect of human dignity – may be considered not to have a monetary value, or even to not be suitable to be bought and sold.[137] Secondly, many privacy contracts do not specifically put a monetary value on the personal data portion of the performance. This incompatibility manifests itself at the level of individual transactions (the microeconomic level) and

135 Directive 95/46/EC (Data Protection Directive), Recital 41; Regulation (EU) 2016/679 (GDPR), Recital 63; Sandra Wachter, Brent Mittelstadt and Luciano Floridi, 'Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation' (2017) 7 International Data Privacy Law 76 <https://academic.oup.com/idpl/article/7/2/76/3860948> accessed 13 February 2019.

136 Neil K Komesar, *Imperfect Alternatives: Choosing Institutions in Law, Economics, and Public Policy* (University Of Chicago Press 1997).

137 Corien Prins, 'When Personal Data, Behavior and Virtual Identities Become a Commodity: Would a Property Rights Approach Matter?' (2006) 3 SCRIPT-ed 270, 275 <https://script-ed.org/wp-content/uploads/2016/07/3-4-Prins.pdf> accessed 13 February 2019; Joseph W Jerome, 'Buying and Selling Privacy: Big Data's Different Burdens and Benefits' (2013) 66 Stanford Law Review Online 47, 47–49 <https://www.stanfordlawreview.org/online/privacy-and-big-data-buying-and-selling-privacy/> accessed 13 February 2019.

at the level of privacy as a factor promoting or hindering societal prosperity or wealth (the macroeconomic level).

To facilitate decision-making, law and economic theory attach economic value to privacy in an indirect way. The law solves the microeconomic part of the problem by separating decisions on the legitimate processing of personal data from decisions on privacy. The latter lacks a precise definition. Evaluating conformity with article 8 ECHR requires judicial decisions, which are difficult to achieve, low in volume and come with high transaction costs. The Data Protection Directive and the GDPR simplify this process by allowing a high volume of decisions at a low cost. They achieve this both by making compliance easier for producers (by setting relatively simple standards for the processing of personal data) and by authorising consumers to enter into agreements (art. 7(a-b), Data Protection Directive; art. 6(1)(a-b) of the GDPR). This, by the way, gives two examples of institutional choices made by legislative bodies.

Requiring consumers' agreement enables them to exchange their personal data for services without money changing hands. The business model of Google is a good example. Google provides its users with an e-mail service, online document collaboration, photo and video sharing services, a searchable map of the world and a search engine to the World Wide Web at no cost. Google also generates and populates databases for and from all these activities. Taken together, these are Google's costs. The company then monetises the users' personal data by offering targeted advertising options within its' services to third parties (an example of a two-sided market).[138] The revenues of this operation exceeded the costs by approximately a billion dollars per month in the third quarter of 2013.[139] For consumers, Google's "free" services apparently offer good value. The price they pay is the risk of decreased personal privacy. Many consumers (implicitly) decide that this is a good deal, whether or not they are aware of all the privacy implications of their contracts.

At the macroeconomic level, economic theory provides a concept for comparing costs and benefits. Loss of privacy for individuals or groups can be considered a form of social cost.[140] This cost does not manifest itself in individual transactions and is therefore not suitable for analysing participation in the market. It can, however, be

138 Rochet and Tirole (n 29) 992.

139 Bruce Schneier, '"Stalker Economy" Here to Stay' (*CNN-Opinion*, 26 November 2013) <http://www.cnn.com/2013/11/20/opinion/schneier-stalker-economy/index.html> accessed 13 February 2019.

taken into account in the political process or the courts, for example when considering admissibility of terms, taxation, bona fides or the common good.

2.4 All created unequal: the catalogue of comparisons

This chapter compares markets, the political process and courts at the levels of Member States and the European Union. Institutions at the level of the Council of Europe (CoE) are excluded. They are not open to the same mass of participants – there is no common market in the CoE, and neither its political process nor the European Court of Human Rights (ECtHR) is accessible to consumers or producers participating at the EU or national levels. The ECtHR only handles cases against a State or between States.[141] The officials taking part in the political process of the CoE are not appointed after general elections. The verdicts of the ECtHR do not bind the institutions of the EU, although the EU accepts the ECtHR's interpretation of human rights as its own (article 6, TEU).

Furthermore, national courts cannot be compared to the European Court of Justice. As far as disputes over privacy contracts between consumers and producers are concerned, the latter court is inaccessible to litigants. If litigants reside in the same State, national courts have jurisdiction. If litigants reside in different Member States, the so-called Brussels I regulation decides which national court has jurisdiction.[142] In all other cases, domestic law decides whether the national courts have jurisdiction. If a dispute before a national court requires a uniform interpretation of EU law, the highest national court is required to put the matter before the CJEU (art. 267, TFEU). However, this serves only to interpret EU law, not to decide the case.

Therefore, this chapter only compares:

- At the national level: the internal market, the political process and the courts;
- At the EU level: the internal market and the political process;

[140] Ronald Harry Coase, 'The Problem of Social Cost' (1960) 3 Journal of Law and Economics 1, (passim); Paul Sholtz, 'Transaction Costs and the Social Cost of Online Privacy' (2001) 6 First Monday <http://journals.uic.edu/ojs/index.php/fm/article/view/859> accessed 13 February 2019.

[141] Arts. 33-34, ECHR.

[142] European Parliament and Council Regulation of 12 December 2012 on jurisdiction and the recognition and enforcement of judgments in civil and commercial matters (recast), [2012] OJ L 351/1 (Brussels I Regulation).

- Between the national and the EU levels: the market and the political process.

2.5 National and European institutions compared

The following comparisons assume that privacy contracts do not specify the processing of personal data for monetary compensation as the primary performance of one of the parties. This is probably a safe assumption given permission for collection of personal data is usually part of a contract with a wider scope. Many of these contracts do not involve payment, e.g. the terms of use of social networks or other web services. In those contracts that do involve monetary exchange (telephone contracts, bank accounts), the privacy aspect is usually not the main consideration.

2.6 Privacy contracts at the national level

2.6.1 In the market

The terms and conditions of a privacy contract are usually not negotiable. As a result, the decision of the market is effectively the decision of the party who drafts the contract. In a transaction between a consumer and a producer, this will be the producer.

The dynamics of participation in this decision making process can be described as follows. As for the benefits in terms of monetary exchange, the stakes per contract are roughly equivalent for consumers and producers. The stakes per contract are supposedly symmetrical. Competition in the market will presumably cause the price the consumer pays to be near cost, so the price for the services will only be marginally different between producers for an equivalent level of service. If no money is changing hands, the "price" one consumer pays is close to the marginal cost of operating the service for one more consumer. If the number of consumers is large enough the same is probably true for the increase in profits that a producer can obtain by making the personal data of one more consumer available for personalised advertising services.

The large number of contracts for producers makes their per capita stakes much higher than consumers'. Producers have the combined value of all their privacy contracts at stake. The earlier example of Google's 2013 quarterly profits indicates that these stakes can indeed be considerable.

In this chapter, the distribution of stakes over consumers and producers as a group is assumed to be close to homogenous. For consumers, this assumption is based on the idea that consumers as a group have roughly the same need for these services. For the scope of this chapter, this seems a safe assumption given there is no discernible group of consumers that has a lot more depending on a phone contract, a loyalty card or a Facebook account, than the average of all consumers combined. For producers engaging in privacy contracts, distribution of stakes is assumed to be homogenous as a starting assumption in the absence of sufficient data. Producers' stakes are not easily determined, with some enterprises standing to gain more from privacy contracts than others; depending on their size, their business case and whether they wield specific market power. Determining the per capita stakes and distribution of stakes for producers is further complicated by the recent tendency of actors in traditionally one-sided markets, to make their market into a platform working as a two-sided market.[143] To keep the comparison manageable, this subject is not explored further.

On the costs side, consumers face higher costs of participation than producers do. The costs of information, for example of reading the terms of a privacy contract, can be significant.[144] These terms will tell a consumer whether he will incur an extra cost for using IM apps on his smartphone, or what liberties a service provider reserves for himself in sharing and using personal data. Every consumer incurs this cost for every contract he or she considers entering into. He or she is a one-shotter every time because the terms are different for every contract. A producer on the other hand only has to draft the contract once. The cost of information can thereby be spread out over a large number of contracts, resulting in a very low cost of information per transaction. A producer is the repeat player in the market.[145]

143 Rochet and Tirole (n 29); Nick Jue, 'ING En Het Gebruik van Klantgegevens. Open Brief van ING Aan Haar Klanten' (maart 2014) <https://www.ing.nl/nieuws/nieuws_en_persberichten/2014/03/ing_en_het_gebruik_van_klant_brief.html> accessed 13 February 2019 (ING and the use of customer data. Open letter from ING to her Customers).

144 Aleecia M McDonald and Lorrie Faith Cranor, 'The Cost of Reading Privacy Policies' (2008) 4 I/S: A Journal of Law and Policy for the Information Society 540, 540; Rainer Böhme and Jens Grossklags, 'The Security Cost of Cheap User Interaction', *Proceedings of the 2011 workshop on New security paradigms workshop* (ACM 2011) s 2.3.2 <https://dl.acm.org/citation.cfm?id=2073284> accessed 13 February 2019.

145 Galanter (n 91) 98.

Likewise, the cost of organisation is lower for producers.[146] A producer of even moderate size can organise collective bargaining force by forming alliances with only a few other producers to combine the interests of a relatively large portion of the market, for example in a trade association (presuming this is possible without violating article 101 TFEU). Although it is true that consumers can – and do – organise themselves into consumer rights associations, the large number, the low per capita stakes, and the homogenous nature of the mass of consumers, means that even these associations usually have to divide their attention between many subjects. Consumer associations focusing on privacy concerns are regularly struggling for money, indicating that resource pooling does not always raise sufficient funds.[147]

All in all, when bargaining for a privacy contract in the market, consumers can be said to be at a disadvantage when compared to producers. They have only small stakes per contract and the presumed homogenous distribution of stakes makes it more difficult to organise buying power to negotiate better terms. They also tend to qualify as one-shotters, with little opportunities to gain significant experience. Producers on the other hand, have the opportunity to benefit from repeat-player status by using the same contract repeatedly. The value of all privacy contracts combined raise their stakes and costs of organisation are low because of their low numbers. This encourages them to invest larger amounts of resources into the contents of privacy contracts.

2.6.2 In the political process

In the political process, the per capita stakes for producers are still determined by the value of all their privacy contracts combined. However, for consumers acting as a society with the ability to affect the contents of all privacy contracts at once by means of elected representatives, per capita stakes can also increase. Combining the value of all voters, the stakes are increased from the value of one privacy contract per transaction, to the social cost of privacy.[148] These increased stakes encourage resource

146 ibid 100.

147 Digital Rights Ireland, 'We Need Your Help to Keep Working for European Digital Rights in 2014' (*Digital Rights Ireland*, 1 January 2014) <http://www.digitalrights.ie/support-us-in-2014/> accessed 13 February 2019; Marie-José Klaver, 'Bits of Freedom Staakt Strijd Op Web; Oprichter: Digitale Burgerrechtenbeweging Harder Nodig Dan Ooit' *NRC Handelsblad* (5 August 2006) 26 26 <https://www.nrc.nl/nieuws/2006/08/05/bits-of-freedom-staakt-strijd-op-web-11172796-a826780> accessed 13 February 2019.

148 Sholtz (n 140).

pooling for voters, which helps in levelling the playing field on the benefit side when compared to the market. Participants also take on a different roles, consumers generally acting as voters and producers active as – or via – lobbyists.

Although votes usually have identical weight for all voters, the distribution of stakes in the political process is possibly less homogenous than in the market. For example, if a political party does not consider privacy contracts an issue for the past or coming elections platform, voters and legislators for this party can be said to have lower stakes in privacy contracts. Producers with large stakes in a certain outcome (or status quo) can exploit differences between political parties by concentrating their lobbying efforts, or donations, to benefit parties with favourable viewpoints. Unevenly distributed stakes will have a positive effect on the formation of pressure groups or political action committees.

Consumers may not stand to profit too much from this change in dynamics. It is not easy to attach a monetary value to privacy as a social cost. Therefore the increased stakes do not easily translate to voters' increased willingness to spend money on participation. Producers do not have to deal with this uncertainty to determine how much they want to spend.

The costs of participating in the political process are usually the cost of gathering information, popular campaigning and the influencing of legislators. These processes are used to organise and spread information among voters to promote public awareness, to gather support among voters and to bring specific viewpoints across to legislators, to influence their political activities (e.g. by hiring lobbyists). These costs can be high, which – at least in theory – again works in favour of producers, who have higher per capita stakes and correspondingly larger resources at their disposal.

When compared to the market, the cost of participation in the political process for consumers is either lower or higher, depending on a number of circumstances. Their costs of organisation can be reduced if groups of voters are pre-organised in political parties or interest groups by reducing sunk costs. This lowers the cost of activities for subsequent issues.[149] In the DPI/net neutrality example, the involvement of existing political parties and interest groups made all the difference, even if the parties pushing for the amendment did not form a majority in Parliament. On the other

[149] "Sunk costs are those costs that have to be incurred to enter or be active on a market but that are lost when the market is exited." European Commission, 'Guidelines on Vertical Restraints' (European Commission 2010) OJ 2010/C 130/01 26.

hand, if none of the political parties see privacy contracts as an important issue, the low per-capita stakes for voters work to their disadvantage as they do in the market. Producers can use this as an opportunity to lobby all parties in Parliament, increasing the costs of organisation for consumers. In those cases, pressure groups or political action committees have to find another way to influence legislators, increasing the cost of organisation. Seeking press coverage is one possible way to reduce these costs. As was seen in the DPI/net neutrality example, press coverage (triggered by statements of interest groups) was the second important factor in influencing politicians. It increased awareness among voters and legislators for privacy aspects of DPI and probably alerted a large number of voters who were unaware of KPN's plans. This added to the effectiveness of organised lobbying from consumers' action committees.

Press coverage can reduce the opportunities for legislators to vote against their constituents' interests: it reduces their slack.[150] Legislators may use their slack to align their political activity to lobbying efforts of producers in return for money, career opportunities or other benefits. Assuming that fundamental rights in a business context remain relevant to the press establishment, engaging the press can significantly lower monitoring costs. As a result, the outcome of the political process is less likely to be a producers' interest-group policy.[151] Interest-group policy (or regulatory capture) is an example of minoritarian bias: decision making dominated by the influence of the concentrated interests of any high-stakes minority, such as the producers in the market for privacy contracts.[152]

All in all, the dynamics of participation in the national political process are more favourable to consumers than they are in the market, due to reduced sunk costs and monitoring costs. However, producers can also have considerable clout in the legislative process, mainly as a result of their high per capita stakes.

2.6.3 In the national courts

Before a court, the dynamics of participation change once again. On the benefit side, the stakes for a producer can be significant, if the verdict can affect a large number of contracts. On the consumer side, the stakes depend on whether the case is a matter of collective redress, or a dispute over a single contract. Collective redress allows for

[150] Levine and Forrence (n 131) 167–176.
[151] ibid 176.
[152] Komesar, *Law's Limits* (n 93) 60–70.

resource pooling among plaintiffs acting as one single litigator, increasing the per capita stakes by "rolling many capita into one". However, collective redress is not always available.[153] Furthermore, seeking collective redress may eliminate any advantages that domestic procedural law would grant individual consumers. In an individual action, the stakes for the consumer are once more no higher than the value of the contract, putting him at a disadvantage when compared to a high stakes producer.

Costs of participation in the courts are not fixed. Both parties are free to spend as much as they want on information or organisation. Rational litigants will not spend more than the stakes of the case as spending more will result in a net loss, even if the case is won. Under this assumption, producers will generally be prepared to spend more as a result of their higher per capita stakes. The party with more willingness to spend can usually rally more influential allies and produce more expert opinions.[154] Information produced by one party may still increase both parties' costs, even if it is available to the opposing party at no extra cost. For example, if judges act as 'passive umpires', they may regard information that is not countered as being undisputed by the opposing party.[155] In such a case, producing excessive amounts of information can exhaust the means of the lower-stakes litigant.

The costs of organisation are also in the hands of the litigants, apart from the minimum costs associated with court fees and counsel. Once more, the party with the most resources at its disposal (e.g. more lawyers) could exhaust the resources of other parties and force them to give up, by using every available legal avenue and by

153 European Commission, 'Towards a European Horizontal Framework for Collective Redress COM(2013) 401 Final' (European Commission 2013) 4–5 <https://eur-lex.europa.eu/procedure/EN/202773> accessed 13 February 2019.

154 Samuel Issacharoff, 'Group Litigation of Consumer Claims: Lessons from the U.S. Experience' (1999) 34 Texas International Law Journal 135, 145.

155 In common law jurisdictions Galanter (n 91) 120. For the United Kingdom Neil Andrews, 'Fundamental Principles of Civil Procedure: Order Out of Chaos' in Xandra Ellen Kramer and others (eds), *Civil litigation in a globalising world* (TMC Asser Press ; Springer 2012) 29. In the Netherlands, the 'lijdelijke rechter'; in Germany, 'Parteibetrieb', in France, the 'juge passive': Regine Genin-Meric, 'Droit de la preuve: l'Example Français' in José Lebre de. Freitas (ed), *The law of evidence in the European Union = Das Beweisrecht in der Europaischen Union = Le droit de la preuve dans l'Union Europeenne* (Kluwer Law International 2004) 140–141; CH van Rhee, 'De Ontwikkeling van het Burgerlijk Procesrecht in het Twintigste-Eeuwse Europa: Een Terugblik' in D Heirbaut, G Martyn and R Opsommer (eds), *De Rechtsgeschiedenis Van De Twintigste Eeuw. the Legal History of the Twentieth Century: Handelingen van het contactforum* (Peeters Bvba 2006) 1–5.

complicating or lengthening proceedings to the maximum extent possible. The level of proficiency in the task at hand also determines the costs of organisation.[156] By this metric, repeat players are at an advantage because they can use their experience in organising their work more efficiently. Producers are more likely to be repeat players because of the large number of contracts they engage in.

All this puts consumers at a clear disadvantage before the courts. To quote Prof. Giesen on Dutch law: 'Private law [...] does not deal in affirmative action for the weaker party (rather the opposite, really)'.[157] The difference in per capita stakes makes it unlikely that an individual consumer can effectively challenge a privacy contract if there is no clear-cut legislation on which the Court can easily decide the case. The individual consumer is more likely to be a one-shotter. Compared to a producer with high stakes, the consumer also qualifies as a have-not.[158] Producers are more likely to be repeat players in the court system. Given their increased stakes and proportionally larger resources, they qualify as the haves in this context.

Collective redress opportunities may compensate this disadvantage: it may increase the stakes, allow for resource pooling and reduce the costs of information. In high-profile test cases, publicizing fundraising efforts may lower the cost of organisation and engage experts that will regard increased publicity as a form of compensation, thus lowering the costs of information. However, this reintroduces some of the factors we already saw working against consumers in the market. Both test cases and collective redress rely on contributions from a large class of consumers. In a large class, the benefits of participation may again be low because of the evenly distributed stakes.

Substantive and procedural law offer another way to reduce consumers' disadvantages. For example, Dutch consumers are usually entitled to a procedure before a court that does not require legal representation – a measure aimed specifically at reducing their costs of organisation.[159] European consumers may profit

156 Ivo Giesen, 'Sommige Procespartijen Zijn "More Equal than Others". De Macht van de Tabaksindustrie En de Nederlandse Rechtspleging' in Nienke Doornbos, Nick Huls and Wibo van Rossum (eds), *Rechtspraak van Buiten. Negenendertig door de rechtssociologie geïnspireerde annotaties (Liber Amicorum prof. dr. J.F. Bruinsma)* (Kluwer 2010).

157 '... dat het privaatrecht [...] niet aan "positieve discriminatie" van de zwakkere partij doet (eerder het tegendeel).' ibid 21.

158 Galanter (n 91) 103; Giesen (n 156).

159 Almost certainly applicable to privacy contract cases: art. 93 a and c, Rv (Dutch law of civil procedure).

from ex officio application of EU consumer protection law. As a side effect, the cost of information may be reduced in some cases.[160] The effectiveness of these measures is not explored further in this chapter, as it goes beyond the scope of comparative institutional analysis.

It seems unlikely that these measures are able to compensate consumers' disadvantage completely because producers' per capita stakes remain much higher than consumers'. Galanter's observation still rings true: before the courts, the haves come out ahead.[161]

2.6.4 Comparison at the national level

When seeking decisions on the contents of privacy contracts at the national level, participation opportunities for consumers compare unfavourably to those of producers in the market and the courts, and probably in the political process as well.

Consumer disadvantage is smallest when participating in the political process. Earlier organisational efforts lower the cost of organisation for consumers on new issues, even if these parties organise only a tiny fraction of the mass of consumers. The cost of participation can also turn out lower if privacy contracts continue to be of interest to the press. The macroeconomic aspect of the political process helps raise the stakes, from the value of a single contract to the value of privacy as a social cost of big data. The raised stakes help increase the resources available by encouraging resource pooling. But this smaller disadvantage for consumers is no guarantee for their success in the political process. Per-capita stakes for producers are high. The political process is therefore not immune to the lobbying force of concentrated minority interests. The results of the political process may still have a minoritarian bias.

For producers, the courts and the market are very efficient institutions to achieve their objectives. In the market, they enjoy low costs of information and organisation and the benefits of repeat player status. They are also motivated by high per capita stakes as a result of the large number of contracts. In court, where the stakes per decision may be higher than in the market, this difference is enhanced even further. In the political process, this advantage is reduced.

---

160 Case 618/10 *Banco Español de Crédito SA v Joaquín Calderón Camino* [2012] ECLI:EU:C:2012:349, para 42.

161 Galanter (n 91).

2.7 The market vs. the political process at the EU level

The dynamics of participation at the EU level differ from those at the national level for two reasons. Firstly the institutions themselves are different, and secondly the scope of the EU is much larger. Comparing different institutions at the EU level is therefore not easily separated from comparing their similar institutions at the national and EU level. As a result, this section addresses both comparisons simultaneously. This subsection offers conclusions for the comparison on the EU level. Subsection C briefly summarises the comparison between institutions at the national and the EU-levels.

2.7.1 In the market

For consumers, the dynamics of participation on the European level are partially similar to those on the national level. The effects of being a one-shot player with stakes not exceeding the value of a single contract are the same.

Consumers' benefits of participation are lower in the internal market. The larger number of individuals makes the distribution of stakes even more uniform.[162] This again means there probably is no substantial subgroup of consumers for which the benefits of participation are significantly higher. This lowers the incentives for consumers to organise themselves. The stakes for a single consumer still don't exceed the value of a privacy contract.

Consumers also face higher costs of participation at the European level. A penalty on participation in the market at the European level exists if cross-border transactions bear higher costs than domestic transactions, as may be the case with telephone or wireless Internet contracts.[163] For most consumers, the cost of information rises if this involves acquiring information in a foreign language that needs translation. This applies both to the contents of the contract and to the law of the land. Translation costs may similarly increase the cost of organisation for consumers, for example if

[162] Actually, stakes are most likely spread according to a *normal distribution*: many consumers will have stakes near the mean; very few consumers will have stakes that are much higher or lower.

[163] World Administrative Telegraph and Telephone Conference and International Telecommunication Union, 'International Telecommunication Regulations: Melbourne, 1988 (WATTC-88).', *Final acts of the World Administrative Telegraph and Telephone Conference* (ITU 1989) 4 (art. 1.5).

they want to organise across national borders. Incentives for compensating these increased costs are low because of the low per capita stakes.

Consumers are not alone in thinking (and actually experiencing) that cross-border participation is expensive. The European Commission has undertaken specific efforts to reduce the costs translating the law of sales by proposing a Common European Sales Law.[164] Furthermore, the Commission subsidises consumer organisations, as can be seen in the financial statements of BEUC (Bureau Européen des Unions de Consommateurs). Even though national consumer associations contribute the greatest share of the expenses, the European Commission provided 49% of BEUC's revenues in 2016, indicating that voluntary resource pooling alone cannot cover the increased costs.[165]

Conversely, for producers the dynamics of participation on the internal market are working in their favour. The potential number of contracts a producer can enter into is increased, raising the per-capita stakes. This is the same for all producers engaging in privacy contracts. Therefore, producers can still be seen as a small, more-or-less uniform group with high per-capita stakes.

The opportunity to spread the costs of doing business over a larger number of contracts will encourage producers to participate in the market on the EU level. Trading in the internal market potentially lowers costs of participation per contract for producers, making one more contract more profitable. The internal market is one of the largest economic areas in the world, offering a large potential for additional contracts.[166] As a side effect, this may enhance the benefits of repeat player status.

---

164 European Commission, 'A Common European Sales Law to Facilitate Bross-Border Transactions in the Single Market' (European Commission 2011) Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions COM/2011/0636 final <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52011DC0636> accessed 13 February 2019.

165 BEUC, 'Annual Report 2016' (BEUC 2017) 14 <http://www.beuc.eu/publications/beuc-x-2017-045-annual-report-2016.pdf> accessed 19 March 2019. Subsidies for consumer participation are part of a EU programme of 188,8 million Euros over the 2014-2020 period; European Parliament and Council Regulation (EU) No 254/2014 of 26 February 2014 on a multiannual consumer programme for the years 2014-20 and repealing Decision No 1926/2006/EC, [2014] OJ L 84/42, article 3(1)(b).

166 The EU economy is larger than that of the US. 'Europa- The economy': , <http://europa.eu/about-eu/facts-figures/economy/index_en.htm> accessed 1 November 2014.

The free movement of personal data, guaranteed by the Data Protection Directive article 1(2) and the GDPR article 1(3), lowers compliance costs associated with data processing. Compliance with data protection legislation in one Member State guarantees free movement of these data to other Member States.[167] As a result, large telecom providers and most social network websites tend to work in several Member States simultaneously. This, in turn, helps reduce costs even further, for example by consolidation of processing equipment and administrative functions. Many producers engaging in privacy contracts formally rely on a single point of presence in the EU. For example, LinkedIn and Facebook provide an address in the Republic of Ireland for correspondence on data protection matters.[168]

2.7.2 The political process

For consumers, the political process at the EU level compares unfavourably to its equivalent at the national level. On the benefit side, the distribution of stakes for voters is strongly homogenised by the combination of national seats from 28 countries into 7 larger parties in the European Parliament. This is reflected in voter turnout. Average turnout in all Member States in 2014 was 42,54%, trailing far behind typical voter turnouts in national general elections.

The per capita stakes at the EU level are larger than on the national level. EU decisions affect a much larger number of citizens and businesses, increasing their impact on privacy as a social cost. They also affect legislation in many countries, eliminating opportunities for consumers to find a better legal arrangement in another Member State. The stakes may however appear smaller to voters, because the European Parliament's powers are less pronounced than those of national parliaments. The EP can pass a motion of censure on the entire Commission only with a two-thirds majority whereas national parliaments can usually do so with a simple majority and for single Ministers.[169] Compared to national parliaments, the EP is elected for a longer period of time and it cannot be dissolved to enable voter participation in unresolved disputes between the EP, the Council and the

[167] Data Protection Directive, art. 1(2); GDPR, art. 1(3)

[168] Linkedin, 'Postadres voor vragen over gebruikersovereenkomst of privacybeleid' (in Dutch) <https://www.linkedin.com/help/linkedin/answer/79728?trk=microsites-frontend_legal_privacy-policy&lang=nl> accessed 13 February 2019; Facebook, 'Gegevensbeleid' (in Dutch) <https://nl-nl.facebook.com/about/privacy/> accessed 28 April 2018.

[169] Art . 234 TFEU.

Commission.[170] Instead, the resolution of conflicts between these bodies is referred to a conciliation committee without any possibility of voter input.[171] This makes voting less attractive.

Consumers face higher costs of organisation in the EU than on the national level. A very simple example is the increased physical distance, with most voters needing to travel abroad, to Brussels or Strasbourg, if they want to interact with legislators directly. Another aspect is the large number of legislators and the large number of venues where participation may be required: Commissioners, MEP's and their staff for primary law, and the Commission, agencies and comitology for secondary law.[172] Representatives in the European Parliament are under pressure not only from their political party on the national level, but also from the European parties on the EP level. Furthermore, no single national delegation can dominate a party in the EP. This decreases the possible benefits of "piggybacking" on previous organisation efforts. European Parliament proceedings usually receive far less coverage in the national press, increasing the costs of information and monitoring costs for consumers. This gives legislators more possibilities to decide against voters' interests.[173]

The higher costs and the decreased benefits of participation at the EU level leave consumer interest groups short of money. Participation in the political process by consumer organisations requires external funding from the European Commission.[174] BEUC and the national consumer associations are appointed as members of the European Consumer Consultative Group and in this capacity they are eligible for subsidies.[175] The Commission decision establishing the group states that it is tasked with consumer interests in general, which means that it needs to spread itself thin over all consumer issues. BEUC is thus not a special interest privacy group. Such groups do exist, but even a group like European Digital Rights also had to rely on a

170 Art. 14(3) TEU.

171 Art. 294 (8)(b), (10-14) TFEU.

172 Marinus PCM van Schendelen, *Machiavelli in Brussels: The Art of Lobbying the EU* (2nd edn, Amsterdam University Press 2005) 58.

173 Levine and Forrence (n 131) 173.

174 National members of BEUC may be eligible for subsidies. At least in the Netherlands, these subsidies accounted for much less of the budget (24% for 2016). Consumentenbond, 'Jaarverslag 2016' (Consumentenbond 2017) Jaarverslag 73 <https://www.consumentenbond.nl/binaries/content/assets/cbhippowebsite/over-ons/wie-zijn-we/consumentenbond-jaarverslag-2016.pdf> accessed 13 February 2019.

175 Article 3(1)(a) and annex, Commission Decision of 14 September 2009, setting up a European Consumer Consultative Group (2009/705/EC), [2009] OJ L 244/21

grant from the European Commission of approximately 25% of their annual budget in 2013 although this subsidy was no longer on the books in 2016.[176] Like in the market, resource pooling in the political process at the European level apparently does not sufficiently offset the higher costs of organisation at the EU level.

For producers, the benefits are quite large. The European political process is a high-stakes game because it regulates the internal market. This makes the distribution of stakes almost uniform, but at a very high level. Producers have a strong incentive to organise, even if the costs of organisation are high.

Like consumers, producers face higher costs of information and organisation in Brussels – but in accordance with their increased stakes, this appears to present no extra problems. Producers are the haves in this institution. They are able to pool and spend considerable resources and do not need EC subsidies. If lobbying efforts in the United States are any indication, spending several million on influencing legislation is an acceptable proposition to many individual companies.[177] Producers have formed their own lobbying groups dedicated to data protection and privacy issues.[178] Like consumers at the national level, producers can also benefit from earlier organisation efforts. Many industries that engage in privacy contracts have already formed interest groups for their core business, like the European Competitive Telecommunications Association, Digital Europe and the Euro Banking Association. When lobbying for privacy contracts legislation, this helps them avoid sunk costs.

It is therefore no surprise that high-stakes players are better represented than consumers in the Brussels lobbying circuit. In 2007, commercial and professional interests made up 63% of all permanent representations in Brussels, whilst consumer and human rights interests accounted for 13%.[179] Their apparent abundance of funding puts them at an advantage over consumer special interest groups like EDRi, whilst general-interest groups like BEUC are probably already stretched thin because of their

176 EDRi - European Digital Rights, 'Annual Report, January 2013 - December 2013' (EDRi - European Digital Rights 2014) 31 <https://edri.org/wp-content/uploads/2014/04/EDRi_Annual_Report_2013.pdf> accessed 13 February 2019; EDRi - European Digital Rights, 'Annual Report 2016: January 2016 – December 2016' (EDRi - European Digital Rights 2017) 30 <https://edri.org/files/edri_annual_report_2016.pdf> accessed 13 February 2019.

177 April Dembosky, 'Facebook Spending on Lobbying Soars' *Financial Times* (24 January 2013).

178 James Fontanella-Khan, 'Brussels: Astroturfing Takes Root' *Financial Times* (26 June 2013).

179 van Schendelen (n 172) 50.

wide scope. These differences express themselves in increased opportunities to interact with officials in the EP, the Commission and comitology.

Legislators' large amounts of slack, taken together with their larger exposure to producers than to consumers, increase the odds of special-interest policies being adopted in the EP. The costs and benefits of participation in the political process on the EU level are almost ideal for regulatory capture.

2.7.3 Comparison of the market and the political process at the EU level

Producers are at an advantage both in the market and in the political process at the EU level. This is mainly due to the large cost of participation, combined with the large stakes for producers, giving them a bigger incentive to participate.

For consumers, participating in the market on the European level does not increase their per capita stakes but it does increase their costs. In the political process, the stakes for consumers are actually higher, but the costs of participation are raised even more: lobbying on the EU level is a very expensive undertaking. The comparatively low perceived stakes are an insufficient incentive to achieve sufficient resource pooling on this level. As a result, consumers on the EU level need financial aid to organise effectively in the market as well as in the political process.

The internal market reduces producers' operation and compliance costs. Therefore, the EU political process is capable of decision-making that affects their profitability significantly. For many producers, concluding privacy contracts is not their core activity, but a side effect of other activities already being lobbied for in Brussels. As a result, they qualify as repeat players in the political process when compared to single-issue consumer groups dedicated to privacy issues. Their high stakes, resulting in abundant funding, gives them an advantage over general consumer interest groups that are stretched too thin over many issues. This increases the possibility that EU regulations cater to special interests contrary to the interests of consumers. This risk of regulatory capture is not imaginary, with the effectiveness of producer lobbying in telecom issues already documented.[180]

[180] Breindl (n 122) 354; Kimberlee Weatherall, 'Three Lessons from ACTA and Its Political Aftermath' (2012) 35 Suffolk Transnational Law Review 575, 595.

2.8 Comparison between the national and EU levels

2.8.1 The market

At the national level, the market favours producers of privacy contracts – at the European level, the market favours them even more. The free movement of services, capital and personal data within the internal market can lower the cost of information and organisation significantly. These lowered costs together with the high per-capita stakes provide producers with a powerful incentive for participating in the market on the European level. These benefits offset any extra costs for organisation and information.

Consumers, on the other hand, reap no such benefits at the European level. The increased costs of participation are such that contracts in their home country are significantly more attractive, and consumer organisations need EU subsidies to participate efficiently. Their per capita stakes in privacy contracts are not raised when participating in the European market. Increased costs of information – for example, the costs of translation – make it even less attractive to look for a better privacy contract in another EU Member State.

2.8.2 The political process

Both for producers and consumers, shifting the political process to the European level increases costs of participation significantly. However, consumers may falsely tend to think of the EU political process as having lower stakes than their national process, because the European Parliament has less pronounced powers than the national parliaments and press coverage of other institutions and decisions is relatively scarce. Yet consumers' stakes in the EU political process may actually be higher. A decision at the EU level decreases consumers' opportunities to obtain a better agreement in another Member State. However, lower perceived stakes decrease active voter participation in the political process. Consumer lobbying groups need grants from the European Commission as a result.

Judging by the resources spent on lobbying efforts, producers are very well aware that the stakes in Brussels are higher than in national Parliaments. Their higher stakes, their low numbers and the homogenous nature of producers form an effective incentive to organise.

2.9 Summary of institutional comparisons

Based on the dynamics of participation, the circumstances for consumer participation aimed at influencing the contents of privacy contracts are least unfavourable in the political process at the national level. The costs of information and participation are lowest here, mainly as a result of previous organisation efforts. They are also lower than in the internal market. This doesn't automatically imply that conditions are favourable as consumers' low per capita stakes weaken their opportunities in all institutions. Consumers really have to choose between imperfect alternatives.[181]

The effectiveness of consumer participation at the national level is further reduced or limited due to the fact that data protection legislation is mainly decided in Brussels. A concerted effort on behalf of all consumers at the European level could theoretically be more effective, but the lower perceived stakes and the significantly increased costs of participation in Brussels might make this an unattainable goal for the foreseeable future.

Producers seeking to influence the contents of privacy contracts can effectively achieve their goals in the market, in the national courts or in the European political process. Their advantage is arguably most pronounced in the political process at the EU level, where their stakes are highest. The high costs of participation for consumers in this forum increase the opportunities for legislators to decide against consumers' interests, increasing the risk of regulatory capture.

2.10 Institutional choice and policy objectives

Comparative institutional analysis in itself provides no guidance on which institution is best charged with deciding on privacy contracts. The goals a society wishes to achieve are equally important; not every institutional difference is equally relevant to every policy choice. This chapter explores the margins inside which national or European policy choices must remain. Subsequently it broadly categorizes two possible policy objectives and proposes matches between institutions and these objectives.

[181] Komesar, *Imperfect Alternatives* (n 136).

2.10.1 Two sets of European margins

Both EU law and the European Convention on Human Rights limit the adverse effects of privacy contracts for consumers. They show significant differences, in both their scope and their available enforcement mechanisms.

The Convention requires Member States to respect private and family life.[182] This makes safeguarding consumer privacy primarily a matter for the Member States, even if data protection law is mostly EU law and the EU accepts the rights guaranteed in the Convention as general principles of EU law.[183] After all, the EU Charter of Fundamental Rights is only binding upon the institutions of the EU. Some issues governing privacy contracts – like the interpretation of bona fides – are beyond the scope of EU legislation.

EU law itself provides another set of minimum requirements. The principle of sincere cooperation requires Member States to make EU law effective.[184] The Data Protection Directive and the GDPR (with its provisions for free movement of data) harmonise national requirements to privacy contracts and prevent Member States from using the limiting of the free flow of data as an enforcement mechanism in response to breaches of data protection law. Several further consumer protection directives prevent Member States from giving producers too much free rein. Examples of EU legislation limiting the contents of privacy contracts are the Unfair Terms directive and the Unfair Commercial Practices Directive.[185]

These limits are not enforced equally. EU law has to be applied ex officio in the courts of all Member States and the Commission can employ enforcement mechanisms to guarantee compliance.[186] The ECtHR cannot enforce the Convention directly, since it has no authority to alter decisions by a State. It can only award 'just satisfaction', usually monetary compensation, if a State allows only partial reparation after a violation.[187] Article 8 ECHR is stated in broad terms and the ECtHR interprets it on a case-by-case basis. This limits Member States' abilities to predict whether legal acts are a breach of the Convention, especially in the light of new technological developments like big data.

182 Articles 1 and 8, ECHR.
183 Article 1 ECHR; article 6(3), TEU.
184 Art. 4(3), TEU.
185 Unfair Terms Directive (n 128); Unfair Commercial Practices Directive (n 96).
186 Art. 258 TFEU.
187 Art. 41, ECHR.

The ECtHR will allow Member States considerable leeway when judging cases involving privacy contracts, due to the fact that a privacy contract is a matter between private parties. Traditionally, the right to respect for one's private and family life primarily means that the State should refrain from undue privacy breaches (a negative obligation). If no government body is a party to a privacy contract, any duty of the State will be interpreted as a positive obligation.[188] States may trigger a positive obligation under article 8 ECHR by inaction, by insufficient action and by not taking control when privacy breaches are getting out of hand. On the other hand, Member States are afforded a margin of appreciation that is wider than in cases concerning negative obligations. For example, the needs and the resources of the community may be taken into account.[189]

A wider margin of appreciation is relevant in the case of privacy contracts and big data has large possible benefits that communities may decide not to want to do without. A State is allowed to weigh these benefits against the loss of privacy that these contracts can entail. Societal gains are important, and gains in expensive policy areas, such as healthcare and education, may under certain circumstances outweigh an effective loss of privacy caused by a private contract. For example, it is conceivable (although untested) that Member States would allow the analysis of many students' interactions with teaching materials in an electronic learning environment to provide the best form of education for each student, even if this analysis could also reveal deeply personal traits. A State may consider economic growth, jobs or investments essential for its general welfare. Even private gains may outweigh privacy, if a State counts the right to returns on investments as a property right safeguarded by article 1 of the Protocol to the Convention. All this remains speculative, since the ECtHR until now hasn't given any verdicts on privacy contracts.

2.11 Making a match

A society might want to choose to protect privacy or to promote societal or economic benefits. The analysis performed in section III indicates that institutional choices show different levels of compatibility with these objectives. In choosing institutions

[188] European Union Agency for Fundamental Rights, European Court of Human Rights and Council of Europe (n 90) 16. See also ECtHR Odièvre v France 2003-III 1 para 40; Jean-François Akandji-Kombe, *Positive Obligations under the European Convention on Human Rights* (Directorate General of Human Rights, Council of Europe 2007) 15.

[189] *Powell and Rayner v the United Kingdom* (1990) Series A no 172, para 41; *Johnston and others v Ireland* (1986) Series A no 122, para 55.

for privacy contracts, it is assumed that consumers have a better chance for strong privacy protection if their participation opportunities are better.

2.11.1 Maximizing privacy protection

Considering the comparatively unfavourable circumstances in other institutions, the national political process offers consumers the best opportunities for achieving strong privacy protection in privacy contracts.

An important benefit of the national political process lies in the fact that national Parliaments are not subject to the scope limitations of EU institutions. The scope of EU decisions is limited by the principles of conferral, proportionality and subsidiarity.[190] EU data protection law therefore cannot capture the complete scope of the right to privacy as protected in article 8 of the Convention. For example, it does not address reasonable expectations of privacy or the interpretation of contracts.

2.11.2 Maximizing social or economic benefits

It seems reasonable to expect that the supposedly large economic benefits of the big data revolution[191] will arrive more slowly if restrictions on privacy contracts are stronger. Legal restrictions increase compliance costs and the risk of noncompliance.

The EU political process will probably offer the best opportunities for maximizing the economic and societal benefits of big data. The increased scope of the internal market, when compared to the political process at the national levels, will reduce compliance costs and increase legal certainty for producers active in several Member States. This institution also offers the most benefits of participation to producers, improving the odds of the resulting legislation and not hampering their objectives. This however comes at a cost: the influence of consumers is greatly reduced.

Regulating privacy contracts at the European level, regardless of the outcome, has added significance for the interpretation of article 8 ECHR. A decision at the EU level indicates that Member States agree on the necessary degree of privacy protection. Even if consumers in the EU have limited participation opportunities, any decision on the contents of privacy contracts counts as an indication of common ground between Member States. For the ECtHR, this is a factor in deciding whether a Member State

[190] Art. 5, TEU.

[191] 'The new oil of the internet'? Moerel (n 87) p, 20. Moerel (n 110)

has fulfilled its positive obligation within the margin of appreciation.[192] For the CJEU, allowing producers to retain and use personal data under private contracts is much more likely to be acceptable than the Data Retention Directive.[193] This directive enabled governments to use personal data collected under privacy contracts and it has been held unlawful in both national and European courts.[194]

Of course, common ground between Member States is not in itself a guarantee for compliance with article 8 ECHR. The fact that voters cannot easily hold legislators and regulators to account creates a significant risk of regulatory capture and special interest legislation. The legitimacy of EU legislation may also be significantly reduced if many consumers consider their opportunities for participation insufficient.

2.12 Concluding remarks

Privacy contracts authorise a deep insight into the personal lives of consumers. Therefore, they require sufficient legitimacy. This chapter has focused specifically on the participation aspect of the legitimacy requirement (transparency and accountability aspects were excluded). It has examined consumers' participation options in decision-making processes concerning privacy contracts. Particularly, it has compared the dynamics of participation for consumers and producers in the political process and the market on the national and European levels, and before the national courts, by performing comparative institutional analysis.

Although an important factor, maximizing consumer participation opportunities will not guarantee adequate privacy protection by itself. Choosing between institutions requires both awareness of the strengths and weaknesses of each option and a clear view of the goals that a society wants to achieve. The choice is not an easy one. Optimal privacy protections for consumers may impede important societal or economic benefits, whilst maximizing these benefits may sooner or later trigger positive obligations under article 8 ECHR.

---

192 ECtHR *Rasmussen v Denmark* (1984) Series A no 87, para 40.

193 Art. 4, European Parliament and Council Directive 2006/24/EC of 15 March 2006 on the retention of data generated or processed in connection with the provision of publicly available electronic communications services or of public communications networks and amending Directive 2002/58/EC, [2006] OJ L 105 (Data Retention Directive).

194 BverfG, Urteil des Ersten Senats vom 02. März 2010 1 BvR 256/08 Rn 1 – 345; Joined Cases C-293 and C-594/12 *Digital Rights Ireland v Minister of Communications and Kärntner Landesregierung v Seitlinger and others* [2014] ECLI:EU:C:2014:238.

The political process at the national level offers consumers the best opportunities for participation in the decision-making process on privacy contracts. In the market, the national courts and the political processes at the EU level, opportunities for consumer participation are greatly reduced. As a result, decisions from these institutions are more likely the result of capture or minoritarian bias. The legitimacy of such decisions is possibly insufficient.

This tentatively points towards the EU Directive as the better option for regulating privacy contracts, as it leaves Member States the choice of form and methods to implement them.[195] Transposing directives into national law will allow Member States to debate those aspects of privacy contracts left open due to the limited scope of EU decisions. This will enable, for example, the setting of standards for reasonable expectations of privacy, reasonable interpretation of contracts and unfair terms, within the limits imposed by EU law and the Convention. A regulation such as the GDPR, which became applicable in May 2018, precludes national debates in the legislative process and therefore offers less opportunities for consumers to meaningfully engage in the debate on an important issue.

However, the next step in EU data protection legislation will be a regulation, directly applicable in all Member States.[196] In that case, options for national political debate are limited. On the EU level consumers already have few participation opportunities. Adopting a regulation calls on the members of the European Parliament to honour their duty to their constituents. At the same time, MEP's large amount of slack gives them an opportunity to decide against consumers' interests without suffering any consequences.

If national or EU law insufficiently protect consumer privacy, national courts will eventually be called upon to apply article 8 ECHR to privacy contracts. This could take many years. The transaction costs for consumers before the courts are so high and individual stakes are so low, that individual consumers are unlikely to prevail before the courts against a large producer with high stakes. This result would be undesirable. It could effectively allow for permanent observation of every consumer in a new kind of panopticon, possibly with far-reaching effects on consumers' personal autonomy.

195 Article 288, TFEU.

196 John Bowman, 'EU Data Protection Regulation: A Tipping Point Has Been Reached' (*The Privacy Advisor: The official Newsletter of the IAPP*, 7 November 2014) <https://iapp.org/news/a/eu-data-protection-regulation-a-tipping-point-has-been-reached/> accessed 13 February 2019.

Such a result could very well be within the limits of data protection legislation, but it will almost certainly be a breach of article 8 of the Convention.

The low per capita stakes for consumers will likely continue to impede effective consumer participation. Altering the dynamics of participation in the market or the political process might seem an easy solution, but this requires careful consideration. For example, subsidizing organisation or lobbying efforts may introduce new monitoring costs. Consumer representatives would enjoy an amount of slack similar to members of the European Parliament. It could also distort the stakes for representatives. If subsidies are much higher than the contributions from consumers, these subsidies could effectively become the stakes. Such an arrangement does not necessarily offer more guarantees for effective consumer participation, nor does it necessarily reduce the risk for regulatory capture or special-interest legislation. Innovative new contract types may be a better choice to improve the legitimacy of big data applications. Wauters et al. have already researched a number of possible improvements.[197] Other solutions may lie in optimizing transparency and accountability.

It is not yet clear whether big data will call for merely incremental adaptations to data protection law, or rather a fundamental redesign of privacy law. In any case, legal developments in response to big data need to coherently address privacy contracts, non-contractual relations, the reasonable expectation of privacy and the possible effects of datafication on human autonomy – in other words, an integrated framework for consumer privacy in the age of big data. The limits on EU decisions imposed by the principles of conferral, proportionality and subsidiarity may stand in the way of developing such a framework at the EU level.

The law has to keep up with technology to effectively continue safeguarding consumer privacy in the coming age of big data. The pace of technological developments shows no signs of slowing down. The law had better be ready.

[197] Ellen Wauters, Eva Lievens and Peggy Valcke, 'Social Networking Sites' Terms of Use Addressing Imbalances in the User-Provider Relationship through Ex Ante and Ex Post Mechanisms' (2014) 5 JIPITEC <https://www.jipitec.eu/issues/jipitec-5-2-2014/4001> accessed 13 February 2019.

3 Beyond Consent

Improving data protection through consumer protection law

Prelude

After the previous paper was published, I first wrote a conference paper for the 2015 Amsterdam Privacy Conference. This conference paper was eventually converted into Chapter 4 of this book. However, the following chapter was published earlier and builds on the subject matter of Chapter 2: therefore, it is included as Chapter 3.

When the Internet Policy Review announced its special issue on "Big data, big power shifts?", this seemed relevant in the light of my earlier work using Komesar's method of comparative institutional analysis. Many data subjects are also consumers as defined by EU consumer protection law. Legislation in this area is an important instrument in dealing with the power differences between consumers and commercial actors. The call for papers gave me the opportunity to expand on the theoretical underpinnings for shifting power relations associated with the advent of big data.

This chapter is relatively brief: the Internet Policy Review imposes a very strict limit of 30.000 characters (including spaces). I stayed within that limit, but the margin was thin. To help authors stay within the character limit, the Style Guide offered sage advice: "Most people are not going to reach the end of the article. There is no harm in 'giving the story away' in the first paragraph."

Timeline and citation

The original article was published in the "Big data, big power shifts" special issue of the Internet Policy Review on 31 March 2016 under a "Creative Commons Attribution 3.0 Germany" open access license. The call for papers was published 8 September 2015 and is archived at https://policyreview.info/node/374.

The original version was submitted on 18 December 2016. The Internet Policy Review conditionally accepted the first version on 16 February 2017. This version was finalised 16 March. The Internet Policy Review recommends the following citation:

Rhoen, M. (2016). Beyond consent: improving data protection through consumer protection law. Internet Policy Review, 5(1). DOI: 10.14763/2016.1.404.

3.1 Introduction

Big data is shifting power away from consumers and data subjects towards data controllers. In a legal sense, natural persons often act as both a consumer and a data subject at the same time.[198] Controllers have come to collect data and metadata on an increasing number of common consumer activities like personal communications, online behaviour, shopping, banking and public transport – a trend known as datafication.[199] The Internet of Things (IoT)[200] will soon generate even more data. The collection and analysis of big data streams can amount to consumers' permanent surveillance. This gives controllers the power to influence consumer behaviour through dynamic or discriminatory pricing, filter bubbles or subtly influencing individual decisions (nudging).[201]

Big data's power shift has a significant privacy and data protection dimension. According to "Zimmermann's law",[202] this happens by virtue of technological progress alone. Data is also evolving into new currency. Increasingly, data controllers offer services like games and social networking not for money, but in exchange for the right to collect and use personal data. Consumers often enter these "privacy contracts"[203] if they want to enjoy a service seemingly for free, but even paying customers are not safe

198 A consumer is "any natural person who is acting for purposes which are outside his trade, business or profession" (Art. 2(b), Directive 93/13/EC).

199 Mayer-Schönberger and Cukier (n 115) 29.

200 ITU-T, 'Overview of the Internet of Things' (International Telecommunication Union 2012) Recommendation ITU-T Y.4000/Y.2060 1.

201 "A nudge (...) is any aspect of the choice architecture that alters people's behavior in a predictable way without forbidding any options or significantly changing their economic incentives." Richard H Thaler and Cass R Sunstein, *Nudge: Improving Decisions About Health, Wealth, and Happiness* (Revised & Expanded edition, Penguin Books 2009) 6.

202 "The natural flow of technology tends to move in the direction of making surveillance easier." Om Malik, 'Zimmermann's Law: PGP Inventor and Silent Circle Co-Founder Phil Zimmermann on the Surveillance Society' (*Gigaom*, 11 August 2013) <https://gigaom.com/2013/08/11/zimmermanns-law-pgp-inventor-and-silent-circle-co-founder-phil-zimmermann-on-the-surveillance-society/> accessed 13 February 2019. See also https://en.wikipedia.org/wiki/Phil_Zimmermann#Zimmermann.27s_Law (accessed 2 May 2018).

203 Verhelst (n 119) chap. 3.

from this practice.[204] Nevertheless, the recently accepted General Data Protection Regulation refers to consumer protection law only once (Recital 42). Similarly, the European Commission's 2012 proposal for a consumer agenda mentions data protection efforts only in passing.[205] The European Data Protection Supervisor, however, has stated that consumer protection law has a part to play in data protection, especially on the subject of transparency.[206]

EU data protection law has facilitated the aforementioned power shift since the introduction of the 1995 Data Protection Directive (DPD). It allows the collection and use of personal data based on consumers' consent, or if it is "necessary for the performance of a contract to which the data subject is a party". The GDPR contains a similar provision.[207] Compliance with the directive is then rewarded with the right to freely move this data within the European Union and a number of other jurisdictions, depending on an "adequacy decision" from the European Commission (art. 45, GDPR).[208]

These provisions seemingly empower data subjects, but data subjects acting as consumers lack effective participation options in the market.[209] They can hardly avoid privacy contracts: almost all banks, software and hardware vendors, social networking sites, digital content services, retail loyalty programmes and telecommunications providers employ them. The difficulties consumers face when invoking fundamental

---

204 Nicole Perlroth, 'How Superfish's Security-Compromising Adware Came to Inhabit Lenovo's PCs' *The New York Times* (1 March 2015) <http://www.nytimes.com/2015/03/02/technology/how-superfishs-security-compromising-adware-came-to-inhabit-lenovos-pcs.html> accessed 13 February 2019.

205 European Commission, 'A European Consumer Agenda - Boosting Confidence And Growth' (European Union 2012) Communication from the Commission to the European Parliament, the Council, the Economic and Social Committee and The Committee of the Regions COM(2012) 225 final 3–4 <https://ec.europa.eu/commission/sites/beta-political/files/consumer_agenda_2012_en.pdf> accessed 13 February 2019.

206 European Data Protection Supervisor, 'Privacy and Competitiveness in the Age of Big Data: The Interplay between Data Protection, Competition Law and Consumer Protection in the Digital Economy' (2014) Preliminary Opinion of the European Data Protection Supervisor 2 <https://edps.europa.eu/sites/edp/files/publication/14-03-26_competitition_law_big_data_en.pdf> accessed 13 February 2019.

207 Article 7(a) or (b), DPD; article 6(a) or (b), GDPR.

208 European Commission, 'Adequacy of the Protection of Personal Data in Non-EU Countries' (*European Commission - European Commission*) <https://ec.europa.eu/info/law/law-topic/data-protection/data-transfers-outside-eu/adequacy-protection-personal-data-non-eu-countries_en> accessed 19 March 2019.

209 See section 2.9 *supra*.

rights in court complicate matters further: privacy law does not clearly describe a minimum level of privacy that should always be maintained; instead, it provides criteria for the balancing of individual privacy against other interests. As a result, it does not offer simple rules for courts to decide cases. Claiming damages for privacy breaches is hampered by the fact that consumers "give it away in exchange for so little".[210]

If data controllers become too powerful, the validity of consumers' and data subjects' consent or their autonomy when entering into privacy contracts can be questioned. Therefore, controllers' increasing power should not remain unchecked. Analogous to the notion of due process in United States law, Gutwirth and De Hert have seen the application of three requirements that serve as checks on the unlimited exertion of power: participation, transparency and accountability.[211] In western societies, many safeguards of fundamental rights at every level of government, for example the protection of suspects in criminal proceedings, can be "decomposed" into these three requirements.[212] Now that privacy contracts and datafication give private companies capabilities similar to those of police, prosecutors or national security agencies when it comes to data collection and use, these requirements and their effects on the underlying power dynamics have also become relevant in contractual relations.

Barnett and Duvall define power as "the production, in and through social relations, of effects that shape the capacities of actors to determine their circumstances and fate".[213] They refine the concept further by qualifying the expression of power (either through interaction or constitution) and the specificity of the social relations through which it works (either direct or diffuse). In short: Power is expressed through interaction if it results from what actors do (like drawing a gun during an argument); it is expressed through constitution if it results from what they are (like their authority or identity). Social relations are direct if parties to the relations are in direct communication with each other (like during negotiations); they are diffuse if their interaction happens as a result of previously defined rules (like when parties'

210 Bruce Schneier, *Data and Goliath: The Hidden Battles to Collect Your Data and Control Your World* (1 edition, W W Norton & Company 2015) 227.

211 Citron (n 125) 1256–1257; Gutwirth and de Hert (n 125) paras 12–13.

212 For examples, see article 3(1) of the Convention on Access to Information, Public Participation in Decision-Making and Access to Justice in Environmental Matters (Aarhus Convention); article 8 of the Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data and articles 6(3) and 13 of the European Convention on Human Rights.

213 Barnett and Duvall (n 94) 39, 45–57.

behaviour is bound or prescribed by law).[214] Using this conceptualisation of power, this contribution examines the following research question:

Can consumer protection law assist in attenuating the shift in power from consumers to data controllers caused by big data?

The following sub-questions will help to answer the research question:

- How does big data cause power to shift?
- Are data protection and privacy law effective in preventing this shift?
- What opportunities does EU consumer protection law offer for addressing the shift?

In answering the main question, this chapter considers only cases where a natural person is both a data subject and a consumer at the same time. The text references the final text of the GDPR unless otherwise specified.

3.2 How big data shifts power towards data controllers

Data controllers use their existing structural power over data subjects in the contracting phase to increase their institutional power after the contract is concluded. Structural power (expressed through constitution in direct social relations) follows directly from the roles actors play, i.e. the roles of suppliers and consumers in the market, and enables the powerful party to limit the capacity of the less powerful party to act in their own best interest. Institutional power (expressed through interaction in diffuse social relations) is the power differential resulting from "constraint(s) that human beings devise to share human interaction".[215]

In the contracting phase, structural power expresses itself in the market as a lack of bargaining power on the consumer side, resulting in non-negotiable terms.[216] This reduces consumers' party autonomy and therefore touches on a key element of private law in Europe.[217] The root cause of this smaller bargaining power is asymmetric

[214] ibid 42–43.
[215] ibid 51–55; North (n 92) 4.
[216] Case 240/98 *Océano Grupo Editorial SA* v *Roció Murciano Quintero and others*, [2000] I-4963, para. 25.
[217] Study Group on a European Civil Code and Research Group on EC Private Law (Acquis Group), *Principles, Definitions and Model Rules of European Private Law: Draft Common Frame of Reference (DCFR). Outline Edition* (Sellier European Law Publishers 2009) 123.

information. For consumers, the cost of information per contract is higher than it is for data controllers, mainly because the controllers unilaterally draft the privacy contracts and reuse them many times. This has led Gomez to state that the primary goal of consumer protection law is to overcome information asymmetries.[218] This higher cost of information, in turn, can be explained by analysing the dynamics of market participation.[219] The uneven distribution of the costs of information and organisation favours data controllers when consumers and controllers decide on contract terms. Individual consumers usually lack expertise and have little to gain by pooling their resources to negotiate a better deal on privacy in each separate contract.

Data controllers then use this structural power to increase their institutional power. As noted before, collection and use of personal data is lawful insofar as a data subject has consented to it, or if the processing is necessary for the performance of a contract. If the consumers' consent allows for their permanent observation, the data controller has obtained a method of exerting power over the consumer.[220] Analysis and actual use of the data further increase this power. If the consumer has agreed to contract terms allowing it, the controller can then grant this power to third parties by using his right of free movement of data. Some of the largest of these third parties, data brokers, are not dealing with consumers directly; this makes the scale of the collection and use of their data less transparent to consumers.[221]

The increase in institutional power can express itself in many ways. Exposing a data subject to targeted advertising is an example of a subtle form of control: a data subjects' deeply personal characteristics can be gleaned from seemingly innocuous data. Such advertising is designed to appeal to personal desires which, although deeply and individually felt, are common to most people and therefore easily discovered.[222] The time frame and context in which these desires come into play in a consumer's life can become apparent by analysing data collected under privacy

218 Fernando Gomez, 'EC Consumer Protection Law and EC Competition Law: How Related Are They? A Law and Economics Perspective' in Hugh Collins (ed), *The Forthcoming EC Directive on Unfair Commercial Practices - Contract, Consumer and Competition law implications* (Kluwer Law International 2004) 193 onward; W David Slawson, 'Standard Form Contracts and Democratic Control of Lawmaking Power' (1970) 84 Harvard Law Review 529, 544.

219 Komesar, *Law's Limits* (n 93) 30.

220 Bentham (n 50) 31; Schneier, 'Metadata = Surveillance' (n 49).

221 Federal Trade Commission, 'Data Brokers: A Call for Transparency and Accountability' (2014) 46 <https://www.ftc.gov/reports/data-brokers-call-transparency-accountability-report-federal-trade-commission-may-2014> accessed 13 February 2019.

contracts and comparing it to previously determined patterns in a larger population (using "machine learning"). For example, a controller may determine whether someone is pregnant by observing a change in their buying patterns.[223]

At least one possible effect of this increased institutional power is the further increase of data controllers' (already larger) structural power, for example, if loans to data subjects living in certain neighbourhoods only become available at discriminatory rates. In this way, the power shift could worsen the existing marginalisation of groups of people.[224] Outside the scope of any privacy contract, data controllers can cooperate with governments to further national security interests in a "surveillance-industrial complex".[225] Finally, the resulting power shift may allow controllers to leave consumers in the dark about the effectiveness of security measures against unlawful processing.

3.3 Data and privacy protection law do not prevent the power shift

EDRi (European Digital Rights), an EU-based advocacy group, asserts that the EU has a "strong, comprehensive and enforceable privacy protection framework".[226] This framework consists of EU data protection law, the Charter of Fundamental Rights of the European Union, Member States' national human rights law and the European Convention on Human Rights. Article 5(1) of the GDPR establishes a number of firmly worded principles governing the processing of personal data such as: lawfulness, fairness and transparency, purpose limitation, data minimisation, storage limitation

222 Vance Packard and Mark Crispin Miller, *The Hidden Persuaders* (Reissue Ed, Ig Publishing 2007) 86–93.

223 Duhigg (n 116) 194.

224 Kate Crawford and Jason Schultz, 'Big Data and Due Process: Toward a Framework to Redress Predictive Privacy Harms' (2014) 55 BCL Rev. 93, 99–101 <http://heinonlinebackup.com/hol-cgi-bin/get_pdf.cgi?handle=hein.journals/bclr55&section=5> accessed 19 March 2019; Cynthia Dwork and Deirdre K Mulligan, 'It's Not Privacy, and It's Not Fair' (2013) 66 Stanford Law Review Online 35, 36–37 <http://www.stanfordlawreview.org/online/privacy-and-big-data/its-not-privacy-and-its-not-fair> accessed 19 March 2019.

225 Committee on Legal Affairs and Human Rights and Pieter Omtzigt, 'Mass Surveillance' (Parliamentary Assembly, Council of Europe 2015) Doc. 13734 29 <http://assembly.coe.int/nw/xml/XRef/X2H-Xref-ViewPDF.asp?FileID=21583&lang=en> accessed 13 February 2019.

226 EDRi - European Digital Rights, 'EU: The Global Standard Setter for Privacy and Data Protection' (EDRi - European Digital Rights 2013) Issue 2 1 <http://edri.org/files/eudatap-02.pdf> accessed 13 February 2019.

and accountability. Article 6(1) limits the number of grounds for lawful processing. If performance of a contract depends on consent, article 7(4) stresses the need to carefully consider whether consent was freely given. Data subjects have the right not to be subjected to profiling (art. 21(1)). Controllers must use principles like "data protection by design and by default" (art. 25). And of course, the Strasbourg and Luxembourg courts guard over fundamental rights, including the right to privacy.

But the complex reality of both data protection and privacy law makes these legal protections less effective. Privacy as a human right is a complex issue because every case is different; the European Court of Human Rights in Strasbourg can only decide individual cases based on all relevant facts, in complex and long proceedings requiring expensive legal representation and thus being not very accessible to individuals. Data protection law seems more easily applicable on its face, because it regulates data controllers' behaviour directly to ensure privacy.[227] But these formal requirements contain very complex standards aimed at specialised operators, intended to be enforced by specialised government agencies (Data Protection Authorities or DPAs). This is not necessarily a shortcoming of the GDPR: regulating controllers' behaviour is one way of keeping the GDPR enforceable, effective and relevant as technology progresses.

Even so, this complexity, combined with the increasing number of privacy contracts, makes consumer participation more difficult as it can make the effects of the GDPR unpredictable for consumers.[228] A few examples:

- Fairness means that data is not collected in secret, that the purpose of the collection is made clear and that data subjects have access to their data.[229] This requirement can improve transparency, but following it to the letter can in fact achieve quite the opposite effect. For example, if a controller provides exhaustive information and updates it several times a year, he effectively increases the cost of information since reading privacy statements has a very real cost.[230] As an example, Microsoft's privacy statement amounted to 35 pages in 2016 and has been updated at least three times between June 2015 and January 2016; its length

227 Paul De Hert and Serge Gutwirth, 'Data Protection in the Case Law of Strasbourg and Luxemburg: Constitutionalisation in Action' in Serge Gutwirth and others (eds), *Reinventing Data Protection?* (Springer Netherlands 2009) 44.

228 Komesar, *Law's Limits* (n 93) 28.

229 European Union Agency for Fundamental Rights, European Court of Human Rights and Council of Europe (n 90) 76.

230 McDonald and Cranor (n 144).

amounted to 52 pages in October 2018.[231] If the costs become too high, consumers may choose not to inform themselves.

- Storage limitation only applies when identifiable data is kept on hand for "longer than is necessary" (art. 5(1)(e), GDPR). However, the drafter of a privacy contract can unilaterally define the purpose and the necessity. To discover what this principle means in a specific context, consumers need to carefully examine all contracts they enter, which they often do not.[232]
- Purpose limitation itself is limited: if a controller wants to reuse personal data previously collected for a different purpose, he "shall" take into account a number of complex factors, including the terms of the privacy contract itself (art. 6(4)(b)).
- Opaque contextual parameters, such as "appropriate technical and organizational measures" determine the accountability of controllers and the "protection by design and by default" requirement (art. 24(1) and 25(1)).
- Data protection impact assessments and data breach notifications should be carried out if there is a "high risk to the rights and freedoms of natural persons" (articles 34(1), 35(1)) but what constitutes a high risk is left undefined.
- Consumers enter into agreements and give consent in very simple or almost imperceivable ways. Ticking one of the ubiquitous "I agree" boxes on a website, and even the state of "technical settings for information society services", such as arcane browser or device settings can constitute consent (recital 32). The right to object to profiling may not apply in these cases (art. 21(1)).
- Finally, the complexity of data protection law encourages consumers to rely on enforcement by DPAs. But that DPAs fall short in enforcing existing data protection was apparently an "open secret" in 2014.[233] If consumers are unaware that enforcement is lacking, this reduces transparency for consumers as well as accountability for controllers.

---

231 Microsoft, 'Privacy Statement' <https://privacy.microsoft.com/en-us/privacystatement/> accessed 13 February 2019.

232 Yannis Bakos, Florencia Marotta-Wurgler and David R Trossen, 'Does Anyone Read the Fine Print? Consumer Attention to Standard Form Contracts' (2014) 43 Journal of Legal Studies 20–21, 31 <http://papers.ssrn.com/abstract=1443256> accessed 13 February 2019.

233 Moerel (n 87) 30.

Another reason for privacy and data protection law's reduced effectiveness for privacy contracts is the fact that the ECHR was originally drafted to protect citizens against their governments in the aftermath of World War II. That the ECHR governs relations between citizens, including contractual relations, has been established in case law but states have a very wide margin of appreciation – wider than in cases against governments.[234] Whether the Charter of Fundamental Rights of the European Union applies to contractual relations seems doubtful at the moment.[235]

Finally, data protection and privacy are not the only fundamental rights recognised in Europe. Freedom of contract, party autonomy and freedom to conduct a business are covered by articles 12 and 16 fundamental rights.[236] A consumer's or controller's appeal on these rights may be used to make permanent observation through privacy contracts lawful. For example, if a consumer enters into a loyalty programme, he "performs" by allowing collection and analysis of personal data, whilst the controller performs by proposing "personalized offers" by him and "selected partners" according to art. 6(1)(b). Based on the term "personalized", a controller can arguably justify collecting data for as long as the contract exists, and on anything that can assist in further segmenting the market to further personalise his offerings. Common (and legal) business practices such as tying (offering two different contracts in one transaction, e.g. one for a "regular" service and another for the processing of personal data) further expand these possibilities.

In short, any practical effect of data and privacy protection law on the power shift associated with big data is reduced by the fact that both work through complex standards instead of simpler rules.[237] This complexity, together with the increasing number of privacy contracts, reduces transparency and the opportunities for participation for consumers as well as accountability for data controllers.

234 See section 2.10.1 *supra*.

235 Eleni Frantziou, 'The Horizontal Effect of the Charter of Fundamental Rights of the EU: Rediscovering the Reasons for Horizontality' (2015) 21 European Law Journal 657, 671 <http://doi.wiley.com/10.1111/eulj.12137> accessed 19 March 2019.

236 See articles 12, 16, Charter of Fundamental Rights of the EU.

237 Pierre Schlag, 'Rules and Standards' (1985) 33 UCLA Law Review 379, 381–390.

3.4 Consumer protection law can help shift power from data collectors to consumers

If a data subject is also a consumer, the European Union aims for a "high level of protection" of his economic activities.[238] The object of protection of consumer protection law is similar to that of data and privacy protection law: they both aim to protect the autonomy of the natural person (in the market for consumer protection; in a moral sense for data and privacy protection).[239] But the concept of protection for consumers is more clear. Where privacy and data protection law involve complex balancing of interests in an endless variety of contexts, consumer protection specifically aims to address power differentials based on information asymmetries in the market. Because of this specific applicability, applying EU consumer protection law to privacy contracts could help shift power back to consumers by improving participation and accountability. This follows from two features of EU consumer protection law: the scope of the fairness criterion and opportunities for participation.

Firstly, the scope of the fairness criterion is wider in consumer protection law than in data protection law. Applying consumer protection law would therefore increase the accountability of data controllers. This follows from the legal texts themselves.

The GDPR, as previously noted, mainly considers the processing of personal data unfair if it happens in secret or if profiling methods are faulty.[240] This is the basis for the GDPR's extensive disclosure requirements.[241] And indeed, mandatory disclosure is also an important regulatory technique in EU consumer protection law. But the resulting transparency is not enough to address substantive unfairness.[242] Therefore, in consumer protection law, fairness instead applies to the terms of the contract and to the way the consumer is persuaded to enter into it. The Unfair Commercial Practices Directive considers a practice unfair if it is "contrary to the requirements of professional diligence and materially distorts or is likely materially to distort the economic behaviour of the average consumer with regard to the product."[243] EU unfair commercial practices doctrine shows substantial similarities with Unfair and

238 Article 38, Charter of Fundamental Rights of the European Union.

239 Gomez (n 218) 193 ff; Helen Nissenbaum, *Privacy in Context: Technology, Policy, and the Integrity of Social Life* (Stanford Law Books 2009) 81–84.

240 See recitals 42 and 48 of the GDPR. For profiling, where it concerns the use of adequate mathematical or statistical procedures to prevent errors, data breaches or discriminatory effects, see recital 71.

241 See for example: art. 5(1)(b), art. 12(1, 3, 5) and art. 13(1-2).

242 Weatherill (n 95) 92–93.

Deceptive Commercial Practices doctrine in the United States.[244] The possible usefulness of EU unfair commercial practices doctrine for data protection may therefore be derived from the fact that its US counterpart is the basis for a number of important data protection cases in the United States.[245]

The Unfair Terms Directive (UTD) regards a non-negotiated term in a contract or consent statement as unfair if "contrary to the requirement of good faith, it causes a significant imbalance in the parties' rights and obligations arising under the contract, to the detriment of the consumer."[246] This means that all rights and obligations are included in establishing unfairness, not just those pertaining to the processing of personal data.[247] The original proposal for the GDPR hinted in this direction: it contained a clause making consent invalid if there was a "significant imbalance between the position of the data subject and the controller".[248] However, the scope of the final provision is far more limited; specific consideration is only given to performance of a contract that is depending on consent, and no longer on all cases where consent is given.

Admittedly, like the earlier examples from the GDPR, fairness in consumer protection law is also a complex standard. However, its application is easier for consumers

243 Article 5(2) and 6(1)(a), 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 and amending Council Directive 84/450/EEC, Directives 97/7/EC, 98/27/EC and 2002/65/EC of the European Parliament and of the Council and Regulation (EC) No 2006/2004 of the European Parliament and of the Council ('Unfair Commercial Practices Directive'), OJ L 149, 11.6.2005, p. 22–39

244 Nico van Eijk, Chris Jay Hoofnagle and Emilie Kannekens, 'Unfair Commercial Practices: A Complementary Approach to Privacy Protection' (2017) 3 Eur. Data Prot. L. Rev. 325, 335.

245 ibid 329–332.

246 Article 3, Council Directive 93/13/EEC of 5 April 1993 on unfair terms in consumer contracts ('Unfair Terms Directive'), OJ L 95, 21.4.1993, p. 29–34.

247 Commission des Clauses Abusives, 'Recommandation n° 2014-02 relative aux contrats proposés par les fournisseurs de services de réseaux sociaux' <http://www.clauses-abusives.fr/recommandation/contrats-de-fourniture-de-services-de-reseaux-sociaux-nouveau/> accessed 18 March 2019; Ellen Wauters, Eva Lievens and Peggy Valcke, 'A Legal Analysis of Terms of Use of Social Networking Sites, Including a Practical Legal Guide for Users: "Rights & Obligations in a Social Media Environment"' (iMinds-ICRI 2013) D1.2.4 64 <https://limo.libis.be/primo-explore/fulldisplay?docid=LIRIAS1709099&context=L&vid=Lirias&search_scope=Lirias&tab=default_tab&lang=en_US> accessed 13 February 2019.

248 European Commission, 'Proposal for a General Data Protection Regulation' (n 81) 45 (art. 7[4].

because it relates to circumstances that they participate in every day. Furthermore, annexes to both the Unfair Terms Directive and the Unfair Commercial Practices Directive give concrete examples. Consumer advocacy groups have been giving guidance on their application, and taking offenders to court, since they came into force.[249]

Applying consumer law's fairness criterion to privacy contracts can expand the accountability of data controllers when compared to only applying the GDPR. Consider the hypothetical case of a provider of a smartphone app enabling the user to use his camera flash LED as a flashlight.[250] The provider could present an agreement in which he grants the consumer a license to use the app in return for which the consumer allows the provider, acting as a data controller, to collect location and usage data to provide advertising for as long as the app is installed.

In terms of data protection law, this case could arguably be made GDPR-compliant by presenting all the relevant clauses and obtaining agreement to them in exchange for a software license. Applying article 7(4) of the GDPR, containing a criterion for determining whether consent is freely given, may not improve matters for the consumer. Because the data collected is not one of the special categories of data as defined by article 9(1), explicit consent may be unnecessary. "Necessary for the performance of the contract" probably suffices to make the processing of personal data lawful, because the criterion of necessity is interpreted in the light of the clauses in the contract. Party autonomy dictates that consumers are free to perform their part by offering their personal data, even if this data is not necessary to turn a phone's LED on or off.

However, such a case would almost certainly violate art. 3, UTD. Allowing a party the opportunity for constant surveillance in exchange for the ability to switch an LED on or off seems like such a bad deal, that the "requirement of good faith" has probably

[249] Drjur Friedrich Bultmann, '30 Jahre Praxis Der AGB-Verbandsklage: Kurzfassung Des Gutachtens Im Auftrag Des Verbraucherzentrale Bundesverbandes' para 14 <http://www.vzbv.de/sites/default/files/mediapics/kurzfassung_gutachten_verbandsklage_2008.pdf> accessed 19 March 2019; Verbraucherzentrale Bundesverband, 'Samsung App-Store: Viele Klauseln Unzulässig' (*Samsung App-Store: Viele Klauseln unzulässig*, 6 June 2013) <https://www.vzbv.de/urteil/samsung-app-store-viele-klauseln-unzulaessig> accessed 13 February 2019.

[250] Nicole Vincent Fleming, 'Sharing Your Location… In a Flash' (*FTC Consumer information blog*, 5 December 2013) <https://www.consumer.ftc.gov/blog/sharing-your-location-flash> accessed 21 March 2019.

not been met. Depending on how the app was advertised, offering the app under these conditions could also be called misleading according to article 6(1)(a) of the Unfair Commercial Practices Directive, insofar as it presents an offer with data protection relevance as a standard software license agreement – especially since consumers hardly ever read software licenses.

Secondly, consumer law offers better participation options than the GDPR when seeking a remedy in court or before an administrative authority. This is often burdensome for consumers, especially against an opponent with large resources: usually, "the haves come out ahead".[251] Limited individual stakes in the outcome of costly proceedings may discourage consumers from bringing a matter to court. Article 11(1) of the Unfair Commercial Practices Directive and 7(2) of the Unfair Terms Directive state that EU member states "shall ensure" that consumer rights organisations can bring an action before the national courts. This allows consumers to pool resources, reducing the cost of information and participation; it also allows consumers to build on previous organisation efforts, reducing the cost of organisation. This improves the dynamics of participation for consumers.[252] The GDPR does not require member states to allow complaints by advocacy groups, it merely allows them to do so (art. 80(2)). Nonetheless, the GDPR does require that member states allow these organisations to represent data subjects in individual proceedings, possibly lowering the cost of legal representation (art. 80(1). Forum choice is handled equally for consumers and data subjects: art. 79(2) of the GDPR allows data subjects to bring proceedings before a court in their country of residence, like art. 16(1), Regulation (EC) No 44/2001 does for consumers.

A more indirect way in which consumer protection law offers better participation options stems from the treaties establishing the EU and the levels of harmonisation within the EU that follow from them. EU data protection law is based on conferral of competence by the member states, whereas consumer protection law is based on shared competence.[253] As a result, member states cannot increase the level of protection that EU data protection law provides unless it is expressly allowed, whereas for consumer protection law this is possible unless it is expressly forbidden.[254] This

[251] Galanter (n 91).

[252] Komesar, *Law's Limits* (n 93) p.30.

[253] Article 39, TEU; articles 2(f), 12, 16, 114 and 169, Treaty on the Functioning of the European Union (TFEU).

[254] European Commission, 'Proposal for a General Data Protection Regulation' (n 81) 5–6; see also Case 101/01 *Bodil Lindqvist* [2003] ECR I-12992, para. 66-67. See art. 9(5), GDPR for an

can help consumers: for them, participation in legislation efforts is much harder at the EU level than at the national level.[255]

Thus, applying consumer protection law to privacy contracts can increase accountability for data controllers and offer better participation options for consumers. Both effects will decrease the institutional power of data controllers in favour of consumers.

3.5 Conclusion: improve enforcement of consumer protection law

When compared to the GDPR, existing EU directives regarding unfair contract terms and unfair commercial practices can increase the accountability of data controllers and offer more effective participation options for consumers. This is important in addressing the increase in institutional power that data controllers stand to gain from big data.

However, this possibility can only materialise if consumer protection law is effectively enforced. Determining the effectiveness of the current enforcement regime is not easy, but in 2012 the Commission claimed that "(r)edress and enforcement mechanisms need to be further improved" and launched the European Consumer Agenda, partly to achieve this. The commission also identified perceived low individual stakes as one of the reasons why consumers often do not seek redress.[256] Apparently, lack of effective enforcement and low individual stakes similarly affect the effectiveness of both consumer protection and data protection law. Under these circumstances, expecting beneficial effects from applying consumer protection law without increasing enforcement efforts can only lead to disappointment. Ensuring proper coordination between national and European authorities for data protection and consumer protection may also be needed. Having two or even more competent authorities in each member state on the subject of privacy contracts may not have any

example where member states can increase the level of protection; See art. 8, 8a, UTD for an example of the greater freedom that consumer protection law allows. Recent EU consumer protection law tends to rule out this option. See art. 4, Consumer Rights Directive and art. 3(5), Unfair Commercial Practices Directive.

255 See section 2.9 *supra*.

256 European Commission, 'A European Consumer Agenda - Boosting Confidence and Growth' (n 205) para.3.4.

beneficial effect if this joint competence leads to indecision, turf wars or other intra-governmental inefficiencies.

At the same time, very strict enforcement has its own risks and limits. The power shifts associated with big data are too complicated to be addressed only by applying consumer protection law to privacy contracts. Yes, putting consumers under surveillance will become easier with time according to Zimmermann's law. But big data is also driving important innovations and, in an important way, datafication is the price we pay for automation. Billing, correction of errors and malfunctions, and detection of hacking and crime all rely on data generated by automated processes – "It's impossible to overstate the importance of logging".[257] Any well-intentioned effort to suppress the creation, storage and analysis of event logs – in other words, to suppress datafication – could disempower both consumers and data controllers, as it takes away their opportunity to construct or counter evidence of mistakes or wrongdoing.[258] Data streams are also becoming a way of personal expression, e.g. in the "quantified self" movement, which means that curtailing their creation and use can interfere with yet another fundamental right.[259] Addressing big data's power shifts by narrowly focusing on privacy contracts can cause unforeseen power shifts all by itself.

Nonetheless, spirited enforcement of consumer protection law for privacy contracts seems like the way forward. Both the Unfair Terms Directive and the Unfair Commercial Practices Directive offer open norms with ample possibilities to develop a nuanced approach. Controllers who necessarily have access to data streams on many aspects of consumers' lives, like banks and telecommunications providers, should probably be prevented from seducing consumers to allow permanent observation all too easily. On the other hand, consumers should have a reasonable amount of freedom to enter into contracts with providers of specialised data-intensive services. A nuanced approach has a lower risk of negatively affecting related fundamental rights

[257] Vassilis Prevelakis and Diomidis Spinellis, 'The Athens Affair' (2007) 44 Spectrum, IEEE 26, 31.

[258] This is closely related to the "legitimate interest" ground for lawful processing of personal data (art. 6(1)(f), GDPR).

[259] Dawn Nafus and Jamie Sherman, 'This One Does Not Go Up To 11: The Quantified Self Movement as an Alternative Big Data Practice' (2014) 8 International Journal of Communication 1784, 1793 <http://ijoc.org/index.php/ijoc/article/view/2170> accessed 13 February 2019.

and halting innovation, than blanket bans on the generation, storage and use of data. Of course, improving the enforcement of data protection law will also help.

Increasing enforcement efforts will certainly have a cost. Member states will have to provide additional funding; they also need to strengthen co-ordination between consumer- and data protection authorities, both at the national and the EU level. It may also be necessary to improve the dynamics of participation for consumers. For example, targeted subsidies for consumer and privacy advocacy groups at the national level, aimed at representation both in the lawmaking process, in civil society and in the national courts, could somewhat offset consumers' costs of information and organisation. This would complement similar subsidies at the EU level in member states that do not currently subsidise these efforts.[260]

But the benefits may very well outweigh the costs. The more consumers feel their privacy really is protected and enforceable, the faster industrialised societies can collectively benefit from datafication. Addressing the power shift associated with big data will therefore be an important part of Europe's economic future.

[260] Art. 3(1)(b), European Parliament and Council Regulation (EU) No 254/2014 of 26 February 2014 on a multiannual consumer programme for the years 2014-20 and repealing Decision No 1926/2006/EC, [2014] OJ L84/42.

4 Rear view mirror, crystal ball

Exploring the future of data protection through recent developments of environmental protection law

Prelude

The following chapter originated as a conference paper. My education and expertise as a fire officer and risk management policy advisor from 1996 until the present day have served as the basis for the approach in this chapter. A preliminary version was presented at the 2015 Amsterdam Privacy Conference on 25 October as "Privacy: The cost of doing business" in the "Commercial value of Privacy" track. The call for papers for this conference was published February 2015 and is archived at http://apc2015.net/content/call-papers-0.

The proposal for my contribution was submitted 11 March. The proposal was accepted on 1 May 2015. The conference paper was submitted 14 July.

Acting on the advice of my supervisors and after consulting with the Editor of the Computer Law and Security Review, the conference paper was extensively rewritten to make it suitable for publication in a peer-reviewed journal.

Timeline and citation

The first version of the revised paper was submitted to the Editor-in-Chief of the Computer Law and Security Review (CLSR) 16 December 2016. The paper was conditionally accepted January 17, 2017. The revised version was submitted 15 March 2017. It was accepted 24 March and published online on 30 May 2017. It appeared in print in the October 2017 edition of CLSR. The article was not published under an open access license.

The OSCOLA reference to the original article is.

> Rhoen M, 'Rear View Mirror, Crystal Ball: Predictions for the Future of Data Protection Law Based on the History of Environmental Protection Law' (2017) 33 Computer Law & Security Review 603

4.1 Introduction: Looking into a rear view mirror

Data protection law and environmental protection law are both a response to technological development. Environmental law aims to mitigate the adverse effects of industrialisation and its origins can be traced back to nuisance law from Roman times.[261] These effects can manifest themselves at different levels, from the immediate surroundings of industrial activity, to the entire world. Similarly, data protection law is societies' response to automation. The first legal response to the processing of personal data stems from the 1970s and 1980s.[262] Until now, data protection appears to have been treated mainly as an individual fundamental human right, whereas environmental law originated as a means for protecting communities and societies.

Environmental law has been expanding towards human rights law. Environmental protection has been added to the EU catalog of fundamental rights in 2010. It is now not only a duty of care for societies but also an individual right.[263] In contrast, expansion of the scope of data protection law towards protecting entire societies is not so easily visible. In the recently adopted General Data Protection Regulation (GDPR), the right to data protection is still mainly seen as an individual right, even though the advent of "big data" (resulting from the seemingly unstoppable trend of "datafication")[264] can be expected to have a serious impact on entire societies. An explanation may be that the risk to individual privacy is easily visible. Many of these risks occur when data subjects act as consumers: the permanent collection of metadata on telecommunications, shopping, media streaming and social networks all rely on consumer contracts for lawfulness, and can reveal many deeply personal aspects of one's life.[265] But the risk of large-scale surveillance to a free and democratic society may be equally or even more important. Datafication has been called "the pollution problem of the digital age."[266] Big data will cause power to shift from data

261 Ulpianus 17 ad edictum, D. 8,5,8,5-7.

262 See, for example, the Hessische Datenschutzgesetz of 7 October, 1970, the Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data, done at Strasbourg (Strasbourg Convention), 28 January 1981 (articles 1 and 8) and the German Volkszählungsurteil [1983] BVerfGE 65,1 (Bundesverfassungsgericht)

263 Article 37, Charter of Fundamental Rights of the European Union, [2010] OJ C 83/02 ("the Charter"). Note that this right is stated in general terms and not in individual terms.

264 Mayer-Schönberger and Cukier (n 115) 29.

265 Richards and King (n 118) 44.

266 Bruce Schneier, 'The Battle for Power on the Internet' [2013] *Internet and Security* 19 <https://www.theatlantic.com/technology/archive/2013/10/the-battle-for-power-on-the-internet/280824/> accessed 13 February 2019.

subjects to data controllers, regardless of whether the controller is a government or a private party.[267] Big data can increase or cement social inequalities.[268] It risks reintroducing the chilling effects on personal freedoms historically associated with police states.

According to Baldwin, regulation can be seen as "being inherently about the control of risks".[269] Social theories about risk appear to have influenced environmental law much more than data protection law. The EU's recent data protection legislation efforts were almost completely focused on individual rights.[270] But to achieve a balanced data protection policy, considering the risks and benefits of datafication to society is just as necessary. Like in environmental law, the rights of data subjects will always necessarily be weighed against the rights of other data subjects, data controllers and the interests of society as a whole.[271] Already in 2006, Hirsch indicated a number of analogies between environmental law and data protection law, mainly focusing on regulatory strategies.[272]

This chapter focuses on two theories from the social sciences that have discernibly influenced environmental law and policy: Beck's theory of the risk society and Perrow's theory of normal accidents. Because the risks of datafication and industrialisation both result from technological progress and both pose risks at the

---

267 See article 3(1), Strasbourg Convention; Bruce Schneier, 'The Myth of the "Transparent Society"' (*WIRED*, 3 June 2008) <https://www.wired.com/2008/03/securitymatters-0306/> accessed 20 March 2019 and section 3.3 *supra*.

268 Cathy O'Neil, *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy* (Crown 2016) (passim); FJ Zuiderveen Borgesius, 'Improving Privacy Protection in the Area of Behavioural Targeting' (Universiteit van Amsterdam 2014) 118 <http://hdl.handle.net/11245/1.434236>.

269 Baldwin, Cave and Lodge (n 82) 83.

270 See, for example, European Commission, 'Impact Assessment, Accompanying the Document "Regulation of the European Parliament and of the Council on the Protection of Individuals with Regard to the Processing of Personal Data and on the Free Movement of Such Data (General Data Protection Regulation)" and "Directive of the European Parliament and of the Council on the Protection of Individuals with Regard to the Processing of Personal Data by Competent Authorities for the Purposes of Prevention, Investigation, Detection or Prosecution of Criminal Offences or the Execution of Criminal Penalties, and the Free Movement of Such Data"' (European Commission 2012) Accompanying document SEC(2012) 72 final 29: "In a free and democratic society, the individual must have reassurance that fundamental rights are respected."

271 Bundesverfassungsgericht, *Volkszählungsurteil* [1983] BVerfGE 65,1 para C II 1 a.

272 Dennis D Hirsch, 'Protecting the Inner Environment: What Privacy Regulation Can Learn from Environmental Law' (2006) 41 Ga. L. Rev. 1.

societal level, considering the application of these social theories to data protection seems appropriate. In this context, this contribution aims to answer the following question.

What avenues for the future development of consumer data protection law can be gleaned from the application of risk society theory and normal accident theory in environmental protection law?

This question will be approached through the following sub-questions.

- How is risk society theory relevant to data protection law and policy?
- How is normal accident theory relevant to data protection law and policy?
- To what extent does the General Data Protection Regulation apply both theories in its risk management model?
- What does this analysis indicate for the future of the consumer data protection debate?

The discussion will be limited to the jurisdiction of the European Union and will consider the text of the General Data Protection Regulation (GDPR), which shall apply from May 25, 2018. This chapter deals only with consent and performance of a contract (article 6(1)(1–2)) as the basis for lawfulness of processing, excluding consent and contracting by minors. The risks associated with the processing of sensitive personal data as defined in article 9(1), or profiling based on sensitive data (article 22(4)), will be considered in chapter 5 below.

4.2 Big data and the risk society

4.2.1 Risk society theory

The essence of risk society theory is best summed up by Beck himself. The first sentences of his seminal work read:

> "In advanced modernity the social production of wealth is systematically accompanied by the social production of risks. Accordingly, the problems and conflicts relating to distribution in a society of scarcity overlap with the problems and conflicts that arise from the production, definition and distribution of techno-scientifically produced risks."[273]

[273] Ulrich Beck, *Risk Society: Towards a New Modernity* (Sage Publications 1992) 19.

Beck describes how, starting somewhere around the 1970s, society has progressed from the so-called "first modernity" into "second modernity". In first modernity, scientific innovations were mainly aimed at removing natural hazards; in the current second modernity (or "advanced modernity"), risks are mainly the result of scientific progress itself ("reflexive modernisation"). Environmental risks such as the depletion of the ozone layer, mass extinction of species and anthropogenic climate change find their origins in scientific successes resulting from, for example, attempts to defeat hunger or to protect humans from the elements. Beck calls these risks "modern risks". These risks carry extreme catastrophic potential, but they cannot be measured accurately. The place of their origin is not always exactly known. These characteristics can make traditional risk coping mechanisms like insurance less suitable.[274] Not all reflexive risks are created equal: Klinke and Renn have distinguished six "risk classes", based on the different degrees to which the probability and the effects of a risk can be known or estimated. All of their risk classes are applicable to modern risks. They are named after figures from Greek mythology, asserting that many of these stories illustrate the transition from an economy of hunter-gatherers to an agricultural economy, and that this transition had similar effects on the perception of risk as our transition to an industrialised society.[275]

The second, and (at least according to Beck) often overlooked aspect of risk society theory is institutionalised individualism.[276] Beck asserts that in second modernity, social groupings like class, gender and family have gradually become less important in the exercise of fundamental rights, paid employment and education. The individual is increasingly detached from social groups. But because inequalities do not disappear, the individual is forced into shifting alliances for each particular issue.[277] The advent of the second modernity therefore entails a "profound systemic transformation of the political." Centralised political culture will lose its power in the enforcement of civil rights, and the organisation of social structure will increasingly happen through sub-

---

274 ibid 22.

275 Andreas Klinke and Ortwin Renn, 'A New Approach to Risk Evaluation and Management: Risk-Based, Precaution-Based, and Discourse-Based Strategies' (2002) 22 Risk Analysis 1071, s 3.4 <http://onlinelibrary.wiley.com/doi/10.1111/1539-6924.00274/full> accessed 13 February 2019.

276 Ulrich Beck and Elisabeth Beck-Gernsheim, *Individualization: Institutionalized Individualism and Its Social and Political Consequences* (1st edition, SAGE Publications Ltd 2002) xxi; Beck, *Risk Society* (n 273) 85–150.

277 Beck, *Risk Society* (n 273) 100.

politics.[278] This means that decisions affecting civil rights will increasingly take place outside the institutions of democracy. Beck sees this not so much as an undermining of the individual's participation opportunities, but as an enhancement — a view that appears to be at odds with (for example) Galanter's analysis of how decision-making processes tend to favour repeat players with large resources.[279]

From this, two of Beck's conclusions are relevant. Firstly, science alone can no longer sufficiently decide on the acceptability of risk.[280] Because science is not capable of providing accurate insights in the risks that science itself generates, Beck suggests that science itself must install "brakes and a steering wheel" into its processes.[281] Secondly, political acceptance becomes essential in deciding about modern risk.[282] Because of the lack of enduring social cohesion in the risk society, political acceptance means that the individual must have the right to have his say in which risks are acceptable and which are not.

4.2.2 Risk society and environmental law

Three environmental law instruments show distinctive features derived from risk society theory: the Rio Declaration of 1992, the Aarhus Convention of 1998 and the Charter of Fundamental Rights of the European Union of 2010.[283]

Principles 9 and 10 of the Rio Declaration encapsulate both elements of the risk society thesis. Principle 9 states that improving and sharing of scientific understanding and technology is essential for endogenous capacity-building; principle 10 states that environmental protection issues are best dealt with, not merely by applying science and technology, but "with participation of all concerned

278 ibid 190.

279 Galanter (n 91) 125.

280 Beck, *Risk Society* (n 273) 156.

281 ibid 180. This suggestion may not be achievable in the context of the scientific method as it is currently being practiced. The scientific method does, however, seem suitable to investigate the possible effectiveness of enforcement mechanisms for political decisions.

282 ibid 168.

283 United Nations Conference on Environment and Development, 'Rio Declaration on Environment and Development' <http://www.unesco.org/education/pdf/RIO_E.PDF> accessed 13 February 2019; United Nations Economic Commission for Europe, 'Convention on Access to Information, Public Participation in Decision-Making and Access to Justice in Environmental Matters Done At Aarhus, Denmark, On 25 June 1998'; Charter of Fundamental Rights of the European Union, [2010] OJ C 83/02.

citizens". This participation should be facilitated by access to information concerning the environment, and "effective access to justice."

The Aarhus Convention, mentioning principle 10 in its second recital, provides for public access to environmental information in general (article 5), public participation in decisions on specific activities (article 6); public participation in policy development and "generally applicable legally binding normative instruments" (articles 7-8) and individual access to justice (article 9). The Convention reflects individualisation: the right to access to justice applies to every member of the public or to "the public concerned", with the status of non-governmental organisations and pressure groups derived from the status of individuals. The Aarhus Convention links "scientific, industrial and governmental elites and the ordinary public/citizen affected by scientific, industrial and/or economic change."[284]

The Charter of Fundamental Rights of the European Union has incorporated environmental protection in its catalog (under title IV, "Solidarity"), as a duty of care of the institutions of the EU (article 37). This makes determining whether article 47 of the Charter (granting the right to an effective remedy and a fair trial) applies, a bit less straightforward. However, the EU is a signatory to the Aarhus Convention, making it clear that the EU must provide access to justice in the case of any individual decision or policy decision on environmental protection in accordance with article 9 of the Convention.[285]

The adoption of requirements for transparency, participation and accountability (access to justice) for environmental decisions underscores the development of the right to environmental protection towards an individual, fundamental right. Applying these requirements is firmly entrenched legal practice in western societies in matters where fundamental rights are subject to interference.[286] It is also a way to prevent the possibility that individuals have to rely on sub-politics as their primary means of participation — an approach that Beck seemed to have been in favour of, considering it a "more direct route to political engagement".[287] In the case of the Aarhus

[284] Deiniol Jones, 'Solidarity and Public Participation: The Role of the Aarhus Convention in Containing Environmentally Induced Social Conflict' (2008) 20 Global Change, Peace & Security 151, 151, 154 <http://dx.doi.org/10.1080/14781150802079706> accessed 13 February 2019.

[285] [2005] OJ L124/4.

[286] See, for example, the European Convention on Human Rights. articles 6, 8 and 13. See also Gutwirth and de Hert (n 125) paras 12-13; Citron (n 125) 1256–1257.

[287] See Mythen (n 100) 160–161.

Convention, their application empowers individuals: it reduces the possible subversion of democracy's decision-making processes by expanding sub-politics with access to the institutions of the Rechtsstaat.

4.2.3 Reflexive modernisation in action: The Seveso III-Directive

A clear example of the EU applying the Aarhus Convention, and thus of reflexive modernisation is discernible in the directive named after the Seveso environmental disaster of July 10, 1976, aimed at the prevention of major accidents.

Like its predecessor from 1996, the Seveso III Directive aims for a high degree of transparency regarding the operation of chemical plants where major accidents can occur. It specifies a generous minimum of information that must be made available to the public. Public participation is required whenever government decisions are given on either zoning or plant operation, regardless who initiates the decision making process. Every government decision can be brought before a judge for review (article 23(b)). Requirements for public participation and transparency can be found with respect t.

- the drafting of an emergency response plan (article 12(5)).
- the establishment's major accident prevention plans and policies (article 14(2)(b)).
- the keeping of records by the competent authority (article 14 and annex V),
- decisions regarding permits and land use for these establishments (article 15),
- disclosing the occurrence of a major accident (article 17(e)).

Assuming that knowledge of the risks of major accidents and the possibilities for emergency response gives citizens a sense of personal control, the Seveso III Directive can help increase society's trust in institutions: Ter Huurne and Gutteling have established that "institutional trust and perceived personal control are strongly correlated."[288] Nevertheless, the Directive appears not to have completely satisfied society's need for participation. Although access to relevant information for the public is deemed sufficient, experts outside industry and government have stated that

[288] Ellen FJ ter Huurne and Jan M Gutteling, 'How to Trust? The Importance of Self-efficacy and Social Trust in Public Responses to Industrial Risks' (2009) 12 Journal of Risk Research 809, 819 <http://dx.doi.org/10.1080/13669870902726091> accessed 13 February 2019.

the quality of the dialogue between members of the public, their organisations, and the establishments and public authorities must still be improved. Still, many actors in the chemical industry and competent authorities seem to agree that the Directive (actually, its predecessor, the Seveso II-Directive) has had a positive effect on safety in the affected establishments.[289]

4.2.4 Big data and the risk society

There seems to be consensus among scholars that big data is a possible source of risk. Among many others, Mayer-Schönberger and Cukier, Schermer and O'Neil have given examples each with their own emphasis.[290] Kerr and Earle qualify the use of sorting algorithms as a way for data controllers to reduce risk for themselves, but this sorting can introduce new risks for data subjects. Citron and Pasquale identify three problems with social sorting based on the processing of personal data: opacity to data subjects, arbitrary assessments of data subjects and a disparate impact on traditionally disadvantaged groups.[291] Using advanced algorithms for the analysis of large datasets can reveal more than individual behaviour or personal thoughts.[292] It can also reveal "special categories of data" without processing them directly, thereby weakening the protection that data protection law aims to offer.[293] As an example, grouping people with similar shopping habits or observing shifts in buying patterns can easily reveal special categories of personal data like religious observance or

[289] O Salvi and others, 'F-Seveso: Study of the Effectiveness of the Seveso II Directive (Final Report)' (European Virtual Institute for Integrated Risk Management (EU-VRi) 2008) Contract n°070307/2007/476000/MAR/A3 37, 41 and 54 <https://relevant.nl/download/attachments/4096340/seveso_report.pdf> accessed 13 February 2019.

[290] Mayer-Schönberger and Cukier (n 1) ch 8; Bart Willem Schermer, 'Risks of Profiling and the Limits of Data Protection Law' in Bart Custers and others (eds), *Discrimination and Privacy in the Information Society* (Springer-Verlag 2013); O'Neil (n 268) (passim).

[291] Ian Kerr and Jessica Earle, 'Prediction, Preemption, Presumption: How Big Data Threatens Big Picture Privacy' (2013) 66 Stanford Law Review Online 65, 69 <http://www.stanfordlawreview.org/online/privacy-and-big-data/prediction-preemption-presumption> accessed 13 February 2019; Danielle Keats Citron and Frank A Pasquale, 'The Scored Society: Due Process for Automated Predictions' (2014) 89 Wash. UL Rev. 1, 10–16 <http://papers.ssrn.com/abstract=2376209> accessed 19 March 2019.

[292] Zeeshan Aleem, 'All the Secret Ways You're Being Tracked That You Don't Even Realize' (*Mic*, 23 March 2015) <https://mic.com/articles/113078/all-the-secret-ways-you-re-being-tracked-that-you-don-t-even-realize> accessed 13 February 2019.

[293] Article 9(1), GDPR.

pregnancy.[294] Permanent observation of large parts of the populace can therefore affect entire societies: it could deter eccentric behaviour, lead to social domination and limit free expression.[295]

The risks associated with permanent and ubiquitous surveillance are hardly catastrophic in the physical sense. But pervasive and permanent surveillance accompanied by social sorting based on datafication would still pose a severe threat to a democratic society. Kerr and Earle have indeed linked big data to Beck's notion of modern risk, especially when the prediction potential of algorithms is used for "various new forms of social preemption", possibly resulting in discriminatory presumption.[296] Seen in this way, the risks of datafication fit Beck's description of a modern risk: they are a by-product of the highly successful digital revolution, but they come hand-in-hand with risks for entire societies.

As is typical for modern risks, the risks of datafication do not have a clear point of origin. In essence, datafication is merely the keeping track of events occurring in automated processes. Not keeping track of automated processes is not an option: "It's impossible to overstate the importance of logging":[297] without it, billing, correction of errors and malfunctions, and detection of hacking and crime become almost impossible. The possibility of surveillance is essentially a side-effect. The qualification of this side-effect as a modern risk is underscored by the fact that the effects of these risks are not easily quantifiable and that they amount to a new kind of individualisation of society. It is no longer necessary to rely on "average" or approximate qualifications of data subjects if access to their personal data can reveal a person's exact properties.[298]

Assuming that datafication poses modern risks, Klinke and Renn then offer additional insights for categorisation. Applying their catalog of risk classes, the risks of big data could fall in several categories. If one considers both the probability and the extent of damage that big data can cause to be uncertain, it would be an example of the "Pythia"-class. If one would rate the disaster potential as high, but the probability of disaster actually occurring as uncertain, big data poses a risk of the Cyclops-class. If,

294 Duhigg (n 116) ch 7; René Bogaarts and Wilco Dekker, 'De Dagelijkse Volkstelling' *De Volkskrant* (21 February 1998) 50.

295 Neil M Richards, 'The Dangers of Surveillance' (2013) 126 Harvard Law Review 1934; Bentham (n 50) 35–37.

296 Kerr and Earle (n 291) 69.

297 Prevelakis and Spinellis (n 257) 31.

298 Mayer-Schönberger and Cukier (n 115) 30–31.

like the CJEU's Advocate General Cruz Villalón has asserted, the adverse effects of datafication are inherent in the collection and storing of data per se, this may increase the probability of adverse effects to "certain"; datafication could then be an example of a "Cassandra"-class risk, where the delay between trigger and effect is very long so that the risks are being downplayed.[299]

Consistent with Beck's primary assertion, the distribution of these risks then becomes an important question. At least in the European Union since the invalidation of the Data Retention Directive, the general principles relating to the processing of personal data appear quite equitable and adhering to them will arguably result in a fair distribution of risk.[300] The principles of lawfulness, fairness, transparency, purpose limitation, data minimisation, accuracy, storage limitation, integrity and confidentiality, and accountability provide useful guidelines in situations where the distribution of power and limits to institutional behaviour are already clearly defined, as is the case in employment law, administrative law and criminal law.

However, in settings where legal constraints on actors are less prominent, data subjects may not enjoy the same level of protection. In the consumer market for example, European data protection law appears to shift the risk of data processing towards the data subject. Specifically, personal data regarding consumers tends to be processed based on the necessity for the performance of a contract to which the data subject, acting as a consumer, is a party.[301] As a result of the cost of information and organisation working in favour of producers, consumers only have very limited bargaining power when they enter into these "privacy contracts".[302] Data controllers can therefore unilaterally set the terms of these contracts, undermining the party autonomy of consumers despite the EU's extensive consumer protection framework.[303] Avoiding privacy contracts is difficult. They are linked to enjoying the ownership of a "smart" or connected modern device (a personal computer, a mobile telephone, a television set or a car) and to using modern services like social networks, instant

[299] Joined Cases C-293 and C-594/12 Digital Rights Ireland v Minister of Communications and Kärntner Landesregierung v Seitlinger and others [2014] ECLI:EU:C:2014:238, opinion of AG Cruz Villalòn paras 65-66; Klinke and Renn (n 275) 1080–1081.

[300] The use of data sets by police and national security services is a possible exception. See Joined Cases C-293 and C-594/12 *Digital Rights Ireland v Minister of Communications* and *Kärntner Landesregierung v Seitlinger and others* [2014] ECLI:EU:C:2014:238 [51]

[301] Art. 6(1)(a-b), GDPR. A consumer is "any natural person who is acting for purposes which are outside his trade, business or profession" (Art. 2(b), Directive 93/13/EC).

[302] Verhelst (n 119) ch 3.

[303] See section 3.2 *supra*.

messaging, streaming entertainment or retail loyalty programs. Consistent with both Beck's prediction and Galanter's analysis, decision-making on important civil rights is effectively relegated to sub-politics — i.e. the market — where consumers have little power and few opportunities for participation.

The general principles of data protection law are therefore less effective in the consumer market. Privacy contracts are usually drafted "by lawyers for lawyers",[304] either in very wide terms, or in very lengthy detailed terms. In both cases, it is difficult to imagine any form of primary or secondary use of personal data that would fall outside of the scope of such an agreement, even if explains the rights of the consumer in "clear and plain language" (art. 12 GDPR).[305] Thus, the principle of lawfulness is easily met.[306] The principles of purpose limitation, data minimisation, storage limitation and accountability may be less effective, because they are heavily influenced by the terms of the agreement. The effectiveness of one's right to erasure (article 17, GDPR) may also be limited. Firstly, asking for erasure can be a difficult task because personal data is typically stored in hundreds of databases;[307] secondly, now

304 Fred H Cate and Viktor Mayer-Schönberger, 'Notice and Consent in a World of Big Data' (2013) 3 International Data Privacy Law 67, 67 <http://idpl.oxfordjournals.org/content/3/2/67> accessed 19 March 2019.

305 For example, Microsoft's "Privacy statement", linked to the license agreement of all their commercial software, consists of more than 23.000 words. See also Brendan van Alsenoy and others, 'From Social Media Service to Advertising Network: A Critical Analysis of Facebook's Revised Policies and Terms, v.1.3' (ICRI - The Interdisciplinary Centre for Law & ICT, Katholieke Universiteit Leuven 2015) <https://www.law.kuleuven.be/citip/en/news/item/facebooks-revised-policies-and-terms-v1-3.pdf> accessed 13 February 2019 for a 100-page analysis of Facebook's terms and conditions.

306 Lawfulness, in this example, should follow article 6(1)(b) GDPR, which is based on the freedom to associate in civil matters and the freedom to conduct a business (articles 12 and 16, Charter of Fundamental Rights of the European Union). Judicial review of such contracts could involve a collision between fundamental rights. Assuming the absence of precedence among fundamental rights, a fair balance must be struck in each individual case: Axel Springer AG v. Germany (2012), ECLI:CE:ECHR:2012:0207JUD003995408, para 84. In this balance, the right to data protection does not automatically trump the right of a natural person or a business to contract for services that are based on the processing of personal data and knowledge creation. Opinions of Independent Supervisory Authorities, the Board or its closest predecessor, the Article 29 Working Party, do not meet the criteria for judicial review of article 47, Charter or article 13, ECHR. This does, of course, not imply that the terms of the contract or the commercial practice through which it was concluded are therefore fair.

307 In 2009 in the Netherlands, Schermer and Wagemans estimated that natural persons were represented in 250 to 500 databases 'Onze Digitale Schaduw. Een Verkennend Onderzoek

that datafication covers wide areas of daily life, a consumer will generate a significant amount of data again soon after the original data is erased.

The combined effect of these factors will make it difficult for consumers to escape surveillance, social sorting or profiling; where data processing is based on a consumer agreement, the right to object against profiling may not even apply. The scope of the important principle of "data protection by design and by default" (article 25) will be determined to a large extent by the terms of the agreement. If, for example, this agreement specifies that a consumer 'performs' by giving access to personal data,[308] this is usually to enable effective profiling for targeted advertising and personalisation of commercial offers. In this context, efficient segmentation of the market could preclude data minimisation or pseudonymisation; the reasonable expectation that newer algorithms could make better use of older data would make long-term storage of personal data "appropriate", and a sufficiently wide definition of "commercial offers" would make this storage appropriate for "each specific purpose". Stating that profiling data would only be shared with "selected commercial partners" could satisfy the requirement that the personal data is not made accessible to an "indefinite number of natural persons".[309]

4.3 Big data and normal accident theory

4.3.1 Normal accident theory

Perrow's Normal Accident Theory (NAT) characterises systems along two axes: whether their interactions are linear or complex, and whether they are loosely or tightly coupled.[310] Even though the emphasis of Perrow's original book was on

Naar Het Aantal Databases Waarin de Gemiddelde Nederlander Geregistreerd Staat.' (Considerati 2009) 40 <https://autoriteitpersoonsgegevens.nl/sites/default/files/downloads/rapporten/rap_2009_onze_digitale_schaduw.pdf> accessed 13 February 2019.

308 Recital 13, European Commission, 'Proposal for a Directive of the European Parliament and of the Council on Certain Aspects Concerning Contracts for the Supply of Digital Content' (European Commission 2015) COM(2015) 634 final; 2015/0287 (COD) <https://ec.europa.eu/transparency/regdoc/rep/1/2015/EN/1-2015-634-EN-F1-1.PDF>.

309 A more optimistic view can be found with Georgia Skouma and Laura Léonard, 'On-Line Behavioral Tracking: What May Change After the Legal Reform on Personal Data Protection' in Serge Gutwirth, Ronald Leenes and Paul de Hert (eds), *Reforming European Data Protection Law* (2015 edition, Springer 2014) 56.

310 Perrow, 'Accidents, Normal' (n 103) 33–34.

technological hazards (risks, as Beck would call them), he made it clear that the categorisation of systems along these two axes is applicable to "the organisational world" in its entirety: his examples included universities and the post office as well as nuclear and chemical plants and space missions.[311]

Chemical and nuclear plants are examples of complex, tightly coupled systems. They entail a great number of interactions that are not easily visible (complexity); due to this complexity, designers of these systems have to expect that operators follow very strict rules to prevent anomalous behaviour. Deviation from these rules can quickly lead to uncontrollable escalation with little chance of correction: they are strongly time-dependent (tightly coupled).[312] This is not necessarily a design flaw. Perrow concludes that it is in the nature of tightly coupled, complex systems to have incidents escalate to accidents with catastrophic potential: they are "system accidents" (in contrast, if systems are more linear or less tightly coupled, they typically display "component failures".) Perrow asserts that, even though system accidents are rare events, they are nevertheless inherent and unpreventable in complex and tightly coupled systems. Therefore, system accidents are normal accidents.[313]

Many accident investigations may conclude that an accident is caused by "operator error", but Perrow offers a different view. Operators are usually more directly exposed to the risk of a system than anyone else and they are often under pressure to ignore safety precautions to work faster — but if this causes injury or an accident, this may be considered "their own fault." In complex, tightly coupled systems, operators are rarely the root cause of accidents. Much more likely, they are the first-party victims.[314]

Perrow also asserts that personal control over our environment is being steadily eroded by systems that we participate in, or are passively affected by.[315] Similar to Beck, he concludes that social decision-making on risks posed tends to be determined by "the power to impose risks on the many for the benefit of the few."[316] In response to the increase of industrial and technological risks, risk assessment has emerged as a

311 Perrow, *Normal Accidents: Living with High-Risk Technologies* (n 101) 96–100 and figure 3.1; see also Karl E Weick, 'Normal Accident Theory As Frame, Link, and Provocation' (2004) 17 Organization & Environment 27, 29.

312 Perrow, *Normal Accidents: Living with High-Risk Technologies* (n 101) 93–96.

313 ibid 8.

314 ibid 67.

315 ibid 313.

316 ibid 306.

science mainly to "legitimise the decisions of elites in private and public sectors."[317] To make risks acceptable to the authorities, to the public and to the employees of hazardous systems, risk assessors produce "fantasy documents" detailing how high-risk systems are safe because layers and layers of safety measures are in place to avert or mitigate accidents.[318] Fantasy documents "are designed to be maximally persuasive", and never conclude that the effects of an accident will be unrecoverable.

Society can choose different strategies for dealing with high-risk systems. Perrow distinguishes absolute rationality, bounded rationality and social and cultural rationality, with the third as his strategy of choice. He proposes a form of limited rationality where fears of the public are taken into account, even if these fears may be less well-informed than scientific knowledge or statistics.[319] He suggests that societies could choose to abandon the riskiest systems with high catastrophic potential, restrict technologies where benefits are such that some risk is acceptable, or tolerate and improve less risky technologies.[320] In this respect, Perrow arrives at the same conclusion as Beck: the acceptance of certain types of technology is based at least in part on political decision-making.[321]

Seventeen years after his initial publication of Normal Accident Theory, Perrow concluded that — contrary to his original predictions — system accidents in high-risk systems remain rare events, although the frequency of the more preventable component failures appears to be increasing even in high-risk systems. For this he proposes a number of explanations. Firstly, it is very hard to have a catastrophe because so many elements need to fall into place — apparently, "the world is not as tightly coupled as many of us thought."[322] Secondly, even though component failures can be prevented, they continue to happen in high risk systems because of economic pressures, lack of government enforcement, biased accident investigations and new mechanisms to let the public at large suffer the costs of these accidents for the benefit of a small number of decision-makers.[323] However, where Beck imagines these to be

317 ibid 307 Compare also Beck (1992), p. 64-69.

318 Lee Clarke and Charles Perrow, 'Prosaic Organizational Failure' (1996) 39 American Behavioral Scientist 1040, 1053 <http://abs.sagepub.com/content/39/8/1040.short> accessed 13 February 2019.

319 Perrow, *Normal Accidents: Living with High-Risk Technologies* (n 101) 315–324. See also Klinke and Renn (n 275).

320 Perrow, *Normal Accidents: Living with High-Risk Technologies* (n 101) 349.

321 Beck, *Risk Society* (n 273) 168.

322 Perrow, 'Accidents, Normal' (n 103) 37.

323 ibid 36.

implemented in science itself, Perrow expects societies' institutions to brake and steer the course of technology — as becomes apparent from his proposed choice between abandon, restrict, and tolerate/improve.

Perrow underscores that in human history, many systems started in the complex/coupled quadrant, but societies have found ways to make most risky systems more linear and less tightly coupled over time. This may explain why system accidents remain extremely rare, even though Normal Accident Theory predicts that they are inevitable. Indeed, the apparent absence of large numbers of normal accidents is a weakness of Perrow's theory.[324] Nevertheless, Normal Accident Theory provides important insights that can be used at all levels — individual operators, plant supervisors as well as public authorities — to reduce risk and meaningfully investigate accidents. He explicitly sees a role for government to play, stating that regulatory efforts remain necessary to provide for a fair distribution of technological risk.

4.3.2 Normal accidents and Environmental law

Environmental regulation predates Normal Accident Theory by many years. Nevertheless, NAT can help to increase insight into a large number of provisions that have found their way into many environmental law instruments over the years. Choosing between banning, restricting and tolerating (and improving) technology is at the essence of environmental law.

Abandoning entire technologies is usually the result of a society-wide debate and political discourse. The best example, already indicated by Perrow, is the decision to phase out nuclear power in several countries in response to the nuclear disasters in Chernobyl (1986) and Fukushima (2011).[325] Persistent chemicals are another example: as per 2016, at least 605 chemical compounds have been banned from use and trade in the European Union.[326] Interactions in ecosystems are of course complex, and because emissions of persistent chemicals may be irreversible, events are also tightly coupled. This is especially true if the release is gradual, if it can remain unnoticed for a long

[324] Nancy Leveson and others, 'Moving beyond Normal Accidents and High Reliability Organizations: A Systems Approach to Safety in Complex Systems' (2009) 30 Organization Studies 227, 229 <https://journals.sagepub.com/doi/abs/10.1177/0170840608101478> accessed 20 March 2019.

[325] A nuclear power phase-out has been decided in Belgium and Switzerland and is under discussion in Germany and Spain.

[326] See https://echa.europa.eu/information-on-chemicals/pic/chemicals (retrieved on 12 October 2016).

time, or if it has taken many years to establish the harmfulness of a substance. A third example is the phasing out of incandescent bulbs in the European Union.[327]

Restriction usually happens by means of administrative decisions. Examples at the EU level are prior authorisation of pesticides and plant protection products, maximum residue levels of pesticides on agricultural products, measures against the release of genetically modified organisms, and a "prior informed consent" regime for a large number of chemicals.

"Tolerate and improve" is a strategy usually reserved for technologies that have been in use for a long time or whose risks are clearly understood. Environmental regulations use the term "best available techniques" to indicate that risks that are acceptable now, may be considered excessive in the future.[328] "Tolerate" is used for technology that is not only well understood, but whose risks are also considered almost universally acceptable when balanced against other interests of society. In those cases, explicitly stating the hazards by means of labelling and giving explicit instructions for safety is often considered sufficient. Specifically for the environmental protection context, this approach is visible on almost every household container for chemical products.[329] Another policy area where the EU legislator has chosen a "tolerate and improve" approach to an environmental risk is with the Energy Label Directive,[330] introducing labels to inform buyers of the energy-efficiency of many "energy-related products". Verplanken and Weenig have suggested that these labels can have a positive effect when consumers are not under time pressure during selection of household appliances.[331]

---

327 See: http://europa.eu/rapid/press-release_MEMO-09-368_en.htm.

328 Article 2(12), Directive 2008/1/EC of the European Parliament and of the Council of 15 January 2008 concerning integrated pollution prevention and control, [2008] OJ L 24 of 29 January 2008, p. 8-29

329 Regulation (EC) no 1272/2008 (CLP Regulation), [2008] OJ L353/1.

330 Recital 15 and article 2(a), Directive 2010/30/EU (Energy Label Directive), [2010] OJ L 153/1.

331 Bas Verplanken and Mieneke WH Weenig, 'Graphical Energy Labels and Consumers' Decisions about Home Appliances: A Process Tracing Approach' (1993) 14 Journal of Economic Psychology 739, 749 <http://linkinghub.elsevier.com/retrieve/pii/016748709390019H> accessed 13 February 2019; Labels appear to be less effective for home buyers, who may take more than energy efficiency into account: Kirsten Gram-Hanssen and others, 'Do Homeowners Use Energy Labels? A Comparison between Denmark and Belgium' (2007) 35 Energy Policy 2879, 2887 <http://linkinghub.elsevier.com/retrieve/pii/S0301421506004071> accessed 13 February 2019.

4.3.3 Normal accident theory in action: again, the Seveso III-Directive

The aforementioned Seveso III Directive contains several methods compatible with Normal Accident Theory, although there is no clear evidence on whether the Directive was drafted with Perrow's theory in mind. A strong example is article 9 of the Directive, stating that competent authorities must investigate whether different establishments where major accidents can occur are tightly coupled through "Domino effects".

Furthermore, the Directive reflects Perrow's assertion that blaming accidents on operator errors diverts attention from underlying causes and that continuous efforts are necessary to move systems away from the tightly coupled/complex quadrant. The "Safety management system" described in Annex III to the Directive must consist of technical as well as organisational elements. The system is part of a mandatory major-accident prevention policy which must be reviewed and, if necessary, updated at least every five years, creating a Plan-Do-Check-Act cycle ("Deming circle") for technical, organisational and managerial systems. The category of the largest establishments must maintain a "safety report" that has to be renewed every five years, or sooner if a major accident has occurred or if "new facts or new technological knowledge" justify a re-evaluation. This mandatory Deming cycle helps moving systems out of the complex/tightly coupled quadrant by requiring continuous improvements and codifying the ever-increasing knowledge and expertise that is accrued by scientific development, accident investigation and regular operation of the establishment. Taking the process one step further, article 12 of the Directive also requires that Member States' competent authorities provide for emergency response plans for measures to be taken outside of the establishment. Emergency plans must be reviewed every three years.

Seveso III aims to prevent fantasy documents by requiring several cycles of verification, control and enforcement at different levels of authority. During inspection of the establishment, the competent authorities must examine technical, organisational as well as managerial systems to verify whether the operator can demonstrate that the appropriate measures are taken to prevent major accidents.[332] Competent authorities are required to perform their inspections according to a plan that must be drafted based on a systematical appraisal of the hazards in all the establishments over which a competent authority has jurisdiction; the inspection plan

[332] Article 20, Seveso III Directive.

must be reviewed and updated regularly. Member States must provide reports on their activities to the European Commission. The Commission has several enforcement mechanisms at its disposal, mainly based on articles 258 and 260 of the Treaty on the Functioning of the European Union.

4.3.4 Big data and normal accidents

In a sense, data protection law aims to reduce the risks of complex, tightly coupled systems. The big data ecosystem consists of millions of devices, each generating or processing hundreds of transactions every day; every transaction can generate personal data. This data is then transmitted, stored, analysed, sorted, queried and delivered for reuse over numerous pathways, relayed through many different nodes, to large numbers of endpoints. Every pathway, node and endpoint is an avenue for the introduction of risk where errors can occur, or where data can be intercepted, corrupted or abused. Protection mechanisms are "brittle": a single failure, sometimes in another area of a system, may expose large quantities of data. These failures are not always accidental: many data leaks are a result of targeted hacking attempts or calculated "leaks from the inside." Indeed, especially in the big data ecosystem, protection mechanisms can introduce new pathways for error.[333]

"Normal accidents" concerning personal data appear to occur frequently. A very concrete example of complexity and tight coupling is the Diginotar case, where the hacking of a root certificate authority in the Netherlands resulted not only in the supposed spying on individual Gmail users in Iran and the potential spying on all Gmail users, but also to the revocation of security credentials for almost all Dutch government servers.[334] In another case, a vendor of PC hardware intercepted all network traffic from new laptops to serve personalised advertisements to its customers, exposing all data sent to and from the computer to third parties. Consumers had actually "agreed" to this practice by clicking through a license

[333] Daniel Nunan and Marialaura di Domenico, 'Big Data: A Normal Accident Waiting to Happen?' [2015] Journal of Business Ethics 1, 5.

[334] Axel Arnbak and others, 'Security Collapse in the HTTPS Market' (2014) 57 Communications of the ACM 47, 48–49 <http://dl.acm.org/citation.cfm?doid=2661061.2660574> accessed 19 March 2019; Hans Hoogstraaten and others, 'Black Tulip: Report of the Investigation into the DigiNotar Certificate Authority Breach' (Fox-IT BV 2012) 45 <https://roselabs.nl/files/audit_reports/Fox-IT_-_DigiNotar.pdf> accessed 13 February 2019.

agreement for the software that provided this function.[335] Other possibilities for attacking data security, confidentiality or integrity are deep packet inspection (at the network level) or the leaking of sensitive login information, like in the Ashley Madison server hack.[336] Nunan and Di Domenico give more examples of these "data accidents" and call them "normal accidents waiting to happen".[337] The information processing industry appears to agree on the necessity of systematically applying "best practices" to prevent these incidents: information security is the subject of a series of ISO standards, describing best known practices for information security management systems.[338]

Besides providing a framework for accident analysis, NAT provides a vocabulary that can be used for preventive analysis at all levels, possibly including the level of entire societies, for example to analyse whether free speech, equality and a democratic culture may be affected by permanent surveillance.[339] Seen in Perrow's terms, datafication is causing data subjects to increasingly live in a complex, tightly coupled environment: it can "bring classification to an increasing range of human activity."[340] Especially through profiling, it can have effects in seemingly unrelated areas: Havard and O'Neil have indicated education, advertising, predictive policing, employment, credit, insurance and the effects on civil society as areas of concern.[341] In this way, it can cement or exacerbate the discrimination of marginalised groups. Peppet is wary about the onset of a "signalling economy", where citizens feel pressure to reveal ever more details about their personal lives to avoid giving the impression that they are

335 Tim Ring, 'Keeping Tabs on Tracking Technology' (2015) 2015 Network Security 5, 6 <http://www.sciencedirect.com/science/article/pii/S1353485815300477> accessed 13 February 2019; Perlroth (n 204).

336 Christopher Mele, 'No Anonymity for Plaintiffs Suing Ashley Madison Over Hack, Judge Rules' *The New York Times* (21 April 2016) <http://www.nytimes.com/2016/04/22/technology/no-anonymity-ashley-madison-hack-case.html> accessed 20 March 2019.

337 Nunan and di Domenico (n 333).

338 See, for example, NEN-ISO/IEC 27001:2013 en: Information technology – Security techniques – Information security management systems – Requirements and the standards referred therein.

339 Weick (n 311) 29; Neil M Richards, *Intellectual Privacy: Rethinking Civil Liberties in the Digital Age* (Oxford University Press 2015) 153.

340 Dwork and Mulligan (n 224) 35.

341 Cassandra Havard, '"On the Take": The Black Box of Credit Scoring and Mortgage Discrimination' (2011) 20 Boston University Public Interest Law Journal 241, 287 <http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1710063> accessed 13 February 2019; O'Neil (n 268) chapters 3-10.

hiding embarrassing or disqualifying information.[342] "Scores can become self-fulfilling prophecies, creating the financial distress they claim merely to indicate" or they could "trap us in patterns that perpetuate the basest or narrowest versions of ourselves."[343]

Complexity is further increased due to the fact that many data controllers, especially operators of online platforms like social media and search engines, operate in two-sided markets: a data controller can offer a service or product to consumers for free or at a reduced price, making a profit from selling user information to third parties. Individual behaviours having effects in unrelated areas of life is a clear indicator of complexity, and — like in the case of ongoing release of persistent pollutants — the permanent "emission" of personal data in the context of datafication means that there may be no real possibility to "recover" from profiling, indicating tight coupling. This complexity and tight coupling and the possibility of large-scale accidents can be fertile grounds for fantasy documents to persuade lawmakers, data subjects and data controllers that the risks are acceptable.

4.4 Application of Risk Society Theory and Normal Accident Theory in the GDPR

Due to environmental law's long and convoluted history, national and European legislation is interspersed with examples of both Risk Society Theory and Normal Accident Theory. In contrast, the General Data Protection Regulation forms a single body of law addressing the risks of the processing of personal data as comprehensively as possible.[344] This affords an opportunity to evaluate the degree in which EU data protection law incorporates concepts of both sociological theories. A brief outline of the risk management framework of the GDPR will illustrate this evaluation.

342 Scott R Peppet, 'Unraveling Privacy: The Personal Prospectus & the Threat of a Full Disclosure Future' [2010] Northwestern University Law Review 28 <http://papers.ssrn.com/abstract=1678634> accessed 3 November 2015.

343 Citron and Pasquale (n 291) 18 and note 106; Dwork and Mulligan (n 224) 40.

344 The GDPR frequently requires that certain aspects of processing (such as processing special data or for journalistic purposes, or in the context of employment relations) are allowed or required by Union or Member State law (see also section 2.3 supra). Apart from that, delegated acts and Member state law authorized by the GDPR is relatively scarce. See Chapter X and articles 6(2), 8(1), 9(4), 84(1), 85, 87-88 and 90, GDPR.

4.4.1 Risk management model of the GDPR

The GDPR intends to address the risks to fundamental rights and freedoms of data subjects associated with the processing of personal data. It recognizes that data subjects may perceive these risks differently than the EU legislator.[345] Fundamental rights and freedoms are defined broadly to prevent accidental exclusion of protections under new technologies or policies.[346]

The risk management model of the Regulation is not stated explicitly. From the structure and content of the Regulation, the following model becomes visible:

- Primarily, the GDPR empowers data subjects to manage their own risks through the lawfulness requirement of article 5(1)(a). This covers both sub-politics and politics. By entering into agreements or consenting to processing, data subjects can determine the risks they are willing to take. As citizens of a Member State, they can participate in decision-making on whether a legal obligation, a public interest, or an official authority for processing should exist. Without their participation, lawfulness has to meet stricter criteria: the controller should then demonstrate his own legitimate interest or the vital interest of the data subject.
- Many aspects of processing are not visible to data subjects. To reduce the risks of these aspects, the GDPR assigns accountability for compliance to controllers and sets rules for data transfers outside the EU. It relies on independent oversight by specialised supervisory authorities for verification and enforcement.
- Article 5 provides technologically neutral, broadly worded underlying principles and recital 4 states the principle of proportionality. These principles serve as final touchstones for the acceptability of risk.
- The European Data Protection Board (article 68 GDPR) may publish guidelines, recommendations and best practices indicating what constitutes the "state of the art" of acceptable risk.

345 See recitals 9 and 51 for examples. See also Paul Schwartz, 'Risk and High Risk: Walking the GDPR Tightrope' (*Privacy Perspectives - Ideas and Insights on Data Protection*, 29 March 2016) <https://iapp.org/news/a/risk-and-high-risk-walking-the-gdpr-tightrope/> accessed 13 February 2019.

346 See, for example, recital 75 and article 1(2)). The GDPR also identifies "risks of circumvention" of data protection provisions, which it addresses by giving rules that are technologically neutral (recital 15) and the risk of data being insufficiently secure or confidential (recital 83 and article 32), which is addressed by (among others) a breach notification obligation (article 33).

In addition to the application of proportionality for evaluating the necessary measures that controllers need to implement, a limited number of obligations is triggered if the processing of personal data is likely to result in a "high risk to the rights and freedoms" of data subjects:

- In the case of data breaches, a high risk triggers an "undue delay"-requirement in notifying the data subjects of the breach (article 34(1));
- A data protection impact assessment is necessary for types of processing that cause a high risk; the assessment must be completed prior to the processing (article 35(1)).[347]

4.4.2 Identifying the underlying assumptions of the risk management model

A look at Klinke and Renn's figure 3 indicates that the EU lawmaker appears to have evaluated the risks associated with the processing of personal data as a Cassandra-type risk. The damage potential is considered to be well known, and both the probability of adverse effects and delay effects are high. This assumption can be derived from the chosen risk management method, "strengthening of responsibility" (articles 5(2), 83(4-5)). The implied GDPR risk management model may reveal some weaknesses when applied to consumer contracts, because some underlying assumptions may not be valid.

- It assumes that data subjects can assess the risks of processing. Whether this is true is far from certain.[348]
- It assumes that data subjects have bargaining power as consumers in the market and that the legislative process will ensure compliance with data protection principles, lowering the risk of processing. Both assumptions merit skepticism.[349]

347 The GDPR gives three cases that always pose a high risk in article 35(3).The "high risk"-threshold also triggers obligations for supervisory authorities. See the useful explanation and enumeration in Gabriel Maldoff, 'The Risk-Based Approach in the GDPR: Interpretation and Implications' <https://iapp.org/media/pdf/resource_center/GDPR_Study_Maldoff.pdf> accessed 13 February 2019.

348 McDonald and Cranor (n 144).

349 TNS Opinion and Social, 'Special Eurobarometer 431: Report' (European Commission 2015) DS-02-15-415-EN-N 30–32 <https://ec.europa.eu/COMMFrontOffice/publicopinion/index.cfm/ResultDoc/download/DocumentKy/66372> accessed 13 February 2019; see also: Joined Cases C-293 and C-594/12

- It assumes that managing risks for individual data subjects is sufficient to protect fundamental rights and freedoms, whereas datafication can have emergent effects at the societal level (e.g., discriminatory effects on marginalised groups). These risks have not been evaluated, which means their probability is unknown. This would indicate another risk classification and another management strategy (Cyclops-type and ascertaining probability, respectively).
- It assumes effective enforcement, where this effectiveness is uncertain, possibly due to a lack of funding for supervisory authorities.[350]

The remaining part of this section examines the extent to which the GDPR implements Risk Society and Normal Accident theories.

4.4.3 Risk society theory in the GDPR

Comparing the risks of industrialisation and datafication shows that both qualify as modern risks. This tentatively points to the conclusion that Beck's statements about decision-making in the risk society can also be applied to the data protection context. This paragraph aims to identify articles of the GDPR that reveal a degree of compatibility with three elements of Beck's ideas: identification of the processing of personal data as a modern risk, institutionalised individualism and political acceptance of risk.

Identification of processing as a modern risk

Like the Rio Declaration and the Aarhus Convention, the GDPR contains provisions indicating that the processing of personal data poses modern risks, originating from unknown sources and causing unquantifiable effects:

Digital Rights Ireland v *Minister of Communications* and *Kärntner Landesregierung* v *Seitlinger and others* [2014] ECLI:EU:C:2014:238, which invalidated the Data Retention Directive. The Dutch national government proposed to reinstate the effects of the Data Retention Directive through national law, arguing that it cannot be known in advance who will become a suspect later. KST II 2014-2015, 33542, nr. 16, p. 8 (Dutch parliamentary documents).

350 European Agency for Fundamental Rights, 'Fundamental Rights Report 2016' (European Union Agency for Fundamental Rights 2016) 129 <http://fra.europa.eu/en/publication/2016/fundamental-rights-report-2016> accessed 19 March 2019.

- The GDPR aims to provide rules that are technologically neutral. This indicates that no particular (use of) technology has been isolated as the source of these risks;[351]
- The GDPR aims to address broadly defined "threats to fundamental rights and freedoms", indicating that the EU legislator does not venture to predict where the adverse effects of processing can materialize;
- In a more abstract sense, the high administrative fines that independent supervisory authorities can impose (10 million Euros or 2% of annual turnover, see article 83(4)), indicate that the damage that the processing of personal data can impose is ill-suited to private compensation. This corroborates with the difficulties associated with attaching a monetary value to privacy.[352]

Individualism and sub-politics

Unlike the Seveso III-Directive, which has exclusively organized individual participation through traditional government institutions, the GDPR partially relies on individualisation and sub-politics. Controllers can achieve lawfulness of their processing by directly contracting with data subjects in the market (article 6(1)(a-b)) and controllers shall seek the views of data subjects or their representatives when conducting a data protection impact assessment "where appropriate" (art. 35(9)).

But the GDPR does not completely rely on sub-politics. Several mechanisms are in place to compensate for power differences between controllers and data subjects. Independent supervisory authorities subject controllers to more traditional oversight and enforcement mechanisms to ensure their compliance with specific obligations (art. 51(1)). Consumers may indeed be ill-equipped to evaluate the extent to which controllers adhere to e.g. the "by design and by default"-requirement or the general and specific requirements for data protection impact assessments. Likewise, the GDPR relies on judicial authorities to resolve disputes between data subjects, supervisory authorities and controllers (articles and 77-80 and 83(8)).[353]

---

[351] See paragraph 2.1 supra.

[352] "(c)ourts have been reluctant to find a value in privacy, because people willingly give it away in exchange for so little." Schneier, *Data and Goliath* (n 210) 227.

[353] Already under the DPD, individual access to justice was proven to be of significance. Case 362/14 *Maximillian Schrems v Data Protection Commissioner* [2015] ECLI:EU:C:2015:650.

Political acceptance of risk
Unlike several examples of environmental law, the GDPR does not preemptively abandon any type of processing, does not explicitly forbid any adverse effects for data subjects, nor does it specify an acceptable level of risk for these effects. Instead, it aims for proportionality between risk and protection measures. See, for example, art. 35(7)(d), requiring that a controller describes measures to "address the risks" (and presumably, reducing them to acceptable levels) instead of eliminating them.

In a more general sense, the measures in the GDPR are aimed almost exclusively at the rights and freedoms of individual data subjects. Precaution-based or discourse-based decision-making on whether pervasive surveillance is compatible with a free society was not applied, casting doubt on whether the residual risks of datafication for a democratic society have been politically accepted.[354]

4.4.4 Normal accident theory

EU data protection law does not yet explicitly display a systems approach to data protection. Instead, it focuses on the activity of processing (article 2.1 juncto 4(2) GDPR). Still, from a controller standpoint, a systems approach is probably useful: well-chosen definitions of processing (sub-)systems permit e.g. the reuse of previous impact assessments if only certain parts of processing systems are modified (article 35(1) GDPR). Additionally, datafication is a possible source of system accidents. This paragraph aims to identify articles of the GDPR that reveal a degree of compatibility with three elements of Perrow's ideas: moving out of the complex/tightly coupled quadrant, fantasy documents and social and cultural rationality.

Moving out of the Complex/Tightly coupled quadrant
As was proposed in section 4.3.4 above, the processing of personal data can have the characteristics of a complex, tightly coupled system for data subjects. Contrary to the example of the Seveso III-Directive, the GDPR does not contain provisions for the mandatory performance of periodical risk analysis; contrary to the Energy Label Directive, it does not contain provisions for the mandatory labelling of products or services. The GDPR does, however, charge public authorities with "encouraging" the use of data protection seals and marks and empowers the Commission to develop standardised icons and determine the information they should represent (arts. 43, 12(7-8) GDPR). The European Parliament proposed a number of icons after the vote

354 Klinke and Renn (n 275).

on the original Commission proposal, but these icons did not make it into the final version of the GDPR.[355]

The principles of fairness and transparency may fail to reduce tight coupling and perceived complexity, not only because taking in all the relevant information requires significant efforts from data subjects,[356] but also because transparency must be balanced against the right to protect intellectual property and trade secrets (recital 63).

Fantasy documents

Measures to prevent the creation of fantasy documents in the GDPR are possibly not as thorough as in the Seveso III-Directive. Fantasy documents emanate from new systems lacking a verifiable track record, they cover a large array of possible accidents, are designed to persuade, and they never doubt that any crisis can be resolved.[357] The following documents could eventually turn out to be fantasy documents:

- Agreements and consent statements providing lawfulness of processing and enumerating authorised forms of processing, in cases where consumers can be expected not to read them;[358]
- Data protection impact assessments, where controllers may have large amounts of leeway in interpreting whether the assessment is necessary. Even when they are necessary, they may not be very useful: according to Moerel, the large number of requirements for impact assessments may result in a "tick box list for compliance measures regardless of their actual impact on compliance."[359] Hempel and Lammerant conclude that the drafters of impact assessments tend to assume that their purpose is mainly to increase knowledge and much less to offer opportunities to discuss issues of power, indicating the potentially limited effectiveness of the intended sub-politics;[360]

355 Jan Philipp Albrecht, 'GDPR Inofficial Consolidated Version after LIBE Committee Vote Provided by the Rapporteur' (2013) Annex I <http://www.janalbrecht.eu/wp-content/uploads/2018/05/DPR-Regulation-inofficial-consolidated-LIBE.pdf> accessed 13 February 2019.

356 McDonald and Cranor (n 144) 562.

357 as explained in Clarke and Perrow (n 318) 1053.

358 McDonald and Cranor (n 144) 562.

359 Moerel (n 87) 52 and note 200.

360 Leon Hempel and Hans Lammerant, 'Impact Assessments as Negotiated Knowledge' in Serge Gutwirth, Ronald Leenes and Paul de Hert (eds), *Reforming European Data Protection Law* (2015 edition, Springer 2014) s. 5.2.

- Binding corporate rules governing data transfers, where the "enforceable rights" that they confer to data subjects may be expected not to be effectively known or understood by data subjects;
- Certification mechanisms and data protection seals and marks signalling compliance with the GDPR, where it is uncertain that data subjects can fully grasp what this compliance, in conjunction with their privacy contracts, actually entails;
- The GDPR itself, especially when it states that it respects all fundamental rights and therefore implies that the same is true for all processing of personal data that complies with it (recital 4), where knowledge of the risks of datafication and of the processing of personal data may be incomplete and where there can be reasonable uncertainty regarding the effectiveness of enforcement.[361]

Mechanisms against fantasy documents in the GDPR are not yet formally embedded in plan-do-check-act cycles,[362] nor are they designed in the form of interlocking systems of oversight. The supervisory authorities are expressly intended to be completely independent, which makes "watching the watchers" difficult; only the European Data Protection Board – consisting of members of national supervisory authorities and one EU appointee[363] – has the authority to monitor the effectiveness of enforcement activities. This independence is useful to shield Independent Supervisory Authorities from political influence, which is certainly appropriate: supervising the processing of personal data by government bodies may not be effective if it is politically influenced or motivated. However, the GDPR does not require a planned, risk-based, cyclic execution of the authorities' duties and tasks aimed at continuous improvement, nor does it empower the Board to impose such a system.[364]

---

361 Omer Tene, 'For Privacy, European Commission Must Be Innovative | Center for Democracy & Technology' (*Center for Democracy & Technology*, 28 February 2011) <https://cdt.org/blog/for-privacy-european-commission-must-be-innovative/> accessed 13 February 2019; European Agency for Fundamental Rights (n 350) 129; European Agency for Fundamental Rights, 'Data Protection in the European Union: The Role of National Data Protection Authorities. Strengthening the Fundamental Rights Architecture in the EU II' (European Union Agency for Fundamental Rights 2010) 42–43 <doi:10.2811/47216>.

362 The GDPR does not require data protection impact assessments to be periodically reviewed; only the processing needs to be reviewed to verify accordance with the existing assessment (art. 35(11)).

363 Art. 68(3), GDPR; art. 42(1), Regulation (EC) No 45/2001.

364 The Board is authorized to issue "guidelines, recommendations and best practices" (art. 70(1)).

Social and Cultural rationality

Perrow suggests that societies "abandon, restrict or tolerate and improve" risky technology. The GDPR contains no provisions that aim to abandon any forms of processing. It does, however, impose restrictions, for example on the processing of special categories of data. As was noted before, the recitals consider the effects of datafication on entire societies only very briefly; even when discussing the necessity to prevent social disadvantage, this is aimed at individual data subjects (recitals 75, 85). However, through its technologically neutral approach, the GDPR holds some promise to improve currently tolerated forms of processing, especially through the general principles in article 5(1). Technological progress has a tendency to raise the bar for technologies to be politically acceptable.[365] Similarly, the publishing of guidelines, recommendations and best practices (art. 70) could reflect changing social norms and codify technological progress. However, the GDPR offers no strict mechanism to ensure compliance with these instruments.

4.5 Looking into the crystal ball

Contrary to the European Commission's stated intention, it appears safe to say that the General Data Protection Regulation is not yet "future proof for the decades to come".[366] Looking into a rear view mirror, it is clear that Beck's and Perrow's theories have played a significant role in addressing the risks of technology for industrialised societies through environmental law. Although datafication poses similar risks to societies, data protection law appears to not yet have considered these risks beyond the scope of individual rights. Risk Society Theory and Normal Accident Theory therefore offer an opportunity to transfer insights from environmental law to the data protection law domain. Looking into a crystal ball, I expect that opportunities for useful application of both theories can be found when addressing the following issues:

Better understanding of the risks of big data

During the last few decades, fundamental research has increased our insights in the effects of human activity on ecosystems, reducing the "modern risk"-qualities of these effects. Likewise, for data protection law to effectively address the risks that datafication and big data can pose to individuals and societies, our understanding of the effects of the processing of personal data needs to improve. Stated in Beck's terms,

[365] See the reference to "best available techniques" in section 3.2 supra.

[366] European Commission, 'Proposal for a General Data Protection Regulation' (n 81) 104.

brakes and a steering wheel will be more effective if we have a speedometer, a roadmap and a compass. Research efforts into algorithms and modelling, psychology and the social sciences could hopefully offer useful insights and reduce the probability that the risks of big data are obscured by fantasy documents. A better understanding of the underlying mechanisms is essential to increasing the effectiveness of regulation and enforcement. The European Framework Programme for Research and Innovation could play a part in the funding of relevant research efforts.[367]

Better implementation of sub-politics

A free market is arguably one of the better suited mechanisms to find big data's best applications, both for the common good and for private enterprise. The GDPR relies on the participation of data subjects in the market and in the drafting of data protection impact assessments. But Galanter's warning still holds: usually, the "haves" come out ahead. Consumers face high costs of information and organisation, which could stand in the way of effective participation. Without additional protections and incentives for consumers engaging in privacy contracts, the GDPR's implementation of sub-politics may not achieve its goals. Consumer protection law may have a part to play in addressing these matters. It may also be necessary to create formal participation options for consumers, for example vis-a-vis national Independent Supervisory Authorities and the European Data Protection Board.

Better transparency

For better sub-politics to have a positive effect, data subjects need better insight into the effects of their privacy contracts. The GDPR's seals and marks could provide insights to consumers, but only if they signal meaningful information. Demonstrating compliance with the GDPR may not help to move the processing of personal data out of the complex/tightly coupled quadrant, because data subjects would need to understand the GDPR itself (counting over 50.000 words) as well as their own privacy contracts, to know what this compliance does or does not accomplish. Like energy labels classifying appliances on a scale from A to E, seals and marks could encourage comparison between controllers. Especially for consumer contracts, the EU could consider programs empowering and subsidising consumer organisations to collect and publish meaningful comparisons between different controllers and data-relevant

[367] Regulation (EU) No 1291/2013 of the European Parliament and of the Council of 11 December 2013 establishing Horizon 2020 – the Framework Programme for Research and Innovation (2014-2020) and repealing Decision No 1982/2006/EC, OJ L 347/104.

products and services, like it already subsidises consumer participation in the standardisation and lawmaking process.[368]

Improving social and cultural rationality

If we acknowledge the possibility of a "surveillance market" coming about through consumer contracts, an open discussion about whether its effects are desirable needs to take place. The GDPR's articles 6(1)(1–2) offer no opportunities for the discourse-based or precaution-based risk management strategies proposed by Klinke and Renn for the Pythia-, Cyclops-, or Cassandra-type risks that several authors associate with datafication. The collection of census data resulting in the Volkszählungsurteil was enough to trigger a far-reaching court case in 1983;[369] it is therefore remarkable that the risks of datafication appear to have escaped public or political deliberation on whether – and how – society should "abandon, restrict or tolerate and improve" this technology.

This omission is understandable to some extent: as was stated earlier, datafication is the result of, and an indispensable by-product of, automation. Over the past decades, iteratively computerising an ever-increasing number of ordinary tasks may not have been cause for public concern. But now, as almost all daily activity is facilitated through automated platforms, big data has altered the landscape. The impact of datafication is by no means smaller than the impact of a census. All the information that citizens refused to provide to their government is now easily discerned from the analysis of all the data points that they provide, as consumers, to data controllers.

Therefore, society still needs to discuss the future of datafication in terms of acceptable risk, regulation, compliance and especially the desired level of enforcement. Such a discussion will be of a political nature. Considering the scale of the EU internal market or even a worldwide economy, the effects of having these discussions at the national level may be limited. Determining where this discussion should take place is therefore not easy. All options have significant imperfections.[370] In accordance with Perrow's notion of "tolerate and improve", the results of these discussions may need periodical re-evaluation for the foreseeable future.

---

368 European Parliament and Council Regulation (EU) No 254/2014 of 26 February 2014 on a multiannual consumer programme for the years 2014-20 and repealing Decision No 1926/2006/EC, [2014] OJ L 84/42, art 3(1)(b).

369 Volkszählungsurteil [1983] *BVerfGE* 65,1 para C II 1 a.

370 Komesar, *Law's Limits* (n 93) 30–31; see also section 2.1.4 *supra*.

4.6 In conclusion

Big data and the associated social risks have been said to be an example of the Collingridge dilemma: the need to regulate technology sometimes becomes apparent only after the point of no return has been passed.[371] The current EU data protection framework, supposedly among the strongest in the world, is proof that this dilemma is not a universal truth. But technology tends to outpace legislation, and it is probably unwise to assume that the GDPR will be an exception. This chapter aims to illustrate that Risk Society Theory and Normal Accident Theory have contributed significantly to environmental law and that they similarly have a lot to offer to the future development of EU data protection law. These theories indicate some areas of research, social discourse and legal development that can help to keep the EU framework effective and future-proof.

[371] "Regulators having to regulate emerging technologies face a double-bind problem: the effects of new technology cannot be easily predicted until the technology is extensively deployed. Yet once deployed they become entrenched and are then difficult to change." David Collingridge, *The Social Control of Technology* (Frances Pinter 1980); quoted in Moerel (n 87) 4.

5 Why the "*Computer says no*"

Illustrating big data's discrimination risk through complex systems science

Prelude

Like Chapter 4 before it, this chapter also started as a conference paper. On 20 September 2016 I presented a paper called "Decomposing contractual privacy: using complexity theory for preventing indirect discrimination" during the Law and Complexity satellite session of the 2016 Conference on Complex Systems in Amsterdam.

In the aftermath of the conference, I was introduced to dr. Qing Yi Feng of Utrecht University. She could see the points I made in the paper, but offered insightful comments on the effects that applying complex systems science could achieve. She accepted my proposal to rewrite the conference paper and submit it to a scholarly journal in the legal field. Dr. Feng wrote section 5.2. Dr. Feng and I co-wrote sections 5.7 and 5.8, for which I provided the first drafts.

This paper builds on chapter 4 in the way that it also deals with risk. However, this paper is more specifically focused on the risks associated with the processing of sensitive personal data.

Timeline and citation

The first version of this paper was submitted to International Data Privacy Law on 29 November 2017. It was conditionally accepted 10 February 2018. A new version was submitted 22 March; this version was accepted 22 April 2018. The article was not published under an open access license. IDPL recommends the following citation.

> Michiel Rhoen, Qing Yi Feng; Why the 'Computer says no': Illustrating big data's discrimination risk through complex systems science, International Data Privacy Law, Volume 8, Issue 2, 1 May 2018, Pages 140–159, https://doi.org/10.1093/idpl/ipy005.

5.1 Introduction

Due to the nature of human society as a complex system,[372] big data threatens to undercut the anti-discrimination efforts in the EU data protection framework. This is especially relevant in the contexts of the processing of sensitive personal data and profiling. It is far from certain whether the General Data Protection Regulation (GDPR)[373] can effectively reduce this risk.

In this chapter, we aim to present some of the insights that complex systems science has to offer to the study of EU data protection law. A settled definition of a complex system is not yet available, nor do we aim to propose such a definition here. We will follow the analysis of Ladyman et al: complex systems are ensembles of many elements that have the possibility to engage in interactions with other elements with a certain degree of disorder (or non-predictability); furthermore we assume that these interactions can lead to ordered patterns ("robust order") and that these patterns can persist for a relevant period of time ("memory").[374] Natural persons, and the groups, communities and societies that they form, all qualify as complex systems in this context.

We use the term "big data" to refer to the high volume, variety and velocity of data streams demanding cost-effective and innovative means of processing for enhanced insight and decision-making, especially profit-seeking.[375] Big data results from datafication – "the ability to render into data many aspects of the world that have never been quantified before."[376] Datafication has affected many previously ephemeral behaviours. For example, private communication, reading the news, listening to

372 For an example of a definition of "complex system", see Yaneer Bar-Yam, *Dynamics of Complex Systems* (Addison-Wesley 1997) 12.

373 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.

374 James Ladyman, James Lambert and Karoline Wiesner, 'What Is a Complex System?' (2013) 3 European Journal for Philosophy of Science 33, s 6.1 <http://link.springer.com/article/10.1007/s13194-012-0056-8> accessed 20 March 2019.

375 Lemi Baruh and Mihaela Popescu, 'Big Data Analytics and the Limits of Privacy Self-Management' (2017) 19 New Media & Society 579 <http://dx.doi.org/10.1177/1461444815614001> accessed 19 March 2019; Jacques Bughin, 'Big Data, Big Bang?' (2016) 3 Journal of Big Data 2, 9–10 <https://journalofbigdata.springeropen.com/articles/10.1186/s40537-015-0014-3> accessed 19 March 2019.

376 Zikopoulos and Eaton (n 7) 5; Mayer-Schönberger and Cukier (n 115) 30.

music, watching video, consuming electrical power in the home, and moving about in physical space are now continuously generating personal data. Datafication shares many features with surveillance or permanent observation. Log files of infrastructure "choke points"[377] enable anyone with suitable access to put almost anyone under surveillance through "big data analytics".[378]

This opportunity for individual surveillance on a massive scale introduces new types of risk, or "modern risks", to industrialised societies, as Beck predicted when he proposed his theory of the Risk Society.[379] According to Baldwin et al, regulation "can be seen as being inherently about the control of risks".[380] One of the risks that the GDPR expressly addresses is discrimination (recital 75). Discrimination – and other risks – are addressed in a technologically neutral fashion (recital 15). To this end, the GDPR uses criteria like "appropriate safeguards" (article 6(4)(e)) or "appropriate technical and organisational measures" (art. 24(1)). These are examples of the principle of proportionality (recital 4) and, indeed, a "risk-based" approach.

The GDPR is not the only European legislative effort aimed at preventing discrimination. In the Council of Europe, the European Convention on Human Rights (ECHR) prohibits discrimination in the exercise of the rights and freedoms it guarantees "on any ground" (art. 14).[381] In EU law, non-discrimination efforts differ in

377 A "choke point" is a network node connected to many other nodes. For example, a Facebook server is a choke point for a large number of Facebook users, a cell tower is a choke point for all its connected mobile telephones and a mobile switch is a choke point for a large number of cell towers. See Assane Gueye and others, 'Defensive Resource Allocations with Security Chokepoints in IPv6 Networks' in Pierangela Samarati (ed), *Data and Applications Security and Privacy XXIX* (Springer, Cham 2015) 262 <https://link.springer.com/chapter/10.1007/978-3-319-20810-7_19> accessed 20 March 2019.

378 Zikopoulos and Eaton (n 7) 48.

379 Beck, *Risk Society* (n 273) 19–20; 88, see also Dwayne Winseck, 'Netscapes of Power: Convergence, Network Design, Walled Gardens, and Other Strategies of Control in the Information Age' in David Lyon (ed), *Surveillance as social sorting: Privacy, risk and digital discrimination* (Routledge 2003) 176, 188. Surveillance concerns were also instrumental to invalidation of the Data Retention Directive and the EU-U.S. safe harbor agreement: see Joined Cases C-293 and C-594/12 *Digital Rights Ireland v Minister of Communications and Kärntner Landesregierung v Seitlinger and others* [2014] ECLI:EU:C:2014:238 paras 32, 37 and 56-57; C-362/14, *Maximillian Schrems v Data Protection Commissioner*, ECLI:EU:C:2015:650.

380 Baldwin, Cave and Lodge (n 82) 83.

381 European Convention for the Protection of Human Rights and Fundamental Freedoms, Sept. 3, 1953, ETS 5, 213 UNTS 221.

scope.[382] This text focuses on discrimination based on individual attributes that the GDPR qualifies as "special categories of personal data" or "sensitive data (in relation to fundamental rights and freedoms)" (recitals 10 and 51). Article 9(1) of the GDPR applies this moniker to the following types of information: "racial or ethnic origin, political opinions, religious or philosophical beliefs, (…) trade union membership, 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". Article 22(4) refers to article 9(1) in relation to profiling.

Many individuals suffering social, political or economic discrimination share one or more protected traits from the categories listed in article 9(1). Because these traits are deeply personal, they can be expected to drive individual behaviour. Therefore, pervasive datafication could result in data sets that can reveal these traits for individual data subjects if the data is analysed. For example, it has been proven possible to discover expectant mothers by analysing shopping behaviour;[383] ethnicity, religion and sexual preference can be inferred from what individuals "like" on Facebook.[384] Several authors have underscored the possible discriminatory effects of big data through automated decision-making or profiling.[385] New examples seem to appear regularly.[386] These developments have given rise to the notion that algorithms can be biased and engage in discrimination against protected groups, cementing or

[382] Raphael Gellert and others, 'A Comparative Analysis of Anti-Discrimination and Data Protection Legislations' in Bart Custers and others (eds), *Discrimination and Privacy in the Information Society - Data Mining and Profiling in Large Databases* (Springer 2013) s 4.3.

[383] Duhigg (n 116) chap. 8; Eric Siegel, *Predictive Analytics: The Power to Predict Who Will Click, Buy, Lie, or Die* (2 edition, Wiley 2016) 38–40.

[384] Michal Kosinski, D Stillwell and T Graepel, 'Private Traits and Attributes Are Predictable from Digital Records of Human Behavior' (2013) 110 Proceedings of the National Academy of Sciences 5802, 5803 <http://www.pnas.org/cgi/doi/10.1073/pnas.1218772110> accessed 20 March 2019.

[385] Richards (n 295) 1956–1958; Omer Tene and 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, 251–252 <http://papers.ssrn.com/abstract=2149364> accessed 21 March 2019; Bart Willem Schermer, 'The Limits of Privacy in Automated Profiling and Data Mining' (2011) 27 Computer Law & Security Review 45, 47 <http://linkinghub.elsevier.com/retrieve/pii/S0267364910001767> accessed 20 March 2019; Brent Daniel Mittelstadt and others, 'The Ethics of Algorithms: Mapping the Debate' (2016) 3 Big Data & Society 2053951716679679, 8–9 <http://journals.sagepub.com/doi/abs/10.1177/2053951716679679> accessed 20 March 2019.

[386] Siegel (n 383) ch 3.

increasing their marginalisation: "(it) (t)urns out algorithms are racist", but also sexist or a vector for homophobia, and more dangerous than killer robots.[387]

The GDPR was originally proposed in 2012. The EC intends it to be future proof for the decades to come.[388] However, it was largely drafted before the effects and risks of datafication became prominent in public consciousness.[389] Since discrimination is one of these risks, this contribution aims to answer the following question:

How can complex systems science increase the understanding of the risks and remedies associated with datafication for the efficacy of anti-discrimination efforts in the GDPR?

This question is answered by addressing the following sub-questions:

- How does complex systems science help explain that datafication enables the discerning of sensitive data?
- How does the processing of special categories of data pose a risk?
- How can datafication adversely affect the GDPR protections against the processing of sensitive data?
- How does datafication undermine the GDPR protections against discriminatory profiling?
- How could possible adverse effects of datafication be mitigated or remedied?

---

387 Committee of Experts on Internet Intermediaries (MSI-NET), 'Study on the Human Rights Dimensions of Algorithms (Second Draft)' (Council of Europe 2017) MSI-NET(2016)06 rev 17 <https://rm.coe.int/16806fe644> accessed 19 March 2019; Navneet Alang, 'Turns Out Algorithms Are Racist' [2017] *The New Republic* <https://newrepublic.com/article/144644/turns-algorithms-racist> accessed 19 March 2019; Julia Angwin and others, 'Machine Bias' (*ProPublica*, 23 May 2016) <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing> accessed 19 March 2019; Yilun Wang and Michal Kosinski, 'Deep Neural Networks Are More Accurate than Humans at Detecting Sexual Orientation from Facial Images' (2017) (in press) Journal of Personality and Social Psychology <https://osf.io/fk3xr/> accessed 21 March 2019; Will Knight, 'Google's AI Chief Says Forget Elon Musk's Killer Robots, and Worry about Bias in AI Systems Instead' (*MIT Technology Review*, 3 October 2017) <https://www.technologyreview.com/s/608986/forget-killer-robotsbias-is-the-real-ai-danger/> accessed 20 March 2019.

388 European Commission, 'Proposal for a General Data Protection Regulation' (n 81) 104.

389 Mayer-Schönberger and Cukier (n 115).

This chapter focuses primarily on the effects of datafication for data subjects who are also consumers.[390] The issue of discrimination is limited to the grounds listed as "special categories of personal data" in the GDPR. Since this analysis is primarily focused on the legal aspects of big data, theorems and terminology from complex systems theory are presented colloquially; their definitions, theoretical underpinnings, formal proofs or experimental results are provided as references.

5.2 Emergence as a fundamental property of complex systems

Our society is a typical complex system composed of many components, interacting on different scales and levels.[391] For example, human beings, the main components of our society, are complex systems themselves.[392] They communicate with others in different forms of conversations, such as calling and emailing. Humans form larger-scale composites (organizations, communities and societies) with specific complexities; these composites then associate and communicate with each other and with humans for both fun and profit.[393] As a result of datafication, more and more such interactions are recorded and digitized by ubiquitous sensors (e.g., in smart phones), log files of internet infrastructure elements, commercial transactions etc., basically turning many aspects of the lives of people and societies into big data. This data provides near real-time measures of the aforementioned complex systems. It contains patterns that can be used to find attributes that were previously difficult to detect.[394]

Many complex systems display so-called emergent properties. An emergent property is a property displayed by a complex system, that is not directly predictable from the

390 "(A)ny natural person who, in contracts covered by this Directive, is acting for purposes which are outside his trade, business or profession": article 2(b), Council Directive 93/13/EEC on unfair terms in consumer contracts (Unfair Terms Directive), [1993] OJ L 95, p. 29-34.

391 Claudio J Tessone, 'The Complex Nature of Social Systems' in Bernardo Alves Furtado, Patricia AM Sakowski and Marina H Tóvolli (eds), *Modeling Complex Systems for Public Policies* (IPEA 2015); John H Holland, *Complexity: A Very Short Introduction* (Oxford University Press 2014) ch 3.

392 Simon A Levin, 'Ecosystems and the Biosphere as Complex Adaptive Systems' (1998) 1 Ecosystems 431, 432.

393 Orlando Gomes, 'The Economy as a Complex Object' in Bernardo Alves Furtado, Patricia AM Sakowski and Marina H Tóvolli (eds), *Modeling Complex Systems for Public Policies* (IPEA 2015).

394 Bar-Yam (n 372) 11–12.

properties of that systems' elements.[395] In the context of this chapter, relevant examples of an emergent property are distinguishable patterns like a behavioural convention among members of a group, the response of the human body to illness that resembles other individuals' responses, or the segregation of individuals into groups. Emergent properties of a complex system can become apparent by observing the entire system or by observing interactions between the elements in the system. The interactions within human societies are carried out through the behaviour of its individual elements: human beings. Human behaviour in turn is based on both our cognition and our biology.[396] Human cognition perceives our environment, including others' behaviour, values, beliefs, attitudes and intentions, and then to a large extent shapes our traits and social conventions, mainly through implicit or even automated processes; our biology guides other responses to our environment.[397] Evidence shows that our cognition works "effortlessly, and even unintentionally".[398] Human behaviour is an emergent property of the human organism as a complex system; it lies at the root of conventions and segregation just as human biology lies at the root of our response to temperature changes or illness. Therefore, as datafication covers more and more aspects of our lives and society, behavioural patterns of cognition and biology are encoded in data.

One of the fastest-developing techniques for the processing of data is Artificial Intelligence (AI).[399] The performance of AI has increasingly been proven to beat human performance in certain fields, like playing Go and poker, making predictions,

395 Holland (n 391) ch 6; Here, "emergence" is used as shorthand for "higher-level order". Note that emergence has been called a "notoriously murky notion". It is undecided whether it is related purely to human understanding or to underlying causality. Still, "If a system doesn't exhibit higher-level order (...), it is not complex." Ladyman, Lambert and Wiesner (n 374) 40–41, 58–59.

396 Biological and psychosocial systems of humans count as complex systems. Bar-Yam (n 372) 2–4.

397 James S Uleman, S Adil Saribay and Celia M Gonzalez, 'Spontaneous Inferences, Implicit Impressions, and Implicit Theories' (2008) 59 Annual Review of Psychology 329, 330 <http://www.annualreviews.org/doi/10.1146/annurev.psych.59.103006.093707> accessed 21 March 2019 and the referenced literature.

398 James S Uleman, Leonard S Newman and Gordon B Moskowitz, 'People as Flexible Interpreters: Evidence and Issues from Spontaneous Trait Inference' (1996) 28 Advances in experimental social psychology 211, 211 <http://www.sciencedirect.com/science/article/pii/S0065260108602397> accessed 5 August 2017.

399 Nils J Nilsson, *The Quest for Artificial Intelligence* (1 edition, Cambridge University Press 2009).

and judging human character.[400] Machine learning, a major subfield of AI, provides a number of cost-effective algorithms aiming to make sense of big data.[401] Such algorithms are roughly divided into three different categories: supervised learning, unsupervised learning and reinforcement learning.[402] Supervised learning comprises techniques that predict the value of a target variable given an input variable. An example is the automated recognition of handwriting in US ZIP codes: since each element of a ZIP code is a digit, predictions for each digit can be limited to an integer with a target value between 0 and 9. In unsupervised learning, the aim is to find patterns in the data such that certain variables can be identified. An example is the analysis of a large customer database to find groups of "similar" customers for achieving market segmentation, without identifying those segments in advance. Finally, in reinforcement learning, a certain goal is pursued in a dynamic process without knowing beforehand whether or not the approach will lead to reaching the goal, and the learning process is driven by feedbacks. An example would be developing an algorithm predicting the best possible next move in a turn-based game like Go by playing a large number of games to their conclusion.

Just like the complex systems they analyse and represent, machine learning algorithms can exhibit emergent properties in their output. This appears to be the underlying cause of algorithms' perceived bias: if an algorithm is trained using biased

400 AR Guess, 'Artificial Intelligence Had a Breakthrough Year in 2015' (*DATAVERSITY*, 9 December 2015) <http://www.dataversity.net/artificial-intelligence-had-a-breakthrough-year-in-2015/> accessed 20 March 2019; Tonya Riley, 'Artificial Intelligence Goes Deep to Beat Humans at Poker' (*Science*, 3 March 2017) <http://www.sciencemag.org/news/2017/03/artificial-intelligence-goes-deep-beat-humans-poker>; BBC News, 'Artificial Intelligence: Google's AlphaGo Beats Go Master Lee Se-Dol' (*BBC News*, 12 March 2016) <http://www.bbc.com/news/technology-35785875> accessed 19 March 2019; Navin Sharma and others, 'Predicting Solar Generation from Weather Forecasts Using Machine Learning', *Smart Grid Communications (SmartGridComm), 2011 IEEE International Conference on* (IEEE 2011) 551 <http://ieeexplore.ieee.org/abstract/document/6102379/> accessed 20 March 2019; Stephen F Weng and others, 'Can Machine-Learning Improve Cardiovascular Risk Prediction Using Routine Clinical Data?' (2017) 12 PLOS ONE e0174944, 9 <http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0174944> accessed 21 March 2019; Wu Youyou, Michal Kosinski and David Stillwell, 'Computer-Based Personality Judgments Are More Accurate than Those Made by Humans' (2015) 112 Proceedings of the National Academy of Sciences 1036, 1039 <http://www.pnas.org/content/112/4/1036> accessed 21 March 2019.

401 Peter Flach, *Machine Learning: The Art and Science of Algorithms That Make Sense of Data* (1 edition, Cambridge University Press 2012).

402 Christopher M Bishop, *Pattern Recognition and Machine Learning* (Springer 2006) 3.

data, this bias can be reflected in its output.[403] This can be understood as follows. Human behaviour is characterised by traits and views at the scale of individuals, groups and entire societies. Shared traits lead people to associations, or the presumption of attributes, in accordance with those traits. These associations can develop conventions that are, in turn, expressed in individual and collective human behaviour. When applied to data sets resulting from datafication, machine learning algorithms can build a mathematical representation of these traits, conventions and behaviours. Therefore, if these algorithms are then used to assign attributes to data subjects, these attributes may reflect these traits, conventions and behaviours. If controllers use the assigned attributes in such a way that data subjects sharing protected traits are treated differently, this can have discriminatory effects. In those cases, algorithms can be said to be "racist", "sexist", or otherwise biased against groups of data subjects sharing protected traits in the sense that these data subjects are treated differently when compared to other data subjects with otherwise similar traits. For example, extending special rebates to consumers who have bought alcohol but not to others, could be seen as discrimination against consumers who do not drink alcohol for religious reasons.

Thus, we propose that the concept of emergence can provide insights relevant to the subject of EU anti-discrimination law in relation to data processing algorithms. These insights are especially relevant to data subjects: being unaware of the information that controllers can obtain through algorithms can make it difficult to detect or escape discriminatory effects. Several instances of the successful deduction of sensitive traits from non-sensitive data have been published. For example, Kosinski et al. found that gender, racial origin, sexual orientation, political opinions and religious beliefs can be predicted by Facebook "likes" with more than 80% accuracy.[404] Seneviratne et al. show that they can predict users' religion (and some non-sensitive traits) with over 90% precision in some cases by taking a snapshot of the apps that data subjects have downloaded to their smartphones.[405]

---

403 Jieyu Zhao and others, 'Men Also Like Shopping: Reducing Gender Bias Amplification Using Corpus-Level Constraints', *arXiv:1707.09457 [cs, stat]* (2017) s 3 <http://arxiv.org/abs/1707.09457> accessed 21 March 2019.

404 Kosinski, Stillwell and Graepel (n 384) 5803.

405 Suranga Seneviratne and others, 'Predicting User Traits from a Snapshot of Apps Installed on a Smartphone' (2014) 18 Mobile Computing and Communications Review 1, 6 <http://dl.acm.org/citation.cfm?id=2636244> accessed 20 March 2019.

Avoiding possibly discriminatory effects of algorithms is difficult for both data subjects and controllers because emergence is a fundamental property of complex systems. It is virtually impossible to eliminate all possible emergent properties from a data set because these properties may not be known in advance and they are captured in innocuous individual data points. External triggers are not necessary for emergence.[406] Extensive efforts have been made to understand emergence, resulting in theories and tools for complex systems such as nonlinear dynamics, fractal theory, and agent-based modelling.[407] Based on these efforts, Feng et al. proposed a theoretical framework to understand why machine learning algorithms can successfully identify patterns from the data containing the information of interactions within the complex system. They argue that by introducing non-linear interactions and optimization, machine learning algorithms themselves are complex systems, assimilating the dynamics of pattern formation from the complex system they represent.[408] They also pointed out that the more exhaustive the available data, the more accurately the patterns will be identified: increasing datafication will therefore increase the risk of the discovery of patterns coinciding with sensitive traits.

Thus in the era of datafication, it may be unavoidable that large data sets will contain patterns coinciding with sensitive traits because they cover "activities resulting from (protected) opinions or beliefs".[409] If these patterns can be found in collections of non-sensitive data, the safeguards against the processing of sensitive data may become less effective.

---

406 Damon Centola and Andrea Baronchelli, 'The Spontaneous Emergence of Conventions: An Experimental Study of Cultural Evolution' (2015) 112 Proceedings of the National Academy of Sciences 1989, 1989 <http://www.pnas.org/content/112/7/1989> accessed 19 March 2019 and S5 in the Supporting Information.

407 For examples, see Stephen H Kellert, *In the Wake of Chaos: Unpredictable Order in Dynamical Systems* (University of Chicago press 1994); Jean-Francois Gouyet and B Mandelbrot, *Physics and Fractal Structures* (1 edition, Springer 1996); Volker Grimm and others, 'Pattern-Oriented Modeling of Agent-Based Complex Systems: Lessons from Ecology' (2005) 310 Science 987 <http://science.sciencemag.org/content/310/5750/987> accessed 20 March 2019.

408 Qing Yi Feng and others, 'An Exploratory Statistical Approach to Depression Pattern Identification' (2013) 392 Physica A: Statistical Mechanics and its Applications 889, 894 <http://linkinghub.elsevier.com/retrieve/pii/S0378437112009211> accessed 19 March 2019. For a formal treatment of the same thesis: Bar-Yam (n 372) ch 2 (especially pages 296-297).

409 Council of Europe (n 72) para 44.

5.3 Exploring the risk of sensitive data

The French "Loi n° 78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés" is likely the first instance where special categories of personal data were codified in Europe. Article 31 requires consent for the processing of personal data regarding "racial origin, political, philosophical or religious opinion or trade union membership". The preparatory report accompanying the legislative proposal offered the following rationale: "The essential idea is not so much to prohibit the processing of [sensitive data], as to open a right to direct control to the citizen on the use of the data he has provided."[410] From this, it seems reasonable to infer that the consent requirement was introduced to reduce the risk that data subjects lose control over their sensitive data.

Article 6 of the 1980 Council of Europe Convention no. 108 ("Strasbourg Convention") added data regarding health and sexual life to this list.[411] The explanatory report associated with the Convention states that "there are exceptional cases where the processing of certain categories of data is *as such* likely to lead to encroachments on individual rights and interests" (emphasis added). The COE Member States agreed that this was true for all the characteristics listed in article 6.[412] This agreement is likely based on the persecution of religious and ethnic minorities, disabled persons and trade union members in Europe during the period 1933-1945. Even though the COE Member States agreed on the types of data of which processing could lead to encroachments on individual rights, they did apparently not agree on the risk, as article 6 contains no remedy. It leaves the choice of the appropriate safeguards to the Member States.

Article 8 of the 1995 Data Protection Directive harmonised the special categories of data for all EU member states and codified a number of exceptions and safeguards regarding its processing.[413] Article 8 allows processing sensitive data based on the data subject's consent (art. 8(2)(a)). Additionally, the directive provides a number of

410 M Foyer, 'Projet de Loi (No 2516) et Propositions de Loi (Nos 1004 et 3092)' (Assemblé Nationale 1977) 3125 13 <http://www.senat.fr/rap/l77-3125/l77-31251.pdf> accessed 19 March 2019.

411 Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data, done at Strasbourg 28 January 1981, ETS 108.

412 Council of Europe (n 72) 9.

413 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.

possible cases where the benefits may outweigh the risks, like the processing of health data for public health purposes (art. 8(3)). It also provides common-sense exceptions, like the processing of membership information by trade unions and religious institutions (8(2)(d)) and the processing of information that data subjects themselves have "manifestly made public" (8(2)(e)). It also opens an avenue for member states to codify exceptions in national law "for reasons of substantial public interest" (art. 8(4)). Article 9 of the GDPR is essentially equivalent, with a small number of additions. In EU law, the processing of sensitive data is generally forbidden, unless it is specifically allowed; at the same time, the "substantial public interest" provision offers considerable discretion to Member States.

Article 9 of the GDPR and its predecessors offer a typically European view on what kinds of personal data count as sensitive. The 1980 OECD Recommendation "Concerning Guidelines Governing the Protection of Privacy and Transborder Flows of Personal Data" does not contain a similar provision. Instead, it states: "[...] it is debatable to what extent people belonging to a particular group (i.e. mentally disabled persons, immigrants, ethnic minorities) need additional protection against the dissemination of information relating to that group."[414] Similarly, the 2013 revised OECD recommendation does not contain a provision resembling art. 6 of the Strasbourg Convention.[415] The United States, for example, does not proscribe the processing of data regarding ethnicity or religious beliefs.[416]

If Baldwin is right that regulation is "inherently about the control of risks"[417], the open-ended number of possible exceptions makes it difficult to pinpoint the risk that article 9(1) GDPR aims to control. One risk easily associated with the processing of sensitive data is that of discrimination and unfair treatment. However, discrimination is already directly addressed in European (EU and COE) and national non-

414 Council of the OECD (n 69) para 32.

415 Council of the OECD, 'OECD Guidelines Governing the Protection of Privacy and Transborder Flows of Personal Data - C(80)58/FINAL, as Amended on 11 July 2013 by C(2013)79' (Organization for Economic Cooperation and Development 2013) 17 <http://www.oecd.org/sti/ieconomy/2013-oecd-privacy-guidelines.pdf> accessed 19 March 2019.

416 'The US has no special category of "sensitive data" but US privacy law does protect certain forms of data more stringently (health, financial).' Daniel J Solove, 'What Is Sensitive Data? Different Definitions in Privacy Law' (*Privacy + Security Blog*, 31 July 2014) <https://teachprivacy.com/sensitive-data-different-definitions-privacy-law/> accessed 20 March 2019.

417 Baldwin, Cave and Lodge (n 82) 83.

discrimination legislation with its own doctrine and an associated body of human rights case-law in the area of administrative, commercial and criminal law.[418] The original rationale for article 21 of the aforementioned "Law n° 78-17" was to offer an opportunity for control to the data subject; this indicates that the perceived risk was inherent in collection, processing and dissemination of sensitive personal data unbeknownst to the data subject.[419] In circumstances where information asymmetries are relatively small, for example where a limited number of controllers process simple lists of names and attributes, consent arguably offers an adequate opportunity for control. But in the era of datafication, individual data subjects have to deal with a large number of controllers[420] processing large volumes of data, which is processed by algorithms unknown to the data subject. Furthermore, data subjects may not be aware that algorithmic processing of non-sensitive personal data can reveal sensitive traits. This information asymmetry may limit the effectiveness of the "opportunity of control" that consent can provide. Data subjects face high costs of information if they want to carefully assess all the situations where they want to specifically allow the processing of their data.[421]

The risk assessment behind the general prohibition of art. 9(1) GDPR appears to embody some ambiguity. It seems unlikely that the EU legislator intended to allow discrimination based on ethnic origin, religious beliefs or health status, if only a data subject had "direct control" over this use of personal data. At the same time, compliance with the GDPR (for example by obtaining consent) is considered to offer enough protection to allow the free movement of sensitive data within the EU and towards a number of other jurisdictions where the European Commission finds data protection law "adequate" (articles 1(3) and 45(1)).[422] This seems to represent a justifiable trade-off. The processing of sensitive data may have risks, but it also has useful public and private sector applications. This indicates that datafication is another example of the risks associated with Beck's notion of the Risk Society:

---

418 European Union Agency for Fundamental Rights and European Court of Human Rights (n 90).

419 Foyer (n 410) 13. See also Daniel J Solove, 'A Taxonomy of Privacy' (2006) 154 U. Pa. L. Rev 477, 483–558.

420 Schermer and Wagemans (n 307) 40.

421 McDonald and Cranor (n 144) 562.

422 European Commission, 'Adequacy of the Protection of Personal Data in Non-EU Countries' (n 208).

> "a central contradiction of risk society results from the fact that the world is confronted with large-scale threats whose origin lies in the triumphs of modern society (more industry, new technologies), threats which, in view of the institutionalized state promise of security, can nevertheless neither be adequately confirmed nor attributed, nor compensated, nor (preventively) managed in accordance with prevailing legal, scientific and political principles."[423]

On point as Beck's analysis is, it offers no concrete perspective on risk management. To this end, Klinke and Renn have proposed a classification method for the evaluation and management of modern risks.[424] This method is useful to assess whether the GDPR's management methods relating to sensitive data are in line with the risk. From the fact that sensitive data are treated more strictly than other data, we can assume that the EU legislator estimated the risks emanating from the processing of this data to be higher than those from the processing of other types of personal data. But from the trade-off visible in the strategy associated with sensitive data, it seems reasonable to presume that the EU legislator has no clear view of the damage potential. For those risk types, Klinke and Renn have chosen the name "Pythia" after the ambiguous prophecies of the Oracle at Delphi in ancient Greece.[425] This risk class requires a precaution-based management strategy, consisting of:

> "Containment of application in space and time, constant monitoring of potential side effects, development of functional equivalents, promoting diversity and flexibility and capacity building for organizational competence (...)"[426]

However, the consent requirement of article 9 GDPR suggests that the EU legislator also aims to enable data subjects to participate in the decision-making process regarding the processing of their sensitive data. This strategy corresponds with the "Medusa" risk class, where the damage potential is known, the disaster potential is low, but social mobilisation is high. This risk class requires a discourse-based management strategy, consisting of:

423 Ulrich Beck, *World at Risk* (Polity Press 2009) 30.
424 Klinke and Renn (n 275) s 3.4.
425 ibid 1081–1083.
426 ibid 1088.

> "building up consciousness, building confidence, strengthening trustworthiness in regulatory bodies, and initiating collective efforts of institutions for taking responsibility."[427]

When dealing with the risks of the processing of sensitive data, the GDPR appears to incorporate a majority of the elements of a precautionary strategy, consistent with a Pythia-type risk, although some elements have a stronger presence than others.[428] The principle of data minimisation (art. 5(1)(c)) promotes containment in space and time. Independent supervisory authorities (arts. 51-59) provide constant monitoring. The less stringent criteria for non-sensitive data could be construed as an incentive for controllers to find alternatives for the processing of sensitive data because it encourages the development of functional equivalents. The pressures of competition in the free market can be seen as an incentive for "investments in diversity and flexibility" and the obligation in specific cases to employ a data protection officer (art. 37-39 GDPR) aims to promote organisational competence.

Elements of a discourse-based strategy are less prominently visible in the GDPR. "Building up consciousness" for data subjects appears to be limited to giving consent. The GDPR aims to achieve the trustworthiness of regulatory bodies mainly by requiring independence for the Independent Supervisory Authorities and avoiding discourse: these Authorities need not consult data subjects for policy decisions. Similarly, the introduction of seals and marks (article 42-43 GDPR) does not require controllers or supervisory authorities to involve data subjects; data protection impact assessments could involve the participation of data subjects or their representatives, but only "where appropriate" and "without prejudice to the protection of commercial or public interests or the security of processing operations" (art. 36(9)). Supervisory authorities have no obligation to engage in collective efforts with other regulatory bodies in the field of anti-discrimination law or consumer law.

In conclusion, the GDPR appears to see the processing of sensitive data as posing a higher risk than the processing of other personal data. The associated management strategy is both precaution-based and – in a limited way, through the consent requirement – discourse-based. Applying Klinke and Renn's model does therefore not lead to an unambiguous risk classification and the rationale for the proposed remedies does not become totally clear. Additionally, the risk analysis underlying the

[427] ibid 1088–1089.
[428] ibid 1088.

GDPR appears to be incomplete: it does not account for the fact that the risks that article 9 aims to address can also occur as a result of the processing of non-sensitive data. The next sections will offer two examples.

5.4 First example: Discovering protected traits in complex systems

As a first example, three scenarios serve to illustrate the possibility that non-sensitive data can reveal sensitive traits.

Firstly, a number of explicit conventions associated with protected traits can be detected by comparing consumer behaviour to publicly observable features. As an example, if consumers observe religious food and drink prohibitions or religious holidays, their shopping patterns will reflect this convention through the presence or absence of certain products, and through peaks or dips in acquired quantities at predefined intervals. More short-lived conventions are also known to coincide with protected traits: retail brands can acquire temporary popularity among certain ethnic groups.[429] A data set containing dates and timestamps for retail transactions with product codes (and quantities, where applicable) linked to unique personal identifiers suffices to extract the protected traits of "religious or philosophical beliefs" and "ethnic origin" whenever the customer can be linked to the identifier.

In a second scenario, a controller can ask a limited number of consumers to consent with the processing of sensitive data. Based on their consent, the controller can search for emergent patterns coinciding with known sensitive traits. If a pattern is found, a controller can then induce sensitive information about other data subjects without processing their sensitive data. As an example, Duhigg documented how analysis of buying patterns revealed that consumers who voluntarily shared information on their

[429] Center for Disease Control, 'Cigarette Brand Preference Among Middle and High School Students Who Are Established Smokers - United States, 2004 and 2006' (2009) 58 Morbidity and Mortality Weekly Report 112, 112–113 <https://www.cdc.gov/mmwr/preview/mmwrhtml/mm5805a3.htm> accessed 19 March 2019; quoting Office of Applied Studies, 'Cigarette Brand Preferences in 2005' (United States Department of Health and Human Services, Substance Abuse and Mental Health Administration 2007) Short reports <https://www.datafiles.samhsa.gov/study-publication/cigarette-brand-preferences-2005-nid15156> accessed 20 March 2019; Jennifer Cullen and others, 'Seven-Year Patterns in US Cigar Use Epidemiology Among Young Adults Aged 18–25 Years: A Focus on Race/Ethnicity and Brand' (2011) 101 American Journal of Public Health 1955 <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3222378/> accessed 19 March 2019.

pregnancy and due date, bought "unusually large quantities of unscented lotion around the beginning of their second trimester". This information was then used to single out other consumers, who had also started buying unscented lotion at a certain time, and target them with advertising for products related to pregnancy and newborns. Special care was taken to hide the fact that the controller had discovered an indicator for pregnancy.[430]

A third scenario entails observing emergent patterns from direct interactions between humans. Especially where people use automated platforms for communication, these interactions can be observed and analysed because many events on the network are logged by the software controlling the network choke points. This often happens for billing purposes and to detect errors and fraud.[431] Madan et al. have found that analysis of the number of interactions between mobile phones and choke points can reveal patterns that coincide with depression and influenza; knowledge of the contents of the interactions is not necessary. For this analysis, they again processed sensitive data for a limited number of data subjects to match the interaction patterns with sensitive information.[432]

5.4.1 Complex systems theory and the observation of emergence from non-sensitive personal data

This example of discovering sensitive traits in complex systems concerns supervised learning: a controller uses the input data to predict any number of known attributes of data subjects to identify a correspondence with a known (emergent) property. In the scenarios resembling those mentioned at the beginning of this section, finding emergent properties usually happens in three stages. In the training stage, an algorithm is designed or optimised using part of an existing data set to discover relevant patterns that correspond to known properties. In the testing stage, the resulting algorithm is used on the remainder of the existing data to verify its efficiency. In the deployment stage, "new" personal data is mined for occurrence of

[430] Duhigg (n 116) 194–195; 209–210. The technique is known as pattern mining following feature extraction: Bart Custers and others (eds), *Discrimination and Privacy in the Information Society - Data Mining and Profiling in Large Databases* (Springer 2013) 36–39.

[431] "It's impossible to overstate the importance of logging." Prevelakis and Spinellis (n 257) 31.

[432] Anmol Madan and others, 'Sensing the "Health State" of a Community' (2012) 11 IEEE Pervasive Computing 36, 38–39 <http://ieeexplore.ieee.org/document/6072198/> accessed 20 March 2019.

the pattern identified in the training stage.[433] Assuming that the pattern identifies a single property, recognition of the pattern in personal data is equivalent to assigning the property to an identifiable natural person, with bias as a possible result.[434]

However, during the training stage, the processing of sensitive data is often necessary because the emergent property can be known to be sensitive. In such cases, a valid procedural safeguard from art. 9(2) GDPR is required to lawfully process the data. Note that if the patterns coinciding with sensitive traits are public knowledge, as is the case with published scientific results or well-known conventions, discovery may not trigger this prohibition. For example, instead of asking shoppers to disclose their pregnancy status, a controller could be inspired by scientific knowledge to look for shifts in buying patterns toward unscented lotions in order to find pregnant customers.[435] In the testing and deployment stages, recognising these patterns in subsequently acquired data no longer directly requires the processing of sensitive data.

5.4.2 Lawfulness of pattern recognition under the prohibition of article 9(1) GDPR

If controllers use pattern recognition algorithms to directly compile a list of data subjects with their associated protected traits, this constitutes a violation of art. 9(1) GDPR. However, if no such list is compiled, the lawfulness of the processing of such data is less clear. Art. 9(1) reads:

> "Processing of 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 shall be prohibited."

---

433 Toon Calders and Indrė Žliobaitė, 'Why Unbiased Computational Processes Can Lead to Discriminative Decision Procedures' in Bart Custers and others (eds), *Discrimination and Privacy in the Information Society - Data Mining and Profiling in Large Databases* (Springer 2013) 43–45. Some processes include validation and/or testing stages; in those cases, the training data is subdivided in several sets.

434 Zhao and others (n 403); Wang and Kosinski (n 387).

435 Steven Nordin and others, 'A Longitudinal Descriptive Study of Self-Reported Abnormal Smell and Taste Perception in Pregnant Women' (2004) 29 Chemical Senses 391 <https://academic.oup.com/chemse/article/29/5/391/368321/A-Longitudinal-Descriptive-Study-of-Self-reported> accessed 13 February 2019.

However, it is unclear from this text whether "revealing" refers to "processing", to "personal data", or to both. If it refers to "processing" or to both "processing" and "personal data", the GDPR's wide definition of processing would bring pattern recognition in big data under the ambit of the prohibition. After all, "processing" is defined as "any operation or set of operations which is performed on personal data or on sets of personal data (...) (art. 4(2)). Only if "revealing" refers to "personal data", the data itself would need to explicitly enumerate sensitive traits for the prohibition to be directly applicable.

Recital 51 could help to resolve this ambiguity. It appears to indicate that "revealing" in article 9(1) should be interpreted as referring only to "personal data":

> "Personal data which are, by their nature, particularly sensitive in relation to fundamental rights and freedoms merit specific protection as the context of their processing could create significant risks to the fundamental rights and freedoms. Those personal data should include personal data revealing racial or ethnic origin, ..."

However, according to the Court of Justice of the European Union in the *Nilsson and others* case, a non-binding text such as a recital "cannot be relied on as a ground for derogating from the actual provisions of the act in question".[436] A data subject claiming that revealing a sensitive trait through pattern recognition triggers the prohibition of art. 9(1) will experience an infringement of his data protection rights if a court uses the recitals to allow the practice under art. 9(1).

A controller or data subject proficient in other languages than English might study other versions to resolve the ambiguity. But it turns out that not all language versions of the GDPR convey the same message. The French text of art. 9(1) is clear: pattern recognition falls under article 9(1). In the sentence:

> "Le traitement des données à caractère personnel qui révèle l'origine raciale ou ethnique, (...)",

"qui révèle" agrees in number with "traitement" and not with "des données à caractère personnel", indicating that the prohibition applies if the processing reveals sensitive traits, and not just if the data enumerates them. The German text of article 9(1) is also unambiguous, but it states the opposite. In the sentence:

[436] Case 162/97 *Nilsson and Others* [1998] ECR I-07477, para 54.

> "Die Verarbeitung personenbezogener Daten, aus denen die rassische und ethnische Herkunft (...) hervorgehen (...)"

"denen" agrees in number with "Daten" and not with "Verarbeitung", indicating that the prohibition only applies if the data themselves enumerate sensitive traits.

To be fair, in English, French as well as German, recital 51 of the GDPR indicates that the concept of sensitivity applies to the quality of the data. But this may be of little help in the light of the *Nilsson and others* decision. Moreover, no EU language takes precedence over any other. English, French and German are all official languages of the European Union, but they share this status with all other official languages of the EU Member States.[437] Therefore, inconsistencies cannot be resolved by merely comparing different language versions.

Article 9 is the successor of article 8 in the 1995 Data Protection Directive (DPD). This text contains similar – yet different – inconsistencies across the different language versions. Like recital 51 of the GDPR, recital 33 of the DPD indicates that only data – not processing – can have a sensitive nature in English, French and German, but once more the English version of article 8(1) DPD is ambiguous. The German and French version both link "revealing" to sensitive data. This means that the French text of article 8(1) DPD has a different meaning from that of article 9(1) of the GDPR in the same language.

Despite this possible confusion, the case law of the European Court of Justice provides conclusive evidence supporting the thesis that "revealing" refers to "personal data" and not to "processing", at least as indicated by the 1995 Directive. In *Bodil Lindqvist*, the Court considered that the expression "data concerning health" must be given a wide interpretation; it then refers to "the fact that an individual has injured her foot (...)" constitutes medical data. This information was published on a web page, which means that it was also "processed" in the terms of art. 2(b) of the Directive, but this processing was not considered in determining whether art. 8(1) applied.[438] The Article 29 Data Protection Working Party has interpreted article 8 in a similar vein.[439] Until

[437] Art. 1, Regulation no. 1 determining the languages to be used by the European Economic Community, [1958] OJ L17/385, as most recently amended by Council Regulation (EC) No 1791/2006, [2006] OJ L363/1.

[438] Case 101/01 *Bodil Lindqvist* [2003] ECR I-12992, paras 50-51, see also Case T-190/10 *Egan and Hackett v European Parliament* [2012] ECLI:EU:T:2012:165, para 101.

[439] Article 29 Data Protection Working Party, 'Opinion 2/2010 on Online Behavioural Advertising' (2010) 00909/10/EN WP 171 19.

the ECJ adopts a new interpretation for the GDPR, we can therefore conclude that the test whether sensitive traits are revealed, should be applied to properties enumerated in the personal data, and not to forms of processing.

Mining data for sensitive traits may violate the principles of fairness, transparency, purpose limitation and minimisation in article 5(1) GDPR, and controllers who fail to disclose that they are mining personal data for sensitive traits violate data subjects' right to access and information as described in art. 15(1) GDPR. We explore these risks in more detail in section 5.6, because they also apply to the discriminatory effects of profiling, which we discuss in the next section.

5.5 Second example: Discriminatory profiling in complex systems

The 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 particular to analyse or predict aspects concerning that natural person's performance at work, economic situation, health, personal preferences, interests, reliability, behaviour, location or movements" (article 4(4)). The term therefore applies both to creating profiles relating to natural persons and creating profiling information based on anonymous data and subsequently matching natural persons to those profiles.[440]

Profiling serves many useful and legitimate purposes. In education, profiling based on datafication of interactions between students, materials and teachers could help optimise remedial teaching efforts, identify the most effective training materials or methods, and generally offer a learning environment more focused on the individual student without the need to increase the number of educators. In medicine, datafication of vital signs and treatments could help identify patterns that are both relevant to diagnosis or treatment, and non-obvious when observing a single

[440] This definition matches that of Wim Schreurs, Mireille Hildebrandt and Michaël Vanfleteren, 'Cogitas, Ergo Sum. The Role of Data Protection Law and Non-Discrimination Law in Group Profiling in the Private Sector' in Mireille Hildebrandt and Serge Gutwirth (eds), *Profiling the European Citizen: Cross-Disciplinary Perspectives* (Softcover reprint of hardcover 1st ed 2008 edition, Springer 2008) 241: "Profiling is the process of 'discovering' correlations between data in databases that can be used to identify and represent a subject and/or the application of profiles (sets of correlated data) to individuate and represent a subject or to identify a subject as a member of a group or category."

patient.[441] In social security settings, analysis could help assess the right support mechanisms for people in need, whilst reducing the risk of free-riding. In employment settings, datafication could increase productivity. All these opportunities could represent significant reductions in collective expenses or increased private profit. However, profiling could also lead to discrimination based on sensitive traits. Two classes of scenarios serve to illustrate this possibility.

For the first class of scenarios, consider Cathy O'Neil's example of St. George's Hospital Medical School. The screening of applications for eligibility was originally performed by specialised screening staff. In the 1970s, the school wished to partly automate the selection or rejection of applications in order to both increase efficiency and remove some of the capriciousness inherent in human judgment. A large volume of applications and corresponding admission/rejection decisions from previous years was used as training data. This resulted in a number of patterns that could be recognised through automated processing. A new batch of applications was then matched against these patterns. The resulting algorithm was partially based on screeners' tendency to reject applications that contained many grammatical mistakes. As it turned out, "birthplaces and, to a lesser degree, surnames" could predict grammatical correctness to a useful extent – they turned out to be features for grammatical correctness.[442] Therefore, these features were used in the decision making process. This led to a significantly lower chance of acceptance for applicants born abroad or living in immigrant neighbourhoods – two features correlated with sensitive traits like religion or ethnicity. In 1988, St. George's Hospital Medical School was fined for racial and gender discrimination as a result of using the algorithm.[443]

For the second class of scenarios, consider a bank performing risk analysis for loans and mortgages based not on previous risk assessments, but on past performance for previously issued similar loans. Based on geographical similarity, for example, a bank could decide not to offer mortgages based on area codes (ZIP codes), or only offer them under less favourable conditions, because the default rate in those areas is higher. This rule would not be based on the processing of sensitive data. But if the

441 Mayer-Schönberger and Cukier (n 1) ch 4; C McGregor, 'Big Data in Neonatal Intensive Care' (2013) 46 IEEE Computer 54.

442 See Flach (n 401) ch 10 on features and feature selection methods.

443 This scenario is described in O'Neil (n 268) 115–117; See also: Great Britain Commission for Racial Equality, *Medical School Admissions: Report of a Formal Investigation Into St. George's Hospital Medical School* (Commission for Racial Equality 1988).

affected areas include racially segregated neighbourhoods,[444] such a decision making process can be a source of indirect discrimination based on a sensitive trait. Similar examples could be applicable in the context of education, health care, employment and other areas.[445]

5.5.1 Complex systems theory and discrimination through profiling

Profiling with non-obvious discriminatory effects is an example of unsupervised learning: the controller does not necessarily know in advance by which traits data subjects will be grouped, but only that the members of each group will have some undetermined similarity. In the two classes of scenarios above, the cause of the discriminatory outcome is different.

In the example of St. George's, the algorithm that was built on previous acceptance/rejection data reflected human decision-making processes whose results coincided with an emergent property, namely that immigrants and people born in an immigrant community not only have more difficulty writing grammatically correct English, but also tend to live near each other.[446] The algorithm therefore introduced bias in the automated selection process. In the example of assessing the risk of mortgages based on area code, the data itself is not biased but the algorithms used to generate profiling parameters may lead to outcomes indirectly representing protected traits: an emergent property is represented in data capture. If a controller then uses the data for profiling under the assumption the data offers a neutral representation of the relevant properties of data subjects, biased decisions may be the result.

In cases where data is obtained through datafication of a complex system, a sufficiently sophisticated algorithm can present itself to us as a complex system.[447]

---

444 In terms of complex systems, segregation can be an emergent property of a city. Thomas C Schelling, 'Dynamic Models of Segregation' (1971) 1 Journal of mathematical sociology 143, 181.

445 Executive Office of the President, 'Big Data: A Report on Algorithmic Systems, Opportunity, and Civil Rights' (Executive Office of the President 2016) 10–22 <https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/2016_0504_data_discrimination.pdf> accessed 13 February 2019.

446 Dino Pedreschi, Salvatore Ruggieri and Franco Turini, 'The Discovery of Discrimination' in Bart Custers and others (eds), *Discrimination and Privacy in the Information Society - Data Mining and Profiling in Large Databases* (Springer 2013) 95–96; O'Neil (n 268) ch 6.

447 Bar-Yam (n 372) ch 2; Hema R Madala and Aleksei Grigoevich Ivakhnenko, *Inductive Learning Algorithms for Complex Systems Modeling* (CRC Press 1994) 285–287; Ladyman, Lambert and Wiesner (n 374).

This can cause emergence in the results of the algorithm that mirrors properties of the system under datafication, including emergence resulting from behaviour that is induced by sensitive traits.[448] A data controller may not be aware of bias in the resulting profiling decisions because the processing of sensitive data would be required for their verification, and such processing could be unlawful. Ironically, the prohibition in article 9(1) GDPR may therefore prevent the discovery of the resulting indirect discrimination.[449]

Discovering discriminatory effects of algorithms without direct verification is largely an unsolved problem. The opacity of artificial neural networks, a widely used classification method, is well-documented.[450] Datafication results in data sets suffering from the "Curse of dimensionality".[451] This curse makes classification (an important goal of big data analytics) difficult; overcoming this difficulty within a reasonable amount of computing time can involve methods that increase the opacity of the algorithm even further.[452] This makes it almost impossible for humans (data subjects, data controllers and independent supervisory authorities) to detect unintentional indirect discrimination through profiling without directly processing sensitive data of the profiled data subjects. The fact that profiling algorithms deserve protection as intellectual property or trade secrets (recital 63 GDPR) can make discovery by data subjects or their representatives even more difficult.

5.5.2 Lawfulness of profiling based on emergent properties under article 22 GDPR

Data subjects have 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" (Art. 22(1)). In the two sets of scenarios

448 Schermer (n 290) 138.

449 Similarly: Dwork and Mulligan (n 224) 37.

450 AM Wildberger, 'Alleviating the Opacity of Neural Networks', *1994 IEEE International Conference on Neural Networks, 1994 IEEE World Congress on Computational Intelligence* (1994).

451 Flach (n 401) 243; Bernhard Anrig, Will Browne and Mark Gasson, 'The Role of Algorithms in Profiling' in Mireille Hildebrandt and Serge Gutwirth (eds), *Profiling the European Citizen: Cross-Disciplinary Perspectives* (Softcover reprint of hardcover 1st ed 2008 edition, Springer 2008) 83.

452 Jenna Burrell, 'How the Machine "Thinks": Understanding Opacity in Machine Learning Algorithms' (2016) 3 Big Data & Society 2053951715622512, 9 <http://dx.doi.org/10.1177/2053951715622512> accessed 19 March 2019; Flach (n 401) 227.

illustrated above, we assume that a decision based on profiling that mirrors sensitive traits significantly affects the relevant data subjects, because discriminatory effects coinciding with sensitive traits are relevant to their rights and freedoms. Therefore, the scenarios must be tested for lawfulness against the other provisions of article 22.

The right not to be subject to profiling does not apply if such a decision is necessary for entering into, or performance of, a contract; if the profiling is authorised by law, or if decision-making is based on the consent of the data subject. (Art. 22(2)(a-c)). Limiting the analysis to consumers, service providers and retailers can become controllers when they offer consumers personalised services such as recommendations and other offers based on prior behaviour. This could suffice to justify profiling, especially considering the fact that such a contract also provides lawfulness in the sense of article 6(1)(b). Obtaining consent from a consumer can similarly provide a basis for profiling as well as a basis for lawfulness.

In cases where contract or consent form the basis for profiling, "suitable measures" should be implemented to safeguard the rights and freedoms of the data subject (Art. 22(3)); if profiling is authorised by law, of course the law should provide such measures. Suitability is an open norm.[453] Considering that art. 22 addresses the risk that automated decision making traps us 'in patterns that perpetuate the basest or narrowest versions of ourselves'[454] or replaces human contact and the opportunity to enter into negotiations, suitable measures could consist of an opportunity to oppose the decision, and to request a revised decision from an authorised human agent who will meaningfully reconsider the proposed automated decision, possibly based on additional information[455] – instead of merely reiterating that "Computer says no."[456] However, in the scenarios presented above, such measures would require that consumers be aware of the fact that they are being treated unfairly. This is far from certain: it would require the possession of sensitive data of a significant part of the profiled population as well as the output of the algorithm for those individuals. In scenarios where profiling results in the inadvertent creation of "filter bubbles" –

453 European Digital Rights, 'Proceed with Caution: Flexibilities in the General Data Protection Regulation' (EDRi - European Digital Rights 2016) 19 <https://edri.org/files/GDPR_analysis/EDRi_analysis_gdpr_flexibilities.pdf> accessed 19 March 2019.

454 Dwork and Mulligan (n 224) 40.

455 ibid; Arnoud Engelfriet and others, *De algemene verordening gegevensbescherming: artikelsgewijs commentaar* (Ius mentis 2017) 107–108.

456 Computer Says No, 'Computer Says No', *Wikipedia* (2017) <https://en.wikipedia.org/wiki/Computer_says_no> accessed 19 March 2019.

limiting the amount of information that data subjects can easily find – recognising unfairness is even harder.[457]

Article 22(4) prohibits automated decision-making and profiling based on sensitive data, unless it is based on explicit consent of the data subject or a substantial public interest is served (art. 22(4) and 9(2)(a, g)). It is not sure whether this prohibition applies in the above classes of scenarios, because discriminatory effects can occur even when the data used for profiling is not sensitive in and of itself. Furthermore, the discovery of discriminatory effects of profiling requires verification through the use of sensitive data. Conclusive evidence of discriminatory effects may not be obtainable in cases where the output of a machine learning algorithms does not completely correspond to the presence or absence of a sensitive trait.

The next section tests the two examples against the principles of processing.

5.6 Pattern recognition, profiling and the principles of processing

The examples above illustrate that pattern recognition and profiling could introduce a risk in the sense that the GDPR's measures do not result in the achievement of its goals – respect for fundamental rights, observation of freedoms, and possibly offering a form of control to data subjects[458] – even if the letter of the law is followed. The Regulation aims to prevent the creation of this risk by imposing technologically neutral principles on the processing of personal data: lawfulness, fairness, transparency, purpose limitation, data minimisation, accuracy, storage limitation, integrity, confidentiality and accountability. The importance of these principles is stressed by the fact that their infringement carries the highest penalty that the GDPR can impose (up to 20 million Euros or 4% of worldwide annual turnover, art. 83(5)(a)). We will not review the principles of accuracy, integrity and confidentiality, because we assume adherence to these principles to be in the self-interest of the controller, either because they are essential for the effectiveness of processing (accuracy and integrity) or because they could serve to maintain a competitive advantage (confidentiality). However, it is relevant to consider whether pattern recognition and profiling as described in the above examples would violate the principles of lawfulness, fairness, transparency, purpose limitation, data

457 Pariser (n 59) ch 8.
458 Recital 4, GDPR.

minimisation, and accountability, because these principles mainly serve to protect the interests of data subjects.

For most consumers, lawfulness for the processing of personal data related to economic activities is based on art. 6(1)(b), making processing lawful if it is necessary for the performance of a contract. Arguably, this provision enables controllers to lawfully store and analyse logs of customer transactions for billing and "personalised offers", either through case-by-case pattern recognition or profiling. This is a lawful exercise of the fundamental rights of freedom of contract and freedom to conduct a business: consumers are free to enter into contracts offering personalised services and commercial enterprises are free to customise their services if they believe it is good for business.[459] In a legal sense, a controller can provide for fairness and transparency as required by article 13, particularly 13(2)(f), by offering general contract terms and conditions, and/or privacy statements. Verhelst has coined the term "privacy contracts" for these types of contracts.[460] Data subjects are not necessarily aware of the contents of their privacy contracts: terms, conditions and their associated privacy policies take very long to read and many consumers do not bother reading them.[461] Even if data subjects do read them, it takes significant effort to evaluate privacy contracts for lawfulness. This is also true for supervisory authorities. For example, the Windows 10 operating system terms and conditions had been in use for two years, governing more than 4 million agreements in the Netherlands alone, before the Dutch supervisory authority examined them in 2017 and opined that the processing of personal data by Microsoft was not in agreement with applicable law.[462] Even if the

459 For another view on this matter, including an analysis of this class of cases regarding purpose limitation, see Article 29 Data Protection Working Party, 'Opinion 03/2013 on Purpose Limitation' (2013) 00569 /13/EN WP 203 21–27 <https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2013/wp203_en.pdf> accessed 20 March 2019.

460 Verhelst (n 119) chap 3.

461 Robert P Bartlett III and Victoria C Plaut, 'Blind Consent? A Social Psychological Investigation of Non-Readership of Click-through Agreements.' (2012) 36 Law and human behavior 293, 297 <http://psycnet.apa.org/journals/lhb/36/4/293/> accessed 19 March 2019; Rich Parris, 'Online T&Cs Longer than Shakespeare Plays – Who Reads Them? - Online T&Cs Word Counts Compared to Famous Books' (*Which? Conversation*, 23 March 2012) <https://conversation.which.co.uk/technology/length-of-website-terms-and-conditions/> accessed 20 March 2019.

462 Autoriteit Persoonsgegevens, 'Microsoft Windows 10 - De Verwerking van Persoonsgegevens via Telemetrie' (Autoriteit Persoonsgegevens 2017) Rapport definitieve bevindingen met correcties 148–160 <https://autoriteitpersoonsgegevens.nl/sites/default/files/01_onderzoek_microsoft_windo

terms comply with the law, the effects of algorithms may remain hidden due to the right to consider pattern recognition algorithms as a trade secret. This should not result in "a refusal to provide all information to the data subject" (recital 63), but the extent of this provision has not been tested; Wachter et al. concluded that the GDPR does not legislate a right to a full explanation of how algorithms work.[463] If a controller wishes to transfer personal data to third parties according to art. 13(1)(e), the recipients need not necessarily be disclosed if the "categories of recipients" are indicated, which can make the use of algorithms even less transparent.

The processing of special categories of data is lawful only if it is based on explicit consent; similarly, profiling based on sensitive data requires explicit consent. Asking a data subject for explicit consent is a means to offer data subjects an opportunity to participate in the decision-making process; the associated information is meant to reduce information asymmetries between the controller and the data subject.[464] Requesting consent can indeed serve this purpose, but data subjects may not easily distinguish terms and conditions from consent statements. Consequently, data subjects can often be expected to consent to processing without knowing the contents of their consent statement. They may choose to use "technical settings for information society services" based on considerations of convenience, especially in cases where "smart" devices like their mobile phone or tablet promise the best user experience if settings are such that consent is automatically given in the future (recital 32). Furthermore, pattern recognition or profiling may not require the processing of sensitive data even though the results may coincide with special traits. In those cases, consent is not required but the risk associated with the processing of sensitive data is still present. Data subjects may therefore be unaware of processing that can introduce discriminatory effects. The legal concepts and provisions associated with "sensitive data" (that can also differ between different languages) are therefore not necessarily sufficient to prevent discrimination based on sensitive traits.

The same may be true for the remaining principles of processing. For example, data minimisation and accountability in the context of pattern recognition offer a conundrum for controllers and supervisory authorities. In theory, applying these principles could reduce the risks associated with pattern recognition and discriminatory profiling. Data minimisation could require controllers to discard data

ws_10_okt_2017.pdf> accessed 19 March 2019.

463 Wachter, Mittelstadt and Floridi (n 135) 90–91.

464 Foyer (n 410) 13; See also Gomez (n 218) 198–207.

before emergent patterns become recognisable, and accountability could require controllers to prove that their processing does not introduce discriminatory effects. But this may not be straightforward. For controllers, minimisation can undermine the effectiveness of algorithms. "More data beats a cleverer algorithm"[465] – retaining data sets for recalibration or refinement in the future may be a legitimate reason to store them for a longer time, for example, to reduce discriminatory effects once they are discovered, or to provide proof of compliance with other provisions in the GDPR. Retaining this data is therefore in accordance with the principle of accountability (art. 5(2)), which makes this form of processing lawful "because it is necessary for the purposes of the legitimate interests pursued by the controller" or a supervisory authority (article 6(1)(f) GDPR). Furthermore, minimisation may not meaningfully benefit data subjects: in a world characterised by datafication, they will soon generate new data to be analysed. Finally, both pattern recognition and profiling may also be driven by features derived from anonymised data sets, which could remove the training stage of the machine learning algorithm from the purview of supervisory authorities if the anonymisation were performed effectively. The resulting algorithm could have discriminatory outcomes without offering any possibility to verify its effects. This may make it difficult for supervisory authorities hold controllers accountable for their adherence to the principles of art. 5.

To increase transparency regarding the risks of processing, article 15 grants data subjects the right to access to their personal data and additional information on how it is processed. Arguably, pattern recognition would be a part of "the purpose of the processing" (art. 15(1)(a)). Regarding the risks of profiling, article 15(1)(h) requires that data subjects are informed "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." From this, a controller could conclude that "meaningful information about the logic involved" need not be disclosed in all cases. For example, a retailer may not believe that personalised offers "significantly affect" a data subject, which would mean that the processing is not "profiling" in the sense of art. 22. The right to obtain "meaningful information about the logic involved" is once again moderated by recital 63. Furthermore, it is not clear what information should be disclosed in cases where

[465] Pedro Domingos, 'A Few Useful Things to Know about Machine Learning' (2012) 55 Communications of the ACM 78, 84 <http://dl.acm.org/citation.cfm?id=2347755> accessed 19 March 2019.

profiling is based on patterns built with anonymised datasets. No case law appears to yet exist on what constitutes "meaningful information" about logic, significance and envisaged consequences, even though the Data Protection Directive contained a similar provision (article 12(a), third element).

But discrimination based on sensitive traits – deliberately or not, either direct or indirect[466] – tends to be harmful, unlawful and unfair, even if the discovery of sensitive traits does not infringe art. 9(1) directly and the processing of personal data happens lawfully and transparently:

- Differential treatment accidentally coinciding with a sensitive trait could result in indirect discrimination if it has an adverse effect on "a far greater number" of data subjects sharing the protected trait[467] than other data subjects, or if "in percentage terms considerably less" data subjects sharing such traits enjoy some beneficial effect when compared to data subjects not sharing that trait.[468] This effect can occur if the controller uses unsupervised learning algorithms. Discrimination against data subjects sharing a certain race or ethnic origin in offering access to goods and services violates the principle of equal treatment required by the Racial Equality Directive, unless there is a justification of differential treatment (or incidentally, a defence to discrimination).[469]

- Any other forms of discrimination on sensitive traits may result in a violation of article 14 of the European Convention on Human Rights[470] if this discrimination affects the rights and freedoms protected in the Convention. In the cases of discovery of sensitive traits and profiling, this regards the right to personal and family life (art. 8, ECHR) since this right has horizontal application.[471] Protocol 12

466 See art. 2(2) of Council Directive 2000/43/EC of 29 June 2000 implementing the principle of equal treatment between persons irrespective of racial or ethnic origin (Racial Equality Directive), [2000] OJ L180/22; see also European Union Agency for Fundamental Rights and European Court of Human Rights (n 90) 22–31.

467 ECtHR *D.H. and Others v. the Czech Republic* 2007-IV 241 para. 87, quoting the European Court of Justice Case 170/84 *Bilka-Kaufhaus GmbH* v *Karin Weber von Hartz* [1986] ECR I-1620, para. 31.

468 Case C-171/88 *Rinner-Kühn* v *FWW Spezial-Gebäudereinigung*, [1989] ECR I-2743, para. 11-12.

469 European Union Agency for Fundamental Rights and European Court of Human Rights (n 90) s 2.6.

470 European Convention for the Protection of Human Rights and Fundamental Freedoms, Sept. 3, 1953, ETS 5, 213 UNTS 221.

471 Akandji-Kombe (n 188) 14–16.

to the Convention extends this prohibition to "any right set forth by law', but not all EU Member States are signatories to this Protocol.[472]

- There may be some extreme cases where controllers could expressly aim to discern sensitive traits in order to discriminate against data subjects. Such processing is not in accordance with the principle of purpose limitation (art. 5(1)(b)) because the intent to discriminate based on sensitive traits is not a *legitimate* purpose, unless it meets the specific criteria of articles 4 or 5, Racial Equality Directive.[473] In all other cases, such processing can be expected to be contrary to public policy or to accepted principles of morality in all EU Member States.

In spite of all that, these legal protections could lose their strength in the context of datafication. Especially in the course of profiling, discrimination could happen as an unintended side-effect of providing personalised offers or credit risk management. Furthermore, controllers with no intention to discriminate could at any time unwittingly collect or otherwise obtain biased data sets or have their algorithms stumble upon an emergent property coinciding with a sensitive trait. Those controllers can only detect discrimination by processing sensitive data, which – according to the Strasbourg Convention – introduces new risks of discrimination. Discriminatory effects may be inadvertently obfuscated by code or in biased data sets. Additionally, the nature of human society as a complex system guarantees that new ways to stumble upon sensitive traits will continuously become available as new patterns emerge.

This makes enforcement difficult for supervisory authorities: without specialised knowledge of the data set and the algorithms, it will be hard to distinguish neutral processing from accidental discrimination. Indirect discrimination may be difficult to distinguish from legitimate market segmentation or risk management; it may also be difficult to provide conclusive evidence for the use of sensitive data in intermediate steps of processing. The free movement of personal data within the EU (art. 1(3)) and

472 Protocol No. 12 to the Convention for the Protection of Human Rights and Fundamental Freedoms, done at Rome on 4 November 2000, ETS no. 177; List of signatories found at http://www.coe.int/en/web/conventions/full-list/-/conventions/treaty/177/signatures.

473 European Union Agency for Fundamental Rights and European Court of Human Rights (n 90) 43.

between the EU and other jurisdictions can complicate matters further.[474] Data can be transmitted to other controllers whose algorithms may have unknown properties.

This undermines the principle of accountability: unless supervisory authorities can detect such processing and enforce the corresponding penalty provisions – Tene seemed hesitant to assume this in 2011[475] – articles 9(1) and 22(4) may prove to be dead letters.

5.7 Potential remedies

The fact that the risk of processing sensitive data is not clearly defined makes choosing a remedy more difficult. Still, a number of possible measures could reduce the risks, but they all will have only limited effect.

Strengthening the consent requirement

A first suggestion for a possible measure could be to increase the number of situations where consent is required for processing, through a reinterpretation of art. 9(1). Requiring consent for all forms of processing revealing sensitive traits increases transparency and could improve informed decision-making before entering into privacy contracts. But effectiveness of this measure may well be limited. An important reason for this is that a consent requirement shifts the burden towards the data subject. Most consumers do not read their contracts.[476] Given the large number of privacy contracts that consumers engage in and the pervasiveness of datafication, data subjects might be asked for their consent so many times per year that they could eventually stop distinguishing between low-risk and high-risk situations. Sheer lack of time to evaluate all the associated privacy contracts and the desire to enjoy the benefits of new products and services built on datafication could lead to the erosion of consent statements until they are no longer distinguishable from general terms and conditions, which need not always be read to be applicable. This would defy the

474 See: Commission decisions on the adequacy of the protection of personal data in third countries and the EU-U.S. Privacy Shield at https://ec.europa.eu/info/law/law-topic/data-protection/data-transfers-outside-eu/adequacy-protection-personal-data-non-eu-countries_en and https://ec.europa.eu/info/law/law-topic/data-protection/data-transfers-outside-eu/eu-us-privacy-shield_en, respectively, accessed 13 February 2019.

475 Tene (n 361).

476 Nili Steinfeld, '"I Agree to the Terms and Conditions": (How) Do Users Read Privacy Policies Online? An Eye-Tracking Experiment' (2016) 55 Computers in Human Behavior 992, 995 <http://linkinghub.elsevier.com/retrieve/pii/S0747563215301692> accessed 20 March 2019; McDonald and Cranor (n 144) 562.

purpose of the consent requirement. Offering consent as a way to open an avenue for control becomes less effective in the context of datafication: information asymmetry makes effective control difficult for data subjects, and they have to deal with large numbers of controllers. If we agree with Klinke and Renn that "act[ing] prudently under the condition of uncertainty"[477] is advisable for Pythia-class risks, the burden of acting prudently should be carried by controllers.

Anonymisation

If personal data is anonymised, it no longer counts as personal data (recital 26). In theory, anonymisation could therefore limit the possibilities to discover sensitive traits from non-sensitive data. In an opinion issued in 2014, the Article 29 Working Party identified three risks of insufficient anonymisation:

- "Singling out: [...] the possibility to isolate some or all records which identify an individual in the dataset;
- Linkability: [...] the ability to link, at least, two records concerning the same data subject or a group of data subjects (either in the same database or in two different databases)
- Inference: [...] the possibility to deduce, with significant probability, the value of an attribute from the values of a set of other attributes."[478]

When dealing with anonymised data, associating an emergent property with a sensitive trait relates to the risks of linkability (a known sensitive trait of a particular data subject or group of subjects can be linked to supposedly anonymised data from another context) or inference.

Avoiding the risks of de-anonymisation is a hard problem for at least two reasons. Firstly, Mascetti et al. point out that "providing data utility and data subject's privacy are contrasting objectives."[479] Controllers may lack incentives to achieve effective anonymisation if it prevents them from fully reaping the benefits from the data they

477 Klinke and Renn (n 275) 1086.

478 Article 29 Data Protection Working Party, 'Opinion 05/2014 on Anonymisation Techniques' (2014) 0829/14/EN WP216 11–12 <https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf> accessed 13 February 2019.

479 Sergio Mascetti and others, 'Anonymity: A Comparison Between the Legal and Computer Science Perspectives' in Serge Gutwirth and others (eds), *European Data Protection: In good health?* (Springer 2013) 95.

collected. Brickell and Shmatikov analysed a number of data sets for the risk of inference (attribute disclosure); they found that privacy and utility were incompatible goals of data mining for the data sets they analysed. The loss of utility was so acute that they were doubtful that any data set could be anonymised and useful at the same time.[480] Using a different methodology, Li and Li claim to have found less dramatic results, although they still see that utility decreases with increased privacy.[481] Secondly, even if the loss of utility does not discourage anonymisation, there is no guarantee for success since it requires a non-trivial effort aimed at specific and well laid out goals. Current anonymisation techniques may not prevent the attribution of sensitive traits for all cases based on emergent properties as a result. For example, in its opinion on anonymisation techniques, the Article 29 Working Party concludes that only two out of seven evaluated techniques (K-anonymity and L-diversity) are capable of eliminating only one out of three risks of re-identification (singling out).[482] Not one of the seven techniques can rule out inference.

Data protection by design and by default

Article 25 of the GDPR, titled "Data protection by design and by default" embodies the precautionary approach that Klinke and Renn associate with the Pythia risk class ("containment in space and time (…), constant monitoring, development of equi-functional replacements, and investments in diversity and flexibility").[483] It calls on controllers to continuously consider the rights and freedoms of data subjects and to apply "appropriate technical and organisational measures". Recital 78 lists a number of possible measures:

> "minimising the processing of personal data, pseudonymising personal data as soon as possible, transparency with regard to the functions and processing of personal data, enabling the data subject to monitor the data processing, enabling the controller to create and improve security features."

480 Justin Brickell and Vitaly Shmatikov, 'The Cost of Privacy: Destruction of Data-Mining Utility in Anonymized Data Publishing', *Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining* (ACM 2008) 76–78.

481 Tiancheng Li and Ninghui Li, 'On the Tradeoff between Privacy and Utility in Data Publishing', *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining* (ACM 2009) 524–525.

482 Article 29 Data Protection Working Party, 'Opinion 05/2014 on Anonymisation Techniques' (n 478) 24.

483 Klinke and Renn (n 275) 1086.

It is difficult to predict whether these types of measures are specifically useful to prevent the association of emergent properties with sensitive traits. Technical measures like limitations on data storage quantity or duration, or limits on the possible coupling of data sets, appear not likely to be effective in preventing discriminatory effects. Firstly, such measures run counter to the economic and social promise of big data, which limits their feasibility. For example, to make analysis of consumer buyer patterns useful, it would typically be necessary to store and analyse at least one year's worth of data points to help account for cyclic behaviours associated with seasonal activities and other life events. But one year of data would also be more than sufficient to find emergent patterns that reveal sensitive traits. Pseudonymisation could be useful where real names can indicate sensitive traits. However, in the examples given in this chapter, real names are not required for discriminatory effects to occur. In cases where algorithms are opaque – as they often are, even to controllers – beneficial effects of monitoring by data subjects may be possible by means of verification. If a large group of data subjects can compare the effects of processing, discriminatory effects can become clear. However, this would require that a large number of data subjects be willing to share their sensitive information with other data subjects and possibly the controller. Otherwise, discovering or remedying discriminatory effects would not be possible. In more general terms, our analysis indicates that such measures have limited feasibility because they could only partially solve the problems associated with datafication or solve them only for a limited time. New emergent properties and new algorithms will continue to increase the possibilities to discover sensitive traits and introduce new risks of accidental discrimination. Keeping personal data secure would not likely prevent discriminatory effects, because it is not directly related to the discovery of sensitive traits.

Improving enforcement by independent supervisory authorities

Ramping up enforcement activities could reduce the risks by increasing controllers' accountability as part of a "constant monitoring" effort. Enforcement could be aimed at the material effects of algorithms as well as controllers' adherence to the principles and other requirements of processing, for example by stressing the need for data protection impact assessments (article 35). An important caveat is that discriminatory effects of algorithms cannot be evaluated by traditional data security practices and formal statements of controllers alone, or even by examining the source code of the algorithms. Datta observes that "traditional preventive access control and information flow control mechanisms are not sufficient for enforcing all privacy policies". External

audits of policy, code and data are necessary to discover discriminatory effects.[484] New emergent properties can arise at any moment and algorithms will be continuously tweaked: audits will provide only a limited form of certainty for only a limited time. Accidental discriminatory effects may only be discovered if additional sensitive data is available, which can pose an unsolvable dilemma for controllers and supervisory authorities. Another caveat is the interpretation of "similar significant effect" in the context of data protection impact assessments (art. 35(3)(a)). An extensive interpretation of this notion could dictate that these assessments are almost always necessary for large datasets, which increases the risk that they become 'a "tick box" list for compliance measures regardless of their actual impact on compliance.'[485]

Raising public awareness and participation

Several authors have noted the "privacy paradox". Schneier states that "(c)ourts have been reluctant to find a value in privacy, because people willingly give it away in exchange for so little".[486] It is not certain whether improved awareness of possible discriminatory effects of data mining and profiling will decrease the risk of these effects occurring. Sensitive traits are strong drivers for individual behaviour, which means that data subjects cannot easily hide or alter them. The free market would probably be the best proving ground for the effects of increased awareness. For example, offers containing a promise not to use data for personalisation and to delete data as soon as possible could be attractive for data subjects who place a high value on privacy. Such effects are not unthinkable: the fact that end-to-end encryption is now commonplace in instant messaging apps indicates that app developers feel pressure to increase consumer data security.[487] Indeed, Cofone found that the privacy paradox is best explained by consumers adapting their behaviour to the result of "uncertainty-

484 Anupam Datta, 'Privacy through Accountability: A Computer Science Perspective' in Raja Natarajan (ed), *Distributed Computing and Internet Technology* (Springer International Publishing 2014) 46.

485 Moerel (n 87) 52; Article 29 Data Protection Working Party, 'Guidelines on Data Protection Impact Assessment (DPIA) and Determining Whether Processing Is "Likely to Result in a High Risk" for the Purposes of Regulation 2016/679' (Article 29 Data Protection Working Party 2017) 17/EN WP248 8.

486 Schneier, *Data and Goliath* (n 210) 227. See Ignacio Nicolás Cofone, 'Privacy Tradeoffs in Information Technology Law' (Erasmus University 2015) 98–104 for additional references.

487 Andy Greenberg, 'Whatsapp Just Switched on End-to-End Encryption for Hundreds of Millions of Users' <https://www.wired.com/2014/11/whatsapp-encrypted-messaging/> accessed 20 March 2019; Katriel Cohn-Gordon and others, 'A Formal Security Analysis of the Signal Messaging Protocol', *Security and Privacy (EuroS&P), 2017 IEEE European Symposium on* (IEEE 2017).

based discounting", suggesting that reducing uncertainty could solve the paradox.[488] The use of specialised seals and marks (in accordance with art. 42 GDPR) could help consumers recognise services that offer increased data protection. Another possible transparency measure is a requirement to open up data and algorithms for public scrutiny (which is also a part of "data protection by design and default", see above). However, controllers have a legitimate interest in protecting their algorithms as trade secrets and/or intellectual property rights. This could limit the effectiveness of increased transparency.

5.8 Concluding remarks

Ubiquitous datafication and big data analytics are relatively new phenomena, the risks of which are not yet fully clear. This chapter aims to paint a clearer picture of two risks associated with datafication and analytics: that controllers can induce sensitive personal data from non-sensitive data, and that algorithms can show bias based on sensitive traits even when the requirements of art. 9 and art. 22 GDPR are met. This is possible because the traits that count as sensitive in the European context, are strongly driving data subjects' individual behaviour. Complex systems science identifies emergent properties as root causes for these risks. It also demonstrates that addressing them requires considerable effort and may not always be effective. Inconsistency between different language versions of the GDPR can cause confusion among both data subjects and controllers regarding their rights and obligations in this matter. At the same time, controllers wishing to actively prevent discriminatory effects from their processing may run afoul of the GDPR if they want to verify specific effects for data subjects sharing sensitive traits. Several measures may help clarify and reduce the risks, but there is no one-size-fits-all, comprehensive solution yet.

Due to the information asymmetries between controllers and data subjects, independent oversight by specialised supervisory authorities is indispensable and the GDPR wisely requires it. But compliance with data protection law alone may not be enough to prevent discriminatory effects. A data subjects' consent is too easily acquired, the consequences of processing are too uncertain, and algorithms are too opaque to rule out discriminatory effects even when the GDPR is followed. Effective prevention of the discrimination risk could require additional enforcement of consumer protection law (concerning possible unfair contract terms and unfair commercial practices), competition law (concerning possible abuse of a dominant

[488] Cofone (n 486) 115–120, 127.

position) and non-discrimination law. All these areas of law have specialised enforcement mechanisms already in place in EU countries. Enforcement efforts also need to strike a balance between the positive and negative effects of processing. Big data promises greater efficiency, e.g. in the field of health care, education and the market for goods and services. Processing and profiling have many lawful and useful applications in these areas and EU Member States cannot easily afford to abstain from them.

Considering that the risk of the processing of personal data remains ambiguous and that all available remedies have significant shortcomings, we believe that the risk management strategy of the GDPR for sensitive personal data was not optimally chosen. Designating deeply personal properties as "sensitive" denies the fact that these properties can appear as emergent properties in many different contexts, and are therefore expressed in countless ways: in today's datafied world, even non-sensitive data can easily reveal them. Furthermore, the usefulness of the consent requirement – carried over from one of the earliest examples of data protection law – decreases in the context of omnipresent datafication. Consent as an avenue of control over processing can become meaningless if a data subject has to give it many times every day in circumstances where the effects of processing are not clear to the controllers nor the data subjects themselves. The technologically neutral terms in which the GDPR is drafted makes the legal text durable, but at the same time it does not always offer useful guidance to data subjects and controllers.

But that no comprehensive solution is available does not at all mean that we believe that nothing can be done. Following Klinke and Renn's classification, two courses of action present themselves.

Firstly, as long as the risks of datafication remain ambiguous (the potential and damage potential of datafication remain unknown), risk management should strike a more effective balance between precaution-based and discourse-based measures.[489] Precaution alone, which calls for containment of applications, constant monitoring and more, may be effective in preventing adverse effects, but all-too effective precaution can also make the benefits of datafication unattainable. Applications departing from precautionary practice should therefore be possible. But to prevent a situation where the risks of datafication being borne exclusively by the data subjects and the benefits enjoyed exclusively by controllers,[490] such deviations should be

[489] Klinke and Renn (n 275) 1088.
[490] Beck, *World at Risk* (n 423) 142.

agreed on using discourse-based strategies, which "require strategies building up consciousness, building confidence, strengthening trustworthiness in regulatory bodies, and initiating collective efforts of institutions for taking responsibility. These are social goals that cannot be accomplished by risk experts or regulators alone."[491] Avenues for effective discourse may need to be designed in a bottom-up fashion; alternatively, underrepresented parties – consumers and data subjects – could be encouraged and subsidised for their efforts to improve their opportunities for participation in current institutions, similar to the European Commission's current programmes for consumer participation in lobbying and standardisation.[492] Some of these strategies could be designed for specific areas of commerce, or at the level of Member States, to allow for diversity resulting from differences at the national level. Collective efforts of supervisory authorities should at least consist of joint programmes for enforcement of data protection-, consumer protection-, competition-, and non-discrimination law.

Secondly, to increase the development of effective regulation and enforcement strategies, knowledge of the risk potential and damage potential of big data must be increased. Increased knowledge can support the correct assignment of responsibilities. This makes enforcement easier and possibly more efficient, especially if it can reduce the need for discourse-based strategies. We propose that complex systems science and the associated understanding of emergent properties can eventually provide insights in *why* the "computer says no".[493] It will be essential both in achieving a better understanding of the GDPR's shortcomings in the context of datafication and in resolving them. Increased knowledge can also enable controllers in remedying discriminatory effects. For example, if an algorithm can replicate society's bias, it can probably also be designed to remove it – something that humans may not be able to accomplish by mere instruction. However, complex systems science is a fairly new discipline.[494] It may take some time before the mechanisms underlying emergence in groups of individuals are sufficiently understood, and complex systems science may provide only a part of the insights necessary to understand the risks of datafication. Still, the examples we presented aim to illustrate

491 Klinke and Renn (n 275) 1088–1089.

492 European Parliament and Council Regulation (EU) No 254/2014 of 26 February 2014 on a multiannual consumer programme for the years 2014-20 and repealing Decision No 1926/2006/EC, [2014] OJ L 84/42, art 3(1)(b).

493 Computer Says No (n 456).

494 Bar-Yam (n 372) 1.

that this avenue of further research offers a real opportunity for increased understanding.

Contrary to the assertion of the European Commission that the GDPR is "future proof for decades to come",[495] it seems reasonable to assume that new legal and practical problems associated with the risks of datafication will continue to arise. We strongly recommend that legal scholarship build lasting alliances with the exact sciences (data science, complex systems science, computer science) and the social sciences to evaluate and remedy the risks of datafication, its effects on society and the effectiveness of the law.

[495] European Commission, 'Proposal for a General Data Protection Regulation' (n 81) 104.

6 In conclusion

6.1 Answering the research question

This research project aimed at answering the following question:

To what extent does the GDPR reflect or employ theories of power relations and risk management presented by Komesar, Barnett and Duvall, Beck, Perrow, Klinke and Renn, and complex systems science?

The question focused mainly on cases where consent or the performance of a contract provided the lawfulness of processing. The research question was approached using four sub-questions:

- How do the decision-making mechanisms in the GDPR itself, and in the EU lawmaking process that produced the GDPR, compare to other available decision-making mechanisms with regards to opportunities for effective participation by data subjects?
- How do the GDPR's protections for data subjects giving consent or entering into a contract compare to the protections in EU consumer protection law?
- To what extent were existing insights from the social sciences and environmental law applied in the GDPR insofar as it deals with the identification of risks of big data or with the addressing of new or unknown risks?
- Is the GDPR's protection of sensitive personal data adequate in the context of big data and relevant insights in the field of Complex Systems Science?

Considering the answers to the four sub-questions, the answer to the research question must be that the GDPR does not appear to reflect any of the models researched. Additionally, the GDPR appears to have used no other models originating in relevant scientific fields aimed at evaluating or addressing the risks of large-scale processing of personal data.

In **chapter 2**, a comparative institutional analysis was applied to the contents of consumer contracts to answer the first sub-question. The goal was to determine whether article 6(1)(1-2) would result in any bias favouring either data subjects or controllers in terms of participation opportunities. The fact that consent and the performance of a contract count as grounds for lawfulness of processing, without any further qualification except the general principles of processing, indicates that no explicit measures have been taken to reduce the risks of a bias favouring controllers in the outcome of the decision-making process in the market. Additionally, data subjects encounter high costs of information and organisation when entering into privacy contracts, where controllers can spread these costs over many transactions. This means that controllers can be expected to determine the terms governing the collection and use of personal data. The opportunities for data subjects to successfully challenge the terms of a privacy contract are small.

The GDPR was finalised within the EU legislative process, which offers individual data subjects and their associations relatively unfavourable opportunities for participation. The EU legislative process compares unfavourably for data subjects because the costs of information and organisation in the EU legislative proceedings are higher than in their own Member States' political systems. For controllers, these costs are probably lower due to controllers' lower numbers and higher stakes. This means that controllers have had better opportunities than data subjects in the EU legislative process to have the text of the GDPR reflect their own interests, e.g. through lobbying.

In **chapter 3**, Barnett and Duvall's model of power in social relations was used to conceptualise the power differentials resulting from big data. Controllers of personal data that use the performance of a contract as the basis for lawfulness have *structural power* over consumers, allowing them to dictate the terms of a contract; the data that is then collected based on the contract is used to increase a controller's *institutional power,* for example the power to observe a consumer and to share the results of observation with other controllers.

Power differentials can lead to unfair treatment if they remain unchecked. This applies both in the context of consumer protection and the context of data protection. An analysis using comparative law methods shows that the GDPR does not account for these power differentials as effectively as EU consumer protection legislation does. For example, by forbidding unfair commercial practices and unfair contract terms, EU consumer protection law can offer wide-ranging protections that consumers can understand relatively easily. In comparison, the standards governing the GDPR are

more difficult to understand. Consumer protection law also offers lower costs of organisation for collective redress than the GDPR does. This could put data subjects at a disadvantage, offering them less effective protections in cases where the GDPR – possibly a *lex specialis* in such contexts – would derogate from consumer protection law.

Chapter 4 has shown that, when compared to EU instruments of environmental protection law, the GDPR's instruments of risk evaluation and risk mitigation are less well-rounded. Instruments in the GDPR regulating contracts and consent do not clearly reflect the insights gained from models of risk management originating in the social and exact sciences. The GDPR nor its predecessors display a clear qualification of the risks associated with the processing of personal data. In the future, it will therefore be relatively difficult to determine the degree to which the GDPR will have achieved its stated goals: protecting data subjects' rights and freedoms. In contrast, several models from the social sciences, including qualifications of technological risk, can be recognised in a number of EU instruments of environmental law.

In the terms of Beck's model of the risk society, big data is a *modern risk*: its origins and effects cannot be exactly located. Perrow's theory of Normal Accidents predicts that big data, which arguably is based on complex and tightly coupled technology, will cause inevitable *normal accidents* at a system-wide scale. Unlike EU environmental protection legislation, the GDPR makes no obvious choices in the risks it wants to address in the cases of consent or contract, nor does it require that controllers or supervisory authorities employ a cyclic improvement process that encourages the development of better understanding and incremental improvement of risk mitigation measures.

The GDPR focuses on the rights of individual data subjects. This denies the possibility that adverse effects of big data can express themselves at the level of communities or societies. Seen in this way, the GDPR compares unfavourably to EU environmental protection law, which aims to protect the environment as a whole as well as individual humans, communities, societies and ecosystems. Consequently, the GDPR may not necessarily *by itself* prevent that a relatively small number of controllers – especially platform providers like, currently, "tech's frightful five" – can offload the risks of processing onto data subjects or society at large. If not addressed through other means, this could count as a power shift all by itself.

Finally, regarding the processing of special categories of personal data, the GDPR does not apply or reflect relevant insights from complex systems science. **Chapter 5** has shown that the processing of data covering many aspects of human behaviour has made the GDPR's anti-discrimination provisions less effective. The limits on the processing of, and profiling based on, sensitive data in articles 9 and 22 appear to be based on a predominantly *precautionary* strategy. This strategy offers safeguards against enumerating sensitive traits in personal data. But, even if not enumerated, sensitive properties can also be detected as *emergent properties* (as defined in complex systems science), which means that their expression as robust order in human behavioural patterns can be distinguished by algorithms if datafication is sufficiently pervasive. If sensitive traits are not specified during the processing of personal data, such processing may not violate the GDPR directly; at the same time, the principles of processing may not be consistently effective in protecting data subjects against adverse effects. When potentially discriminatory algorithms are available on dominant platforms, adverse effects could be expected in many aspects of commerce and communications despite precautionary measures. Furthermore, due to the inscrutability of these algorithms, automated discrimination based on sensitive traits can happen accidentally and it can remain undetected, both by controllers and supervisory authorities. Klinke and Renn's approach to risk management suggests that in a case like this where the risk is not yet completely understood, a discourse-based approach would be more suitable than a precautionary one.

6.2 Discussion

As it appears, the models examined in the preceding chapters have not been applied or integrated into the GDPR clauses regarding processing based on contract or consent. Additionally, no other model regarding risk management or power imbalances appears to lie at the basis of the provisions regarding contract and consent. In this regard, the GDPR differs from the Seveso III-directive, the Unfair Terms Directive and the Unfair Commercial Practices directive: these directives contain standards that display some degree of coherence with models of risk management and power distribution from the social sciences, even though these models are not expressly applied. This finding appears to confirm the criticisms from Moerel, Koops and Zarsky.[496]

[496] See section 1.7 above.

The absence of an underlying testable model in legislation could adversely affect the expected efficacy of the GDPR in least the following ways:

- It could reduce the effectiveness of enforcement strategies by Independent Supervisory Authorities if a lack of focus in enforcement efforts could prevent a cycle of permanent improvement in enforcement or processing practice.
- It could make the outcome of court cases less predictable. This could make developing a coherent doctrine in jurisprudence more difficult, which could, in turn, complicate the solving of any problems with the GDPR that may become evident in the coming years.

A possible explanation for why modelling the effects of big data could have escaped the attention of legislators was seen in section 4.5. Datafication was never the focus of regulation, as it may have often been regarded as a mere by-product of welcomed innovation through automation. Another possible cause was hinted at in the introduction and in section 5.2: in step with Moore's, Keck's and Kryder's laws, the number and the complexity of systems under datafication has increased exponentially. In this context, "surveillance capitalism" in its current form may be an emergent property of a complex system – a society under datafication – where personal data has become ubiquitous. Foreseeing the emergence of the current market for personal data may have been impossible, since emergent properties are unpredictable from the properties of the elements in the system.[497]

Regulating an emergent phenomenon almost at the same time as its emergence could be an impossible task for any legislative effort. Under these circumstances, Coase's prediction that the effects of *any* possible form of regulation would not resemble "anything an economist would call optimal" seems almost self-evident.[498] But no matter why a clear assessment of the associated risks and power shifts is absent, it is proposed here that the challenge that big data presents for society needs to be identified before it can be addressed. This research has shown that models from the social and the exact sciences can be of assistance in better understanding this challenge.

Legislative intervention to address risks and power imbalances seems appropriate. A relatively small number of actors, "Tech's Frightful Five" prominently among them, have managed to increase their structural power by achieving a dominant position in

[497] See section 5.2 above.
[498] See section 1.6 above and note 86.

the market for platforms through the collection of both capital and personal data. This gives them great institutional power: the mere size of the most successful platforms implies that the platforms themselves can be leveraged to unilaterally dictate terms and thereby increase this structural power even further in almost any market that can be accessed through telecommunications – including the marketplace of ideas. This increases the risk for unfair treatment.

For many consumers, escaping the Frightful Five's collection of personal data has become almost impossible in the course of their regular economic, social and intellectual activities. And if escaping it were possible, it remains questionable whether consumers would want to do without the benefits that these platforms are offering. Due to the large amounts of capital needed to build a platform with similar capabilities and market presence, new platforms that aim to compete with the frightful five for dominance may take considerable time to emerge. Even if they do, it is uncertain whether these platforms would aim to relinquish any power or carry any risk that their competitors have accrued and offloaded. Competition, at least in the sense that consumers are free to choose whether or not and with whom to share their personal data, is much more than "a click away".[499]

The European Commission has cited technological developments as an important trigger for drafting the GDPR in 2012.[500] But somewhat surprisingly, the evaluation of the risks associated with these developments is very similar to the risk evaluations in the French *Loi n° 78-17 du 6 janvier 1978* and the Council of Europe's *Convention 108* from 1981. The supporting documents of these three legal instruments, as well as the legal texts themselves, focus on the risks for individual rights and freedoms. These risks are obviously important, but no specific reasoning is provided as to how the proposed measures are expected to reduce the risks or mitigate the effects once the risks have materialised. At the same time, risks at the level of groups, communities or societies should not be ignored.

6.3 Understanding big data better: considerations for future legislation

Sections 1.3 and 1.4 presented a number of particular risks of datafication for individuals and societies. In the future, new risks and power shifts are likely to emerge

499 Per Strömbäck, 'Digital Myth: Competition Is Only One Click Away' (*Netopia*, 23 August 2016) <http://www.netopia.eu/competition-one-click-away/> accessed 21 March 2019.

500 European Commission, 'Proposal for a General Data Protection Regulation' (n 81) 1.

due to the progress of technology. Considering the conclusions of the previous chapters, I propose that future legislative efforts regarding the processing of personal data acknowledge the following propositions as a the foundation for regulating contracts and consent.

- Big data has emerged as a new *logic of accumulation* resulting in a new divide between haves and have-nots: a small number of actors have "all the data on everyone" and a much larger number of actors have gathered much less data, even about themselves. Controllers of personal data use their existing structural power over consumers to increase their institutional power. This carries with it the risk of unfair treatment.
- The resulting information asymmetries and power shifts in the market for goods and services have become so large that the market is losing its effectiveness as a decision-making mechanism. Consent statements or contracts no longer reliably prove that data subjects have formed and expressed their will to accept the effects of datafication in a particular context, or in related contexts.
- The size, complexity and tight coupling of big data systems makes it harder to identify the risks – meaning both the probability and the effects – of system accidents, whilst at the same time making it easier to maintain an illusion of inherent safety through *fantasy documents*. Such documents are of limited value in the context of a contract or consent and can serve to offload risks to consumers.
- The datafication of complex systems (such as the human body but also communities and societies) has become pervasive enough that *emergent properties* of human activity are encoded in the resulting data sets. These properties include insights in possibly intimate details about individuals and groups, such as opinions, social interactions and sensitive traits like health status and political or religious beliefs. An exclusive focus on individual rights and freedoms may therefore no longer be sufficient to address all the relevant effects associated with big data. Larger-scale effects need to be accounted for, especially in the case of consumer contracts.
- Information asymmetries and power differentials in the market limit the opportunities for consumers to *participate in meaningful discourse* on risk evaluation and mitigation. Limiting discourse-based risks management options could perpetuate information asymmetries and fantasy documents. Applying similar standards as provided by EU consumer protection law, especially the

directives on unfair terms and unfair commercial practices, should be considered when interpreting the GDPR.

When seen in this light, a number of features of the GDPR merit reconsideration if it were to be revised in the future:

- The GDPR casts a wide net, but the wide range of fields that it covers may not be suitable for omnibus legislation. Health care, education, employment, executive government, commerce and the marketplace of ideas may be better served with a more granular approach, as they vary in the risks and power relations at play. Especially in addressing risks and power shifts, these fields of legislation have already developed a vocabulary of concepts, models and opportunities for participation, often at the national level. A set of EU principles for data protection, with a uniform interpretation enforced through judicial decisions from the EU Court of Justice, could result in setbacks for achievements at the national level.
- The complexities of the processing of personal data, combined with a desire of national governments to commandeer big data accumulated by private parties, could promote *regulatory capture*. Regulators could have both an incentive and an opportunity to pass or maintain laws that benefit data controllers but that are at odds with the interests of data subjects, consumers and voters. This would be especially true for legislation drafted in the EU legislative process: Article 294, TFEU and especially the conciliation committee procedure could provide circumstances where regulatory capture can occur. Due to the lower costs of information and organisation, national parliaments offer better opportunities for preventing captured legislation than the European legislative apparatus does.
- Technologically neutral legislation encourages the enacting of *complex standards* as opposed to simpler rules. This makes enforcement more complex and possibly less efficient. It can also lead to decision-making on important issues being shifted towards the judicial system, where consumers have limited options for participation and where – as Galanter reminds us – the "haves" could come out ahead.

However, these points of reconsideration should not be regarded as policy proposals *per se*. Alternatives may exist, and the approach chosen for the GDPR is not without merit. Maintaining the GDPR's character of omnibus legislation, keeping discourse on

lawmaking and judicial decision-making at the EU level and applying a technologically neutral approach has a number of advantages:

- The GDPR's character of omnibus legislation reflects the reality that the processing of personal data is ubiquitous. Arguably, different definitions or levels of data protection for different applications could lead to confusion and uncertainty – not only for data subjects, but also for entities that wish to offer processing services to a wide range of corporate and government customers. *General* data protection legislation will make it easier for such entities to increase their market, become profitable and innovate. Such innovations could also turn out to initiate a cycle of continuous improvements as required in environmental law. Indeed, it is by no means certain that such improvements would happen more quickly or more reliably if each field of application would employ its own set of data protection principles.
- Likewise, harmonising data protection law at the EU level has advantages over developing and maintaining bodies of data protection legislation at the national level. Such diversity between Member States could interfere with the Four Freedoms in the internal market, especially since the processing of data is now inseparable from almost any commercial or administrative endeavour. Agreeing on a "consistent and high level of protection of natural persons" (recital 10) can be a suitable way for Member States to give substance to their obligations arising from Article 1, ECHR while respecting the sovereignty of other Member States by relying on their Independent Supervisory Authorities.
- Independent Supervisory Authorities can help relieve the data *have-nots* from the need to litigate contract cases against controllers.
- And even though standards are more complex than rules, they can offer long-term stability that allows undisputed forms processing to continue without interference, while judicial decisions eliminate controllers' excesses.

These advantages are not to be dismissed lightly. But our understanding of the risks of big data may not yet be sufficient to leave their mitigation to *one size fits all*-legislation. At this time, tailor-made legislation within specific areas (eg. health care, news media, education, employment) arguably stands a better chance of identifying and reducing risks and addressing power asymmetries resulting from big data. Consider the example of the Seveso III-directive: an abstract standard to reduce system accidents could probably not have been effective before our knowledge of

industrial risks (as codified in chemistry, materials science, long-term statistics et cetera) had sufficiently developed. Without a similar degree of understanding, the GDPR risks a degree of ineffectiveness that is not compatible with the threat that big data carries.

Furthermore, as long as insights regarding the risks and power imbalances associated with big data are still developing, societies need a forum where citizens and consumers have the best possible opportunities for participation in discourse-based risk management. The EU legislative and judicial decision-making processes present citizens with higher costs of information and organisation than their counterparts at the national level. It is also necessary to allow Member States to choose their own priorities when addressing the risks of big data – it is far from certain that, for example, Finland would arrive at the same solutions as Portugal when addressing a risk in the field of health care.

Therefore, effective participation opportunities for data *have-nots* is necessary to make sure that Member States can use their margin of appreciation in the best interest of their citizens. This would be a good match with Klinke and Renn's recommendation to apply discourse-based risk management. The political process at the level of the Member States is arguably the better forum for discourse-based risk management due to the lower costs of organisation and information for consumers. National governments also stand a better chance of responding quickly to new threats, for instance by delegation of legislative powers to the executive for well-defined classes of cases. EU legislation can then serve to consolidate a baseline among Member States in areas when the understanding of the risk has increased.

A possible side effect of such an approach could be that the free movement of personal data between Member States – one of the few unreservedly precautionary measures in the GDPR – would require some qualifications. But if the protection of natural persons' rights and freedoms is indeed an important goal of the GDPR, such qualifications might be unavoidable. As long as the risk associated with big data remains fluid, and as long as new ways to turn the processing of personal data into an advantage over data subjects or a tool to achieve commercial dominance, national legislation could offer better forms of risk management than the GDPR in its current form.

6.4 Broadening the knowledge base

"Unringing the bell" with regards to advent of big data is not possible: datafication, when seen as a byproduct of automation, is as irreversible as it is indispensable. Over time, consumers, data subjects and regulators can be expected to improve their understanding of the risks surrounding the use of large quantities of personal data. The gradual and ongoing improvements in, for example, the chemical industry and commercial aviation have shown that shifting technologies away from Perrow's *complex/tightly coupled quadrant* is certainly possible and must be pursued in proportion to the risk.

Considering the preceding chapters, changing the focus of EU data protection law should be seriously considered, at least in the area of contract, consent and special categories of data. Several models developed in the social and the exact sciences are suitable to inform legislative choices that make a cycle of permanent improvement of data protection law possible.

Nevertheless, the model proposed in section 6.3 above is still incomplete. After all, the expected impact of big data affects not just contracts for consumers, but also matters of health care, labour relations and the exercise of public authority. The issued raised by big data are probably not solvable through legislation alone. There is a number of areas where an interdisciplinary approach could lead to better solutions, based on an improved understanding of the effects of big data:

An observatory at the European level. The most important step in reducing risks is the promotion of knowledge and awareness of their existence. In a field developing as rapidly as big data, this awareness requires constant monitoring to minimise the risk that regulatory efforts get stuck in the *Collingridge dilemma*.[501] Different developments will be relevant for different stakeholders. Data controllers, journalists, the academic community, consumer groups and human rights organisations, as well as government institutions, the EU and the Council of Europe are all indispensable for forming a complete picture that can support thoroughness and coherence in policy development. Analogous to European observatories in the field of energy poverty, intellectual property, employment, cultural diversity and audiovisual media, a multidisciplinary "Big Data Observatory" could be charged with discerning and contextualising relevant developments in big data applications. An observatory could also serve as a place where research programmes into the effects of datafication, as

[501] See note 371 above.

well as remedies for their adverse effects, could be commissioned or evaluated, and provide participation opportunities for civil society.

Promotion of shared values as guiding principle. Addressing the risks and power shifts associated with big data requires that legislators, data subjects and controllers understand that the focus of the GDPR is too narrow to address all the risks associated with it. Tackling social, economical and political effects, like the for-profit fragmentation of the marketplace of ideas or rent-seeking behaviour by platform providers, requires that other areas of regulation be explored. Based on the preceding chapters, competition law and consumer protection law could be well-suited to address certain issues relating to power shifts. Indeed, laws in these fields were enacted with similar power shifts in mind. An exclusive focus on the rights and freedoms of natural persons will have inherent limits when addressing effects at larger scales. Similar to the findings of Van der Sloot, this research suggests that regulatory efforts specifically designed for the promotion of shared values will be essential in addressing big data's risks and power shifts.[502]

Reducing information asymmetries and power differentials. Based on the preceding chapters, it is also advisable to work towards reducing the information asymmetry resulting from datafication. Data scientists are working on several aspects of this problem. Two examples: Harkous et al. are aiming to reduce the opacity of privacy policies using machine learning algorithms, and Sandvig et al. are designing new methods for the visualisation of the effects of algorithms.[503] These efforts can help reduce the shifting of institutional power towards controllers by increasing the reciprocity of interactions and by making it easier for data subjects to make informed decisions. Addressing structural power shifts through research and discourse may require new forms of *sub-politics* or the extension of existing formal forums to enable the participation of parties lacking structural power.[504] An Observatory, existing

[502] Bart Van Der Sloot, 'Privacy as Virtue: Moving beyond the Individual in the Age of Big Data' (Universiteit van Amsterdam 2017) 196–197.

[503] Hamza Harkous and others, 'Polisis: Automated Analysis and Presentation of Privacy Policies Using Deep Learning' [2018] arXiv preprint arXiv:1802.02561; Christian Sandvig and others, 'Auditing Algorithms: Research Methods for Detecting Discrimination on Internet Platforms' (2014); Ke Yang and others, 'A Nutritional Label for Rankings', *Proceedings of the 2018 International Conference on Management of Data* (ACM 2018).

[504] See Nissenbaum's suggestion in Jaron Lanier and E Glen Weyl, 'A Blueprint for a Better Digital Society' (*Harvard Business Review*, 26 September 2018) <https://hbr.org/2018/09/a-blueprint-for-a-better-digital-society> accessed 20 March 2019.

standardisation forums, or specialised institutions could all offer such a forum at different levels of centralisation.

Permanent improvement through enforcement, preparedness and response. Finally, addressing the power shifts and risks associated with big data requires a long-term commitment from controllers, data subjects and their organisations, legislators, supervisory authorities and the courts. Big data can become a prime area where conflicts between fundamental rights will develop. The identification of new risks will be more useful when followed up by suitable prevention measures and response capabilities. Judicial decisions will have to clarify how big data affects the position of the right to private life between other fundamental rights like the freedom to conduct a business, freedom of expression or the freedom to enter into a contract.

System accidents involving big data lack the visibility and physical turmoil shared by many disasters in the physical world, and liability for these accidents may be difficult to determine when there is no physical damage. The GDPR's penalty provisions provide a partial solution, but their effectiveness will depend on the interpretation of complex standards. Therefore, requiring that controllers implement a plan-do-check-act cycle aimed at accident prevention is essential for continuous improvement. Finally, well-rounded risk management implies that systems for incident response are available. Similar to the requirements of the Seveso III directive, controllers and supervisory authorities should identify and test appropriate measures to mitigate the effects of foreseeable incidents and provide the response capabilities necessary to implement them when the next normal accident occurs.

6.5 Further research

This research was concerned with the shifts in risk and power associated with big data and focused on the ways that the GDPR addressed these shifts. In the preceding chapters, several possible courses of action were recommended based on the research into the sub-questions. Apart from these possible suggestions, two questions stand out to me as particularly relevant.

The first question is whether societies need new legal instruments to impose limits on the consumers' freedom of contract to let themselves be observed in the digital panopticon, and on governments' powers to commandeer the resulting data. Can consumers agree to any form of surveillance or should a minimal level of remaining privacy be required by law? And if such a minimum level of privacy is to be

maintained, who should decide on that minimum? On the one hand, European states cannot stand idly by while fundamental rights and freedoms are eroded by permanent surveillance through private platforms: positive obligations under the ECHR might arise within a decade. On the other hand, national governments – especially their police and national security forces – will need *some* powers of observation now that digital platforms have become another space where illegal activity can be planned and carried out. Public discourse on this matter should take place before regulatory capture can obscure the issue from view.

The second question is whether societies need new legal instruments to ensure and maintain dynamic competition in the marketplace of ideas. Private platforms can – and do – enforce their own notions of acceptable discourse through their terms and conditions and through visible enforcement, but they can also create filter bubbles. If a platform has a dominant position in the marketplace of ideas, it may not need to abuse this position to have harmful effects: the dominant position may be leveraged by other information providers using the platform. Users engaging in discourse on these platforms may lose their voice if the platform provider disagrees with them or if an intransigent minority of vigorous opponents flag their expression as "violating community standards". At the same time, the intended audience may be unaware that they never receive a message that they were meant to hear, if they are not aware that the information that reaches them is automatically filtered and skewed towards their calculated preferences in order to maximise the revenue of the platform. At the moment, the GDPR as well as competition law and consumer protection law appear insufficiently equipped to address these effects of market dominance.

The civil exchange of ideas between disagreeing parties is vitally important for an informed populace in a democracy. Competition law is of only limited use in the marketplace of ideas, since it is geared towards behaviour that interferes in the functioning of the price mechanism. If the marketplace of ideas is dominated by automated platforms with a profit motive, European democracies may need to develop new safeguards to maintain pluralism in civil discourse.

7 Bibliography

Akandji-Kombe J-F, *Positive Obligations under the European Convention on Human Rights* (Directorate General of Human Rights, Council of Europe 2007)

Alang N, 'Turns Out Algorithms Are Racist' [2017] *The New Republic* <https://newrepublic.com/article/144644/turns-algorithms-racist> accessed 19 March 2019

Albrecht JP, 'GDPR Inofficial Consolidated Version after LIBE Committee Vote Provided by the Rapporteur' (2013) <http://www.janalbrecht.eu/wp-content/uploads/2018/05/DPR-Regulation-inofficial-consolidated-LIBE.pdf> accessed 13 February 2019

Aleem Z, 'All the Secret Ways You're Being Tracked That You Don't Even Realize' (*Mic*, 23 March 2015) <https://mic.com/articles/113078/all-the-secret-ways-you-re-being-tracked-that-you-don-t-even-realize> accessed 13 February 2019

Andrews N, 'Fundamental Principles of Civil Procedure: Order Out of Chaos' in Xandra Ellen Kramer and others (eds), *Civil litigation in a globalising world* (TMC Asser Press; Springer 2012)

Angwin J and others, 'Machine Bias' (*ProPublica*, 23 May 2016) <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing> accessed 19 March 2019

Anrig B, Browne W and Gasson M, 'The Role of Algorithms in Profiling' in Mireille Hildebrandt and Serge Gutwirth (eds), *Profiling the European Citizen: Cross-Disciplinary Perspectives* (Softcover reprint of hardcover 1st ed 2008 edition, Springer 2008)

Aoki M, *Toward a Comparative Institutional Analysis* (1st edition, The MIT Press 2001)

Arnbak A and others, 'Security Collapse in the HTTPS Market' (2014) 57 Communications of the ACM 47 <http://dl.acm.org/citation.cfm?doid=2661061.2660574> accessed 19 March 2019

Arora N, 'Seeds Of Apple's New Growth In Mobile Payments, 800 Million ITune Accounts' (*Forbes*, 24 April 2014) <https://www.forbes.com/sites/nigamarora/2014/04/24/seeds-of-apples-new-growth-in-mobile-payments-800-million-itune-accounts/> accessed 19 March 2019

Article 29 Data Protection Working Party, 'Opinion 2/2010 on Online Behavioural Advertising' (2010) 00909/10/EN WP 171

——, 'Opinion 03/2013 on Purpose Limitation' (2013) 00569 /13/EN WP 203 <https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2013/wp203_en.pdf> accessed 20 March 2019

——, 'Opinion 05/2014 on Anonymisation Techniques' (2014) 0829/14/EN WP216 <https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf> accessed 13 February 2019

——, 'Guidelines on Data Protection Impact Assessment (DPIA) and Determining Whether Processing Is "Likely to Result in a High Risk" for the Purposes of Regulation 2016/679' (Article 29 Data Protection Working Party 2017) 17/EN WP248

Autoriteit Persoonsgegevens, 'Microsoft Windows 10 - De Verwerking van Persoonsgegevens via Telemetrie' (Autoriteit Persoonsgegevens 2017) Rapport definitieve bevindingen met correcties <https://autoriteitpersoonsgegevens.nl/sites/default/files/01_onderzoek_microsoft_windows_10_okt_2017.pdf> accessed 19 March 2019

Bakos Y, Marotta-Wurgler F and Trossen DR, 'Does Anyone Read the Fine Print? Consumer Attention to Standard Form Contracts' (2014) 43 Journal of Legal Studies <http://papers.ssrn.com/abstract=1443256> accessed 13 February 2019

Baldwin R, Cave M and Lodge M, *Understanding Regulation: Theory, Strategy, and Practice* (2nd edition, Oxford University Press 2013)

Barnett M and Duvall R, 'Power in International Politics' (2005) 59 International Organization 39 <http://www.jstor.org/stable/3877878> accessed 13 February 2019

Bartlett III RP and Plaut VC, 'Blind Consent? A Social Psychological Investigation of Non-Readership of Click-through Agreements.' (2012) 36 Law and human behavior 293 <http://psycnet.apa.org/journals/lhb/36/4/293/> accessed 19 March 2019

Baruh L and Popescu M, 'Big Data Analytics and the Limits of Privacy Self-Management' (2017) 19 New Media & Society 579 <http://dx.doi.org/10.1177/1461444815614001> accessed 19 March 2019

Bar-Yam Y, *Dynamics of Complex Systems* (Addison-Wesley 1997)

BBC News, 'Artificial Intelligence: Google's AlphaGo Beats Go Master Lee Se-Dol' (*BBC News*, 12 March 2016) <http://www.bbc.com/news/technology-35785875> accessed 19 March 2019

Beck U, *Risikogesellschaft. Auf Dem Weg in Eine Andere Moderne* (1st ed., Suhrkamp Verlag 1986)

——, *Risk Society: Towards a New Modernity* (Sage Publications 1992)

——, *World at Risk* (Polity Press 2009)

Beck U and Beck-Gernsheim E, *Individualization: Institutionalized Individualism and Its Social and Political Consequences* (1st edition, SAGE Publications Ltd 2002)

Bentham J, *The Panopticon Writings* (Miran Božovic ed, Verso 1995)

Berghel H, 'Equifax and the Latest Round of Identity Theft Roulette' (2017) 50 IEEE Computer 72

BEUC, 'Annual Report 2016' (BEUC 2017) <http://www.beuc.eu/publications/beuc-x-2017-045-annual-report-2016.pdf> accessed 19 March 2019

Bishop CM, *Pattern Recognition and Machine Learning* (Springer 2006)

Blessing M, 'Het Verzet Tegen de Volkstelling van 1971' (2005) 15 *Historisch Nieuwsblad* <https://www.historischnieuwsblad.nl/nl/artikel/6697/het-verzet-tegen-de-volkstelling-van-1971.html> accessed 19 March 2019

Bogaarts R and Dekker W, 'De Dagelijkse Volkstelling' *De Volkskrant* (21 February 1998) 50

Böhme R and Grossklags J, 'The Security Cost of Cheap User Interaction', *Proceedings of the 2011 workshop on New security paradigms workshop* (ACM 2011) <https://dl.acm.org/citation.cfm?id=2073284> accessed 13 February 2019

Boucher P, 'Safari Ou La Chasse Aux Français' *Le Monde* (Paris, 21 March 1974) <http://rewriting.net/2008/02/11/safari-ou-la-chasse-aux-francais/>

Bowman J, 'EU Data Protection Regulation: A Tipping Point Has Been Reached' (*The Privacy Advisor: The official Newsletter of the IAPP*, 7 November 2014) <https://iapp.org/news/a/eu-data-protection-regulation-a-tipping-point-has-been-reached/> accessed 13 February 2019

Breindl Y, 'Promoting Openness by "Patching" European Directives: Internet-Based Campaigning during the EU Telecoms Package Reform' (2011) 8 Journal of Information Technology & Politics 346 <http://dx.doi.org/10.1080/19331681.2011.595326> accessed 13 February 2019

Brickell J and Shmatikov V, 'The Cost of Privacy: Destruction of Data-Mining Utility in Anonymized Data Publishing', *Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining* (ACM 2008)

Briscoe B, Odlyzko A and Tilly B, 'Metcalfe's Law Is Wrong - Communications Networks Increase in Value as They Add Members-but by How Much?' (2006) 43 Spectrum, IEEE 34

Bughin J, 'Big Data, Big Bang?' (2016) 3 Journal of Big Data 2 <https://journalofbigdata.springeropen.com/articles/10.1186/s40537-015-0014-3> accessed 19 March 2019

Bultmann Dr jur. F, '30 Jahre Praxis Der AGB-Verbandsklage: Kurzfassung Des Gutachtens Im Auftrag Des Verbraucherzentrale Bundesverbandes' <http://www.vzbv.de/sites/default/files/mediapics/kurzfassung_gutachten_verbandsklage_2008.pdf> accessed 19 March 2019

Bundeskartellamt, 'Bundeskartellamt Eröffnet Verfahren Gegen Facebook Wegen Verdachts Auf Marktmachtmissbrauch Durch Datenschutzverstöße' (*Meldung*, 3 March 2016) <https://www.bundeskartellamt.de/SharedDocs/Meldung/DE/Pressemitteilungen/2016/02_03_2016_Facebook.html> accessed 19 March 2019

Burrell J, 'How the Machine "Thinks": Understanding Opacity in Machine Learning Algorithms' (2016) 3 Big Data & Society 2053951715622512 <http://dx.doi.org/10.1177/2053951715622512> accessed 19 March 2019

Calders T and Žliobaitė I, 'Why Unbiased Computational Processes Can Lead to Discriminative Decision Procedures' in Bart Custers and others (eds), *Discrimination and Privacy in the Information Society - Data Mining and Profiling in Large Databases* (Springer 2013)

Cate FH and Mayer-Schönberger V, 'Notice and Consent in a World of Big Data' (2013) 3 International Data Privacy Law 67 <http://idpl.oxfordjournals.org/content/3/2/67> accessed 19 March 2019

Center for Disease Control, 'Cigarette Brand Preference Among Middle and High School Students Who Are Established Smokers - United States, 2004 and 2006' (2009) 58 Morbidity and Mortality Weekly Report 112 <https://www.cdc.gov/mmwr/preview/mmwrhtml/mm5805a3.htm> accessed 19 March 2019

Centola D and Baronchelli A, 'The Spontaneous Emergence of Conventions: An Experimental Study of Cultural Evolution' (2015) 112 Proceedings of the National Academy of Sciences 1989 <http://www.pnas.org/content/112/7/1989> accessed 19 March 2019

Chun R, 'Big In... China: Machines That Scan Your Face' [2018] *The Atlantic* <https://www.theatlantic.com/magazine/archive/2018/04/big-in-china-machines-that-scan-your-face/554075/> accessed 19 March 2019

Citron DK, 'Technological Due Process' (2007) 85 Wash. UL Rev. 1249

Citron DK and Pasquale FA, 'The Scored Society: Due Process for Automated Predictions' (2014) 89 Wash. UL Rev. 1 <http://papers.ssrn.com/abstract=2376209> accessed 19 March 2019

Clarke L and Perrow C, 'Prosaic Organizational Failure' (1996) 39 American Behavioral Scientist 1040 <http://abs.sagepub.com/content/39/8/1040.short> accessed 13 February 2019

Coase RH, 'The Problem of Social Cost' (1960) 3 Journal of Law and Economics 1

Cofone IN, 'Privacy Tradeoffs in Information Technology Law' (Erasmus University 2015)

Cohen JE, 'Turning Privacy Inside Out' (2019) 20 Theoretical Inquiries in Law (forthcoming) <https://papers.ssrn.com/abstract=3162178> accessed 19 March 2019

Cohn-Gordon K and others, 'A Formal Security Analysis of the Signal Messaging Protocol', *Security and Privacy (EuroS&P), 2017 IEEE European Symposium on* (IEEE 2017)

Collingridge D, *The Social Control of Technology* (Frances Pinter 1980)

Commission des Clauses Abusives, 'Recommandation n° 2014-02 relative aux contrats proposés par les fournisseurs de services de réseaux sociaux' <http://www.clauses-abusives.fr/recommandation/contrats-de-fourniture-de-services-de-reseaux-sociaux-nouveau/> accessed 18 March 2019

Committee of Experts on Internet Intermediaries (MSI-NET), 'Study on the Human Rights Dimensions of Algorithms (Second Draft)' (Council of Europe 2017) MSI-NET(2016)06 rev <https://rm.coe.int/16806fe644> accessed 19 March 2019

Committee on Legal Affairs and Human Rights and Omtzigt P, 'Mass Surveillance' (Parliamentary Assembly, Council of Europe 2015) Doc. 13734 <http://assembly.coe.int/nw/xml/XRef/X2H-Xref-ViewPDF.asp?FileID=21583&lang=en> accessed 13 February 2019

Computer Says No, 'Computer Says No', *Wikipedia* (2017) <https://en.wikipedia.org/wiki/Computer_says_no> accessed 19 March 2019

Constantiou ID and Kallinikos J, 'New Games, New Rules: Big Data and the Changing Context of Strategy' (2015) 30 Journal of Information Technology 44

Consumentenbond, 'Review: ANWB Veilig Rijden' (*Consumentenbond*, 21 July 2016) <https://www.consumentenbond.nl/autoverzekering/anwb-veilig-rijden> accessed 19 March 2019

——, 'Jaarverslag 2016' (Consumentenbond 2017) Jaarverslag <https://www.consumentenbond.nl/binaries/content/assets/cbhippowebsite/over-ons/wie-zijn-we/consumentenbond-jaarverslag-2016.pdf> accessed 13 February 2019

Council of Europe, 'Explanatory Report to the Convention for the Protection of Individuals with Regard to Automatic Processing of Personal Data' (Council of Europe 1981) Explanatory Report 108

——, 'Convention 108 and Protocol: Background' (*Data Protection*) <https://www.coe.int/en/web/data-protection/convention108/background> accessed 19 March 2019

Council of the OECD, 'Recommendation of the Council Concerning Guidelines Governing the Protection of Privacy and Transborder Flows of Personal Data - C(80)58/FINAL' <http://www.oecd.org/internet/ieconomy/oecdguidelinesontheprotectionofprivacyandtransborderflowsofpersonaldata.htm> accessed 19 March 2019

——, 'OECD Guidelines Governing the Protection of Privacy and Transborder Flows of Personal Data - C(80)58/FINAL, as Amended on 11 July 2013 by C(2013)79' (Organization for Economic Cooperation and Development 2013) <http://www.oecd.org/sti/ieconomy/2013-oecd-privacy-guidelines.pdf> accessed 19 March 2019

Crawford K and Schultz J, 'Big Data and Due Process: Toward a Framework to Redress Predictive Privacy Harms' (2014) 55 BCL Rev. 93 <http://heinonlinebackup.com/hol-cgi-bin/get_pdf.cgi?handle=hein.journals/bclr55&section=5> accessed 19 March 2019

Cullen J and others, 'Seven-Year Patterns in US Cigar Use Epidemiology Among Young Adults Aged 18–25 Years: A Focus on Race/Ethnicity and Brand' (2011) 101 American Journal of Public Health 1955 <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3222378/> accessed 19 March 2019

Custers B and others (eds), *Discrimination and Privacy in the Information Society - Data Mining and Profiling in Large Databases* (Springer 2013)

Datta A, 'Privacy through Accountability: A Computer Science Perspective' in Raja Natarajan (ed), *Distributed Computing and Internet Technology* (Springer International Publishing 2014)

De Hert P and Gutwirth S, 'Data Protection in the Case Law of Strasbourg and Luxemburg: Constitutionalisation in Action' in Serge Gutwirth and others (eds), *Reinventing Data Protection?* (Springer Netherlands 2009)

de Tocqueville A, *Democracy in America* (JP Mayer and George Lawrence eds, Harper Perennial Modern Classics 2006)

Dembosky A, 'Facebook Spending on Lobbying Soars' *Financial Times* (24 January 2013)

Diebold FX, 'A Personal Perspective on the Origin(s) and Development of "Big Data": The Phenomenon, the Term, and the Discipline' (2012)

<http://www.ssc.upenn.edu/~fdiebold/papers/paper112/Diebold_Big_Data.pdf> accessed 20 March 2019

Digital Rights Ireland, 'We Need Your Help to Keep Working for European Digital Rights in 2014' (*Digital Rights Ireland*, 1 January 2014) <http://www.digitalrights.ie/support-us-in-2014/> accessed 13 February 2019

Digital transformation, 'Digital Transformation', *Wikipedia* (2018) <https://en.wikipedia.org/wiki/Digital_transformation> accessed 19 March 2019

DLA Piper, 'Part 4: The Future of Online Privacy and Data Protection' (European Union 2009) SMART 2007/0037 <http://ec.europa.eu/newsroom/dae/document.cfm?doc_id=833> accessed 13 February 2019

Domingos P, 'A Few Useful Things to Know about Machine Learning' (2012) 55 Communications of the ACM 78 <http://dl.acm.org/citation.cfm?id=2347755> accessed 19 March 2019

Duhigg C, *The Power of Habit: Why We Do What We Do in Life and Business* (Random House Trade Paperback Edition, Random House Trade Paperbacks 2014)

Dür A and De Bièvre D, 'The Question of Interest Group Influence' (2007) 27 Journal of Public Policy 1 <https://www.cambridge.org/core/product/identifier/S0143814X07000591/type/journal_article> accessed 17 May 2019

Dwork C and Mulligan DK, 'It's Not Privacy, and It's Not Fair' (2013) 66 Stanford Law Review Online 35 <http://www.stanfordlawreview.org/online/privacy-and-big-data/its-not-privacy-and-its-not-fair> accessed 19 March 2019

EDRi - European Digital Rights, 'EU: The Global Standard Setter for Privacy and Data Protection' (EDRi - European Digital Rights 2013) Issue 2 <http://edri.org/files/eudatap-02.pdf> accessed 13 February 2019

——, 'Annual Report, January 2013 - December 2013' (EDRi - European Digital Rights 2014) <https://edri.org/wp-content/uploads/2014/04/EDRi_Annual_Report_2013.pdf> accessed 13 February 2019

——, 'Annual Report 2016: January 2016 – December 2016' (EDRi - European Digital Rights 2017) <https://edri.org/files/edri_annual_report_2016.pdf> accessed 13 February 2019

Engelfriet A and others, *De algemene verordening gegevensbescherming: artikelsgewijs commentaar* (Ius mentis 2017)

European Agency for Fundamental Rights, 'Data Protection in the European Union: The Role of National Data Protection Authorities. Strengthening the Fundamental Rights Architecture in the EU II' (European Union Agency for Fundamental Rights 2010) <doi:10.2811/47216>

——, 'Fundamental Rights Report 2016' (European Union Agency for Fundamental Rights 2016) <http://fra.europa.eu/en/publication/2016/fundamental-rights-report-2016> accessed 19 March 2019

European Commission, 'Proposal for a Council Directive on the Protection of Individuals with Regard to the Processing of Personal Data' (European Commission 1990) COM(90) 314 final

——, 'Amended Proposal for a Council Directive on the Protection of Individuals with Regard to the Processing of Personal Data and on the Free Movement of Such Data' (European Commission 1992) COM (92) 422 final

——, 'Guidelines on Vertical Restraints' (European Commission 2010) OJ 2010/C 130/01

——, 'A Common European Sales Law to Facilitate Bross-Border Transactions in the Single Market' (European Commission 2011) Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions COM/2011/0636 final <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52011DC0636> accessed 13 February 2019

——, 'Impact Assessment, Accompanying the Document "Regulation of the European Parliament and of the Council on the Protection of Individuals with Regard to the Processing of Personal Data and on the Free Movement of Such Data (General Data Protection Regulation)" and "Directive of the European Parliament and of the Council on the Protection of Individuals with Regard to the Processing of Personal Data by Competent Authorities for the Purposes of Prevention, Investigation, Detection or Prosecution of Criminal Offences or the Execution of Criminal Penalties, and the Free Movement of Such Data"' (European Commission 2012) Accompanying document SEC(2012) 72 final

——, 'Proposal for a Regulation of the European Parliament and of the Council on the Protection of Individuals with Regard to the Processing of Personal Data and on the Free Movement of Such Data (General Data Protection Regulation) COM(2012) 11 (FINAL)' (European Commission 2012) <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52012PC0011> accessed 19 March 2019

——, 'Safeguarding Privacy in a Connected World. A European Data Protection Framework for the 21st Century' (European Commission 2012) Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions COM(2012) 9 final <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2012:0009:FIN:EN:PDF> accessed 20 March 2019

——, 'A European Consumer Agenda - Boosting Confidence And Growth' (European Union 2012) Communication from the Commission to the European Parliament, the Council, the Economic and Social Committee and The Committee of the Regions COM(2012) 225 final <https://ec.europa.eu/commission/sites/beta-political/files/consumer_agenda_2012_en.pdf> accessed 13 February 2019

——, 'Antitrust: Commission Fines Microsoft for Non-Compliance with Browser Choice Commitments' (6 March 2013) <http://europa.eu/rapid/press-release_IP-13-196_en.htm> accessed 19 March 2019

——, 'Towards a European Horizontal Framework for Collective Redress COM(2013) 401 Final' (European Commission 2013) <https://eur-lex.europa.eu/procedure/EN/202773> accessed 13 February 2019

——, 'Proposal for a Directive of the European Parliament and of the Council on Certain Aspects Concerning Contracts for the Supply of Digital Content' (European Commission 2015) COM(2015) 634 final; 2015/0287 (COD) <https://ec.europa.eu/transparency/regdoc/rep/1/2015/EN/1-2015-634-EN-F1-1.PDF>

——, 'Antitrust: Commission Fines Google €2.42 Billion for Abusing Dominance as Search Engine by Giving Illegal Advantage to Own Comparison Shopping Service' (27 June 2017) <http://europa.eu/rapid/press-release_IP-17-1784_en.htm> accessed 19 March 2019

——, 'State Aid: Commission Refers Ireland to Court for Failure to Recover Illegal Tax Benefits from Apple Worth up to €13 Billion' (4 October 2017) <http://europa.eu/rapid/press-release_IP-17-3702_en.htm> accessed 19 March 2019

——, 'Adequacy of the Protection of Personal Data in Non-EU Countries' (*European Commission - European Commission*) <https://ec.europa.eu/info/law/law-topic/data-protection/data-transfers-outside-eu/adequacy-protection-personal-data-non-eu-countries_en> accessed 19 March 2019

European Data Protection Supervisor, 'Privacy and Competitiveness in the Age of Big Data: The Interplay between Data Protection, Competition Law and Consumer Protection in the Digital Economy' (2014) Preliminary Opinion of the European Data Protection Supervisor <https://edps.europa.eu/sites/edp/files/publication/14-03-26_competitition_law_big_data_en.pdf> accessed 13 February 2019

European Digital Rights, 'Proceed with Caution: Flexibilities in the General Data Protection Regulation' (EDRi - European Digital Rights 2016) <https://edri.org/files/GDPR_analysis/EDRi_analysis_gdpr_flexibilities.pdf> accessed 19 March 2019

European Union Agency for Fundamental Rights and European Court of Human Rights, *Handbook on European non-discrimination law* (Publications Office of the European Union 2011) <http://fra.europa.eu/sites/default/files/fra_uploads/1510-fra-case-law-handbook_en.pdf> accessed 19 March 2019

European Union Agency for Fundamental Rights, European Court of Human Rights and Council of Europe, *Handbook on European Data Protection Law* (Publications Office of the European Union 2014)

Executive Office of the President, 'Big Data: A Report on Algorithmic Systems, Opportunity, and Civil Rights' (Executive Office of the President 2016) <https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/2016_0504_data_discrimination.pdf> accessed 13 February 2019

Experian, 'Determine the Best Offer: Make Credit Decisions That Yield the Best Results' (2018) <http://www.experian.com/business-services/customer-leads.html> accessed 19 March 2019

Federal Trade Commission, 'Data Brokers: A Call for Transparency and Accountability' (2014) <https://www.ftc.gov/reports/data-brokers-call-transparency-accountability-report-federal-trade-commission-may-2014> accessed 13 February 2019

Feng QY and others, 'An Exploratory Statistical Approach to Depression Pattern Identification' (2013) 392 Physica A: Statistical Mechanics and its Applications 889 <http://linkinghub.elsevier.com/retrieve/pii/S037843711200921l> accessed 19 March 2019

Ferenstein G, 'Google's Cerf Says "Privacy May Be An Anomaly". Historically, He's Right.' (*TechCrunch*, 20 November 2013) <http://social.techcrunch.com/2013/11/20/googles-cerf-says-privacy-may-be-an-anomaly-historically-hes-right/> accessed 19 March 2019

Fisher AW and McKenney JL, 'The Development of the ERMA Banking System: Lessons from History' (1993) 15 IEEE Annals of the History of Computing 44

Flach P, *Machine Learning: The Art and Science of Algorithms That Make Sense of Data* (1 edition, Cambridge University Press 2012)

Floridi L, 'Philosophical Conceptions of Information' in Giovanni Sommaruga (ed), *Formal theories of information: from Shannon to semantic information theory and general concepts of information* (Springer 2009)

Fontanella-Khan J, 'Brussels: Astroturfing Takes Root' *Financial Times* (26 June 2013)

Foucault M, *Discipline and Punish: The Birth of the Prison* (Second Vintage Books edition, Vintage 1991)

Foyer M, 'Projet de Loi (No 2516) et Propositions de Loi (Nos 1004 et 3092)' (Assemblé Nationale 1977) 3125 <http://www.senat.fr/rap/l77-3125/l77-31251.pdf> accessed 19 March 2019

Frantziou E, 'The Horizontal Effect of the Charter of Fundamental Rights of the EU: Rediscovering the Reasons for Horizontality' (2015) 21 European Law Journal 657 <http://doi.wiley.com/10.1111/eulj.12137> accessed 19 March 2019

Galanter M, 'Why the Haves Come out Ahead: Speculations on the Limits of Legal Change' (1974) 9 Law & Society Review 95 <https://www.jstor.org/stable/3053023> accessed 13 February 2019

Ganley P and Allgrove B, 'Net Neutrality: A User's Guide' (2006) 22 Computer Law & Security Review 454 <http://www.sciencedirect.com/science/article/pii/S0267364906000902> accessed 19 March 2019

Gellert R and others, 'A Comparative Analysis of Anti-Discrimination and Data Protection Legislations' in Bart Custers and others (eds), *Discrimination and Privacy in the Information Society - Data Mining and Profiling in Large Databases* (Springer 2013)

Genin-Meric R, 'Droit de la preuve: l'Example Français' in José Lebre de. Freitas (ed), *The law of evidence in the European Union = Das Beweisrecht in der Europaischen Union = Le droit de la preuve dans l'Union Europeenne* (Kluwer Law International 2004)

Giesen I, 'Sommige Procespartijen Zijn "More Equal than Others". De Macht van de Tabaksindustrie En de Nederlandse Rechtspleging' in Nienke Doornbos, Nick Huls and Wibo

van Rossum (eds), *Rechtspraak van Buiten. Negenendertig door de rechtssociologie geïnspireerde annotaties (Liber Amicorum prof. dr. J.F. Bruinsma)* (Kluwer 2010)

Gomes O, 'The Economy as a Complex Object' in Bernardo Alves Furtado, Patricia AM Sakowski and Marina H Tóvolli (eds), *Modeling Complex Systems for Public Policies* (IPEA 2015)

Gomez F, 'EC Consumer Protection Law and EC Competition Law: How Related Are They? A Law and Economics Perspective' in Hugh Collins (ed), *The Forthcoming EC Directive on Unfair Commercial Practices - Contract, Consumer and Competition law implications* (Kluwer Law International 2004)

Gouyet J-F and Mandelbrot B, *Physics and Fractal Structures* (1 edition, Springer 1996)

Graham-Harrison E and Cadwalladr C, 'Revealed: 50 Million Facebook Profiles Harvested for Cambridge Analytica in Major Data Breach' (*the Guardian*, 17 March 2018) <http://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election> accessed 20 March 2019

Gram-Hanssen K and others, 'Do Homeowners Use Energy Labels? A Comparison between Denmark and Belgium' (2007) 35 Energy Policy 2879 <http://linkinghub.elsevier.com/retrieve/pii/S0301421506004071> accessed 13 February 2019

Grassegger H and Krogerus M, 'Ich Habe Nur Gezeigt, Dass Es Die Bombe Gibt' [2016] *Das Magazin* <https://www.dasmagazin.ch/2016/12/03/ich-habe-nur-gezeigt-dass-es-die-bombe-gibt/> accessed 20 March 2019

Great Britain Commission for Racial Equality, *Medical School Admissions: Report of a Formal Investigation Into St. George's Hospital Medical School* (Commission for Racial Equality 1988)

Greenberg A, 'Whatsapp Just Switched on End-to-End Encryption for Hundreds of Millions of Users' <https://www.wired.com/2014/11/whatsapp-encrypted-messaging/> accessed 20 March 2019

Greenfield A, 'China's Dystopian Tech Could Be Contagious' [2018] *The Atlantic* <https://www.theatlantic.com/technology/archive/2018/02/chinas-dangerous-dream-of-urban-control/553097/> accessed 20 March 2019

Grimm V and others, 'Pattern-Oriented Modeling of Agent-Based Complex Systems: Lessons from Ecology' (2005) 310 Science 987 <http://science.sciencemag.org/content/310/5750/987> accessed 20 March 2019

Guess AR, 'Artificial Intelligence Had a Breakthrough Year in 2015' (*DATAVERSITY*, 9 December 2015) <http://www.dataversity.net/artificial-intelligence-had-a-breakthrough-year-in-2015/> accessed 20 March 2019

Gueye A and others, 'Defensive Resource Allocations with Security Chokepoints in IPv6 Networks' in Pierangela Samarati (ed), *Data and Applications Security and Privacy XXIX* (Springer, Cham 2015) <https://link.springer.com/chapter/10.1007/978-3-319-20810-7_19> accessed 20 March 2019

Guibault LMCR and others, *Digital Consumers and the Law. Towards a Cohesive European Framework* (Kluwer Law International 2012)

Guiness M, 'France Maintains Long Tradition of Data Protection' (*DW.COM*, 26 January 2011) <http://www.dw.com/en/france-maintains-long-tradition-of-data-protection/a-14797711> accessed 20 March 2019

Gutwirth S and de Hert P, 'Een Theoretische Onderbouw Voor Een Legitiem Strafproces. Reflecties over Procesculturen, de Doelstellingen van de Straf, de Plaats van Het Strafrecht En de Rol van Slachtoffers" (2001) 31 Delikt & delinkwent 1048 <http://www.vub.ac.be/LSTS/pub/Gutwirth/006.pdf> accessed 13 February 2019

Hammond G, 'Your Phone Dial Computes Your Bill' (1949) 154 *Popular Science* 135

Harkous H and others, 'Polisis: Automated Analysis and Presentation of Privacy Policies Using Deep Learning' [2018] arXiv preprint arXiv:1802.02561

Havard C, '"On the Take": The Black Box of Credit Scoring and Mortgage Discrimination' (2011) 20 Boston University Public Interest Law Journal 241 <http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1710063> accessed 13 February 2019

Hecht J, 'Is Keck's Law Coming to an End?' (2016) 2016 IEEE Spectrum 11 <https://spectrum.ieee.org/semiconductors/optoelectronics/is-kecks-law-coming-to-an-end>

Hempel L and Lammerant H, 'Impact Assessments as Negotiated Knowledge' in Serge Gutwirth, Ronald Leenes and Paul de Hert (eds), *Reforming European Data Protection Law* (2015 edition, Springer 2014)

Hill R, *The New International Telecommunication Regulations and the Internet* (Springer Berlin Heidelberg 2014)

Hirsch DD, 'Protecting the Inner Environment: What Privacy Regulation Can Learn from Environmental Law' (2006) 41 Ga. L. Rev. 1

Holland JH, *Complexity: A Very Short Introduction* (Oxford University Press 2014)

Holvast J, 'Op weg naar een risicoloze maatschappij? De vrijheid van de mens in de informatie-samenleving' (Leiden University 1986)

Hoogstraaten H and others, 'Black Tulip: Report of the Investigation into the DigiNotar Certificate Authority Breach' (Fox-IT BV 2012) <https://roselabs.nl/files/audit_reports/Fox-IT_-_DigiNotar.pdf> accessed 13 February 2019

Hopkins A, 'Discussion: The Limits of Normal Accident Theory' (1999) 32 Safety Science 93 <https://linkinghub.elsevier.com/retrieve/pii/S0925753599000156> accessed 16 May 2019

Horowitz M, 'Visualizing Big Data: Bar Charts for Words' (2008) 16 Wired Magazine <https://www.wired.com/2008/06/pb-visualizing/>

Houlder V, Barker A and Beesley A, 'Apple's EU Tax Dispute Explained' *Financial Times* (London, 30 August 2016) <https://www.ft.com/content/3e0172a0-6e1b-11e6-9ac1-1055824ca907> accessed 20 March 2019

International Telecommunication Union, *International Telecommunication Regulations. Final Acts of the World Administrative Telegraphe and Telephone Conference, Melbourne, 1988, (WATTC-88).* (ITU 1989)

Issacharoff S, 'Group Litigation of Consumer Claims: Lessons from the U.S. Experience' (1999) 34 Texas International Law Journal 135

ITU-T, 'The International Public Telecommunication Numbering Plan - Recommendation ITU-T E.164' (Telecommunication Standardization Sector of ITU 2010) Recommendation E 36438

——, 'Overview of the Internet of Things' (International Telecommunication Union 2012) Recommendation ITU-T Y.4000/Y.2060

Jerome JW, 'Buying and Selling Privacy: Big Data's Different Burdens and Benefits' (2013) 66 Stanford Law Review Online 47 <https://www.stanfordlawreview.org/online/privacy-and-big-data-buying-and-selling-privacy/> accessed 13 February 2019

Jones D, 'Solidarity and Public Participation: The Role of the Aarhus Convention in Containing Environmentally Induced Social Conflict' (2008) 20 Global Change, Peace & Security 151 <http://dx.doi.org/10.1080/14781150802079706> accessed 13 February 2019

Jue N, 'ING En Het Gebruik van Klantgegevens. Open Brief van ING Aan Haar Klanten' (maart 2014) <https://www.ing.nl/nieuws/nieuws_en_persberichten/2014/03/ing_en_het_gebruik_van_klant_brief.html> accessed 13 February 2019

Kannan PK and Kopalle PK, 'Dynamic Pricing on the Internet: Importance and Implications for Consumer Behavior' (2001) 5 International Journal of Electronic Commerce 63 <https://doi.org/10.1080/10864415.2001.11044211> accessed 20 March 2019

Kellert SH, *In the Wake of Chaos: Unpredictable Order in Dynamical Systems* (University of Chicago press 1994)

Kerr I and Earle J, 'Prediction, Preemption, Presumption: How Big Data Threatens Big Picture Privacy' (2013) 66 Stanford Law Review Online 65 <http://www.stanfordlawreview.org/online/privacy-and-big-data/prediction-preemption-presumption> accessed 13 February 2019

Klaver M-J, 'Bits of Freedom Staakt Strijd Op Web; Oprichter: Digitale Burgerrechtenbeweging Harder Nodig Dan Ooit' *NRC Handelsblad* (5 August 2006) 26 <https://www.nrc.nl/nieuws/2006/08/05/bits-of-freedom-staakt-strijd-op-web-11172796-a826780> accessed 13 February 2019

Klinke A and Renn O, 'A New Approach to Risk Evaluation and Management: Risk-Based, Precaution-Based, and Discourse-Based Strategies' (2002) 22 Risk Analysis 1071 <http://onlinelibrary.wiley.com/doi/10.1111/1539-6924.00274/full> accessed 13 February 2019

Knight W, 'Google's AI Chief Says Forget Elon Musk's Killer Robots, and Worry about Bias in AI Systems Instead' (*MIT Technology Review*, 3 October 2017) <https://www.technologyreview.com/s/608986/forget-killer-robotsbias-is-the-real-ai-danger/> accessed 20 March 2019

Komesar NK, *Imperfect Alternatives: Choosing Institutions in Law, Economics, and Public Policy* (University Of Chicago Press 1997)

——, *Law's Limits: The Rule of Law and the Supply and Demand of Rights* (Cambridge University Press 2001)

——, 'Governance, Economics and the Dynamics of Participation' in Neil Komesar and others (eds), *Understanding global governance: institutional choice and the dynamics of participation* (European University Institute 2014)

Koops B-J, 'The Trouble with European Data Protection Law' (2014) 4 International Data Privacy Law 250 <https://academic.oup.com/idpl/article-abstract/4/4/250/2569063/The-trouble-with-European-data-protection-law> accessed 20 March 2019

Kosinski M, Stillwell D and Graepel T, 'Private Traits and Attributes Are Predictable from Digital Records of Human Behavior' (2013) 110 Proceedings of the National Academy of Sciences 5802 <http://www.pnas.org/cgi/doi/10.1073/pnas.1218772110> accessed 20 March 2019

Kristensen V, Aven T and Ford D, 'A New Perspective on Renn and Klinke's Approach to Risk Evaluation and Management' (2006) 91 Reliability Engineering & System Safety 421 <https://linkinghub.elsevier.com/retrieve/pii/S0951832005000785> accessed 16 May 2019

Krugman P and Wells R, *Economics* (Third Edition, Worth Publishers 2012)

Kumar V and Reinartz W, *Customer Relationship Management: Concept, Strategy, and Tools* (Springer Science & Business Media 2012)

Ladyman J, Lambert J and Wiesner K, 'What Is a Complex System?' (2013) 3 European Journal for Philosophy of Science 33 <http://link.springer.com/article/10.1007/s13194-012-0056-8> accessed 20 March 2019

Lanier J and Weyl EG, 'A Blueprint for a Better Digital Society' (*Harvard Business Review*, 26 September 2018) <https://hbr.org/2018/09/a-blueprint-for-a-better-digital-society> accessed 20 March 2019

Leveson N and others, 'Moving beyond Normal Accidents and High Reliability Organizations: A Systems Approach to Safety in Complex Systems' (2009) 30 Organization Studies 227 <https://journals.sagepub.com/doi/abs/10.1177/0170840608101478> accessed 20 March 2019

Levin SA, 'Ecosystems and the Biosphere as Complex Adaptive Systems' (1998) 1 Ecosystems 431

Levine ME and Forrence JL, 'Regulatory Capture, Public Interest, and the Public Agenda: Toward a Synthesis' (1990) 6 Journal of Law, Economics, and Organization 167 <http://jleo.oxfordjournals.org/cgi/doi/10.1093/jleo/6.special_issue.167> accessed 13 February 2019

Li T and Li N, 'On the Tradeoff between Privacy and Utility in Data Publishing', *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining* (ACM 2009)

Lichter A, Loeffler M and Siegloch S, 'The Economic Costs of Mass Surveillance: Insights from Stasi Spying in East Germany' (IZA Discussion Papers 2015)

List of public corporations by market capitalization, 'List of Public Corporations by Market Capitalization', *Wikipedia* (2018)

<https://en.wikipedia.org/wiki/List_of_public_corporations_by_market_capitalization> accessed 20 March 2019

Lohr S, 'How Big Data Became So Big - Unboxed' *The New York Times* (11 August 2012) <https://www.nytimes.com/2012/08/12/business/how-big-data-became-so-big-unboxed.html> accessed 20 March 2019

Lynch C, 'How Do Your Data Grow?' (2008) 455 Nature 28 <http://dx.doi.org/10.1038/455028a>

Madala HR and Ivakhnenko AG, *Inductive Learning Algorithms for Complex Systems Modeling* (CRC Press 1994)

Madan A and others, 'Sensing the "Health State" of a Community' (2012) 11 IEEE Pervasive Computing 36 <http://ieeexplore.ieee.org/document/6072198/> accessed 20 March 2019

Maldoff G, 'The Risk-Based Approach in the GDPR: Interpretation and Implications' <https://iapp.org/media/pdf/resource_center/GDPR_Study_Maldoff.pdf> accessed 13 February 2019

Malik O, 'Zimmermann's Law: PGP Inventor and Silent Circle Co-Founder Phil Zimmermann on the Surveillance Society' (*Gigaom*, 11 August 2013) <https://gigaom.com/2013/08/11/zimmermanns-law-pgp-inventor-and-silent-circle-co-founder-phil-zimmermann-on-the-surveillance-society/> accessed 13 February 2019

Manjoo F, 'Tech's Frightful Five: They've Got Us' *The New York Times* (10 May 2017) <https://www.nytimes.com/2017/05/10/technology/techs-frightful-five-theyve-got-us.html> accessed 20 March 2019

Mascetti S and others, 'Anonymity: A Comparison Between the Legal and Computer Science Perspectives' in Serge Gutwirth and others (eds), *European Data Protection: In good health?* (Springer 2013)

Mayer-Schönberger V and Cukier K, *Big Data: A Revolution That Will Transform How We Live, Work, and Think* (Houghton Mifflin Harcourt 2013)

——, 'The Rise of Big Data: How It's Changing the Way We Think about the World' (2013) 92 Foreign Affairs 28

McAfee RP and Te Velde VL, 'Dynamic Pricing in the Airline Industry' in Terrence Hendershott (ed), *Economics and Information Systems* (1st edition, Elsevier 2006)

McDonald AM and Cranor LF, 'The Cost of Reading Privacy Policies' (2008) 4 I/S: A Journal of Law and Policy for the Information Society 540

McGregor C, 'Big Data in Neonatal Intensive Care' (2013) 46 IEEE Computer 54

Mele C, 'No Anonymity for Plaintiffs Suing Ashley Madison Over Hack, Judge Rules' *The New York Times* (21 April 2016) <http://www.nytimes.com/2016/04/22/technology/no-anonymity-ashley-madison-hack-case.html> accessed 20 March 2019

Microsoft, 'Privacy Statement' <https://privacy.microsoft.com/en-us/privacystatement/> accessed 13 February 2019

Mittelstadt BD and others, 'The Ethics of Algorithms: Mapping the Debate' (2016) 3 Big Data & Society 2053951716679679 <http://journals.sagepub.com/doi/abs/10.1177/2053951716679679> accessed 20 March 2019

Moerel L, *Big Data Protection: How to Make the Draft EU Regulation on Data Protection Future Proof. Oratie 14 Februari 2014* (Tilburg University 2014)

Moore GE, 'Cramming More Components onto Integrated Circuits' (1965) 38 Electronics Magazine 114 ff <https://ieeexplore.ieee.org/abstract/document/4785860> accessed 20 March 2019

Morgan, Glenn and others, 'Introduction' in Glenn Morgan and others (eds), *The Oxford Handbook of Comparative Institutional Analysis* (Oxford University Press 2010)

Mythen G, *Ulrich Beck: A Critical Introduction to the Risk Society* (Pluto Press 2004)

Nafus D and Sherman J, 'This One Does Not Go Up To 11: The Quantified Self Movement as an Alternative Big Data Practice' (2014) 8 International Journal of Communication 1784 <http://ijoc.org/index.php/ijoc/article/view/2170> accessed 13 February 2019

Nilsson NJ, *The Quest for Artificial Intelligence* (1 edition, Cambridge University Press 2009)

Nissenbaum H, *Privacy in Context: Technology, Policy, and the Integrity of Social Life* (Stanford Law Books 2009)

Nordin S and others, 'A Longitudinal Descriptive Study of Self-Reported Abnormal Smell and Taste Perception in Pregnant Women' (2004) 29 Chemical Senses 391 <https://academic.oup.com/chemse/article/29/5/391/368321/A-Longitudinal-Descriptive-Study-of-Self-reported> accessed 13 February 2019

North DC, *Institutions, Institutional Change and Economic Performance* (Cambridge University Press 1990)

Nunan D and di Domenico M, 'Big Data: A Normal Accident Waiting to Happen?' [2015] Journal of Business Ethics 1

OECD, 'Data-Driven Innovation for Growth and Well-Being: Interim Synthesis Report' (OECD 2014)

—— (ed), *Data-Driven Innovation: Big Data for Growth and Well-Being* (OECD 2015)

Office of Applied Studies, 'Cigarette Brand Preferences in 2005' (United States Department of Health and Human Services, Substance Abuse and Mental Health Administration 2007) Short reports <https://www.datafiles.samhsa.gov/study-publication/cigarette-brand-preferences-2005-nid15156> accessed 20 March 2019

O'Neil C, *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy* (Crown 2016)

Oxford Dictionaries, 'Tweet Geekery and Epic Crowdsourcing: An Oxford English Dictionary Update' (*OxfordWords blog*, 13 June 2013) <https://blog.oxforddictionaries.com/2013/06/13/oed-june-2013-update/> accessed 20 March 2019

Packard V and Miller MC, *The Hidden Persuaders* (Reissue Ed, Ig Publishing 2007)

Pariser E, *The Filter Bubble: What the Internet Is Hiding from You* (Penguin Press 2011)

Parliamentary Assembly, 'Human Rights and Modern Scientific and Technological Developments' (Council of Europe 1968) Recommendation 509 (1968)

Parris R, 'Online T&Cs Longer than Shakespeare Plays – Who Reads Them? - Online T&Cs Word Counts Compared to Famous Books' (*Which? Conversation*, 23 March 2012) <https://conversation.which.co.uk/technology/length-of-website-terms-and-conditions/> accessed 20 March 2019

Pedreschi D, Ruggieri S and Turini F, 'The Discovery of Discrimination' in Bart Custers and others (eds), *Discrimination and Privacy in the Information Society - Data Mining and Profiling in Large Databases* (Springer 2013)

Peissl W, 'Information Privacy in Europe from a TA Perspective' in Serge Gutwirth, Yves Poullet and Paul De Hert (eds), *Data protection in a profiled world* (Springer 2010)

Peppet SR, 'Unraveling Privacy: The Personal Prospectus & the Threat of a Full Disclosure Future' [2010] Northwestern University Law Review <http://papers.ssrn.com/abstract=1678634> accessed 3 November 2015

Perlroth N, 'How Superfish's Security-Compromising Adware Came to Inhabit Lenovo's PCs' *The New York Times* (1 March 2015) <http://www.nytimes.com/2015/03/02/technology/how-superfishs-security-compromising-adware-came-to-inhabit-lenovos-pcs.html> accessed 13 February 2019

Perrow C, *Normal Accidents: Living with High-Risk Technologies* (Basic Books 1984)

——, 'The Limits of Safety: The Enhancement of a Theory of Accidents' (1994) 2 Journal of Contingencies and Crisis Management 212 <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1468-5973.1994.tb00046.x> accessed 16 May 2019

——, 'Accidents, Normal', *International Encyclopedia of the Social & Behavioral Sciences* (Elsevier 2001) <http://linkinghub.elsevier.com/retrieve/pii/B0080430767045095> accessed 13 February 2019

Posner RA, 'The Social Costs of Monopoly and Regulation' (1975) 83 Journal of Political Economy 807 <https://www.jstor.org/stable/1830401> accessed 20 March 2019

Postel J, 'NCP/TCP Transition Plan' (1981) Request for Comments RFC 801 <https://tools.ietf.org/html/rfc801> accessed 20 March 2019

Prevelakis V and Spinellis D, 'The Athens Affair' (2007) 44 Spectrum, IEEE 26

Prins C, 'When Personal Data, Behavior and Virtual Identities Become a Commodity: Would a Property Rights Approach Matter?' (2006) 3 SCRIPT-ed 270 <https://script-ed.org/wp-content/uploads/2016/07/3-4-Prins.pdf> accessed 13 February 2019

Purtova N, 'Who Decides on the Future of Data Protection? Role of Law Firms in Shaping European Data Protection Regime' (2014) 28 International Review of Law, Computers & Technology 204 <http://dx.doi.org/10.1080/13600869.2013.801591> accessed 20 March 2019

Richards NM, 'The Dangers of Surveillance' (2013) 126 Harvard Law Review 1934

——, *Intellectual Privacy: Rethinking Civil Liberties in the Digital Age* (Oxford University Press 2015)

Richards NM and King JH, 'Three Paradoxes of Big Data' (2013) 66 (2013) Stanford Law Review Online

Riley T, 'Artificial Intelligence Goes Deep to Beat Humans at Poker' (*Science*, 3 March 2017) <http://www.sciencemag.org/news/2017/03/artificial-intelligence-goes-deep-beat-humans-poker>

Ring T, 'Keeping Tabs on Tracking Technology' (2015) 2015 Network Security 5 <http://www.sciencedirect.com/science/article/pii/S1353485815300477> accessed 13 February 2019

Rochet J-C and Tirole J, 'Platform Competition in Two-Sided Markets' (2003) 1 Journal of the European Economic Association 990 <http://onlinelibrary.wiley.com/doi/10.1162/154247603322493212/abstract> accessed 20 March 2019

——, 'Two-Sided Markets: An Overview' (Institut d'Economie Industrielle working paper 2004) <https://pdfs.semanticscholar.org/1181/ee3b92b2d6c1107a5c899bd94575b0099c32.pdf> accessed 20 March 2019

Rudgard S, 'Origins and Historical Context of Data Protection Law' in Eduardo Ustaran and others (eds), *European Privacy: Law and Practice for Data Protection Professionals* (International Association of Privacy Professionals 2012)

Sahota D, 'Global SMS Revenue Declines for First Time' (*Telecoms.com*, 14 January 2014) <http://telecoms.com/212062/global-sms-revenue-declines-for-first-time/> accessed 20 March 2019

Salvi O and others, 'F-Seveso: Study of the Effectiveness of the Seveso II Directive (Final Report)' (European Virtual Institute for Integrated Risk Management (EU-VRi) 2008) Contract n°070307/2007/476000/MAR/A3 <https://relevant.nl/download/attachments/4096340/seveso_report.pdf> accessed 13 February 2019

Sandvig C and others, 'Auditing Algorithms: Research Methods for Detecting Discrimination on Internet Platforms' (2014)

Schelling TC, 'Dynamic Models of Segregation' (1971) 1 Journal of mathematical sociology 143

Schermer BW, 'The Limits of Privacy in Automated Profiling and Data Mining' (2011) 27 Computer Law & Security Review 45 <http://linkinghub.elsevier.com/retrieve/pii/S0267364910001767> accessed 20 March 2019

——, 'Risks of Profiling and the Limits of Data Protection Law' in Bart Custers and others (eds), *Discrimination and Privacy in the Information Society* (Springer-Verlag 2013)

Schermer BW and Wagemans T, 'Onze Digitale Schaduw. Een Verkennend Onderzoek Naar Het Aantal Databases Waarin de Gemiddelde Nederlander Geregistreerd Staat.' (Considerati 2009)
<https://autoriteitpersoonsgegevens.nl/sites/default/files/downloads/rapporten/rap_2009_onze_digitale_schaduw.pdf> accessed 13 February 2019

Schlag P, 'Rules and Standards' (1985) 33 UCLA Law Review 379

Schneier B, *Secrets and Lies: Digital Security in a Networked World* (1 edition, Wiley 2004)

——, 'The Eternal Value of Privacy' (*WIRED*, mei 2006)
<http://archive.wired.com/politics/security/commentary/securitymatters/2006/05/70886> accessed 20 March 2019

——, 'The Myth of the "Transparent Society"' (*WIRED*, 3 June 2008)
<https://www.wired.com/2008/03/securitymatters-0306/> accessed 20 March 2019

——, 'The Battle for Power on the Internet' [2013] *Internet and Security*
<https://www.theatlantic.com/technology/archive/2013/10/the-battle-for-power-on-the-internet/280824/> accessed 13 February 2019

——, '"Stalker Economy" Here to Stay' (*CNN-Opinion*, 26 November 2013)
<http://www.cnn.com/2013/11/20/opinion/schneier-stalker-economy/index.html> accessed 13 February 2019

——, 'Metadata = Surveillance' (2014) 12 IEEE Security & Privacy 84
<http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6798571> accessed 20 March 2019

——, *Data and Goliath: The Hidden Battles to Collect Your Data and Control Your World* (1 edition, W W Norton & Company 2015)

Schreurs W, Hildebrandt M and Vanfleteren M, 'Cogitas, Ergo Sum. The Role of Data Protection Law and Non-Discrimination Law in Group Profiling in the Private Sector' in Mireille Hildebrandt and Serge Gutwirth (eds), *Profiling the European Citizen: Cross-Disciplinary Perspectives* (Softcover reprint of hardcover 1st ed 2008 edition, Springer 2008)

Schulte Nordholt E and others, *Dutch Census 2011: Analysis and Methodology.* (Statistics Netherlands 2014)

Schwartz P, 'Risk and High Risk: Walking the GDPR Tightrope' (*Privacy Perspectives - Ideas and Insights on Data Protection*, 29 March 2016) <https://iapp.org/news/a/risk-and-high-risk-walking-the-gdpr-tightrope/> accessed 13 February 2019

Seneviratne S and others, 'Predicting User Traits from a Snapshot of Apps Installed on a Smartphone' (2014) 18 Mobile Computing and Communications Review 1
<http://dl.acm.org/citation.cfm?id=2636244> accessed 20 March 2019

Shapiro C and Varian HR, *Information Rules: A Strategic Guide to the Network Economy* (Harvard Business Press 1998)

Sharma N and others, 'Predicting Solar Generation from Weather Forecasts Using Machine Learning', *Smart Grid Communications (SmartGridComm), 2011 IEEE International Conference on* (IEEE 2011) <http://ieeexplore.ieee.org/abstract/document/6102379/> accessed 20 March 2019

Sholtz P, 'Transaction Costs and the Social Cost of Online Privacy' (2001) 6 First Monday <http://journals.uic.edu/ojs/index.php/fm/article/view/859> accessed 13 February 2019

Siegel E, *Predictive Analytics: The Power to Predict Who Will Click, Buy, Lie, or Die* (2 edition, Wiley 2016)

Skouma G and Léonard L, 'On-Line Behavioral Tracking: What May Change After the Legal Reform on Personal Data Protection' in Serge Gutwirth, Ronald Leenes and Paul de Hert (eds), *Reforming European Data Protection Law* (2015 edition, Springer 2014)

Slawson WD, 'Standard Form Contracts and Democratic Control of Lawmaking Power' (1970) 84 Harvard Law Review 529

Solove DJ, 'A Taxonomy of Privacy' (2006) 154 U. Pa. L. Rev 477

——, 'What Is Sensitive Data? Different Definitions in Privacy Law' (*Privacy + Security Blog*, 31 July 2014) <https://teachprivacy.com/sensitive-data-different-definitions-privacy-law/> accessed 20 March 2019

statista.com, 'Gmail: Global Active Users Worldwide 2016' (*Statista*, 2017) <https://www.statista.com/statistics/432390/active-gmail-users/> accessed 20 March 2019

——, 'Facebook Users Worldwide 2018' (*Statista*, 2018) <https://www.statista.com/statistics/264810/number-of-monthly-active-facebook-users-worldwide/> accessed 20 March 2019

Steinfeld N, '"I Agree to the Terms and Conditions": (How) Do Users Read Privacy Policies Online? An Eye-Tracking Experiment' (2016) 55 Computers in Human Behavior 992 <http://linkinghub.elsevier.com/retrieve/pii/S0747563215301692> accessed 20 March 2019

Strömbäck P, 'Digital Myth: Competition Is Only One Click Away' (*Netopia*, 23 August 2016) <http://www.netopia.eu/competition-one-click-away/> accessed 21 March 2019

Study Group on a European Civil Code and Research Group on EC Private Law (Acquis Group), *Principles, Definitions and Model Rules of European Private Law: Draft Common Frame of Reference (DCFR). Outline Edition* (Sellier European Law Publishers 2009)

Tene O, 'For Privacy, European Commission Must Be Innovative | Center for Democracy & Technology' (*Center for Democracy & Technology*, 28 February 2011) <https://cdt.org/blog/for-privacy-european-commission-must-be-innovative/> accessed 13 February 2019

——, 'Vint Cerf Is Wrong. Privacy Is Not An Anomaly' (*Center for Internet and Society at Stanford Law School - Other writing*, 22 November 2013) <https://cyberlaw.stanford.edu/publications/vint-cerf-wrong-privacy-not-anomaly> accessed 21 March 2019

Tene O and Polonetsky J, 'Big Data for All: Privacy and User Control in the Age of Analytics' (2013) 11 Northwestern Journal of Technology and Intellectual Property 239 <http://papers.ssrn.com/abstract=2149364> accessed 21 March 2019

ter Huurne EFJ and Gutteling JM, 'How to Trust? The Importance of Self-efficacy and Social Trust in Public Responses to Industrial Risks' (2009) 12 Journal of Risk Research 809 <http://dx.doi.org/10.1080/13669870902726091> accessed 13 February 2019

Tessone CJ, 'The Complex Nature of Social Systems' in Bernardo Alves Furtado, Patricia AM Sakowski and Marina H Tóvolli (eds), *Modeling Complex Systems for Public Policies* (IPEA 2015)

Thaler RH and Sunstein CR, *Nudge: Improving Decisions About Health, Wealth, and Happiness* (Revised & Expanded edition, Penguin Books 2009)

TNS Opinion and Social, 'Special Eurobarometer 359: Attitudes on Data Protection and Electronic Identity in the European Union' (European Commission 2011) Survey

——, 'Special Eurobarometer 431: Report' (European Commission 2015) DS-02-15-415-EN-N <https://ec.europa.eu/COMMFrontOffice/publicopinion/index.cfm/ResultDoc/download/DocumentKy/66372> accessed 13 February 2019

Turner JS, 'New Directions in Communications (or Which Way to the Information Age?)' (2002) 40 IEEE Communications Magazine 50

Uleman JS, Adil Saribay S and Gonzalez CM, 'Spontaneous Inferences, Implicit Impressions, and Implicit Theories' (2008) 59 Annual Review of Psychology 329 <http://www.annualreviews.org/doi/10.1146/annurev.psych.59.103006.093707> accessed 21 March 2019

Uleman JS, Newman LS and Moskowitz GB, 'People as Flexible Interpreters: Evidence and Issues from Spontaneous Trait Inference' (1996) 28 Advances in experimental social psychology 211 <http://www.sciencedirect.com/science/article/pii/S0065260108602397> accessed 5 August 2017

United Nations Conference on Environment and Development, 'Rio Declaration on Environment and Development' <http://www.unesco.org/education/pdf/RIO_E.PDF> accessed 13 February 2019

United Nations Economic Commission for Europe, 'Convention on Access to Information, Public Participation in Decision-Making and Access to Justice in Environmental Matters Done At Aarhus, Denmark, On 25 June 1998'

van Alsenoy B and others, 'From Social Media Service to Advertising Network: A Critical Analysis of Facebook's Revised Policies and Terms, v.1.3' (ICRI - The Interdisciplinary Centre for Law & ICT, Katholieke Universiteit Leuven 2015) <https://www.law.kuleuven.be/citip/en/news/item/facebooks-revised-policies-and-terms-v1-3.pdf> accessed 13 February 2019

van der Kroft D, 'Persbericht: Bits of Freedom Roept KPN-Abonnees Op Om Aangifte Te Doen Tegen Aftappen' (12 May 2011) <https://www.bof.nl/2011/05/12/persbericht-bits-of-freedom-roept-kpn-abonnees-op-om-aangifte-te-doen-tegen-aftappen/> accessed 13 February 2019

Van Der Sloot B, 'Privacy as Virtue: Moving beyond the Individual in the Age of Big Data' (Universiteit van Amsterdam 2017)

van Eijk N, Hoofnagle CJ and Kannekens E, 'Unfair Commercial Practices: A Complementary Approach to Privacy Protection' (2017) 3 Eur. Data Prot. L. Rev. 325

van Rhee CH, 'De Ontwikkeling van het Burgerlijk Procesrecht in het Twintigste-Eeuwse Europa: Een Terugblik' in D Heirbaut, G Martyn and R Opsommer (eds), *De Rechtsgeschiedenis Van De Twintigste Eeuw. the Legal History of the Twentieth Century: Handelingen van het contactforum* (Peeters Bvba 2006)

van Schendelen MP, *Machiavelli in Brussels: The Art of Lobbying the EU* (2nd edn, Amsterdam University Press 2005)

Verbraucherzentrale Bundesverband, 'Samsung App-Store: Viele Klauseln Unzulässig' (*Samsung App-Store: Viele Klauseln unzulässig*, 6 June 2013) <https://www.vzbv.de/urteil/samsung-app-store-viele-klauseln-unzulaessig> accessed 13 February 2019

Verhelst EW, *Recht Doen Aan Privacyverklaringen: Een Juridische Analyse van Privacyverklaringen Op Internet* (Kluwer 2012)

Verplanken B and Weenig MWH, 'Graphical Energy Labels and Consumers' Decisions about Home Appliances: A Process Tracing Approach' (1993) 14 Journal of Economic Psychology 739 <http://linkinghub.elsevier.com/retrieve/pii/016748709390019H> accessed 13 February 2019

Vincent Fleming N, 'Sharing Your Location… In a Flash' (*FTC Consumer information blog*, 5 December 2013) <https://www.consumer.ftc.gov/blog/sharing-your-location-flash> accessed 21 March 2019

Visser M, 'KPN Investor Day: Consumer Wireless. Strengthen - Simplify - Grow' (KPN Investor Day, London, 10 May 2011) <https://ir.kpn.com/download/companies/koninkpnnv/Presentations/KPN_Investor_Day_-_Selective_topics.pdf> accessed 21 March 2019

Von Lewinski K, 'Zur Geschichte von Privatsphäre Und Datenschutz-Eine Rechtshistorische Perspektive' [2012] Datenschutz: Grundlagen, Entwicklungen und Kontroversen, Bundeszentrale für politische Bildung, Bonn 23

Wachter S, Mittelstadt B and Floridi L, 'Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation' (2017) 7 International Data Privacy Law 76 <https://academic.oup.com/idpl/article/7/2/76/3860948> accessed 13 February 2019

Walter C, 'Kryder's Law' (2005) 293 Scientific American 32

Wang M, 'China's Chilling "Social Credit" Blacklist' *Wall Street Journal* (11 December 2017) <https://www.wsj.com/articles/chinas-chilling-social-credit-blacklist-1513036054> accessed 21 May 2019

Wang Y and Kosinski M, 'Deep Neural Networks Are More Accurate than Humans at Detecting Sexual Orientation from Facial Images' (2017) (in press) Journal of Personality and Social Psychology <https://osf.io/fk3xr/> accessed 21 March 2019

Watson Marketing, '10 Key Marketing Trends for 2017 and Ideas for Exceeding Customer Expectations' (IBM Marketing Cloud 2017) WRL12345USEN <https://www-01.ibm.com/common/ssi/cgi-bin/ssialias?htmlfid=WRL12345USEN> accessed 21 March 2019

Wauters E, Lievens E and Valcke P, 'A Legal Analysis of Terms of Use of Social Networking Sites, Including a Practical Legal Guide for Users: "Rights & Obligations in a Social Media Environment"' (iMinds-ICRI 2013) D1.2.4 <https://limo.libis.be/primo-explore/fulldisplay?docid=LIRIAS1709099&context=L&vid=Lirias&search_scope=Lirias&tab=default_tab&lang=en_US> accessed 13 February 2019

——, 'Social Networking Sites' Terms of Use Addressing Imbalances in the User-Provider Relationship through Ex Ante and Ex Post Mechanisms' (2014) 5 JIPITEC <https://www.jipitec.eu/issues/jipitec-5-2-2014/4001> accessed 13 February 2019

Weatherall K, 'Three Lessons from ACTA and Its Political Aftermath' (2012) 35 Suffolk Transnational Law Review 575

Weatherill S, *EU Consumer Law and Policy* (Edward Elgar Publishing 2013)

Weick KE, 'Normal Accident Theory As Frame, Link, and Provocation' (2004) 17 Organization & Environment 27

Weinberger S, 'Son of TIA: Pentagon Surveillance System Is Reborn in Asia' (*WIRED*, 22 March 2007) <https://www.wired.com/2007/03/son-of-tia-pentagon-surveillance-system-is-reborn-in-asia/> accessed 21 March 2019

Weng SF and others, 'Can Machine-Learning Improve Cardiovascular Risk Prediction Using Routine Clinical Data?' (2017) 12 PLOS ONE e0174944 <http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0174944> accessed 21 March 2019

Wildberger AM, 'Alleviating the Opacity of Neural Networks', *1994 IEEE International Conference on Neural Networks, 1994 IEEE World Congress on Computational Intelligence* (1994)

Williams EW and Coase RH, 'Discussion' (1964) 54 The American Economic Review 192 <http://www.jstor.org/stable/1818503> accessed 21 March 2019

Winseck D, 'Netscapes of Power: Convergence, Network Design, Walled Gardens, and Other Strategies of Control in the Information Age' in David Lyon (ed), *Surveillance as social sorting: Privacy, risk and digital discrimination* (Routledge 2003)

World Administrative Telegraph and Telephone Conference and International Telecommunication Union, 'International Telecommunication Regulations: Melbourne, 1988 (WATTC-88).', *Final acts of the World Administrative Telegraph and Telephone Conference* (ITU 1989)

Yang K and others, 'A Nutritional Label for Rankings', *Proceedings of the 2018 International Conference on Management of Data* (ACM 2018)

Youyou W, Kosinski M and Stillwell D, 'Computer-Based Personality Judgments Are More Accurate than Those Made by Humans' (2015) 112 Proceedings of the National Academy of Sciences 1036 <http://www.pnas.org/content/112/4/1036> accessed 21 March 2019

Zarsky T, 'Incompatible: The GDPR in the Age of Big Data' (2017) 47 Seton Hall Law Review 995 <http://scholarship.shu.edu/shlr/vol47/iss4/2>

Zhao J and others, 'Men Also Like Shopping: Reducing Gender Bias Amplification Using Corpus-Level Constraints', *arXiv:1707.09457 [cs, stat]* (2017) <http://arxiv.org/abs/1707.09457> accessed 21 March 2019

Zikopoulos P and Eaton C, *Understanding Big Data: Analytics for Enterprise Class Hadoop and Streaming Data* (McGraw-Hill Osborne Media 2011)

Zuboff S, 'Big Other: Surveillance Capitalism and the Prospects of an Information Civilization' (2015) 30 Journal of Information Technology 75

Zuiderveen Borgesius FJ, 'Improving Privacy Protection in the Area of Behavioural Targeting' (Universiteit van Amsterdam 2014) <http://hdl.handle.net/11245/1.434236>

Zuiderveen Borgesius FJ and others, 'Online Political Microtargeting: Promises and Threats for Democracy' (2018) 14 Utrecht Law Review 82 <https://www.utrechtlawreview.org/article/10.18352/ulr.420/> accessed 21 March 2019

Zwenne G-J, *Diluted Privacy Law* (Leiden University 2013) <http://papers.ssrn.com/abstract=2488486> accessed 13 February 2019

8 Index of cases

Presented in chronological order.
Numbers in parentheses indicate pages where the cases are referenced.

European Court of Human Rights

Court of Justice of the European Union

General Court of the European Union

Bundesverfassungsgericht, Germany

Supreme Court of the United States

9 Curriculum Vitae

Michiel Rhoen (Sittard, August 1969) attended secondary education at the Stella Maris Scholengemeenschap in Meerssen. He obtained his diploma (VWO, Gymnasium B) in 1988. He graduated the Hogeschool Eindhoven (Fontys Hogescholen) with a bachelor's degree in applied physics engineering in 1993, after which he was conscripted for Dutch national service.

In 1994, he was admitted to the 33rd full-time course for professional fire officers at the Rijksbrandweeracademie in Arnhem (Dutch national fire service academy, currently Instituut voor Fysieke Veiligheid).

From 1996 to 2017, he has served as fire prevention officer and risk policy advisor, lastly at Veiligheidsregio Utrecht. From 2006 to 2008, and from 2017 to 2019, he was a senior policy advisor for the Dutch Ministry of the Interior and Kingdom Relations in the areas of Fire Safety and Information Society and Government, respectively. Since April 2019, he works for the Veiligheidsregio Utrecht as an Information Analyst.

While working for the fire service, Michiel started studying Dutch law part-time at Utrecht University in 2005 and graduated with a Master's degree in 2013.

He has been serving as a commissioned fire officer (hazardous chemicals specialist) since 2005.

10 Summary

Big data, big risks, big power shifts: Evaluating the GDPR as an instrument of risk control and power redistribution in the context of big data.

Big data – large amounts of data from automated processes – often constitutes personal data. The processing of large amounts of personal data constitutes a risk and can shift power towards those who collect and control the data. The General Data Protection Regulation (GDPR, Regulation EU 2016/679) acknowledges this risk and aims to regulate the processing of personal data in such a way that these risks are reduced to acceptable levels. Thus, free movement of personal data between EU Member States should become acceptable. However, a number of authors have pointed out a number of shortcomings in the GDPR. They have claimed that the GDPR is not *future proof*, is based on a number of *fallacies* of is *incompatible* with the reality of big data.

Evaluating the GDPR's effectiveness and developing a coherent body of jurisprudence depends on the clarity and testability of the underlying assumptions and expectations. One possible way to achieve such clarity and testability is to use models developed in the sciences. A number of models originating in the social and exact sciences is suitable to evaluate the GDPR. To this end, this research has applied Komesar's theory of comparative institutional analysis, Barnett and Duvall's theory of power in social relations, three theories of risk evaluation and management (Beck, Perrow and Klinke and Renn) and complex systems science.

Within the constraints of these theories, this research answers the following question: **To what extent does the GDPR reflect or employ theories of power relations and risk management presented by Komesar, Barnett and Duvall, Beck, Perrow, Klinke and Renn, and complex systems science?**

This question was answered using comparative methods (comparative institutional analysis and comparative law) and by evaluating two GDPR articles aimed at the prevention of discrimination using complexity theory. The research focuses on the protection of data subjects acting as consumers.

Comparative institutional analysis was used to analyse which group of actors (commercial actors or consumers) would have the best opportunities to influence the outcome of decision-making processes in their favour. Relevant factors in this comparison are the costs and benefits of participation. The comparison shows that in cases where the lawfulness of processing is based on consent or contract, controllers can be expected to have better opportunities for participation and therefore more opportunities to determine the outcome of the decision-making process in their favour.

Barnett and Duvall have proposed a framework for *evaluating power relations in social interactions*. In their terms, decision-making regarding the processing of personal data will be characterised by greater structural power for controllers than for consumers: they can dictate the contract terms. Controllers then use this power differential to increase their institutional power over consumers: controllers can exercise permanent surveillance over consumers. Large power differentials carry the risk of unfair treatment of the party holding less power. Applying comparative law methods shows that two EU consumer protection directives offer better opportunities than the GDPR to counter unfair treatment of data subjects.

Beck, Perrow and Klinke and Renn have mapped the *social consequences of technological risks*. EU environmental protection legislation, like the Seveso III-directive, contains elements tracing back to their analyses and recommendations. The risks associated with big data are comparable to the risks of industrialisation, in the sense that they affect not only individuals but also groups, and in the sense that their origins are not always clearly identifiable. However, contrary to EU environmental protection legislation, the GDPR does not contain a clear identification or evaluation of the risks that it aims to address.

Complex systems science holds that societies display properties that are not easily determined by observing individuals. But if personal data is gathered in sufficient amounts from a sufficient number of data subjects, societal properties can become discernible through analysis, for instance by using machine learning algorithms. In many cases, these properties can be mapped to sensitive traits like religious beliefs, ethnicity or sexual preference. Even though the GDPR aims to offer data subjects special protections when sensitive traits are processed, these protections can become ineffective when controllers are unable to verify the outcome of their algorithms for discriminatory effects. Likewise, Independent Supervisory Authorities may not be equipped to recognise if an algorithm categorises according to sensitive traits.

None of the models researched in the course of this project have been discernibly applied in the GDPR. At the same time, no other underlying conception of the risk and the expected effectiveness have presented themselves. This reflects society's lack of a comprehensive understanding of the risks associated with big data. It can make it more difficult to evaluate whether the GDPR is serving its purpose, it can cause a lack of focus in enforcement efforts and developing a coherent body of jurisprudence could take longer than necessary.

Based on the results of this research, this thesis proposes a framework for understanding the impact of big data on consumer contracts, and three points for reconsideration were the GDPR to be updated.

The framework consists of testable notions intended to serve as guidance for enforcement efforts and judicial decisions. These notions are:

- that the accumulation of personal data concentrates power in the hands of controllers,
- that information asymmetries undermine the assumption that consumers enter freely into contracts regarding personal data,
- that big data can lead to accidents at unexpectedly large scale,
- that the GDPR's protection of special categories of data is insufficient, and
- that data subjects currently cannot effectively take part in the decision-making regarding the risks of big data.

In the light of this research, the following aspects of the GDPR could eventually merit reconsideration:

- The usefulness of a general regulation regarding the processing of personal data, given that many specific areas of legislation already have their own concepts and methods for dealing with risks and power differentials;
- The sufficiency of transparency of the EU lawmaking process, given that the legislative process at the national level tends to offer better opportunities for participation to data subjects and consumers;
- The application of technologically neutral, complex standards for the processing of personal data, given that such forms of legislation tend to be better suited for risks that are well understood.

This research was based on four journal articles published between 2014 and 2018.

11 Samenvatting (Summary in Dutch)

Big data, risico's en schuivende machtsverschuivingen: Beoordeling van de Algemene Verordening Gegevensbescherming als instrument voor risicobeheersing en de regulering van machtsverhoudingen.

Big data – grote hoeveelheden gegevens afkomstig uit geautomatiseerde processen – bestaat vaak uit persoonsgegevens. Het verwerken van grote hoeveelheden persoonsgegevens brengt risico's met zich mee en kan leiden tot het verschuiven van machtsverhoudingen in het voordeel van degene die over grote hoeveelheden gegevens beschikt. De Algemene Verordening Gegevensbescherming (AVG, Verordening EU 2016/679) onderkent dit risico en beoogt zodanige voorschriften te geven voor het verwerken van persoonsgegevens, dat deze risico's tot een aanvaardbaar niveau worden teruggebracht. Deze voorschriften scheppen dan de voorwaarden voor vrij verkeer van persoonsgegevens binnen de Europese Unie. Diverse auteurs hebben echter al gewezen op een aantal tekortkomingen van de AVG: ze zou niet *toekomstbestendig* zijn, er zou een *dwaalleer* aan ten grondslag liggen of ze zou *onverenigbaar* zijn met de praktijk van big data.

Het toetsen van de effectiviteit van de AVG en het ontwikkelen van consistente jurisprudentie is alleen goed mogelijk als de onderliggende veronderstellingen en verwachtingen helder en testbaar zijn. Dit kan bijvoorbeeld door wetgeving te baseren op modellen afkomstig uit de wetenschap. Een aantal modellen dat is ontwikkeld in de sociale wetenschappen en de exacte wetenschappen, is bruikbaar voor het evalueren van de AVG. In dit onderzoek is ingegaan op de vergelijking van besluitvormingsprocessen (Komesar), machtsverhoudingen in sociale betrekkingen (Barnett en Duvall), sociale beoordeling en beheersing van technologische risico's (Beck, Perrow en Klinke en Renn) en de theorie van complexe systemen.

In dat kader beantwoordt dit onderzoek de volgende vraag: **In hoeverre geeft de AVG blijk van de toepassing van de theorieën over machtsverhoudingen en risicobeheersing zoals geformuleerd door Komesar, Barnett en Duvall, Beck, Perrow, Kline en Renn en complexe systemen?**

Deze vraag is beantwoord met behulp van vergelijkende methoden (vergelijkende institutionele analyse en rechtsvergelijking) en door beoordeling van twee AVG-bepalingen gericht op het voorkómen van discriminatie met behulp van de theorie van complexe systemen. Het onderzoek richt zich in het bijzonder op de bescherming van betrokkenen die ook consumenten zijn.

Vergelijkende institutionele analyse (comparative institutional analysis) maakt inzichtelijk welke groepen van partijen (bedrijven, consumenten) de meeste mogelijkheden hebben om de uitkomst van een besluitvormingsproces naar hun hand te zetten. Daarbij spelen de kosten en baten van deelname een rol. Uit de vergelijking blijkt dat wanneer de rechtmatigheid van de verwerking van persoonsgegevens berust op toestemming of een overeenkomst, verwacht kan worden dat verwerkingsverantwoordelijken meer invloed hebben op de overeenkomst of van de toestemmingsverklaring en dus een meer mogelijkheden hebben om de uitkomst van het besluitvormingsproces in hun voordeel te beïnvloeden.

Barnett en Duvall hebben een kader geformuleerd voor de beoordeling van *machtsverhoudingen in sociale interacties*. Volgens dit kader zal besluitvorming over de verwerking van persoonsgegevens worden gekarakteriseerd door een structureel machtsoverwicht voor bedrijven: zij kunnen eenzijdig de bepalingen in het contract dicteren. Bedrijven zetten dit machtsoverwicht daarna in voor het bewerkstelligen van hun institutionele machtsoverwicht, door consumenten te observeren. Te grote machtsongelijkheid kan resulteren in oneerlijke behandeling. Rechtsvergelijking laat zien dat twee richtlijnen uit het Europees consumentenbeschermingsrecht betere bescherming tegen oneerlijke behandeling bieden dan de AVG.

Beck, Perrow en Klinke en Renn hebben de *sociologische aspecten van het omgaan met technologische risico's* in kaart gebracht. Europees milieubeschermingsrecht, zoals de Seveso III-richtlijn, bevat kenmerken die ontleend zijn aan hun analyses en aanbevelingen. De risico's van big data zijn tot op zekere hoogte vergelijkbaar met die van milieurisico's: niet alleen individuen, maar ook (groepen in) samenlevingen kunnen worden blootgesteld aan risico's waarvan de oorsprong vaak moeilijk aanwijsbaar is. De AVG bevat echter geen duidelijke inventarisatie of evaluatie van de risico's die ze tracht te beheersen.

De theorie van *complexe systemen* stelt dat samenlevingen eigenschappen hebben die niet rechtstreeks herleidbaar zijn op de eigenschappen van individuen. Maar als er voldoende persoonsgegevens van een voldoende aantal personen worden verzameld,

kunnen deze eigenschappen ook uit de persoonsgegevens worden afgeleid, bijvoorbeeld met behulp van automatisch lerende algoritmen. In veel gevallen kunnen dergelijke eigenschappen samenvallen met gevoelige persoonskenmerken, zoals religieuze opvattingen, etniciteit of seksuele gerichtheid. Zelfs de bijzondere bescherming die de AVG biedt met betrekking tot deze kenmerken kan tekort schieten, als verwerkingsverantwoordelijken de uitkomsten van algoritmen niet rechtstreeks kunnen beoordelen op discriminatoire effecten. Ook onafhankelijke toezichthoudende autoriteiten kunnen niet in staat blijken om discriminatoire effecten van algoritmen vast te stellen.

De AVG geeft niet duidelijk blijk van de toepassing de modellen die in dit onderzoek aan de orde zijn gekomen. Daarnaast is ook niet gebleken van de uitdrukkelijke toepassing van andere modellen. Dit is een afspiegeling van ons huidige gebrek aan begrip van de risico's van big data. Dit gemis kan het moeilijker maken om de effectiviteit van de AVG te beoordelen, de juiste handhavingsprioriteiten te stellen, en samenhangende jurisprudentie te ontwikkelen.

Op basis van deze resultaten is een kader ontwikkeld voor de beoordeling van de effecten van big data op consumentencontracten, en worden drie punten voorgesteld waarop een eventuele opvolger van de AVG zou kunnen worden heroverwogen.

Het kader bestaat uit toetsbare hypothesen die richting zouden kunnen geven aan toezichthouders en judiciële beslissingen. Deze hypothesen zijn:

- dat het verzamelen van persoonsgegevens leidt tot de concentratie van macht in de handen van verwerkingsverantwoordelijken;
- dat informatie-asymmetrie ertoe leidt dat niet zomaar mag worden verondersteld dat consumenten uit vrije wil overeenkomsten over persoonsgegevens aangaan;
- dat big data kan leiden tot systeemongevallen op onverwacht grote schaal;
- dat de AVG onvoldoende bescherming biedt voor bijzondere persoonsgegevens;
- en dat betrokkenen op dit moment onvoldoende inspraak hebben bij de besluitvorming over de risico's van big data.

De volgende aspecten van de AVG zouden op termijn kunnen worden heroverwogen:

- De bruikbaarheid van een *algemene* verordening voor de verwerking van persoonsgegevens, gezien het feit dat diverse wetgevingsgebieden al beschikken over een vocabulaire en over methoden voor het beschrijven van – en omgaan met – risico's en machtsongelijkheid;

- De al of niet voldoende transparantie van het Europese wetgevingsproces, gezien het feit dat het nationale wetgevingsproces over het algemeen meer mogelijkheden biedt aan consumenten en betrokkenen om hieraan effectief deel te nemen;
- De toepassing van technologisch neutrale, complexe standaarden, met het oog op het feit dat dergelijke standaarden in het algemeen beter op hun plaats zijn bij de beheersing van risico's waarvan de aard en de omvang inzichtelijk zijn.

Dit proefschrift is gebaseerd op vier artikelen, gepubliceerd in de periode 2014-2018.