



Universiteit
Leiden
The Netherlands

H2020 COVR FSTP LIAISON - D2.3 Academic publication featuring the future of robot governance.

Fosch Villaronga, E.; Drukarch, H.G.

Citation

Fosch Villaronga, E., & Drukarch, H. G. (2021). *H2020 COVR FSTP LIAISON - D2.3 Academic publication featuring the future of robot governance. LIAISON*. Leiden: eLaw / Leiden University. Retrieved from <https://hdl.handle.net/1887/3278374>

Version: Publisher's Version

License: [Leiden University Non-exclusive license](#)

Downloaded from: <https://hdl.handle.net/1887/3278374>

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



D2.3 Academic publication featuring the future of robot governance

Type: Report
Access: Public
Date: August 2021 (M9)
Author(s): Eduard Fosch-Villaronga (eLaw, Leiden University)
Hadassah Drukarch (eLaw, Leiden University)
Revisions: 1



Deliverable: D2.3
Grant agreement no: 779966
Date: 2021-08-31

CONTENTS

SUMMARY	3
1. INTRODUCTION	4
2. LIAISON	5
2.1. BACKGROUND	5
2.2. APPROACH	6
2.3. GOALS	7
3. ACADEMIC PUBLICATIONS	8
3.1. INBOTS 2021	8
3.1.1. CONFERENCE DESCRIPTION	8
3.1.2. CONFERENCE PAPER	8
3.2. DATA FOR POLICY	8
3.2.1. CONFERENCE DESCRIPTION	8
3.2.2. CONFERENCE PAPER	9
3.3. FRONTIERS IN CHEMICAL ENGINEERING	9
4. REFERENCES	10
5. ANNEX	11

SUMMARY

While robots should be safe, robot regulatory frameworks do not always frame technology development accurately. LIAISON investigates to what extent compliance tools, in this case, the COVR toolkit, could be used as data generators for policy and standard makers to unravel an optimal regulatory framing for existing and emerging robot technologies and improve robot technology overall safety and market entrance ease. As such, LIAISON aligns with the overall COVR goal to reduce complexity in safety certifying robots.

New technologies sometimes question and challenge existing norms, breathing into existence the need for legal change. While the pace of technology dramatically accelerates, however, legal responsiveness does not always follow as a consequent step. As no formal communication process between robot developers and regulators from which policies could learn has been established yet, a stepback mechanism for robot governance as novel as that introduced by LIAISON is yet to be introduced to all stakeholders involved, in particular robot developers and policy/standard makers. To prove the feasibility and added value of the creation of this link between robot developers and relevant regulators, for the LIAISON project, we focus on three particular standards: ISO 13482:2014 on personal care robots, IEC 80601-2-78:2019 on rehabilitation robots, and EN ISO 18497:2018 on agricultural machinery and tractors. The policy and standard makers involved for this purpose represent both private standardisation organisations and the European Commission.

This report presents a set of academic publications featuring the future of robot governance as part of D2.3. of the FSTP LIAISON project, and the dissemination thereof in three different outlets: the H2020 CSA INBOTS 2021 Conference, the 2021 Data for Policy Conference, and a journal publication.

1. INTRODUCTION

COVR stands for "being safe around collaborative and versatile robots in shared spaces", and is a European H2020 Project which aims to reduce the complexity in safety certifying cobots significantly. In this respect, the project has developed the COVR Toolkit, an online tool that guides developers in their legal compliance process, from helping them find relevant technical standards/directives/protocols to guide them on performing a risk assessment.

Assessing risks through experimentation is essential to ensure robot safety and compliance with existing norms. However, standards do not always frame technology development accurately. LIAISON investigates to what extent compliance tools (tools that help comply with the legislation, such as the COVR toolkit) could be used as data generators for policy and standard makers to unravel an optimal regulatory framing (including change, revise, or reinterpret) for existing and emerging robot technologies. LIAISON is a crucial stepback mechanism to help align robot and regulatory development and improve robot technology's overall safety and market entrance ease. To prove the feasibility and added value of the creation of this link between robot developers and relevant regulators, for the LIAISON project, we focus on three particular standards: ISO 13482:2014 on personal care robots, IEC 80601-2-78:2019 on rehabilitation robots, and EN ISO 18497:2018 on agricultural machinery and tractors. The policy and standard makers involved for this purpose represent both private standardisation organisations and the European Commission. As such, LIAISON aligns with the overall COVR goal to reduce complexity in safety certifying robots by providing policy and standard makers with the necessary knowledge about legal inconsistencies, new categories, or new safety requirements (including psychological) to update existing frameworks where necessary and to ensure that the next generation of robots is 'safe' to the full extent of the word (see figure 1 in the annex). In this way, LIAISON contributes to the COVR mission by adding a link to public and private regulators to complete the cobot value chain.

This report presents the academic and scientific contributions featuring the future of robot governance as part of D2.3. of the FSTP LIAISON project, and the dissemination thereof in three different outlets: the H2020 CSA INBOTS 2021 Conference, the 2021 Data for Policy Conference, and a journal publication.

2. LIAISON

2.1. BACKGROUND

"The art of progress is to preserve order amid change, and to preserve change amid order" – Alfred North Whitehead.

Robot technology is one of the many technologies that challenge the regulatory framework in various ways, including ethics and security for responsible innovation, privacy, and responsibility allocation. As products, robots widely differ in embodiment, capabilities, context of use, intended target users, and many regulations may already apply to them. Having tools such as the COVR Toolkit can be of help. However, new applications may not fit into existing (robot) categories, legislation might be outdated and confusing categories, and technology-neutral regulations may be hard to follow for developers concerned about their particular case. A recent open consultation launched by the European Commission, for instance, acknowledges that current European Harmonized Standards do not cover areas such as automated vehicles, additive manufacturing, collaborative robots/systems, or robots outside the industrial environment, among others (Spiliopoulou-Kaparia, 2017). In light of all the issues this technology arises, part of the literature accentuates the need for an issue manager. Marchant and Wallach (2015) proposed the creation of "Governance Coordinating Committees (GCC)" for the governance of emerging technologies like AI.

Furthermore, the European Parliament proposed creating a European Agency for Robotics and Artificial Intelligence early in 2017, and Schatz put forward the creation of an emerging technology policy lab within the US general services administration in 2018. However, what lacks in robot governance is a backstop mechanism that can coordinate and align robot and regulatory development (Fosch-Villaronga & Heldeweg, 2018). Overlooked in the latest review of "the grand challenges of science robotics," this challenge has already been raised in the literature, albeit only more recently (Yang et al. 2018), and relates to the idea of how policies can frame the rapid development of robotics. LIAISON contributes to these approaches by proposing the *modus operandi* of issue managers, if they were ever to exist, and revolves around the following main research question:

Could the use of compliance tools, such as the COVR Toolkit, as data generators for robot policy purposes reduce emerging robot governance complexity?

LIAISON envisions an iterative regulatory process for robot governance, a theoretical model that represents a practical step forward in the coordination and alignment of robot and regulatory development, called the Iterative Learning Governance Process (ILGP). This research project conceives an effective way to extract compliance and technical knowledge from compliance tools (tools that help comply with the legislation such as the COVR toolkit) and direct it to policy and standard makers to unravel an optimal regulatory framing (including change, revise, or reinterpret) for existing and emerging robot technologies. The primary outcome of the LIAISON Research Project will be the design concept for liaising robot development and policymaking to increase overall robot safety. This design concept will further develop the *Iterative Regulatory Process for Robot Governance*, which was ideated as a theoretical model that links technology impact assessments to legislative ex-post evaluations via shared data repositories intending to create evidence-based policies

that can serve as temporary benchmark for future and new uses or robot developments (Fosch-Villaronga & Heldeweg, 2018, 2019). Part of the 'technical challenge' is to put such a theoretical model into practice and in the context of the COVR project. Explained further in figure 2 (annex), such iterative regulatory process for robot governance stresses that in the light of a new robot development or use, and after assessing all the impacts (and incorporating the findings into the robot itself), it is essential to compile all the Regulation-to-Technology uncovered barriers and constraints that do not allow the roboticists to proceed with their creation. Having collected those constraints in a Technology-to-Regulation manner, the regulator can act thereupon supported by the accountability tool's information, in this case, the COVR Toolkit.

2.2. APPROACH

Seeing regulation (broadly understood) as a tool to advance social goals and subject to adjustments towards this end, LIAISON discusses different regulatory approaches to use iterative governance processes for robot governance. For that purpose, LIAISON aims to engage with representatives from the industry, standardization organizations, and policymakers to present compliance tools as a potential source of information for policy action and understand what information would be helpful to them (e.g., through exploratory meetings, surveys, and workshops). Applying such a novel and interdisciplinary methodology is instrumental in identifying unregulated and underestimated challenges (e.g., over-time integrative and adaptive systems' safety, cyber-physical safety, psychological harm) that regulations should cover, and in gauging the response to, support for, and perceived necessity among relevant stakeholders of the introduction of the LIAISON model.

Following the ideal that lawmaking 'needs to become more proactive, dynamic, and responsive,' LIAISON proposes the formalization of a communication process between robot developers and regulators from which policies could learn, as depicted in figure 1 (annex), thereby channeling robot policy development from a bottom-up perspective towards a hybrid top-down/bottom-up approach. This is novel, as most approaches have been top-down solely, disregarding the richness field knowledge could provide in helping identify gaps and inconsistencies in frameworks governing the technology. In practice, LIAISON builds on the COVR toolkit, a compliance tool built as part of the COVR Project, by envisioning and assessing the usefulness of the proposed model based on the theoretical model of an *Iterative Regulatory Process for Robot Governance*. Following through the COVR Toolkit in the capacity of a robot developer, the Toolkit offers a section on standards and directives, allowing robot developers to filter their search results based on domain and appearance. The Toolkit then presents the relevant regulations, directives, and standards which can be freely accessed or purchased by robot developers. After robot developers have assessed the relevant documents, LIAISON enters into the picture. Focussing specifically on standards in 3 domains of application (rehabilitation, personal care, and agriculture), LIAISON aims to uncover the gaps and inconsistencies in the relevant policy documents. For this purpose, we have created two feedback loops to assess 1) regulatory gaps and inconsistencies in the relevant policy documents; and 2) the usefulness of LIAISON based on Toolkit user feedback and the broader community of stakeholders. To this end, we created a survey to match each feedback loop and distributed these among a predetermined pool of stakeholders through various means and on a variety of platforms.

Concerning the first feedback loop, assessing the identified gaps and inconsistencies in the relevant policy documents was refined through a set of interactive workshops, community engagement, and formal meetings. The data retrieved from these surveys have been channeled to a so-called 'shared data repository,' currently comprising a comprehensive Google sheets file. This shared data repository will be accessible to policymakers in due time, who are encouraged to use the relevant data to change, revise, or reinterpret existing frameworks. Once again, these will be presented in the COVR Toolkit, allowing the iterative regulatory process for robot governance to restart.

2.3. GOALS

We believe that the regulatory cycle is truly closed when it starts — or allows it to be started — again upon new challenges/technologies. LIAISON tests the theoretical model of a dynamic, iterative regulatory process in practice, aiming to channel robot policy development from a bottom-up perspective towards a combined top-down/bottom-up model, leaving the door open for future modifications. The above-envisioned process will clarify what regulatory actions policymakers have to take to provide compliance guidance, explain unclear concepts or uncertain applicability domains to improve legal certainty and inform future regulatory developments for robot technology use and development at the European, National, Regional, or Municipal level. Within this regard, LIAISON takes the lead in tackling the existing regulatory challenge, thereby linking robot development and policymaking to reduce the complexity in robot legal compliance. Moreover, by explicitly shedding light on the standardization activities in the abovementioned domains, LIAISON aims to create awareness about the barrier to access for robot developers and other relevant stakeholders concerning such activities.

In the long-term, the expected project results may complement the existing knowledge on the 'ethical, legal, and societal (ELS)' of robotics by providing clarity on how to address pressing but still uncovered safety challenges raised by robots and represent a practical, valuable tool to advance social goals in a robotized workplace. Overall, advances in safety robot legal oversight will provide a solid basis for designing safer robots, safeguarding users' rights, and improving the overall safety and quality of efficiency delivered by robots.

3. ACADEMIC PUBLICATIONS

3.1. INBOTS 2021

3.1.1. CONFERENCE DESCRIPTION

The INBOTS is an H2020 CSA bringing together experts from the different disciplines involved in the understanding and acceptance of Interactive Robotics INBOTS fosters the collaboration between four pillars and six areas of expertise:

- Business pillar – players that are deeply involved in commercialisation and marketing of robotic systems.
- Technical pillar – academic or industrial experts that are currently releasing efforts in researching and developing new robotic technology.
- Ethical, Legal and Socioeconomic pillar – partners with a background in social and legal science, which have been actively participating in current discussions about robotics and about responsible research and innovation actions.
- End Users, Policy Makers and General Public pillar – guaranteeing that they will be actively involved, informed and trained on current and future developments in Interactive Robotics.

On May 18-20, 2021, they held the conference on Inclusive Robotics in which LIAISON was presented. Click [here](#) for more information about the event.

3.1.2. CONFERENCE PAPER

Drukarch, H., Calleja, C., and Fosch-Villaronga, E. (2021). Building a dynamic framework for evidence utilization in policymaking for robot technologies. *Paper presented* at the H2020 CSA INBOTS 2021 Conference, May 18-19, 2021. The discussion paper will be part of the conference proceedings which will be available at a later stage.

3.2. DATA FOR POLICY

3.2.1. CONFERENCE DESCRIPTION

Data for Policy is a premier forum for interdisciplinary and cross-sector discussions around the impact and potentials of the digital revolution in the government sector. The conference series was established in 2015 with its inaugural edition 'Policy-Making in the Big Data Era'. Its inception was a result of observing that rapidly emerging and developing digital technologies could be utilised to inform and direct public policy. The Data for Policy conference series has created a forum for interdisciplinary, cross-sector discussion for stakeholders from all walks of life. Data for Policy has also launched the Data & Policy Journal in 2019 to archive and disseminate knowledge and to extend the community further. Both venues aim to bridge the gap between research and practice. They involve contributors from academia, government, international organisations, non-profit and commercial sectors.

The conference 'Data for Policy 2021: Lessons for policy-data interactions after Covid-19' offers a set of two tracks:

- Standards stracks
 - Data-driven Transformations in Policy and Governance
 - Data Technologies and Analytics for Policy and Governance
 - Policy Frameworks, Governance and Management of Data-driven Innovations
 - Ethics, Equity and Trust in Policy Data Interactions
 - Algorithmic Governance
 - Data to Tackle Global Issues and Dynamic Societal Threats
- Special tracks
 - Arguments, algorithms and tools: what do we need to shape policy and confront misinformation post-pandemic?’
 - Facilitating Data-Driven Innovation for Sustainability: Policy Frameworks and Measures for Data Governance
 - Systemic engagement: Mobility as a Service (MaaS) and the design challenge of inclusion, sustainability, and data ownership

On September 14-16, 2021. Click [here](#) for more information about the event.

3.2.2. CONFERENCE PAPER

Drukarch, H., Calleja, C., and Fosch-Villaronga, E. (2021). An iterative regulatory process for robot governance. *Accepted for presentation* at the [Data for Policy 2021 Conference](#), Sept 14-16, 2021. The discussion paper is available [here](#), and the conference presentation is available [here](#). Selected contributions may be selected to the conference proceedings in the [Data & Policy journal](#), published by Cambridge University Press.

3.3. FRONTIERS IN CHEMICAL ENGINEERING

Martinetti, A., Chemweno, P., Nizamis, K., & Fosch-Villaronga, E. (2021) Redefining safety in light of human-robot interaction: a critical review of current standards and regulations. *Frontiers in Chemical Engineering*. Research Topic: Artificial Intelligence in Chemical Engineering, forthcoming.

- PLEASE FIND A COPY OF THE ABOVE MENTIONED PAPERS IN THE ANNEX -

4. REFERENCES

A Bill To authorize an emerging technology policy lab within the General Services Administration, and for other purposes, S. 3502, 115th Cong. (2018). <https://www.congress.gov/115/bills/s3502/BILLS-115s3502is.pdf>

European Parliament resolution of 16 February 2017 with recommendations to the Commission on Civil Law Rules on Robotics (2015/2103(INL)).

Fosch-Villaronga, E. (2019) Robots, Healthcare and the Law: Regulating Automation in Personal Care. Routledge.

Fosch-Villaronga, E., & Heldeweg, M. (2018). "Regulation, I presume?" said the robot—Towards an iterative regulatory process for robot governance. *Computer Law & Security Review*, 34(6), 1258-1277.

Marchant, G.E., Allenby, B.R. and Herkert, J.R. eds., 2011. The growing gap between emerging technologies and legal-ethical oversight: The pacing problem (Vol. 7). Springer Science & Business Media.

Marchant, G.E. and Wallach, W., 2015. Coordinating technology governance. *Issues in Science and Technology*, 31(4), p. 43.

Salvini, P., Laschi, C., & Dario, P. (2011). Do service robots need a driving license?. *IEEE robotics & automation magazine*, 18(2), 12-13.

Spiliopoulou-Kaparia M. The evaluation of Directive 85/374/EEC on liability for defective products and Directive 2006/42/EC on machinery. Proceedings of the European Stakeholder Forum – Workshop on Regulatory challenges for a digitizing industry. Essen, 2017.

Yang, G.Z., Bellingham, J., Dupont, P.E., Fischer, P., Floridi, L., Full, R., Jacobstein, N., Kumar, V., McNutt, M., Merrifield, R. and Nelson, B.J., 2018. The grand challenges of Science Robotics. *Science Robotics*, 3(14), p.eaar7650.

5. ANNEX

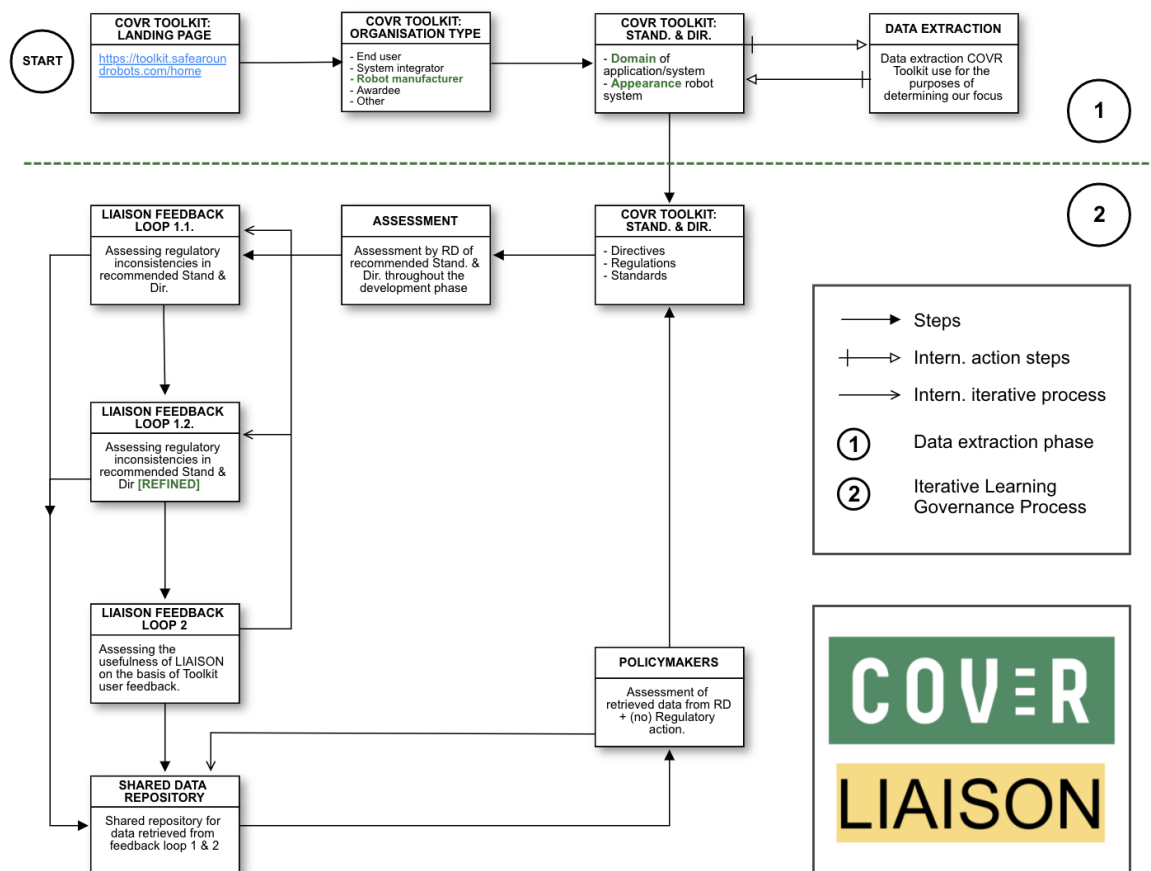


Figure 1: LIAISON Research Project mechanism

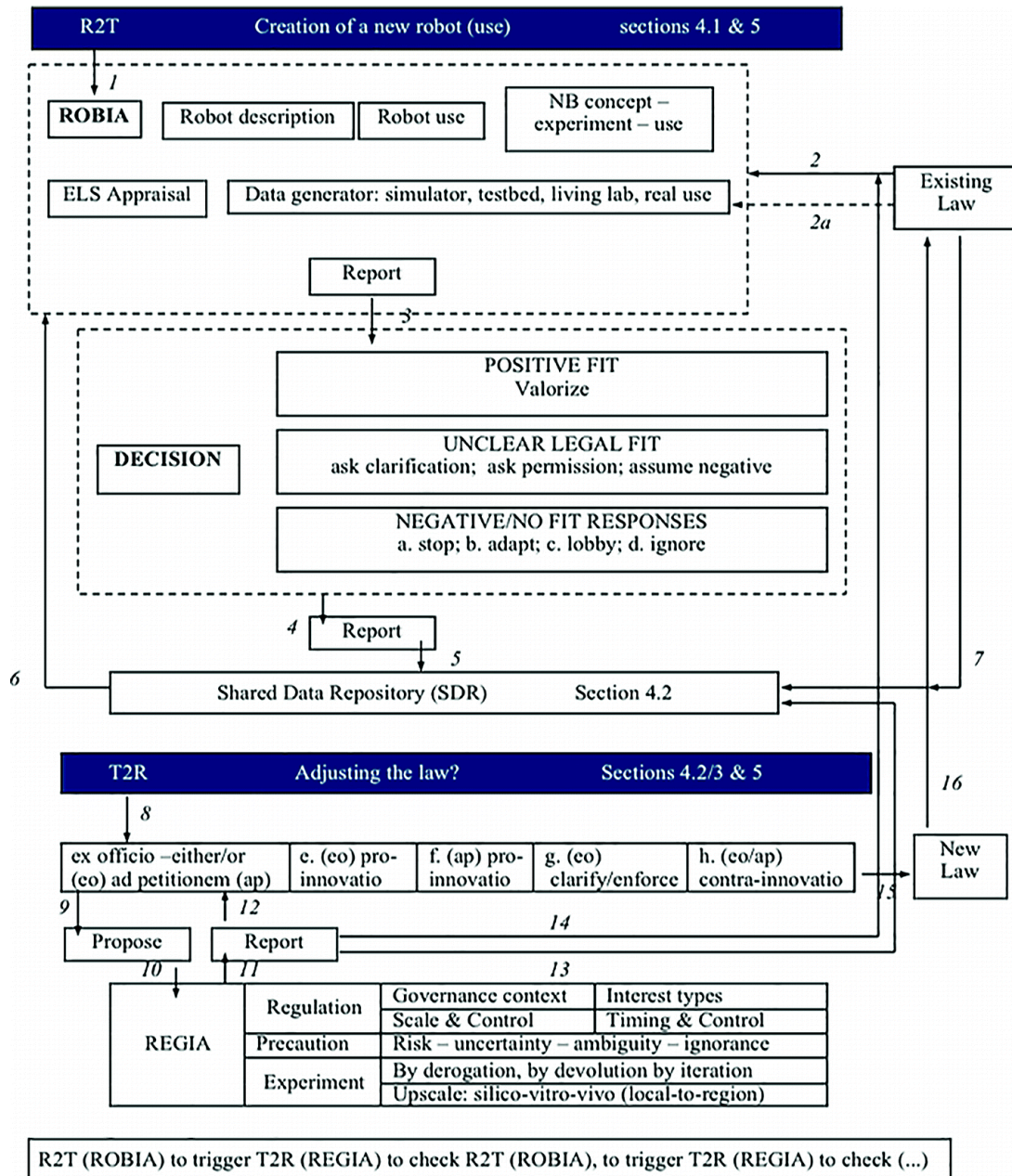


Figure 2: Preliminary iterative process for robot governance¹ (Fosch-Villaronga & Heldeweg, 2018).

¹ As regards the meaning of arrows: #1. signifies that upon the initiative to develop a new robot (use) the ROBIA process commences; #2 and #2a are about information about existing law/legal space being fed into the ROBIA fit to regulation process; #3 outcomes of ROBIA are reported to initiators to decide if and if so, how the development process can be continued; #4 and #5 concern reporting the decision and making information available to the SDR system; #6 is about how (changes in) information in SDR are a source of information to the ROBIA process – as shared learning; #7 is about information about existing law with relevance to robotics is also part of the shared data in SDR (#2 is about specific legal information to a specific ROBIA procedure; #7 about the general updating of legal info in SDR); #8 expresses that upon R2T events a process about possible legal adjustments is started; #9 and #10 when it is decided (ex officio/ad petitionem) that some legal change may be called for, a (basic) proposal is formulated whereupon the REGIA procedure is initiated; #11 and #12 show that outcomes of the REGIA procedure are reported back and feed into the decision on legal change; #13 Information in the report is also fed into SDR to update regulatory information; #14 REGIA report can feed ROBIA without passing via the Existing law> box, as the REGIA report will say something about pros and cons of possible legal change, but should that change follow, then this will communicate via the <New law> box; #15 signifies adjustments in the law; #16 expresses that new law changes and becomes part of existing law.)

An iterative regulatory process for robot governance

Building a dynamic framework for evidence utilization in policymaking for robot technologies

H. Drukarch, C. Calleja, and E. Fosch-Villaronga

Abstract—Assessing risks through experimentation is essential to ensure robot safety and compliance with existing norms. However, standards do not always frame technology development accurately. LIAISON investigates to what extent compliance tools, such as the COVR Toolkit, could be used as data generators for robot policy purposes to reduce the complexity in emerging robot governance. In particular, LIAISON focused on personal care robots (ISO 13482:2014), rehabilitation robots (IEC 80601–2–78–2019), and agricultural robots (ISO 18497:2018). This study brings into view the background and key findings of this project.

I. INTRODUCTION

New technologies sometimes question and challenge existing norms, breathing into existence the need for legal change. While technology's pace dramatically accelerates, an understanding of technology's implications does not always follow in parallel. Thus, current developments fail to provide an adequate level of safety for new technologies to perform well in the wild [1]. Because legal responsiveness does not always follow technological developments as a consequent step [2]–[3], current legislation also lags and does not frame robot technology adequately [4]. This state of affairs increases legal uncertainty, as it is unclear what regulations developers have to follow to comply with the legal system. Given that robotics is a realm characterized by uncertainty on the risks and appropriate safeguards for covering them, there is an urge to establish a coordinated regulatory front matching interests that can easily translate into concrete, practical and widely adopted actions for making robots safe to use [5]. The LIAISON Project investigates to what extent compliance tools, such as the COVR Toolkit, could be used as data generators for policymaking so as to reduce this complexity in the governance of emerging robot technologies. This study brings into view the inconsistencies, dissonances, and inaccuracies of existing frameworks with respect to robot technologies, and the usefulness of LIAISON in better framing these technologies and reducing the complexity in robot legal compliance.

II. A LEGAL 'TRAGEDY OF THE COMMONS'

Regulation is often very complicated, calling for a delicate interplay between various constraints, including the plurality and de-centeredness of our legal systems, the unclear fit of the regulated reality to the new development, and unforeseeable impacts of such emerging technology.

This contribution was written as part of the LIAISON project, an FSTP from the H2020 COVR project, which has received funding from the European Union's Horizon 2020 Research and innovation programme under grant agreement No 779966.

H. Drukarch, is with the eLaw Center for Law and Digital Technologies at Leiden University, Leiden, The Netherlands.

As a result of the inability of regulators to keep up with the fast pace of technological innovation, private actors have increasingly taken the lead to develop standards that aim to mitigate the ethical and legal risks and concerns posed by robotics, implying a shift in the centralization of regulation. Nevertheless, the speed with which robotics advances and the absence of a formal communication process between robot developers and regulators from which policies could learn, significantly hinders the understanding of the need to create technology-specific, long-lasting, and adequately fitting benchmarks, both for private and public policymakers. Moreover, harmonization of existing standards at a European level is also significantly lacking.

In light of all the issues robot technology introduces, several initiatives have been taken so far to reduce this complexity. However, the absence of a backstep mechanism that can coordinate and align robot and regulatory development in robot governance has long been overlooked and has only more recently been raised in the literature.

III. LIAISON: THE MODEL

In 2018, Fosch-Villaronga and Heldeweg built on and complemented these initiatives by proposing a governance process that can serve as a backbone in coordinating and aligning robot and regulatory developers [6]. This theoretical model envisions an iterative regulatory process for robot governance, which is hybrid and combines top-down and bottom-up approaches, but that favors room for improvement, allowing this development to generate knowledge about the adequacy of this framework.

In this respect, the LIAISON Project investigates to what extent compliance tools (here: the COVR Toolkit) could be used as data generators for robot policy purposes and reduce this way emerging robot governance complexity.¹ In particular, we focus on personal care robots (ISO 13482:2014), rehabilitation robots (IEC 80601–2–78–2019), and agricultural robots (ISO 18497:2018). LIAISON's main goal is to bring clarity to what regulatory actions policy and standard makers have to take to provide compliance guidance, explain unclear concepts or uncertain applicability domains to improve legal certainty, and inform future regulatory developments for robot technology use and development at the European, National, Regional, or Municipal level [5]. As such, LIAISON is a crucial stepback mechanism to help align

C. Calleja is with the eLaw Center for Law and Digital Technologies at Leiden University, Leiden, The Netherlands.

Dr. E. Fosch-Villaronga is with the eLaw Center for Law and Digital Technologies at Leiden University, Leiden, The Netherlands.

¹ LIAISON, <https://www.universiteitleiden.nl/en/research/research-projects/law/liaison>.

robot and regulatory development and improve robot technology's overall safety and ease of market entrance.

IV. METHODOLOGY

LIAISON used desk research to investigate robots' safety compliance landscape, and refined the findings resulting from this research through exploratory meetings with robot developers and relevant policymakers (including the EC, ISO and NEN, the Dutch Chapter of ISO). This knowledge formed the basis of the proposed project workflow, including feedback loops and a shared data repository, as depicted in figure 1.

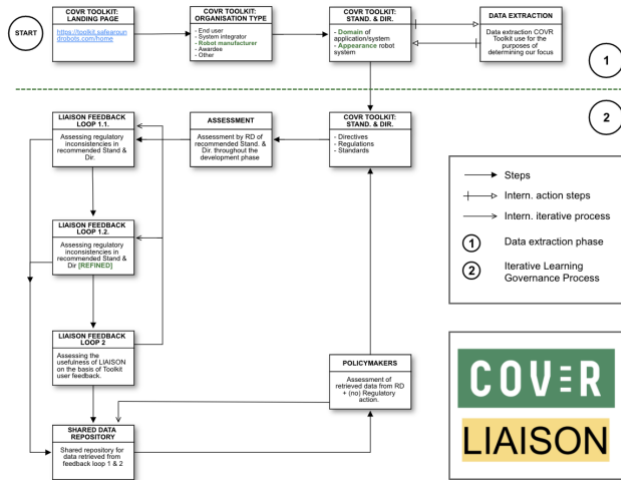


Fig. 1. LIAISON model for an iterative regulatory process for robot governance

The feedback loops took the form of online surveys and were distributed across the networks built by the Digital Innovation Hub (DIH) on Healthcare Robotics, the DIH Agrofood, and the H2020 COVR network, and were posted on social media platforms, including Twitter and LinkedIn. Moreover, the call for participation was featured by the 'Exoskeleton Report' [6]. The findings obtained through these feedback loops were complemented by the community interaction during an organized workshop at the European Robotics Forum 2021, and a European Commission workshops "Trends and Developments in Artificial Intelligence: Standards Landscaping and Gap Analysis on the Safety of Autonomous Robots".

V. KEY FINDINGS

The key LIAISON's preliminary findings we have are:

- 1) Within the context of robot regulation, there is a large ecosystem involving public policymakers, standard organizations, robot developers/manufacturers, and end users. This ecosystem does not have a common platform for better aligned interaction. Bringing all these stakeholders together to align their interests into making current and future robot technologies safe to use is essential.
- 2) Robot manufacturers deal with many different legal frameworks, including standards and regulations. While

compliance tools like the H2020 COVR Toolkit can help in this respect, the platform leaves room for desire. Especially for new robot manufacturers, a clarification of the existing legal landscape and the applicable legal framework would help in reducing the complexity in robot legal compliance. LIAISON can also add to this by providing a mechanism for a more detailed and improved model that can be tailored to different types of robots.

3) While the exact usefulness and feasibility of LIAISON from the perspective of standardisation is difficult to express at this current stage, both robot developers and policymakers believe that the governance model proposed by LIAISON can greatly contribute to reducing the complexity in robot governance and robot legal compliance in the long term.

4) While the data obtained through the above methods throughout the first milestone of the LIAISON Research Project have proved to provide valuable insight into the usefulness of LIAISON, further possibilities should be explored regarding outreach and effective stakeholder involvement in practice.

VI. CONCLUSION

Following the ideal that lawmaking 'needs to become more proactive, dynamic, and responsive' [7], LIAISON tests the theoretical model of a dynamic iterative regulatory process in practice, thereby aiming to channel robot policy development from a bottom-up perspective towards a combined top-down/bottom-up model, leaving the door open for future modifications. In the long-term, the expected project results will complement the existing knowledge on the 'ethical, legal, and societal (ELS)' of robotics by providing clarity on how to address pressing, but still uncovered safety challenges raised by robots, and represent a practical, valuable tool to advance social goals in a robotized workplace.

REFERENCES

- [1] K. Gruber, "Is the future of medical diagnosis in computer algorithms?," in *The Lancet Digital Health*, 1(1), e15-e16.
- [2] G.E. Marchant, B.R. Allenby, and J.R. Herkert, eds., "The growing gap between emerging technologies and legal-ethical oversight: The pacing problem" (Vol. 7), in *Springer Science & Business Media*.
- [3] G. Newlands, C. Lutz, A. Tamò-Larrieux, E. Fosch-Villaronga, G. Scheitlin, and R. Harasgama, "Innovation under Pressure: Implications for Data Privacy during the Covid-19 Pandemic," in *Big Data & Society*, SAGE, 7(2), 1-14.
- [4] E. Fosch-Villaronga, *Robots, Healthcare, and the Law: Regulating Automation in Personal Care*. 1 ed. 2020, Milton: Milton: Routledge.
- [5] E. Fosch-Villaronga, and M. Heldeweg, "Regulation, I presume?" said the robot – Towards an iterative regulatory process for robot governance. The computer law and security report, 2018. 34(6): p. 1258-1277.c
- [6] Exoskeleton Report (2021) LIAISON Project Seeks Feedback on Exoskeleton and Rehabilitation Robotics Standards.
- [7] M. Fenwick, W.A. Kaal, and E.P. Vermeulen, "Regulation tomorrow: what happens when technology is faster than the law," in *Am. U. Bus. L. Rev.*, 6, 561.

An iterative regulatory process for robot governance

Hadassah Drukarh*

Carlos J. Calleja*

Eduard Fosch-Villaronga*

* *eLaw Center for Law and Digital Technologies at Leiden University*, h.g.drukarch@law.leidenuniv.nl

Abstract

There is an increasing gap between the policy cycle's speed and that of technological and social change. This gap is becoming broader and more prominent in robotics, i.e., movable machines that perform tasks either automatically or with a degree of autonomy, since current legislation was unprepared for machine learning and autonomous agents and, as a result, often lags behind and does not adequately frame robot technologies. This state of affairs inevitably increases legal uncertainty. It is unclear what regulatory frameworks developers have to follow to comply, often resulting in technology that does not perform well in the wild, is unsafe, and can exacerbate biases and discrimination. This paper explores these issues and considers the background, key findings, and lessons learned of the LIAISON project, which stands for *Liaising robot development and policymaking*, and aims to ideate an alignment model between robots' legal appraisal channeling robot policy development from a hybrid top-down/bottom-up perspective to solve this mismatch. As such, LIAISON seeks to uncover to what extent compliance tools could be used as data generators for robot policy purposes to unravel an optimal regulatory framing for existing and emerging robot technologies.

Keywords – Robot governance, Policy Cycle; Evidence-based policy; Iterative regulatory process, Robot technology.

1 Introduction

New technologies offer possibilities until recently unimaginable and can solve problems faster, better, and more efficiently than ever before, inevitably disrupting our perception of reality (European Parliament, 2020) and leading us to question and challenge existing norms and call for an adequate regulatory response (Fosch-Villaronga and Heldeweg, 2018). Yet, while the development of new technologies rapidly accelerates, an understanding of technology's implications and consequent assessment and legal responsiveness do not always follow suit (Marchant, 2011; Newlands et al., 2021; Holder et al., 2016), creating an increasing gap between the policy cycle's speed and technological and social change (Sucha and Sienkiewicz, 2020). This gap is becoming broader and more prominent in robotics and AI, and in particular in the field of healthcare robots (Fosch-Villaronga, 2019).

The absence of specific regulations on robots, which differ mainly in embodiment and context of use, in which clear procedures, boundaries, and requirements are set to pose an immense challenge for how robot developers integrate legal considerations into their design to make robots safe (Holder et al., 2016; Fosch-Villaronga, 2019). Moreover, the mosaic of existing frameworks focus on physical safety mainly, while robots interact with humans in various ways, including socially, raising other questions than mere physical safety (Martinetti et al., 2021). As a result, current developments fail to provide an adequate level of safety and do not perform well in the wild (Gruber, 2019). In this state of affairs, there is an urge to establish a coordinated regulatory front matching interests that can easily translate

into concrete, practical and widely adopted actions for making robots safe to use (Fosch-Villaronga and Heldeweg, 2018).

Some European projects try to address this mismatch by creating different compliance tools, i.e., online tools that help comply with the legislation such as the COVR toolkit.¹ The COVR toolkit was developed by the H2020 COVR project, which stands for 'Being safe around collaborative and versatile robots in shared spaces.' This H2020 project compiles safety regulations for collaborative robots or *cobots*, i.e., robots developed to work in close proximity with humans (Surdilovic et al., 2011) in manufacturing, agriculture, and healthcare. The idea of the project is to present detailed safety assessment instructions to *coboteers* and make the safety assessment process clearer and simpler, which allows, in turn, cobots to be used in a more trustworthy and responsible way.

While these compliance tools may be of help to robot developers in their efforts towards robot legal compliance, new applications may nevertheless fail to fit into existing (robot) categories. Moreover, regulatory frameworks resulting from both private and public policymaking activities (Winfield, 2019) might be outdated, contain confusing categories, or directly be technology-neutral which may be hard to follow for developers who are concerned about their particular development (Fosch-Villaronga, 2019). In this sense, compliance tools are often unidirectional and top-down. They circumscribe the regulatory framework based on abstract categories which leads to them missing the input of robot developers, who might find mismatches and dissonances relevant to policymakers.

Bearing this in mind, the LIAISON Project, a subproject from the H2020 COVR project, investigates to what extent compliance tools, such as the COVR Toolkit, could serve as data generators for emerging robot governance purposes. To test its model, LIAISON focuses on personal care robots (ISO 13482:2014), rehabilitation robots (IEC 80601-2-78-2019), and agricultural robots (ISO 18497:2018).

This contribution explains the inner workings and findings of the LIAISON project. After a short introduction, Section 2 frames the inefficiency and inadequacy of emerging robot

governance. Section 3 introduces the LIAISON model and methodology, thereby highlighting the theoretical model envisioned and its practical application. Section 4 highlights the essential findings and lessons learned from building a dynamic framework for evidence-based robot governance as derived throughout the LIAISON project. The paper concludes with a summary of the achieved results and their potential impact on the governance of emerging technologies.

2 Framing the inefficiency and inadequacy of emerging robot governance

Although technology and regulation evolve, they do not always do so simultaneously or in the same direction. Legislation often follows societal changes in framing the rules of power and behavior by establishing rights and obligations for the subjects within society. On the other hand, technological developments are a consequence of scientific outbreaks and pressing needs. A recurrent question is how the law keeps up with such technological advances. The realm of robotics is not an exception, with many developments still largely unregulated. For instance, European Harmonized Standards do not cover automated vehicles, additive manufacturing, collaborative robots/systems, or robots outside the industrial environment (Spiliopoulou-Kaparia, 2017).

Notwithstanding, premature and obtrusive regulation might cripple scientific advancement and prevent potential advantages from materializing (Brundage and Bryson, 2016). This problem might result from ill-informed interventions, where policymakers rush to develop regulatory pieces without sufficient data on the targeted development (Leenes et al., 2017). The lack of a predictable environment and uncertainty regarding the impacts may disincentivize the development and introduction of emergent technologies. However, the ever-present conviction that technological fixes contribute to societal progress usually prevails in the techno-political discourse (Johnston, 2018).

Even in those areas falling within existing regulations, the law often fails to properly guide its addressees' behavior. Robots are complex devices, often combining hardware products, like actuators, with software applications and

¹ See <https://www.safearoundrobots.com/home>.

digital services. Consequently, technology-neutral regulation might fail to squarely frame their development, not to mention the interplay between different instruments and the lack of safeguards for emerging risks. For instance, the impact assessment on the Machinery Directive 2006/42/EC revision highlighted how developers must take several pieces of legislation for the same product to ensure its compatibility with all the applicable norms. Given this room for overlap - the Assessment noted - "there is a risk of applying the wrong piece of legislation and the related voluntary standards, thus negatively influencing safety and compliance of the product." (European Commission, Impact assessment study on the revision of Directive 2006/42/EC on machinery). This remark explains why legal requirements often play a marginal role in robotics. Instead, it is the self-perception of safety risks and economic concerns that often guide developers, who feel pressured to rely on insufficient testing. (see Van Rompaey et al., 2021). Under this uncertainty, neither the regulators nor the addressees know what needs to be done (Sabel et al. 2017), while the users' rights might nevertheless be at stake.

3 LIAISON: Liaising robot development and policymaking

Traditionally, juridical analysis follows top-down approaches to address legal and ethical aspects for robot technologies (Leroux, 2012; Leenes et al., 2017). However, these approaches often presume that the existing laws and norms suffice to understand the consequences of technology. Moreover, there are initiatives that promote reflection upon the consequences of the outcomes of technological research and development (R&D), fostering the incorporation of such reflections into the research or the design process. However, field knowledge could prove useful in identifying gaps and inconsistencies in frameworks

governing relevant technologies (Fosch-Villaronga and Golia, 2018).

Following the ideal that lawmaking 'needs to become more proactive, dynamic, and responsive' (Fenwick, Kaal and Vermeulen, 2016), LIAISON proposes the formalization of a communication process between robot developers and regulators from which policies could learn, as depicted in figure 1, thereby channeling robot policy development from a bottom-up perspective towards a hybrid top-down/bottom-up approach.

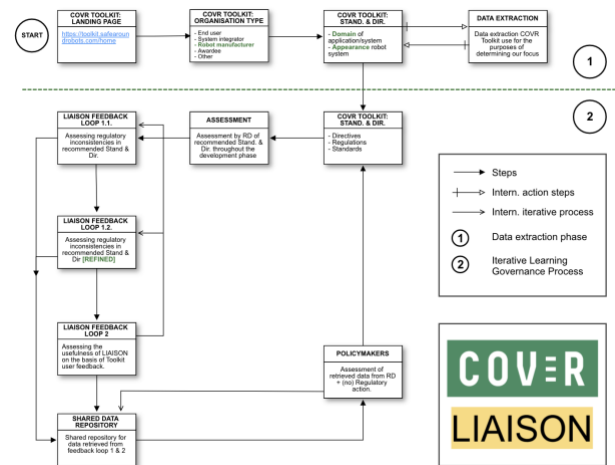


Figure 1. LIAISON model for an iterative regulatory process for robot governance.

By doing so, LIAISON envisions an iterative regulatory process for robot governance, a theoretical model representing a practical step forward in coordinating and aligning robot and regulatory development. It conceives an effective way to extract compliance and technical knowledge from the COVR Toolkit and direct this data to policy and standard makers to unravel an optimal regulatory framing, including decisions to change, revise, or reinterpret existing regulatory frameworks for existing and emerging robot technologies. More practically, the LIAISON project's objective is to clarify what regulatory actions policy and standard makers should take to provide compliance guidance, explain unclear concepts or uncertain applicability domains to improve legal certainty, and inform future regulatory developments for robot technology use and development at the European, National, Regional, or Municipal level (Fosch-Villaronga and Drukarch, 2020).

The LIAISON project brings into view the inconsistencies, dissonances, and inaccuracies of existing regulatory efforts and initiatives towards framing robot technologies. To this end, it focuses on personal care robots (ISO 13482:2014), rehabilitation robots (IEC 80601-2-78-2019), and agricultural robots (ISO 18497:2018), thereby aiming to uncover any gaps and inconsistencies and, additionally, gain insight into the usefulness and perceived value of the novel robot governance mechanism it introduces.

For this purpose, two feedback loops were created, namely for the purpose of assessing 1) regulatory gaps and inconsistencies in the relevant policy documents; and 2) the usefulness of LIAISON based on feedback from COVR Toolkit users and the broader community of stakeholders. The feedback loops depicted in figure 1 took the form of online surveys which were distributed across the networks built by the two major European networks on healthcare and agricultural robotics (the Digital Innovation Hub (DIH) on Healthcare Robots and the DIH on agricultural robots, from now on DIH HERO/DIH AGROBOFOOD), and were posted on social media platforms, including Twitter and LinkedIn. Additionally, the call for participation was featured in the 'Exoskeleton Report'.²

To build the surveys, desktop research was conducted to investigate robots' safety compliance landscape, and the findings resulting from this research were refined through exploratory meetings and formal engagements with representatives from the industry, standardization organizations, and policymakers to present compliance tools as a potential source of data for policy purposes and action, and understand what information would be helpful to them. Moreover, to avoid low response rates and increase the focus of the obtained responses, the LIAISON project organised a set of interactive workshops, community engagement activities, and formal meetings which included the organisation of dedicated workshops and presentations at the online European Robotics Forum 2021³ (13-15 April 2021), and the European Commission workshops on "Trends and Developments in Artificial Intelligence: Standards Landscaping and Gap Analysis on the Safety of

Autonomous Robots" (2-3 March 2021), a set of dedicated webinars for the DIH HERO community⁴, and several formal engagements with the respective DIH communities, the industry, academia, policymakers, and industry associations.

The data retrieved from the relevant surveys have been channeled to a so-called 'shared data repository,' currently comprising a comprehensive Google sheets file. This shared data repository will be made accessible to the relevant policymakers in due time. Any policy changes implemented on the basis of the shared data will, then, be presented in the COVR Toolkit as the presented standards, directives and regulations, allowing the envisioned iterative regulatory process for robot governance to restart.

4 Building a dynamic framework for evidence-based robot governance: key findings, lessons learned, & discussion

4.1 A mere website showing regulations is not enough help to robot developers comply

Robot manufacturers deal with many different legal frameworks, including standards and regulations. While compliance tools like the H2020 COVR Toolkit can help in this respect, the platform leaves room for desire: especially for new robot manufacturers, clarifying the applicable legal framework would help reduce the complexity in robot legal compliance. For instance, the COVR Toolkit does not distinguish between directives, which target EU member states and potentially lead to national legislation, and regulations, which are directly applicable. Furthermore, it does not explain the relationship between standards (usually non-binding) and the law (directives/regulation). Consequently, it is not clear what *binding* legal frameworks robot developers ought to comply with and which *recommended* frameworks concretize abstract legal principles.

² See

<https://exoskeletonreport.com/2021/03/liaison-seeks-feedback-on-exoskeleton-and-rehabilitation-robotics-standards/>.

³ See https://www.eu-robotics.net/robotics_forum/.

⁴ E.g. a dedicated session on medical robot autonomy at [MESROB 2021](#) which was held from 7-9 June 2021, and a dedicated

presentation on gaps and inconsistencies in standards for personal care and rehabilitation robots at the DIH HERO Knowledge Conference 2021 which was held on 20 May 2021. Other engagements included a dedicated conference paper on LIAISON at the [INBOTS conference 2021](#) which was held from 18-20 May 2021 and the COVR Open Day 2021 which was held on 3 June 2021.

Since current laws do not necessarily target specific types of robots, it is crucial to distill abstract standards into particular requirements capable of guiding developers in the development of a specific robot type. At present, the platform is limited to a small number of robot types and lacks a regulatory model personalized to each robot type, missing the opportunity to present developers with a clear picture of which laws apply to their robot technology. Logically, the platform is not intended to provide legal advice and may want to avoid any liability resulting from it.

In this vein, the platform is unidirectional: it merely presents legal information. This has the problem that, first, it needs to be updated and maintained with relevant developments. Most importantly, however, is that there is no engagement or feedback envisioned with the community. Following our engagement with the Digital Innovation Hubs - DIH-HERO and DIH-AgROBOfood, it became clear that LIAISON can strengthen the ecosystem created by these European initiatives by engaging them directly for governance purposes. Such an engagement could create a common platform to identify dissonances, and share lessons learned and tips from developers, which could be more valuable than a static information provision and allow developers to make their voice heard in robot governance activities (Fosch-Villaronga and Drukarch, 2021c). Over time, the generated knowledge could be helpful to policymakers to enact policies more attuned to stakeholder needs and rights. The importance and added value of combining the policy appraisal mechanism introduced through LIAISON with the COVR compliance tool has also been confirmed in the literature, which states that ‘better tools and appraisal design can lead directly to better policy appraisal and hence better policies’ (Adelle, Jordan and Turnpenny, 2012).

4.2 An ecosystem encompassing all stakeholder interests is lacking

It is noticeable that a common platform for channeling the interaction between public policymakers, standard organizations, robot developers/manufacturers, and end-users is currently lacking (Fosch-Villaronga and Drukarch, 2021c). The lack of a global ecosystem encompassing all stakeholder interests may entail many consequences. First, it prevents an active engagement with the affected stakeholders. For instance, ISO and standardization organizations alike welcome many stakeholders to participate in their standardisation activities but they do not

proactively seek any stakeholder involvement e.g., they do not ask user groups to be involved. Over time, this creates power imbalances concerning the creation and production of norms geared towards framing robot development and user rights protection.

LIAISON highlighted how reshaping the mode of interactions yields helpful feedback from the community and thus facilitates the iterative regulatory process (Fosch-Villaronga and Drukarch, 2021c). It is thus essential to devise a mechanism for bringing together all stakeholders to align their efforts into making current and future robots safe to use. To create an optimal regulatory framework for new technologies, the European Commission needs to better communicate with society, standard-makers, and developers. A platform based on the model that LIAISON proposes could be beneficial in improving the communication between all stakeholders involved in the development and regulation (be it through public or private bodies) of new technologies.

4.3 Engaging with robot developers can generate policy-relevant knowledge

Engaging developers and end-users is instrumental in identifying unregulated and underestimated challenges (e.g., over-time integrative and adaptive systems’ safety, cyber-physical safety, psychological harm) that regulatory frameworks should cover. These challenges arose in conversations with developers, and are mainly connected with the concerns and the challenges they find while turning a prototype into a marketable device. In this subsection, we introduce the different findings with respect to personal care, rehabilitation, and agricultural robots (Fosch-Villaronga and Drukarch, 2021b).

4.3.1. Findings concerning personal care robots

Personal care robot developers experience challenges and inconsistencies regarding ISO 13482:2014. While some developers perceive the standard as easy to follow and valuable as it is, others highlight that it could benefit from better guidance and that while useful, the standard could use different and more concrete examples and measures to be simpler and easier to follow. At the same time, the respondents indicate that the standard is clear concerning its language and layout. They nevertheless see room for improvement in the following sections: (i) safety requirements & protective measures, (ii) safety-related

control system requirements, and (iii) verification and validation.

From the standard making side, in the exploratory meeting with a representative of the ISO Technical Committee TC299 (Robotics) Working Group 02 on Service Robot Safety standardization, working on the revision of ISO 13482:2014, it became evident that the TC299 sees potential areas for improvement of the standard in its scope and structure. A concrete example is the need to provide further guidance for specific user types. The importance of considering the elderly, children, and pregnant women under ISO 13482:2014 has been pointed out during the LIAISON workshop at the ERF, where participants generally indicated to believe that such consideration is fundamental. In this regard, the standard stated that the Working Group would revise the definition of personal care robots, taking into account concrete users such as children, elderly persons, and pregnant women. However, the 2020 revised standard shows no changes in this respect (Fosch-Villaronga and Drukarch, 2021a).

More specifically, participants indicated that the standard should consider concrete aspects for these different users: cognitive capabilities, different learning ways, safety, different limit values, mental and physical vulnerability, different body dimensions, interaction requirements, mental culture, physical disabilities, interaction with the robot, situation comprehension, body-reaction time, and size physiology. In addition, participants indicated that there should not be a simplification of specific user groups and standard revision within this context. Moreover, within this debate, the importance of adequate training was also stressed, thereby highlighting the need to reconsider the provided training that considers multiple user types.

Concerning the specific definition of personal care robot, LIAISON's methodology exposed problems associated with a lack of definition for *personal care* within the standard. Without a defined legal scope, engineers might comply with the wrong instruments (e.g., they might avoid medical legislation) and, therefore, be exposed to sanctions and further responsibilities. The LIAISON survey provided an example directed towards this, where a respondent indicated to be dealing with exoskeletons that fall under the physical assistant robots category but was uncertain as to whether she should follow the Medical Device Regulation (MDR).

Moreover, public responses during the ERF workshops indicate that those robot developers who have experience with ISO 13482:2014 run multiple standards for their devices. They also believe that their robot does not fit into the standard category, and they do not know if their robot is a medical device, and have all been confronted with different classifications from public and private policy documents. As a result of incorrect categorization and unclear classification, roboticists might put in place inappropriate safeguards, and users might be put in risky or harmful situations (Fosch-Villaronga, 2019).

In this sense, one of the most notorious confusing categories identified by developers through LIAISON when it comes to healthcare is whether a robot is a medical device or not. While ISO 13482:2014 aimed to bring more clarification to the field, it created many new confusing categories. The focus on personal care robots created an in-between category between service robots and medical devices. This ended up in two standards/categories: those for medical use and well-being and that personal care. However, the MDR states that "devices with both a medical and a non-medical intended purpose shall fulfill the requirements applicable to devices cumulatively with an intended medical purpose and those applicable to devices without an intended medical purpose" on its Art 1.3. This article was meant to avoid treating different devices that presented a similar risk to the user. For instance, colored contact lenses were considered cosmetics while presenting the same risks that contact lenses to replace glasses presented to the human eye. In this regard, the critical question is, to what extent will ISO 13482:2014 provide any room for those robotic devices that flirt the boundary of medical and non-medical.

While ISO 13482:2014 considers the consequences of error in robot autonomous decisions, the wrong autonomous decisions section only states, "a personal care robot that is designed to make autonomous decisions and actions shall be designed to ensure that the wrong decisions and incorrect actions do not cause an unacceptable risk of harm." However, the standard does not clarify the meaning of an "acceptable risk of harm" and a "non-acceptable risk," nor does it define the criteria to decide on this. Silence on these matters, however, does not provide a safeguard baseline.

Finally, as a solution, the standard states that "the risk of harm occurring as an effect of incorrect decisions can be lowered either by increasing the reliability of the decision or by limiting the effect of a wrong decision." This brings

about the question whether a broader range of factors should be considered in the standard in this regard. For instance, it is not clear whether the provisions around safety as stipulated in the standard need to be combined with article 22 - on automated decision-making, including profiling - of the General Data Protection Regulation (GDPR). While the GDPR seems to have been drafted with computer systems in mind, cyber-physical systems have been primarily disregarded in this regard (Felzmann et al., 2018).

4.3.2. Findings concerning rehabilitation robots

Before rehabilitation robots can be made commercially available in the EU, manufacturers need to demonstrate that the device is safe. However, the safety validation of rehabilitation robots is complex. Especially when it comes down to straightforward testing procedures that can be used during robot development, information in regulations and standards is rare or scattered across multiple standards. This is partly because rehabilitation robots are relatively new, which reduces the availability of best practices and applicable safety standards.

Moreover, manufacturers of rehabilitation robots should be aware that article 1.6 of the MDR, in essence, states that devices that can also be seen as machinery (such as a robot) should also meet essential health and safety requirements as set out in Annex I of the Machinery Directive. Similarly, there might be standards from other domains which are more specific than the general safety and performance requirements listed in the MDR and can therefore be relevant for rehabilitation robots (e.g., personal care safety standards). However, the user must consider any restrictions or differences between the domains and be aware that the respective standard is not directly applicable (Bessler et al., 2021). Moreover, in the EU, the legislation for medical devices applies to rehabilitation robots. When a device complies with relevant so-called harmonized standards, the developer can assume that the device complies with the EU legislation. However, for medical devices, the current applicable harmonized standards are harmonized for the MDD. This means that between May 2021 and May 2024, there probably will be no or just a limited number of harmonized standards that can officially be used to demonstrate conformity with the MDR (Bessler et al., 2021). Important to note within this context is that the

familiarization with applicable regulations and standards and the process of safety validation takes much time, which can be a burden, especially for small to medium enterprises and start-ups.

In addition to the documentation of the system and the risks involved, a validation of the risk mitigation strategies is also required. This validation is defined as a set of actions to evaluate with evidence that a bunch of safety functions meet a group of target conditions (Saenz et al., 2020) and is essentially a measurement to prove that a specific system complies with designated operating conditions characterized by a chosen level of risk. There is currently no guidance from standards on how validation measurements should be executed (Bessler et al., 2021). Especially concerning the usefulness of protocols⁵, the majority of the participants in the LIAISON workshop at the ERF indicated that protocols offer a very clear and valuable tool in guiding them through the validation process.

4.3.3. Findings concerning agriculture robots

ISO 18497:2018 specifies principles for designing highly automated aspects of large autonomous machines and vehicles used for agricultural field operations but fails to include small autonomous agricultural robots into its scope. A regulatory framework for small autonomous agricultural robots is yet to be created. A valid question raised within this context relates to the definition the standard attributes to a highly automated agricultural machine - *does this definition also encompass agricultural robotic devices?* The insights provided by participants in the LIAISON workshop at the ERF led to interesting findings, with some participants believing that agricultural robots fall within the scope of this definition, and others thinking that they do not (see figure 2).

Engagement with DIH-AgROBOfood's standards work package leader led to the finding that no safety standard addresses agricultural robots during a rapidly advancing field. As a result, it will become necessary to specifically take this type of robotic field into account in the standard revision or create a standard tailored explicitly to agricultural robots.

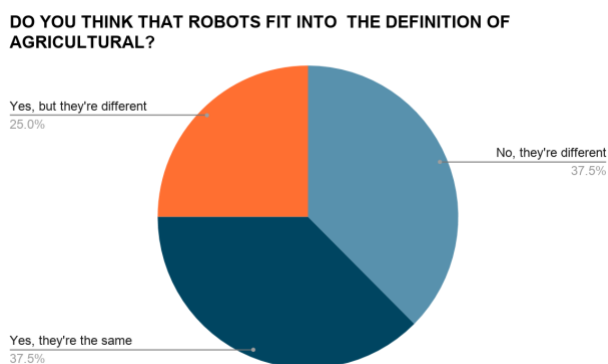
⁵ Protocols are step-by-step guides on how to validate safety of a robotic system. The validation relies on measurements, and the

protocols will tell the robot developer how to conduct these measurements in a correct and reliable way.

On a European level, the Machinery Directive 2006/42/CE is the reference text on the regulation of equipment and machinery, including for agriculture. To observe these requirements, European and international norms and standards (EN and ISO) are applied. However, the emergence of agricultural robots have led to new functionalities and new applications and therefore unknown risks, which must be understood to best comply with the Machinery Directive. Compatibility with the automation of agricultural functions is not always apparent. The Directive stipulates that a machine must not make unexpected movements near a person. This calls into question the automated process that enables a robot to take over on start-up. Other discrepancies between text and practice include operator responsibility, mainly where the operator acts remotely.⁶ Such operators are not always present with the robot but nevertheless may remain legally responsible for the safety of operations and must, therefore, be able to place the machine in safe mode at all times.

A key lesson learned from interacting with the agricultural community is that with risk analysis and the performance of tests and adjustments in the design phase, the key is developing reliable, safe machines within a regulatory context that is ill-suited and inaccurate.

Figure 2. Do you think that robots fit into the definition of agricultural machinery?



Therefore, within agriculture, harmonized standards from other sectors are applied analogously with agricultural robotics. On the remaining points, robot developers explain the risk analysis conducted and set out the solutions

implemented in response, thereby demonstrating the resulting level of performance.

Finally, the involved Digital Innovation Hubs also stress this need for diverse stakeholder involvement. For instance, engagement with DIH AgROBOfood has presented the need for robot developers to pay attention to ethical, legal, and many other issues to determine if a robot will survive in a practical setting. Since agricultural robots barely interact with humans (at least not directly as in personal care or rehabilitation), the community is still not ready to engage with the ethical, legal, and societal (ELS) aspect community yet.

4.3.4. Cross-domain challenges

Safety standards are characterized by a 5-yearly revision, allowing for an evaluation of the adequacy of the relevant standard(s). Concerning the 5-yearly period for revision, the respondents to the LIAISON surveys and participants in the LIAISON workshop at the ERF presented clearly divided opinions on whether this timeframe is too long. Out of a pool of 15 respondents, 40% indicated that this timeframe is too long, while 60% disagreed with that opinion. These results were complemented with arguments from the workshop participants, stating that whether the 5-yearly revision is too long depends on the domain to which the standard relates - is the domain settled or still in the early stages of development? Moreover, it was argued that in some disciplines, there are still too few experts active in ISO, making it impossible to shorten the revision time frame. In addition, it takes time to gain sufficient experience in a particular domain to assess the adequacy of standards properly; revision should not be based on 'single-case experiences.' Moreover, it was argued that standards are supposed to offer a reliable framework for safety. By updating standards more frequently, we might risk undermining the reliability and dependability of standards.

While each of the investigated standards is concerned with physical safety requirements, the legislative system includes many other fundamental rights to be protected - e.g., 1) health, safety, consumer, and environmental regulations; 2) liability; 3) IP; 4) privacy and data protection; 5) capacity to perform legal transactions (Leenes et al., 2014). Concerning

⁶ ISO 18497:2018 defines a *remote operator* as 'a human who is: 1) in primary control of a machine through the supervisory system;

2) receiving data for the purpose of supervising machine activity, and 3) is not on the machine but is located in the field, close to the field, or away from the field.

the adequacy of standards, the involved participants of the LIAISON workshop at the ERF believed that standards should shift from mono-impact to multi-impact, including factors related to ethics, environmental sustainability, liability, accountability, privacy, and data protection, and psychological aspects. As such, it is clear that robots, to be safe, need to comply with the safety requirements set by private standards and include other aspects highlighted within the law to ensure that the rights and protection of the user are not compromised.

The current cross-domain nature of robotics raises a dilemma for roboticists that many other users of the Machinery Directive and related harmonized standards do not encounter. This arises from the fact that the standards focusing on the safety of collaborative robotics are domain-specific, and it is not always clear to a roboticist which standards are applicable to their system. Currently, these standards covering different domains are not synchronized and can have conflicting requirements. This can lead to uncertainty, mainly when robots are used in new fields (such as agriculture) or for multiple domains (i.e., an exoskeleton used for medical purposes or to support workers in manufacturing) (Bessler et al., 2021).

A key cross-domain finding is that robots and AI technologies can impact humans beyond physical safety. Traditionally, the definition of safety has been interpreted to exclusively apply to risks that have a physical impact on persons' safety, such as, among others, mechanical or chemical risks. However, the current understanding is that integrating AI in cyber-physical systems such as robots, thus increasing interconnectivity with several devices and cloud services and influencing the growing human-robot interaction, challenges how safety is currently conceptualized rather narrowly (Martinetti et al., 2021). Thus, to address safety comprehensively, AI demands a broader understanding of safety, extending beyond physical interaction, but covering aspects such as cybersecurity (Fosch-Villaronga and Mahler, 2021) and mental health. Moreover, the expanding use of machine learning techniques will more frequently demand evolving safety mechanisms to safeguard the substantial modifications over time as robots embed more AI features. In this sense, the different dimensions of the concept of safety, including interaction (physical and social), psychosocial, cybersecurity, temporal, and societal, need to be considered for robot development (Martinetti et al., 2021). Revisiting these dimensions may help, on the one side, policy and

standard makers redefine the concept of safety in light of robots and AI's increasing capabilities, including human-robot interactions, cybersecurity, and machine learning; and, on the other hand, robot developers integrate more aspects into their designs to make these robots genuinely safe to use.

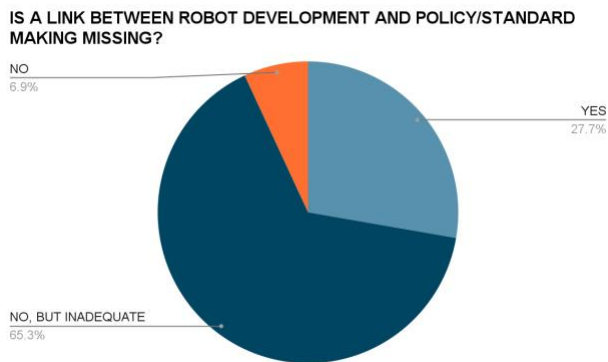
A final cross-domain challenge relates to the autonomy levels of robots. The levels of autonomy define the robot's progressive ability to perform particular functions independently. In other words, 'autonomy' refers to a robot's "ability to execute specific tasks based on current state and sensing without human intervention." For the automotive industry, the Society of Automotive Engineers (2020) established automation levels to clarify the progressive development of automotive technology that would, at some point, remove the human from the driving equation. However, no universal standards have been defined for progressive autonomy levels for personal care, rehabilitation, or agricultural robots.

Yang et al. (2017) proposed a generic six-layered model for medical robots' autonomy levels depicting a spectrum ranging from no autonomy (level 0) to full autonomy (level 5) to bridge this gap in the medical field. The effort is a significant step towards bringing more clarity to the field, especially with respect to the roles and responsibilities resulting from increased robot autonomy. Still, the model needs more detailing on how it applies to specific types of medical robots. Robots' embodiment and capabilities differ vastly across surgical, physically/socially assistive, or agricultural contexts, and the involved human-robot interaction is also very distinctive (Fosch-Villaronga et al., 2021).

4.4 A missing link between robot governance and development

The results gained from the survey on the usefulness of LIAISON and the interactive sessions at the ERF are very revealing. While 28% of the 27 respondents believe that currently there is no link between robot development and policy/standard making, 66% believe that such a link does exist but that this link is either too complex and lacks openness, merely far too complex or only exists between robot development and policy/standard making (Fosch-Villaronga and Drukarch, 2021c). This, while a small 7% of respondents believe that such a link already exists between robot development and policy/standard making.

Figure 3. Is a link between robot development and policy/standard making missing?



More specifically, in response to the question of whether a link between robot development and policymaking is currently missing, a range of responses was provided, namely 1) Yes, currently there is no such a link between robot development and policy/standard making (28%); 2) No, there is already such a link between robot development and policy/standard making, but it is too complex (38%); 3) No, there is already such a link between robot development and policy/standard making, but it is too complex and lacks openness (21%); 4) No, but the link is only between developers and standard organizations (7%); and 5) No, there is already such a link between robot development and policy/standard making (7%) (Fosch-Villaronga and Drukarch, 2021c).

4.5 The need for interdisciplinary cooperation

The results gained from the survey on the usefulness of LIAISON and the interactive sessions at the ERF indicate that for the regulatory approach proposed through LIAISON to be valuable and effective, a diverse group of stakeholders should be involved (Fosch-Villaronga and Drukarch, 2021c). These stakeholders include robot developers, manufacturers, policymakers, standardization organizations, legal scholars, and ethicists. The involved Digital Innovation Hubs also stress this need for diverse stakeholder involvement. For instance, engagement with DIH AgROBOfood has presented the need for robot developers to pay attention to ethical, legal, and many other issues to determine if a robot will survive in a practical setting.

In addition, the exploratory meeting with standard makers clarified the value of LIAISON in liaising standardization

activities and robot development. More specifically, during one of the policy and standard-making institutional meetings, a representative of ISO Technical Committee TC299 (Robotics) Working Group 02 on Service Robot Safety standardization stressed establishing a relationship of cooperation between ISO/TC299/WG2 and LIAISON could be very useful and valuable. On the one hand, ISO/TC299/WG2 could provide LIAISON with the necessary input from standard making. On the other hand, looking at its goal, LIAISON could offer WG2 the relevant knowledge on inconsistencies and gaps in ISO 13482:2014 from the perspective of robot developers (Fosch-Villaronga and Drukarch, 2021d).

Moreover, the results obtained through the survey on the usefulness of LIAISON and the interactive sessions at the ERF also indicate that there is a serious need for cooperation between 1) central policymakers and standardization institutes; 2) major standardization institutes (ISO, BSI, CENELEC); and 3) user group initiatives and policy/standard makers.

4.6 Lack of legal comprehension

The results gained from the survey on the usefulness of LIAISON and the interactive sessions at the ERF indicate an overall lack in legal comprehension among robot developers, thereby adding emphasis to the first finding of a clear missing link between robot developers and policymakers. More specifically, the obtained data highlights this on three points: 1) experience with standards, 2) knowledge about the difference between public and private policymaking, and 3) experience applying standards. 23% of the 33 respondents indicated to have never used a standard before, against 77% who suggested having experience with standards (Fosch-Villaronga and Drukarch, 2021c). At the same time, while all respondents - based on a pool of 15 respondents - indicated being aware of and understanding the difference between standards and the law, approximately 33% showed to be still confused regarding this difference.

Interestingly, with regard to respondents' experience with standards - based on a pool of 26 respondents -, we were presented with a variety of responses, namely: 1) Run multiple standards for my devices (70%); 2) My robot does not fit into the standard category (45%); 3) Do not know if my robot is a medical device (40%), and 4) Do not know the difference between standard and regulation (10%).

It was indicated that smaller and younger companies often lack the necessary knowledge and understanding concerning the applicable legal frameworks and the difference between private and public policymaking within this context. Moreover, various meetings with the Digital Innovation Hubs DIH-HERO and DIH AgROBOfood have indicated this confusion among their community and the lack of legal comprehension. For this reason, these Digital Innovation Hubs stressed the value that LIAISON could also offer and the valuable insights that their community could provide policymakers in this respect. This has led to the organization of domain-specific webinars at a later stage in the LIAISON Research Project.

5 Conclusions

Based on the belief that the regulatory cycle is truly closed when it starts again upon the identification of new challenges, LIAISON puts the theoretical model of a dynamic, iterative regulatory process into practice, aiming to channel robot policy development from a bottom-up perspective towards a hybrid top-down/bottom-up model, leaving the door open for future modifications. As such, LIAISON aims to clarify what regulatory actions policymakers have to take to provide compliance guidance, explain unclear concepts or uncertain applicability domains to improve legal certainty and inform future regulatory developments for robot technology use and development at the European, National, Regional, or Municipal level. Moreover, by explicitly shedding light on the standardization activities in the abovementioned domains, LIAISON has created awareness about the barrier to access for robot developers and other relevant stakeholders concerning such activities.

Overall, LIAISON has proven to be a valuable tool to facilitate effective robot governance, as indicated by relevant stakeholders, because of its all-encompassing nature. Possible avenues for expansion relate to active involvement in standardization organizations, focus on harmonization activities, and legal and educational participation in Digital Innovation Hubs to create more legal awareness among the involved communities of robot developers. A platform based on the model proposed by LIAISON could be thus beneficial in improving the communication between all stakeholders involved in the development and regulation (be it through public or private bodies) of new technologies.

The results presented in this paper highlight the importance of and need for active stakeholder involvement in robot governance. However, currently, the link between robot development and policymaking is complex, and it lacks openness, transparency, and free access. Access to standardization activities is not always accessible due to high costs, and there is a lack of clarity concerning public policymaking activities and their relation to private standard making. This requires a reconsideration of how policy/standard makers engage with stakeholders in the normative process.

Moreover, the above results have indicated the need to seek active participation of affected parties (e.g., NGOs, user group initiatives - e.g., patients organizations and consumer networks -, and other interested groups). These parties should not only be involved at the end of the development and policy/standard-making chain. They form an integral part of the process and should be engaged in these activities from an early stage to provide input and feedback that will consider the needs and concerns of the wider public.

In the long-term, the expected project results may complement the existing knowledge on the 'ethical, legal, and societal (ELS)' of robotics by providing clarity on how to address pressing but still uncovered safety challenges raised by robots and represent a practical, valuable tool to advance social goals in a robotized workplace. Overall, advances in safety robot legal oversight will provide a solid basis for designing safer robots, safeguarding users' rights, and improving the overall safety and quality of efficiency delivered by robots.

Acknowledgements

This paper is part of the H2020 COVR project that has received funding from the European Union's Horizon 2020 research and innovation programme (grant agreement No. 779966).

References

- A Bill To authorize an emerging technology policy lab within the General Services Administration, and for other purposes, S. 3502, 115th Cong. (2018). <https://www.congress.gov/115/bills/s3502/BILLS-115s3502is.pdf>
- Adelle, C., Jordan, A., and Turnpenny, J. (2012). Proceeding in parallel or drifting apart? A systematic review of policy appraisal research and practices. *Environment and Planning C: Government and Policy*, 30(3), 401-415. <https://doi.org/10.1068/c11104>.

- Bessler, J., Prange-Lasonder, G. B., Schaake, L., Saenz, J. F., Bidard, C., Fassi, I., ... and Buurke, J. H. (2021). Safety assessment of rehabilitation robots: A review identifying safety skills and current knowledge gaps. *Frontiers in Robotics and AI*, 8, 33. <https://doi.org/10.3389/frobt.2021.602878>.
- Brundage, M., and Bryson, J. (2016). Smart policies for artificial intelligence. *arXiv preprint arXiv:1608.08196*.
- Collingridge D. (1980) *The Social Control of Technology*. St. Martin's Press, New York.
- European Commission with Valdani Vicari and Associati, Deloitte, The Vienna Institute for International Economic Studies and Ecorys. (2020). Impact assessment study on the revision of Directive 2006/42/EC on machinery.
- European Parliament. (2020). Disruption by technology. Impacts on politics, economics and society. In-Depth Analysis by the EPRS | European Parliamentary Research Service.
- European Parliament resolution of 16 February 2017 with recommendations to the Commission on Civil Law Rules on Robotics (2015/2103(INL)).
- Felzmann, H. (2018). How transparent is your AI? Ethical, Legal and Societal Issues of the Transparency Principle in Cyber-physical Systems. Workshop on AI: Ethical and legal Implications, Center for Cyber Law & Policy, University of Haifa and the European Hub of the Global Network of Internet and Society Research Centers (NoC), 28-30 November 2018.
- Fenwick, M., Kaal, W. A., and Vermeulen, E. P. (2016). Regulation tomorrow: what happens when technology is faster than the law. *Am. U. Bus. L. Rev.*, 6, 561.
- Fosch-Villaronga, E. and Heldeweg, M. A., (2018) "Regulation, I Presume?" Said the Robot. Towards an Iterative Regulatory Process for Robot Governance. *Computer Law and Security Review*, 34(6), 1258-1277.
- Fosch-Villaronga, E., and Heldeweg, M. A. (2019) "Meet me halfway," said the robot to the regulation. Linking ex-ante technology impact assessments to legislative ex-post evaluations via shared data repositories for robot governance. In: Pons-Rovira, J.L. (2019) *Inclusive Robotics for a Better Society*. INBOTS 2018. *Biosystems & Biorobotics*, vol. 25., Springer, Cham., 113-119, https://doi.org/10.1007/978-3-030-24074-5_21.
- Fosch-Villaronga, E. (2019). *Robots, Healthcare and the Law: Regulating Automation in Personal Care*. Routledge.
- Fosch-Villaronga, E. and Drukarch, H. (2020). Introducing the future of robot regulation. *Medium*. Retrieved from <https://eduard-fosch-villaronga.medium.com/introducing-the-future-of-robot-regulation-bf4fe6361a79>.
- Fosch-Villaronga, E., and Drukarch, H. (2021a). On Healthcare Robots: Concepts, definitions, and considerations for healthcare robot governance. *arXiv preprint arXiv:2106.03468*.
- Fosch-Villaronga, E., Khanna, P., Drukarch, H., and Custers, B. H. (2021). A human in the loop in surgery automation. *Nature Machine Intelligence*, 3(5), 368-369.
- Fosch-Villaronga, E., and Mahler, T. (2021). Cybersecurity, safety and robots: Strengthening the link between cybersecurity and safety in the context of care robots. *Computer Law & Security Review*, 41, 105528.
- Fosch Villaronga E. and Drukarch H. (2021b), H2020 COVR FSTP LIAISON – D1.1 Report on safety standards' uncovered challenges LIAISON no. MS1. Leiden: eLaw / Leiden University.
- Fosch Villaronga E. and Drukarch H. (2021c), H2020 COVR FSTP LIAISON – D1.2 Report on usefulness of LIAISON LIAISON no. MS1. Leiden: eLaw / Leiden University.
- Fosch Villaronga E. and Drukarch H. (2021d), H2020 COVR FSTP LIAISON – D1.3 Report on exploratory meetings with relevant policy/standard makers LIAISON no. MS1. Leiden: eLaw / Leiden University.
- George, C., Whitehouse, D., and Duquenoy, P. (2013). Assessing legal, ethical and governance challenges in eHealth. In George, C., Whitehouse, D., and Duquenoy, P. (Eds.). (2012). *eHealth: legal, ethical and governance challenges*. Springer Science & Business Media, 3-22.
- Gruber, K. (2019). Is the future of medical diagnosis in computer algorithms?. *The Lancet Digital Health*, 1(1), e15-e16.
- De Hert, P. (2005). *Biometrics: legal issues and implications*. Background Paper for the Institute of Prospective Technological Studies, DG JRC-Sevilla, European Commission.
- Holder, C., Khurana, V., Harrison, F., and Jacobs, L. (2016). Robotics and law: Key legal and regulatory implications of the robotics age (Part I of II). *Computer law & security review*, 32(3), 383-402.
- Johnston, S. F. (2018). The technological fix as social cure-all: origins and implications. *IEEE Technology and Society Magazine*, 37(1), 47-54.
- Leenes, R., Palmerini, E., Koops, B., Bertolini, A., Salvini, P. and Lucivero, F. (2017). Regulatory challenges of robotics: some guidelines for addressing legal and ethical issues. *Law, Innovation and Technology*, 9:1, 1-44.
- Leroux, P., and Roy, S. (2012). Designing and optimizing swarming in a distributed base station network: Application to power control. *ACM Transactions on Autonomous and Adaptive Systems* (TAAS), 7(2), 1-19.
- Marchant, G.E. and Wallach, W. (2015). Coordinating technology governance. *Issues in Science and Technology*, 31(4), p. 43.
- Marchant, G.E., Allenby, B.R. and Herkert, J.R. eds., (2011). *The growing gap between emerging technologies and legal-ethical oversight: The pacing problem* (Vol. 7). Springer Science & Business Media.
- Martinetti, A., Chemweno, P., Nizamis, K., and Fosch-Villaronga, E. (2021) *Redefining safety in light of human-robot interaction: a*

critical review of current standards and regulations. *Frontiers in Chemical Engineering*, forthcoming.

Newlands, G., Lutz, C., Tamò-Larrieux, A., Fosch-Villaronga, E., Scheitlin, G., and Harasgama, R. (2020). Innovation under Pressure: Implications for Data Privacy during the Covid-19 Pandemic. *Big Data & Society*, SAGE, 7(2), 1-14.

Sabel, C., Herrigel, G., and Kristensen, P. H. (2018). Regulation under uncertainty: The coevolution of industry and regulation. *Regulation & Governance*, 12(3), 371-394. <https://doi.org/10.1111/rego.12146>.

SAE (2020) J3016B: Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles - SAE International. (n.d.). Retrieved from https://www.sae.org/standards/content/j3016_201806/.

Saenz, J., Behrens, R., Schulenburg, E., Petersen, H., Gibaru, O., Neto, P., et al. (2020). Methods for considering safety in design of robotics applications featuring human-robot collaboration. *Int. J. Adv. Manuf. Technol.* 107, 2313–2331. <https://doi.org/10.1007/s00170-020-05076-5>.

Sucha, V., and Sienkiewicz, M. (2020). *Science for Policy Handbook*. European Commission, Joint Research Centre (JRC), Elsevier.

Spiliopoulou-Kaparia M. (2017). The evaluation of Directive 85/374/EEC on liability for defective products and Directive 2006/42/EC on machinery. *Proceedings of the European Stakeholder Forum – Workshop on Regulatory challenges for a digitizing industry*. Essen.

Surdilovic, D., et al. (2011). Human-like variable-impedance control for life-cycle testing. *RO-MAN 2011*. IEEE, 150-155.

Van Rompaey, L., Jønsson, R., Jørgesen, K.E. (2021) Designing lawful machine behaviour. *Robotists' legal concerns*. [Conference presentation]. *EURA Conference 2021 - Regulating UncertAInty*. <https://www.youtube.com/watch?v=GjCMRu86IRU&t=9461s>.

Vayena, E., and Gasser, U. (2016). Between openness and privacy in genomics. *PLoS medicine*, 13(1), e1001937.

Winfield, A. (2019). Ethical standards in robotics and AI. *Nature Electronics*, 2(2), 46-48.

Yang, G.Z., Bellingham, J., Dupont, P.E., Fischer, P., Floridi, L., Full, R., Jacobstein, N., Kumar, V., McNutt, M., Merrifield, R. and Nelson, B.J., (2018). The grand challenges of Science Robotics. *Science Robotics*, 3(14), p.eaar7650.

Yang, G. Z., Cambias, J., Cleary, K., Daimler, E., Drake, J., Dupont, P. E., ... and Taylor, R. H. (2017). Medical robotics—Regulatory, ethical, and legal considerations for increasing levels of autonomy. *Science Robotics*, 2(4), 8638.



Redefining Safety in Light of Human-Robot Interaction: A Critical Review of Current Standards and Regulations

Alberto Martinetti^{1*†}, Peter K. Chemweno^{1†}, Kostas Nizamis^{1†} and Eduard Fosch-Villaronga^{2†}

¹Department of Design, Production and Management, University of Twente, Enschede, Netherlands, ²eLaw Center for Law and Digital Technologies, Leiden University, Leiden, Netherlands

OPEN ACCESS

Edited by:

Maria Chiara Leva,
Technological University Dublin,
Ireland

Reviewed by:

Rajagopalan Srinivasan,
Indian Institute of Technology Madras,
India
Brenno Menezes,
Hamad bin Khalifa University, Qatar
Helen Durand,
Wayne State University, United States

*Correspondence:

Alberto Martinetti
a.martinetti@utwente.nl

[†]These authors have contributed
equally to this work and share first
authorship

Specialty section:

This article was submitted to
Computational Methods in Chemical
Engineering,
a section of the journal
Frontiers in Chemical Engineering

Received: 09 February 2021

Accepted: 14 July 2021

Published: 27 July 2021

Citation:

Martinetti A, Chemweno PK, Nizamis K
and Fosch-Villaronga E (2021)
Redefining Safety in Light of Human-
Robot Interaction: A Critical Review of
Current Standards and Regulations.
Front. Chem. Eng. 3:666237.
doi: 10.3389/fceng.2021.666237

Policymakers need to consider the impacts that robots and artificial intelligence (AI) technologies have on humans beyond physical safety. Traditionally, the definition of safety has been interpreted to exclusively apply to risks that have a physical impact on persons' safety, such as, among others, mechanical or chemical risks. However, the current understanding is that the integration of AI in cyber-physical systems such as robots, thus increasing interconnectivity with several devices and cloud services, and influencing the growing human-robot interaction challenges how safety is currently conceptualised rather narrowly. Thus, to address safety comprehensively, AI demands a broader understanding of safety, extending beyond physical interaction, but covering aspects such as cybersecurity, and mental health. Moreover, the expanding use of machine learning techniques will more frequently demand evolving safety mechanisms to safeguard the substantial modifications taking place over time as robots embed more AI features. In this sense, our contribution brings forward the different dimensions of the concept of safety, including interaction (physical and social), psychosocial, cybersecurity, temporal, and societal. These dimensions aim to help policy and standard makers redefine the concept of safety in light of robots and AI's increasing capabilities, including human-robot interactions, cybersecurity, and machine learning.

Keywords: robots, cobots, artificial intelligence, safety, machinery directive, psychosocial aspects at work

INTRODUCTION

The robotic industry is developing rapidly, affecting different aspects of modern working life. Collaborative robots, so-called cobots, can among other things, support human workers in a shared workspace, nurses working with lifting robots, doctors with intelligent diagnostic systems, and information system designers in public administration. The rate at which these developments occur is faster than ever before (European Agency for Safety and Health at Work, 2020). As pointed out by the Organisation for Economic Cooperation and Development (OECD, 2019), while about 14% of jobs are highly automatable in OECD countries, another 32% of jobs are likely to change radically as individual tasks keep getting automated within these jobs.

While robots help staff extend the professional service they provide, create new opportunities, entail resource efficiency, and increase productivity, it is unclear how such professions adhere and

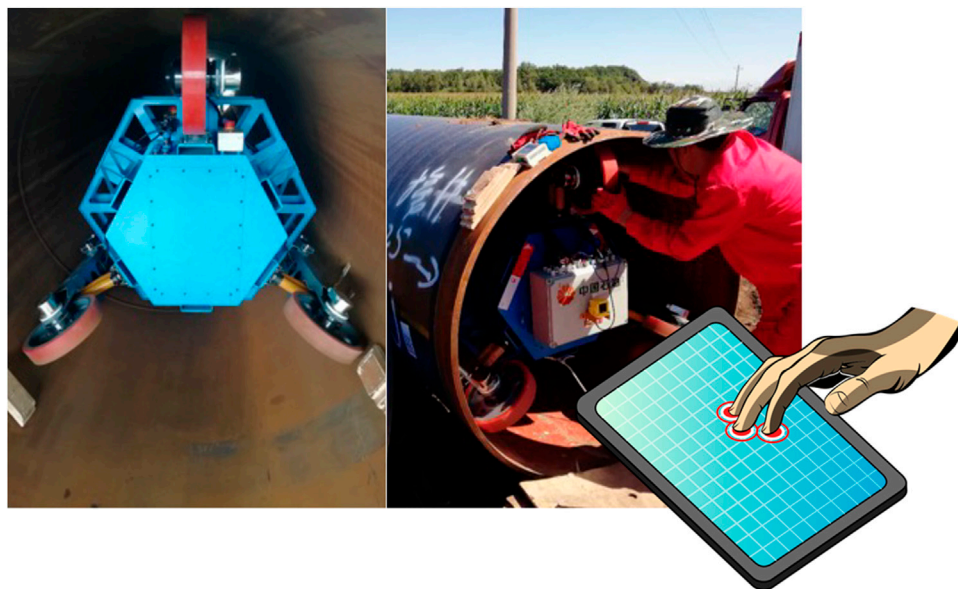


FIGURE 1 | Example of a robot used for petroleum pipelines inspection (modified from Li et al., 2020).

adapt to this new reality. Collaborative robots support different types of interaction, including physical and social, and may evoke social responses from workers or involve psychosocial elements like trust (Di Dio et al., 2020). For the physical elements, robot and AI deployments may increase the risk of collision for the cobots' equipment, negatively impacting workers' safety and short-term health. For the other elements, particularly mental health, which is often neglected and largely underestimated, the human-robot interactions may be sporadic or geared toward supporting long-term engagement over time, often involving emotion and memory adaptations that designers manipulate to combat user interest decline (Ahmad, Mubin and Orlando, 2017). The literature alerts that, given our human tendency to form bonds with the entities with whom we interact and the human-like capabilities of these devices, users may have strong connections with robots that may include dependency, deception, and overtrust (Robinette et al., 2016; Wagner, Borenstein et al., 2018).

In the chemical industry, robots are widely applied for inspection in confined spaces. Some robotic solutions are advancing industrial inspection capabilities with autonomous legged robots, bringing complete visibility and higher-quality data collection to chemical processing plants (Anybotics, 2021). Operators may use inspection robots within confined dangerous spaces to inspect defects in pipelines for inspection robots. To do this, operators may use multiple interaction interfaces, including a "screen" interface, and involving ergonomic constraints (while manoeuvring the robotic agent). Furthermore, the operation may involve high cognitive loads when manipulating the robotic agent to prevent hazards. **Figure 1** illustrates an example of operator-control related hazards, which includes overheating of hardware parts of the robot, due to suboptimal control of speed and orientation of the robot by

the operator. This is usually via interacting with the monitoring pendant (an example shown in **Figure 1**) while controlling the agent to navigate the confined pipeline, which may contain hazardous substances (Li et al., 2020).

As one can imagine, the nature of interactions, and their inherent safety, change with the introduction of these developments. In this respect, regulatory frameworks typically focused on ensuring physical safety by separating the robot from the human operator. However, industrial environments increasingly incorporate robots that interact directly with humans and it is unsure how safety should be addressed in such cases. In this sense, the definition of safety has been traditionally interpreted to exclusively apply to risks that have a physical impact on persons' safety, such as, among others, collision risks. However, the increasing use of service and collaborative robots in shared workspaces that interact with users socially (also known as social robots) challenge the way safety has been addressed (Fosch-Villaronga and Virk, 2016). Moreover, the integration of AI in cyber-physical systems such as robots, the increasing interconnectivity with other devices and cloud services, and the growing human-machine interaction challenges the safety concept's narrowness.

The recent advances in AI demand a broader understanding of safety, covering cybersecurity, and mental health to address safety comprehensively. Moreover, the expanding use of machine learning techniques will more frequently demand evolving safety mechanisms to safeguard the substantial modifications taking place over time. In this sense, this paper puts forward some recommendations to shed light on multiple dimensions of safety in light of AI's increasing capabilities, including human-machine interactions, cybersecurity, and machine learning, to truly insure safety in human-robot interactions.

MAPPING DIFFERENT PERSPECTIVES ON SAFETY IN THE INDUSTRIAL CONTEXT

Definitions: Machinery, Robots and AI

According to the Council Directive 2006/42/EC a machine is defined as an assembly:

- fitted with or intended to be fitted with a drive system other than directly applied human or animal effort, consisting of linked parts or components, at least one of which moves, and which are joined together for a specific application;
- ready to be installed and able to function as it stands only if mounted on a means of transport, or installed in a building or a structure,
- partly completed machinery which, to achieve the same end, are arranged and controlled so that they function as an integral whole;
- of linked parts or components, at least one of which moves and which are joined together, intended for lifting loads and whose only power source is directly applied human effort;

There is no clear and explicit reference to robot/cobot equipment or AI agents within this definition. To have a specific understanding of what a robot is, it is necessary to adopt the definition offered by the ISO 8373:2012 that describes a robot as an automatically controlled, reprogrammable, multipurpose manipulator programmable in three or more axes, which can be either fixed in place or mobile for use in industrial automation applications. To bridge the gap between public policymaking and private setting, other scholars have defined robots as a movable machine that performs tasks either automatically or with a degree of autonomy (Fosch-Villaronga and Millard, 2019).

Having a consensus about the definition of AI is even more challenging. The European Commission has defined it as “systems that display intelligent behaviour by analyzing their environment and taking actions—with some degree of autonomy—to achieve specific goals. AI-based systems can be purely software-based, acting in the virtual world (e.g., voice assistants, image analysis software, search engines, speech and face recognition systems), or AI can be embedded in hardware devices (e.g., advanced robots, autonomous cars, drones, or Internet of Things applications) (HLEG on AI, 2019). Within this complex definition, what is clear is that AI capabilities such as machine learning (Salay et al., 2017) or human-machine interaction have far-reaching implications that challenge many concepts and areas, including standards such as the ISO 26262:2018 on functional safety for road vehicles, and the European Directive for product safety 2001/95/EC.

There are movements to address and define this field on both sides. An example is the establishment of the Sub-Group on AI, connected products, and other new challenges in product safety to the Consumer Safety Network (CSN) at the European Commission; and the creation of an ad hoc working group (SC 42) at the ISO level. In this sense, and once again (Fosch-

Villaronga and Golia, 2018), regulation and standardization approaches are working separately while it is increasingly clear that a joint and more comprehensive approach to safety is deemed necessary (see Figure 2).

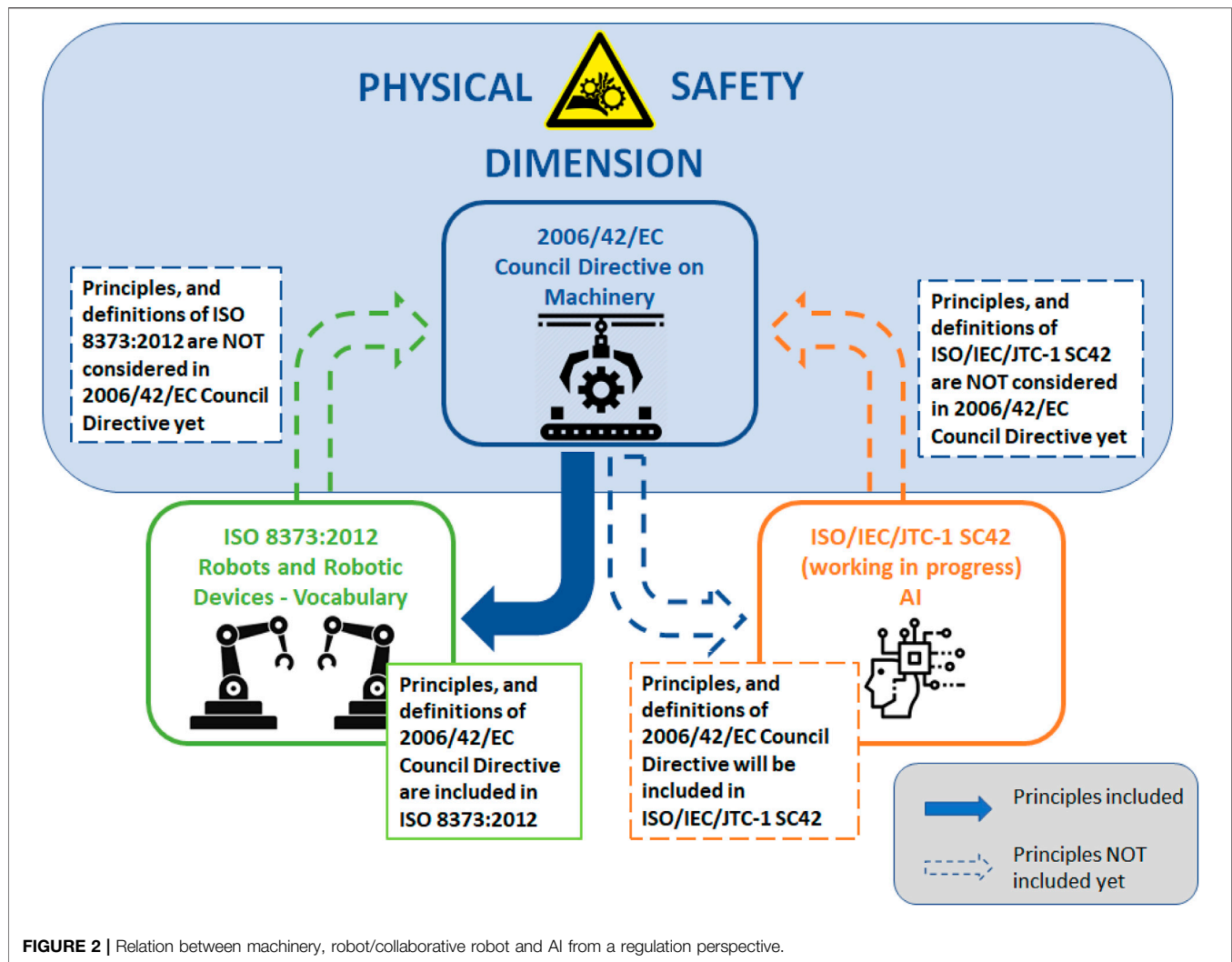
For now, in May 2021, the European Institutions released a “Proposal for a Regulation of the European Parliament and of the Council on machinery products” establishing a regulatory framework for placing machinery on the Single Market (Machinery Proposed Regulation, 2021). The regulation is designedly technology-neutral, laying down essential health and safety requirements to be complied with, without designating any specific technical solution to comply with those provisions. Although very new and not yet binding, it indicates that the concept of safety is not yet clear, as now it switched to a safety component.

The Physicality of the Concept of Safety

Even though safety has an overarching common meaning, laws, regulations and standards have adopted different definitions for different industrial sectors, consequently generating confusion, lack of clarity, and a lack of homogeneity in the understanding of what safety means. Take for instance the definition of *safety* enshrined in the current Directive 2001/95/EC on General Product Safety defines safety on its Art. 2. b): “safe product” shall mean any product which, under normal or reasonably foreseeable conditions of use including duration and, where applicable, putting into service, installation and maintenance requirements, does not present any risk or only the minimum risks compatible with the product’s use, considered to be acceptable and consistent with a high level of protection for the safety and health of persons. The article also establishes different features to consider part of this safety, including the characteristics of the products, the effect on other products, the presentation of the product, and also the different categories of consumers, in particular children and elderly.

This legal definition of safe products is quite broad, as said, and it can be understood as covering all kinds of risks that can, directly or indirectly, cause harm to consumers. However, the concept of safety is also present in other things that are not products directly used by the public (i.e., consumers) but that, still, is used by other humans (i.e., operators). In this sense, to have a whole picture of the evolution in terms of safety regulation in the European market, it is necessary to look at the end of the 1980s where different safety regulations framing safety at the workplace emerged.

The first game-changing pieces of legislation to promote safety in an industrial context were the Council Directive 89/391/EEC of June 12, 1989 on the introduction of measures to encourage improvements in the safety and health of workers at work (OSH Directive) and the Council Directive 89/392/EEC of June 14, 1989 on the approximation of the laws of the Member States relating to machinery. These pieces of legislation were issued simultaneously not accidentally, as both underlined how safety in an industrial setup could be achieved by providing safe workplaces (in terms of work organization and training) and safe equipment (referring to



machinery and tools). Moreover, this framework aimed to create the first European common ground and harmonized regulation to level off the differences between member countries and facilitate the machinery' circulation.

The general principles stated in the Council Directive 89/391/EEC are generally still applicable to foster workplace safety. Still, the definition of safety is not directly defined in the corpus (see Art. 3 Directive 89/391/EEC). The Council Directive 89/392/EEC went through several revisions to include the technology's progress in terms of machinery and safety systems. The in-force Council Directive on machinery is the 2006/42/EC to the date of writing. Despite all the revisions, the concept of safety is neither spelt out in this corpus, only being addressed what is a "safety component" which refers to any component 1) which serves to fulfil a safety function, 2) which is independently placed on the market, 3) the failure and/or malfunction of which endangers the safety of persons, and 4) which is not necessary for the machinery to function, or for which normal components may be substituted for the machinery to function.

However, recent innovations in Artificial Intelligence (AI) and cyber systems, including robots/cobots raise several concerns including whether these robots pose additional risks for workers, lead to unequal treatment, or even commercial exploitation (Vrousalis, 2013; Kapeller et al., 2020). These advancements challenge the Council Directive 2006 42/EC's validity and applicability to these new fields, which led the institutions to go through a thorough revision of this instrument. One of the first aspects that changed over the years is the concept of machinery, as it is currently unclear whether this definition encompasses all the new cyber-physical equipment such as robot/collaborative robot and AI agents. The fact that current standards do not refer to machinery but to robots and robotic devices and the other way round, that enacted/proposed legislation focuses on machinery, confuses this panorama further.

Another aspect was the concept of safety, which had been traditionally interpreted as physical. Although great advancements in this area, still, in February 2021, the regulatory scrutiny board opinion on the "Proposal for a

Regulation of the European Parliament and the Council on machinery products¹ had some reservations concerning the evidence on the scope and magnitude of the problems regarding safety requirements or how the instrument is going to be future-proof given the evolving safety risk that AI entails.

Given the gaps in current legislation, private safety standards developed from international standardization organizations such as ISO/IEC covered different aspects connected to safety, including risk, harm, and hazard:

- Safety: freedom from risk which is not tolerable (ISO/IEC Guide 51:2014, 3.14)
- Risk: combination of the probability of occurrence of harm and the severity of that harm. Note 1 to entry: The probability of occurrence includes the exposure to a hazardous situation, the occurrence of a hazardous event and the possibility to avoid or limit the harm (ISO/IEC Guide 51:2014, 3.9)
- Harm: injury or damage to the health of people, or damage to property or the environment (ISO/IEC Guide 51:2014, 3.1)
- Hazard: Potential source of harm (ISO 13482:2014).

We notice that the definition of harm has been updated over the years. Namely the adjective physical before the injury has been deleted in recent updates. The reasons are that the word injury already means physical damage and that the responsible ISO/IEC Joint Working Group wanted this definition to have a broader interpretation including unreasonable psychological stress.

Psychosocial and Ethical Considerations for the Design of Robot Systems

It can be even claimed that traditionally the definition has been interpreted to exclusively apply to risks that have a physical impact on the safety of persons, such as among others mechanical or chemical risks (Consumer Safety Network, 2020). As previously mentioned, the broad definition of safety has traditionally focused on physical interaction as the focal point of hazard analysis and risk assessment. Only more recently are psychosocial influences on the safety of, especially, collaborative robots gaining more prominence (Chemweno et al., 2020). These include psychosocial influences linked to factors within a collaborative work environment such as high-job demand or nervousness working closely with a robotic agent that, in turn, triggers stressors such as fatigue, fear, or cognitive inattention (Riley, 2015). Often, these stressors likely create pathways that would lead to unsafe interaction with a robotic agent. Owing to the more recent recognition of psychosocial influences on safe work collaboration with robots, researchers have attempted to integrate implicit design features that mitigate stressors such as fear/nervousness when sharing a workspace with a robot (Chemweno et al., 2020). These features are more prominent for designs of social robots, by co-opting humanoid-like features

to create confidence while interacting with the robot (Robinson, et al., 2019).

Additional psychosocial factors that influence safety when interacting with robots include the ergonomic design of the workspace and task design to consider the operator's cognitive load (Chemweno et al., 2020). For instance, more recent research suggests that performing repetitive tasks and those requiring high cognitive capabilities such as complex assembly of a mechanical device often impedes the operator's attention, creating a hazard pathway in shared collaborative workspaces (Bao et al., 2016). Other than interaction safety, psychosocial influences extend to long-term health effects, including depression, burn-outs, and reduced productivity for operators. Therefore, the relatively narrow view on physical human-robot interaction safety is a constraint for developing more robust design safeguards or safety assessment protocols. Moreover, this limitation is seen as discussed in this paper, where largely safety standards including the ISO 10218 (industrial robots), and the more recent ISO 15066 (collaborative robots) focus on techno-centric design safeguards. Primary, these latter safeguards embed hardware and software measures to mitigating collision risks. For instance, "soft" features such as pneumatic muscles limit injury in the event of a collision with a robotic arm (Van Damme et al., 2009).

Substantial techno-centric safeguards for collaborative robots are predominantly addressed by the ISO 15066, which proposes guidelines for designing safeguards, focusing on among other aspects, designing the collaborative workspace considering aspects such as access and a safe separation distance between the operator and robot. The separation here is achieved through passive or active sensing capabilities to prevent hazardous interaction, e.g., by including a laser "eye" to sense a breach of a safe separation distance (Michalos et al., 2015). Additional guidelines for design safeguards mentioned in the ISO 15066 norm include embedding a safety-rated monitored speed feature to limit the robot manipulator's speed as it approaches the operator. Power and force limiting functions are also suggested to prevent hazardous contact and injury of a worker or operator. However, as discussed, the guidelines orient to techno-centric safeguards, thereby largely focusing on mitigating hazardous collisions. Seldomly, the design safeguards suggested in the current standards for robot safety, integrate psychosocial (and behavioural) factors of operators working with robots.

Nonetheless, AI presents quite interesting opportunities to integrate such psychosocial, behavioural factors in the design of robots. Though still largely at a research phase, designing "intelligent" robot systems is gaining traction. For instance, robots embedded with an "anticipatory safety reasoning system" to sense their environment and react to unsafe situations. This reasoning may involve the robot anticipating sudden, dangerous operator movement, by sensing behavioural changes and adjusting parameters such as speed of the robotic arm, or completely stopping (Frederick et al., 2019). Although quite a significant body of literature on anticipatory reasoning focuses on autonomous robots, considering such psychosocial human factors is a positive step, especially for collaborative robots

¹See <https://ec.europa.eu/transparency/regdoc/rep/2/2021/EN/SEC-2021-165-1-EN-MAIN-PART-1.PDF>.

anticipated shortly, to increasingly share the workspace with an operator. This is driven by industry needs to increase productivity by automating highly cognitive tasks. We should draw attention though, that integrating human behavioural factors requires a multidisciplinary effort involving engineering, industrial design, AI engineers, behavioural science, psychology, legal scholars, and safety experts, highlighting the scope of effort required to realize safe robotic systems fully.

However, the lack of established standards and norms to fully optimize multidisciplinary effort remains a critical bottleneck to overcome. More specifically, this relates to the absence of a common set of norms or guidelines robot designers can use to develop robust psychosocial safeguards and respective protocols for verifying the sufficiency of such safeguards as is commonly the case for other safety-critical systems and machinery (Fosch-Villaronga, 2019), e.g., for space exploration, aviation, and nuclear industry. However, there is a considerable body of literature recognising the importance of orienting psychosocial factors for new technologies, including implementing robotisation in workplaces. Although, achieving this orientation is recognised as a considerable challenge owing to demands in the workplace, including increased efficiency, productivity, quality and safety, suggestions that could lead to successful integration are proposed. As an example, Jones (2017) suggests an approach utilising a cooperating team-oriented approach, to implementing new technologies which integrate psychosocial factors while designing new workspaces. Their approach relies heavily on obtaining employee input throughout the implementation process for new technologies, including robots.

Mokarami et al. (2019) propose a structured integration approach, utilising questionnaires to measure psychosocial factors in the workplace, e.g., work-related stress. Based on measurements obtained using the Copenhagen Psychosocial Questionnaire (COPSOQ), the individual's work-stressors are classified on a severity scale. The COPSOQ importantly considers technology factors, including interaction with robots. Predominantly though, psychosocial assessment and integration to technologies still largely concentrate on social robots, for example, where "empathetic features" are embedded on social robots to increase their acceptability, e.g., by patients with dementia (Rouaix et al., 2017). It is worth mentioning that psychosocial measurement is an essential starting basis for developing strategies for integrating human behavioural factors while designing robotic systems to enhance their safety. For instance, we see effort reported in studies towards utilising a measurement of pupillary response as the basis of quantifying cognitive loads for the operator (Strathearn and Ma, 2018; Borghini et al., 2020).

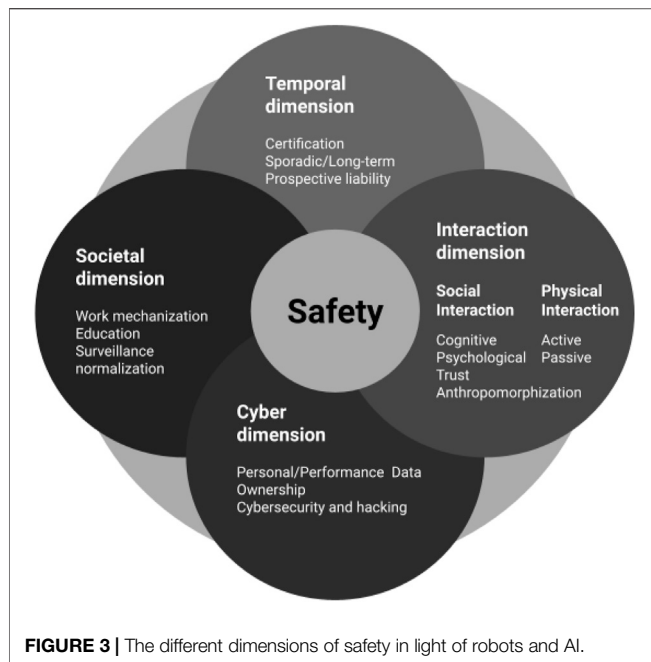
Although psychosocial assessment and measurement of associated human factors are an essential step towards embedding psychosocial design features to enhance safe robot interaction, measurement of such factors raises ethical questions. One may argue that utilising measurement data for designing robotic technologies may be questionable, especially in the absence of ethical standards for guiding psychosocial measurement,

assessment, and use of behavioural measurement data. Bryson and Winfred (2017) argue the importance of standardising ethical design, when designing autonomous systems, including intelligent robots expected to interact more intensely with humans. They highlight ethical issues such as privacy and transparency to build trust when integrating AI into social and collaborative robots.

The development of the BS 8611:2016 focusing on ethical design and application of robot systems may partly address the aforementioned challenge, and an important step towards standardising ethics in AI systems design. Moreso, considering that psychosocial factors will continue to play an essential role in enhancing robotic systems' safety, as designers strive to design more intelligent and interactive robot systems. Important clauses to take note of in the BS 8611:2016 is the requirement for ethical risk assessment, which presents guidance on how designers can integrate ethical considerations within the design of robot systems. The objective of risk consideration is to systematically identify and mitigate ethical risks, grouped into four broad categories: societal, application, commercial, financial, and environmental ethical risks. Particularly interesting for robots and AI is the societal, ethical risks that encompass risks such as loss of trust, infringement of privacy, confidentiality, and employment (Winfield, 2019). From a psychosocial perspective, loss of trust and deception are important considerations when we interact with robots, especially considering research studies focusing on designing "friendly" humanoid features, which may be argued to create trust when interacting with robots. Of course, one may argue that integrating humanoid features to be deceptive as it may mask a robot's intention.

Perhaps also a notable clause of the BS 8611:2016 that relates to AI is the assertions of "humanity first" and "transparency needs." The first assertion views robotics as an enabler of improving human conditions rather than an economic driver. Therefore, this implies the need to involve "all" stakeholders in the intelligent robotic systems' design process creating safe robot systems. As discussed in this section, this involves operators who interact with robots, with psychosocial measurements expected to play an essential role in designing safe robots. Although this may also seem as straightforward, measuring human behaviour to ensure that it is primarily used ethically to improve working conditions at the expense of productivity goals remains unclear. Especially given the fact that pressure for productivity working alongside robots may be an important trigger for work-related stress, leading to long term psychological stress.

The absence of clarity also extends to transparency needs, specifically clarity on how human behavioural measurements are integrated ethically for intelligent robot designs, where AI plays an important role. For instance, Winfield (2019) argues that it should be possible to trace and determine why a robot behaves in a certain way or reaches decisions autonomously. This extends to autonomous behaviour reached by a robot, including collision avoidance when unsafe stressors are sensed from an operator, e.g., fatigue or work-related stress that is likely to cause physical or emotional harm.



TOWARD A RECONCEPTUALISATION OF SAFETY, CONSIDERING DIFFERENT DIMENSIONS OF ROBOTS AND AI

People have interacted and worked alongside machines since the 1st industrial revolution. However, the development of AI and its integration in the industrial setting that is now driving the 4th industrial revolution extends technology beyond just being an inanimate tool under full human control (Gunkel, 2020). In the early times of the introduction of robots in the workplace, the main focus was on physical safety, which was addressed by separating robots from human workers (Bekey, 2009). As robots became more integrated into workplaces, this separation reduced and collaborative robots (cobots) started working in the same space as humans (Bekey, 2009). This led to physical safety being the main dimension of safety addressed by safety regulations for robots (Harper and Virk, 2010; Jansen et al., 2018), and results on a rather narrowed lens through which safety is traditionally evaluated.

Smart robots assume social roles leading to an expansion of the possible dimensions of human-robot interaction (HRI) (Bankins and Formosa, 2020) and consequently safety. Such robots are increasingly employed in the workplace and interact with human workers physically and socially. As discussed previously, this new HRI may create mental, trust-related, affective, long-term health, societal, cybersecurity, and social consequences (Bankins and Formosa, 2020). Moreover, the interaction between robots and humans may hamper their privacy, autonomy, and even dignity (Fosch-Villaronga, 2019). Given the far-reaching dimensions of this phenomenon, it is essential to start reflecting on what this means for the current definitions of safety and discuss the need for a more comprehensive definition of the concept of safety that includes its multiple dimensions. As depicted in **Figure 3**, we can

see that these different dimensions interact with each other, at different moments in times (even simultaneously), and affect one another. Depending on the context of use, the robot embodiment, and the type of HRI involved, the dimension interplay will vary. In the following subsections, we describe the different dimensions in detail for the community.

Interaction Dimension

Physical Interaction

Physical interaction between a robot and a human operator creates a physical safety dimension related to the protection of workers from injuries related to such interaction. One example is the increasing advances in social and industrial robotics and AI, which tremendously increased cobots' autonomy and versatility, yet made them less predictable for human operators who need to collaborate with them (Jansen et al., 2018). For instance, the robot features anticipating human behaviour discussed in *The physicality of the Concept of Safety*. Less predictive robot behaviour makes physical safety challenging, although new sensors and increasingly improved AI and physical HRI (pHRI) algorithms enable humans and robots to work together, ensuring physical safety (Bankins and Formosa, 2020; Bicchi et al., 2009; Keemink et al., 2018; Pulikottil et al., 2021). Still, robots and automated systems at work used by workers may lead to indirect physical deterioration by resulting in awkward body postures, physical inactivity and sedentary work (Stacey et al., 2017). These are associated with heart disease, overweight, certain types of cancers, and psychological disorders (Stacey et al., 2017) and can compromise long-term physical safety.

Social Interaction

Cognitive Component

The introduction of robots in the workplace, among others, also introduces a cognitive and psychosocial behavioural interaction component. For example, more repetitive tasks with low variation may result in a perceived cognitive underload, passivity, and depression. On the other hand, this new smart interaction may lead to cognitive overload due to performance pressure to keep pace with the "perfect" robot, which is a machine that can work 24/7, without pause, and without social benefits. Even if the tasks vary between robots and human operators, their collaboration may lead to synchronization problems not in favor of the human operators. This is especially the case where the organisation is driven by business metrics, such as optimising productivity. The task allocation for the robot and operator should therefore be optimised, to better synchronise their capabilities. This intensification of work may involve additional risk factors too, such as isolation and lack of social interaction, similar to what many people experience due to the COVID-19 (Kniffin et al., 2020). This isolation may impact negatively on teamwork and the workers' self-esteem, leading to over-reliance on technology and deskilling, which may have detrimental effects on job satisfaction and performance (Stacey et al., 2017). This further extends to ethical considerations for robot designs, considering among other factors, the need for ethical risk assessment referred to, in *Psychosocial and Ethical Considerations for the Design of Robot Systems*.

The interplay between physical-cognitive interaction with robots in the workplace, highlights the differences between certified safety and perceived safety (Fosch-Villaronga and Özcan, 2020). Perceived safety is “the user’s perception of the level of danger when interacting with a robot, and the user’s comfort level during the interaction” (Bartneck et al., 2009). Indeed, “a certified robot might be considered safe objectively, but a (non-expert) user may still perceive it as unsafe or scary” (Salem et al., 2015). Being afraid of the device, for instance, affects the device’s adequate performance and the user: heartbeat may accelerate, hands may sweat. Depending on the user’s condition or position at work, these consequences may impact their perception of the device’s overall safety. In a major resolution, the European Parliament (2017) stressed: “you (referring to users) are permitted to make use of a robot without risk or fear of physical or psychological harm.” Because collaborative robots work with users and indissociably physical and cognitive aspects are at play simultaneously, special attention will have to be drawn progressively to both sides to ensure these devices’ safety. Therefore, this reinforces the need to re-evaluate how safety is conceptualised for robots, considering advances in AI-embedded design safeguards.

Psychological Component

Interaction can extend further than the physical dimension. Social robots interact with users socially and it is often zero-contact between the robot and the user, which challenges the applicability of current safeguards focussing solely on pHRI (Fosch-Villaronga, 2019). Recent advances in social and industrial robots allow robots to read social cues, anticipate behaviour, and even predict emotions (Cid et al., 2014; Hong et al., 2020), therefore giving them sophisticated social abilities (Henschel et al., 2020). How individuals interact with robots in a given context and for some time can have adverse short and long-term effects on their well-being and overall behaviour. This may create stressful HRI, especially in the cases of robots employing ruthless AI violating social conventions and expectations [for example in the case of Twitter being racist (Bouvier, 2020)]. This can have severe negative consequences for humans and in the long term compromise their psychological safety (Lasota, Fong and Shah, 2017). Also, the increasing use of emotions in HRI pushes the boundaries of what is considered acceptable or appropriate and in what precise context should be applied and encouraged, and where it should be applied with precaution (Fosch-Villaronga, 2019). For instance, researchers found that jobs with poor psychosocial quality may lead people to lower mental health than unemployment (Stacey et al., 2017). This aligns with ethical considerations, that prioritises the well-being of the worker as opposed to economic concerns of industry, outlined in the BS 8611:2016.

Trust Component

An important social aspect of human-human interaction that also plays a role in workplace acceptance of AI is trust (Brown et al., 2015; Lockey et al., 2021). Workplace performance is favoured in environments where employees enjoy the trust of their superiors (Brown et al., 2015). This trust may be affected by cobots

monitoring their performance or antagonizing human workers, leading to a vicious circle of deteriorating performances (Brown et al., 2015). On the other hand, over-trust of robots’ capabilities and awareness of the workspace can result in high-risk behaviour, compromising physical and mental safety, whereas limited trust can result in acceptance and collaboration problems (Schaefer et al., 2016). Monitoring the performance of workers may lead to performance anxiety, work surveillance, and overlap between work and personal life, enforcing behaviours such as being available to work at any time and result in a lack of personal time and create burnouts, as workers cannot reach the employer’s ambitious targets without working overtime (Wisskirchen et al., 2017). Already at this point, a large percentage of individuals feel compelled to read work emails at home and outside office hours and experience anxiety when away from their electronic devices, a behaviour referred to as “technostress” (Stacey et al., 2017) likely to be replicated with the use of robot technology (Vrousalis, 2013).

Anthropomorphisation Component

Anthropomorphising animate or inanimate objects is something that humans often do to reinforce trust and acceptance (Hooker and Kim, 2019). This human tendency extends from anthropomorphising pets to smart machines (Darling, 2016). However, anthropomorphising robots may lead to complex psychological HRI (Terzioglu et al., 2020). By replacing human-human interaction with something less subtle and humanising machines, we may dehumanize the workplace and therefore compromise the mental and psychological safety of human workers (Hooker and Kim, 2019). A way to avoid such outcomes is proper training of involved workers in AI similar to how our ancestors were trained to collaborate with domesticated animals, and clear ethical and social distinctions between human and robot workers (Hooker and Kim, 2019).

Cyber Dimension

The integration of AI in cyber-physical systems such as robots, the increasing interconnectivity with other devices and cloud services, and the growing human-machine interaction challenges this narrow concept of physical system safety. Cloud services allow robots to offload heavy computational tasks such as navigation, speech, or object recognition on the cloud, and mitigate this way some of the limitations posed by their physical embodiment (Fosch-Villaronga and Millard, 2019). However, “the more functions are performed across interconnected systems and devices, the more opportunities for weaknesses in those systems to arise, and the higher the risk of system failures or malicious attacks” (Michels and Walden, 2018). To date, nonetheless, there is little understanding to what extent an attacker can exploit the computational parts of a robot to affect the physical environment in industrial (Quarta et al., 2017), social (Lera et al., 2017), or medical environments (Bonaci et al., 2015), and what that would entail for the users involved in the interaction (Fosch-Villaronga and Mahler, 2021). Ethical issues of transparency and cybersecurity present also an important concern, and are yet not comprehensively addressed for robots and other cyber-physical systems, despite existing cases

concerning the hacking of autonomous vehicles and healthcare devices such as pacemakers (Miller, 2019).

Data characterizing the performance of human workers can be acquired in real-time in the workplace. This may lead to a rise in productivity, and simultaneously make work more transparent and allow companies to assess employees' performance (Wisskirchen et al., 2017). The collection of personal data either for person-specific performance, analyzing the behaviour of humans collaborating with robots, or research raises multiple ethical questions about how such data is used, their impact on human workers' wellbeing and safety, and who has access to the information. This may create trust issues between employees and employers and create performance anxiety (Brown et al., 2015). Additionally, it is interesting to reflect on what happens to the employee who does not fit in the criteria or working patterns defined by their employer (Bowen et al., 2017).

Temporal Dimension

It is important to mention here that psychological attrition, in contrast with physical, is not limited to proximal interaction and specific time frames, as it can affect workers seamlessly even via remote interactions and outside of working hours (Lasota et al., 2017). This extension provides an instance of the temporal dimension of safety: it is no longer reserved for an immediate physical impact/harm, but such harm, such consequence, can appear at a later stage in time. This relates to the fact that, sometimes, harm appears after the continuous use of a device [in the case of wearable walking exoskeletons, abnormal muscle activation could cause a problem at a later stage causing what is called prospective liability, (Fosch-Villaronga and Özcan, 2020) or in the case of unilateral transfemoral amputation the intact limb joints have a higher prevalence of osteoarthritis, due to amputees not interacting optimally with their artificial limb (Welke et al., 2019)].

Another temporal instance is the increasing use of machine learning algorithms. Machine learning provides machines with the possibility to learn from experience and adapt over time—something that is keeping busy certification agencies and policymakers around the world (FDA, 2021; Consumer Safety Network, 2020). The safety rules should explicitly include protection against risks related to subsequently uploaded software and extended functions acquired by machine learning (Fosch-Villaronga and Mahler, 2021).

The ability to make decisions based on predictive analytics is also a new element added to the temporal dimension that may challenge the user's safety of the user—in case of wrongly predicted or inferred actions (i.e., a wrong future) (Fosch-Villaronga and Mahler, 2021). These capabilities can lead to very unfortunate scenarios that could even widen more the knowledge asymmetry gap, but that have also ulterior consequences at the cybersecurity level.

Societal Dimension

The societal dimension of safety refers to the societal challenges and consequences of introducing robots in the workplace (Frey and Osborne, 2017). Many workers also believe that cobots aim to

replace them rather than reduce their workload. Feeling that one may not be needed in the future and that may be discarded at any time, replaced by a machine, creates societal pressure, and imbalance as it mainly affects manual labour workers and may lead to stress, lower job satisfaction and burnouts (Meissner et al., 2020).

Advancements in AI and virtual/augmented reality, in combination with robotics, may extend the workplace from physical to virtual/remote facilitating the transition towards an Operator 4.0 scenario (Romero et al., 2016). This extended workplace can allow workers to reduce their hours as they approach retirement, better balance work with leisure or education, and work flexibly. Less work supervision and more flexible working schedules may lead to stress and burnout, if not accompanied by workers' training to obtain better skills for managing and organizing their workloads to create a good work-life balance that supports wellbeing (Stacey et al., 2017). Additionally, it can provide more people with access to work and markets that might not have been accessible in the past. This may result in a surplus of job offers in the job market and consequently insecurity, low pay and working quality (Stacey et al., 2017). Employment and social security laws coupled with narrow safety definitions may not have the versatility to deal with these changes in working patterns, and thus create insecurity for workers in the future that may be unprotected and low paid. This may have ulterior consequences in the social security scheme and potentially endanger the whole welfare system.

One societal challenge is how education is changing due to the introduction of these robots, either in factories or in hospitals where the surgery success does not depend just on the surgeon any longer but on the complex interaction and interplay between the doctor, the supporting staff, and the robot (i.e., the manufacturer) (Yang et al., 2017). Latent deskilling has to do with the gradual loss of essential skills due to overreliance on AI and workplace robotics (Rinta-Kahila et al., 2018). That becomes more apparent when a workplace process or a specific product is discontinued, and it may cause severe disruptions in workers' daily work and internal organizational processes. Together with this, another societal dimension of safety is that over-reliance on AI in the workplace may lead to knowledge asymmetries such as the inability of workers to understand the inner workings of these machines and how these impact their lives (including what rights they have and what are the obligations they come with such transformation); or the impossibility of experts to anticipate and understand the impact of their work and communicate it to various societal stakeholders and, more importantly, assume responsibility (Lockey et al., 2021).

CONCLUSION

Working with robots and AI may be equivalent to working with a new species to some extent. The consequences of that interaction demand an open discussion for a more comprehensive safety regulatory framework encompassing and accommodating more dimensions of safety than just physical interaction. However, it may be that the solution to many of the multidimensional safety

challenges introduced by advanced AI in the workplace may be addressed with a more comprehensive view of safety.

Although being a central concept in machinery, robot, AI, and human-robot interaction, the concept of safety is not clearly defined in current norms and legislations. Traditionally, the definition of safety has been interpreted to exclusively apply to risks that have a physical impact on persons' safety, such as, among others, mechanical or chemical risks. However, the current understanding is that the integration of AI in cyber-physical systems such as robots, the increasing interconnectivity with several devices and cloud services, and the growing human-robot interaction challenges the safety concept's narrowness.

The paper, therefore, brought together different dimensions relevant for a re-definition of safety required by the introduction of new technologies. These dimensions relate to interaction humans and these devices have (physical and social), their cyber and intangible components, the temporal nature of such interactions and consequences, and the potential societal effects these may have in the long run. A dimension from an AI perspective is quantifying these interactions. This will potentially yield new anticipatory algorithms that would allow collaborative agents to adapt their behaviour to mitigate hazardous contacts in shared workspaces. Some work in this direction includes, Strathearn and Ma (2018) where the authors explore applying psycho-social indicators to model operator behaviour and anticipate operator actions likely to lead to safety hazards.

We acknowledge though that, to be effective, such a concept needs to be modular and adaptive to the particular needs of each sector. Unfortunately, the proposal for a regulation on safety products does not guide in this respect (Machinery Proposed Regulation, 2021). Therefore, we call on different sectors where robots and AI are well integrated, including chemical engineering, to consider the new multidimensional concept of safety that we propose.

While the paper acknowledges that these dimensions may bring uncertainties concerning what are the assessment methods to ensure safety from a multi-dimensional viewpoint, these dimensions aim to stimulate the discussions among the community and help policy and standard makers redefine the concept of safety in light of robots and AI's increasing capabilities, including human-robot interactions, cybersecurity, and machine learning. Over time, if different sectors share their lessons learned on how they understood and applied the different dimensions of

safety we could generate knowledge that could support further revisions of this legislative instrument (Fosch-Villaronga and Heldeweg, 2018).

As a future direction, the concept of safety should be revised in light of the different dimensions here exposed, mainly physical, psychological, cybersecurity, temporal, and societal. To do so, multidisciplinary conversations and more research need to happen among researchers, legal scholars, and other relevant stakeholders at multiple levels in the public and private fields. One avenue could be to discuss this topic within relevant standardization organizations such as CEN/CENELEC, the European Standardization body, and incorporate such reflections in CENELEC Workshop Agreements (CWA). The authors took the first step in this direction and did so in the context of the H2020 COVR project, in which they presented the ideas developed in this article to the CEN/WS 08 "Safety in close human-robot interaction: procedures for validation tests." The reflections will be successfully incorporated into the CWA. Afterward, the researchers will work on different case studies to see how these theoretical insights revolving around the concept of safety translate into specific case scenarios.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

AM, PC, KN, and EF-V conceived and wrote the paper. AM, EF-V, and KN carried out the mapping on different safety perspectives. PC and KN carried out the reconceptualisation of safety. EF-V attracted part of the funding.

FUNDING

This project is part of LIAISON, a subproject of the H2020 COVR Project that has received funding 673 from the European Union's Horizon 2020 research and innovation programme under grant agreement 674 No 779966.

REFERENCES

- Ahmad, M., Mubin, O., and Orlando, J. (2017). A systematic review of adaptivity in human-robot interaction. *Mti* 1 (3), 14. doi:10.3390/mti1030014
- Anybotics (2021). Animal Robotics. Retrieved from Available at: <https://www.anybotics.com/anybotics-makes-the-case-at-basf-chemical-plant/>.
- Bankins, S., and Formosa, P. (2020). When AI meets PC: exploring the implications of workplace social robots and a human-robot psychological contract. *European Journal of Work and Organizational Psychology* 29 (2), 215–229. doi:10.1080/1359432X.2019.1620328
- Bao, S. S., Kapellusch, J. M., Merryweather, A. S., Thiese, M. S., Garg, A., Hegmann, K. T., and Silverstein, B. A. (2016). Relationships between job organisational factors, biomechanical and psychosocial exposures. *Ergonomics* 59 (2), 179–194. doi:10.1080/00140139.2015.1065347
- Bartneck, C., Kulić, D., Croft, E., and Zoghbi, S. (2009). Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. *Int J of Soc Robotics* 1 (1), 71–81. doi:10.1007/s12369-008-0001-3
- Bekey, G. (2009). "Current Trends in Robotics: Technology and Ethics," in *Robot Ethics: The Ethical and Social Implications of Robotics* (MIT Press), 17–34.
- Bicchi, A., Peshkin, M. A., and Colgate, J. E. (2009). "Safety for Physical Human-Robot Interaction," in *Springer Handbook of Robotics* 57, 1335–1348. doi:10.1007/978-3-540-30301-5_58
- Bonaci, T., Herron, J., Yusuf, T., Yan, J., Kohno, T., and Chizeck, H. J. (2015). *To make a robot secure: An experimental analysis of cyber security threats against*

- teleoperated surgical robots. arXiv preprint arXiv:1504.04339. doi:10.1145/2735960.2735980
- Borenstein, J., Wagner, A. R., and Howard, A. (2018). Overtrust of pediatric health-care robots: A preliminary survey of parent perspectives. *IEEE Robot. Automat. Mag.* 25 (1), 46–54. doi:10.1109/mra.2017.2778743
- Borghini, G., Ronca, V., Vozzi, A., Aricò, P., Di Flumeri, G., and Babiloni, F. (2020). Monitoring performance of professional and occupational operators. *Handbook of clinical neurology* 168, 199–205. doi:10.1016/b978-0-444-63934-9.00015-9
- Bouvier, G. (2020). Racist call-outs and cancel culture on Twitter: The limitations of the platform's ability to define issues of social justice. *Discourse, Context & Media* 38, 100431. doi:10.1016/j.dcm.2020.100431
- Bowen, J., Hinze, A., Griffiths, C. J., Kumar, V., and Bainbridge, D. I. (2017). "Personal data collection in the workplace: Ethical and technical challenges," in HCI 2017: Digital Make Believe - Proceedings of the 31st International BCS Human Computer Interaction Conference, 03 July 2017 (ACM), 1–11. doi:10.14236/ewic/HCI2017.57
- Brown, S., Gray, D., McHardy, J., and Taylor, K. (2015). Employee trust and workplace performance. *Journal of Economic Behavior & Organization* 116, 361–378. doi:10.1016/j.jebo.2015.05.001
- Chemweno, P., Pintelon, L., and Decre, W. (2020). Orienting safety assurance with outcomes of hazard analysis and risk assessment: A review of the ISO 15066 standard for collaborative robot systems. *Safety Science* 129, 104832. doi:10.1016/j.ssci.2020.104832
- Cid, F., Moreno, J., Bustos, P., and Núñez, P. (2014). Muecas: A multi-sensor robotic head for affective human robot interaction and imitation. *Sensors* 14 (5), 7711–7737. doi:10.3390/s140507711
- European Parliament (2017). Civil Law Rules on Robotics European Parliament resolution of 16 February 2017 with recommendations to the Commission on Civil Law Rules on Robotics (2015/2103(INL)). Retrieved from Available at: https://www.europarl.europa.eu/doceo/document/TA-8-2017-0051_EN.pdf.
- Consumer Safety Network (2020). Opinion of the Sub-group On Artificial Intelligence (AI), Connected Products And Other New Challenges in Product Safety to the Consumer Safety Network. Retrieved from Available at: https://ec.europa.eu/consumers/consumers_safety/safety_products/rapex/alerts/repository/content/pages/rapex/docs/Subgroup_opinion_final_format.pdf (last accessed January 27, 2020).
- Council Directive 89/391/EEC 1989 on the introduction of measures to encourage improvements in the safety and health of workers at work.
- Council Directive 89/392/EEC 1989 on the approximation of the laws of the Member States relating to machinery.
- Darling, K. (2016). "Extending legal protection to social robots: The effects of anthropomorphism, empathy, and violent behavior towards robotic objects," in *Robot law*. Editors R. Calo, A. M. Froomkin, and I. Kerr (Edward Elgar Publishing), 213–232.
- Di Dio, C., Manzi, F., Peretti, G., Cangelosi, A., Harris, P. L., Massaro, D., et al. (2020). Shall I Trust You? From Child Human-Robot Interaction to Trusting Relationships. *Front. Psychol.* 11, 469. doi:10.3389/fpsyg.2020.00469
- Directive 2001/95/EC European Parliament and of the Council of 3 December 2001 on general product safety.
- Directive 2006/42/EC European Parliament and of the Council of 17 May 2006 on machinery, and amending Directive 95/16/EC (recast).
- European Agency for Safety and Health at Work (2020). *Digitalisation and occupational safety and health*. Bilbao. doi:10.2802/559587
- HLEG on AI (2019) European Commission Concept of Safety. High-Level Expert Group. Retrieved from Available at: https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=56341 (last accessed 27 January 2020).
- Food and Drug Administration (2021). Artificial Intelligence and Machine Learning in Software as a Medical Device. Retrieved from Available at: <https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-software-medical-device> (last accessed January 27, 2020).
- Fosch Villaronga, E. (2019). "I Love You," Said the Robot: Boundaries of the Use of Emotions in Human-Robot Interactions," in *Emotional Design in Human-Robot Interaction*. Editors H. Ayanoglu and E. Duarte (Cham: Springer), 93–110. doi:10.1007/978-3-319-96722-6_62019
- Fosch Villaronga, E., and Golia, A. J., Jr. (2019). Robots, standards and the law: Rivalries between private standards and public policymaking for robot governance. *Computer Law & Security Review* 35 (2), 129–144. doi:10.1016/j.clsr.2018.12.009
- Fosch-Villaronga, E., and Mahler, T. (2021). *Cybersecurity, safety and robots: Strengthening the link between cybersecurity and safety in the context of care robots*. forthcoming: Computer Law & Security Review.
- Fosch-Villaronga, E., and Millard, C. (2019). Cloud robotics law and regulation. *Robotics and Autonomous Systems* 119, 77–91. doi:10.1016/j.robot.2019.06.003
- Fosch-Villaronga, E., and Özcan, B. (2020). The Progressive Intermittent Between Design, Human Needs and the Regulation of Care Technology: The Case of Lower-Limb Exoskeletons. *Int. J. Soc. Robot.* 12 (4), 959–972.
- Villaronga, E. F., and Virk, G. S. (2016). "Legal Issues for Mobile Servant Robots,". *Advances in Robot Design and Intelligent Control*. RAAD 2016. *Advances in Intelligent Systems and Computing*. Editors A. Rodić and T. Borangiu (Cham: Springer), 540, 605–612. doi:10.1007/978-3-319-49058-8_66
- Frederick, P., Del Rose, M., and Cheok, K. (2019). *Autonomous Vehicle Safety Reasoning Utilizing Anticipatory Theory*.
- Frey, C. B., and Osborne, M. A. (2017). The future of employment: How susceptible are jobs to computerisation?. *Technological forecasting and social change* 114, 254–280. doi:10.1016/j.techfore.2016.08.019
- Gunkel, D. J. (2020). Mind the gap: responsible robotics and the problem of responsibility. *Ethics Inf Technol* 22 (4), 307–320. doi:10.1007/s10676-017-9428-2
- Harper, C., and Virk, G. (2010). Towards the Development of International Safety Standards for Human Robot Interaction. *Int J of Soc Robotics* 2 (3), 229–234. doi:10.1007/s12369-010-0051-1
- Henschel, A., Hortensius, R., and Cross, E. S. (2020). Social Cognition in the Age of Human-Robot Interaction. *Trends in Neurosciences* 43 (6), 373–384. doi:10.1016/j.tins.2020.03.013
- Hong, A., Lunscher, N., Hu, T., Tsuboi, Y., Zhang, X., Alvesdos, S. F. d. R.R., Nejat, G., and Benhabib, B. (2020). "A Multimodal Emotional Human-Robot Interaction Architecture for Social Robots Engaged in Bidirectional Communication/vspace*7pt," in *IEEE Trans. Cybern.* 06 March 2020 (IEEE), 1–15. doi:10.1109/tycyb.2020.2974688
- Hooker, J., and Kim, T. W. (2019). Humanizing Business in the Age of Artificial Intelligence. *Humanizing Business in the Age of AI* 1, 1–24. Retrieved from Available at: <http://public.tepper.cmu.edu/jnh/humanizingBusiness4.pdf> (last accessed February 1, 2021).
- ISO 10218:2011, Industrial Robots Safety
- ISO 15066 2016, Robots and robotic devices — Collaborative robots
- ISO 26262-1:2018 Road Vehicles - Functional Safety - Part 1: Vocabulary.
- ISO 8373:2012 Robots and robotic devices — Vocabulary.
- Jansen, A., van der Beek, D., Cremers, A., Neerinx, M., and van Middelaar, J. (2018). *Emergent Risks To Workplace Safety; Working in the Same Space As a Cobot*.
- Jones, D. R. (2017). "Psychosocial Aspects of New Technology Implementation," in *International Conference on Applied Human Factors and Ergonomics* (Cham: Springer), 3–8. doi:10.1007/978-3-319-60828-0_1
- Kapeller, A., Felzmann, H., Fosch-Villaronga, E., and Hughes, AM (2020). A Taxonomy of Ethical, Legal and Social Implications of Wearable Robots: An Expert Perspective. *Sci Eng Ethics*, 1–19. doi:10.1007/s11948-020-00268-4
- Keemink, A. Q., van der Kooij, H., and Stienen, A. H. (2018). Admittance control for physical human-robot interaction. *The International Journal of Robotics Research* 37, 1421–1444. doi:10.1177/0278364918768950
- Kniffin, K. M., Narayanan, J., Anseel, F., Antonakis, J., Ashford, S. P., Bakker, A. B., Bamberger, P., Bapuji, H., Bhawe, D. P., Choi, V. K., Creary, S. J., Demerouti, E., Flynn, F. J., Gelfand, M. J., Greer, L. L., Johns, G., Kesimal, S., Klein, P. G., Lee, S. Y., Ozcelik, H., Petriglieri, J. L., Rothbard, N. P., Rudolph, C. W., Shaw, J. D., Sirola, N., Wanberg, C. R., Whillans, A., Wilmot, M. P., and Vugt, M. v. (2020). COVID-19 and the Workplace: Implications, Issues, and Insights for Future Research and Action. *American Psychologist* 76, 63–77. doi:10.1037/amp0000716
- Lasota, P. A., Fong, T., and Shah, J. A. (2017). A Survey of Methods for Safe Human-Robot Interaction. *FNT in Robotics* 5 (3), 261–349. doi:10.1561/23000000052
- Lera, F. J. R., Llamas, C. F., Guerrero, Á. M., and Olivera, V. M. (2017). "Cybersecurity of robotics and autonomous systems: privacy and safety," in *Robotics: Legal, Ethical and Socioeconomic Impacts*. *IntechOpen*. Editor

- G. Dekoulis. doi:10.5772/intechopen.69796 <https://cdn.intechopen.com/pdfs/56025.pdf>
- Li, H., Li, R., Zhang, J., and Zhang, P. (2020). Development of a pipeline inspection robot for the standard oil pipeline of China national petroleum corporation. *Applied Sciences* 10 (8), 2853. doi:10.3390/app10082853
- Lockey, S., Gillespie, N., and Holm, D. (2021). *A Review of Trust in Artificial Intelligence: Challenges, Vulnerabilities and Future Directions*, 5463–5472.
- Meissner, A., Trübsetter, A., Conti-Kufner, A. S., and Schmidler, J. (2021). Friend or Foe? Understanding Assembly Workers' Acceptance of Human-robot Collaboration. *J. Hum.-Robot Interact.* 10 (1), 1–30. doi:10.1145/3399433
- Michalos, G., Makris, S., Tsarouchi, P., Guasch, T., Kontovrakis, D., and Chrysosouris, G. (2015). Design considerations for safe human-robot collaborative workplaces. *Procedia CIRP* 37, 248–253. doi:10.1016/j.procir.2015.08.014
- Michels, J. D., and Walden, I. (2018). “How Safe is Safe Enough? Improving Cybersecurity in Europe's Critical Infrastructure Under the NIS Directive,”. Paper No. 291/2018 (Queen Mary School of Law Legal Studies Research Paper).
- Miller, C. (2019). Lessons learned from hacking a car. *IEEE Des. Test* 36 (6), 7–9. doi:10.1109/mdat.2018.2863106
- Mokarami, H., and Toderi, S. (2019). Reclassification of the work-related stress questionnaires scales based on the work system model: A scoping review and qualitative study. *Wor* 64 (4), 787–795. doi:10.3233/wor-193040
- OECD (2019). *OECD Employment Outlook 2019: The Future of Work*. Paris: OECD Publishing. doi:10.1787/9ee00155-en
- Machinery Proposed Regulation (2021) Proposal for a Regulation of the European Parliament and of the Council on machinery products' establishing a regulatory framework for placing machinery on the Single Market. Retrieved from Available at: https://eur-lex.europa.eu/resource.html?uri=cellar:f0f10ee-a364-11eb-9585-01aa75ed71a1.0001.02/DOC_1&format=PDF.
- Pulikottil, T. B., Pellegrinelli, S., and Pedrocchi, N. (2021). A software tool for human-robot shared-workspace collaboration with task precedence constraints. *Robotics and Computer-Integrated Manufacturing* 67, 102051. doi:10.1016/j.rcim.2020.102051
- Quarta, D., Pogliani, M., Polino, M., Maggi, F., Zanchettin, A. M., and Zanero, S. (2017). “An experimental security analysis of an industrial robot controller,” in 2017 IEEE Symposium on Security and Privacy, San Jose, CA, USA, 22–26 May 2017 (IEEE), 268–286.
- Riley, P. (2015). Australian Principal Occupational Health, Safety and Well-Being Survey, Analysis Policy and Observation.
- Rinta-Kahila, T., Penttinen, E., Salovaara, A., and Soliman, W. (2018). “Consequences of Discontinuing Knowledge Work Automation - Surfacing of Deskilling Effects and Methods of Recovery,” in Proceedings of the 51st Hawaii International Conference on System Sciences. doi:10.24251/hicss.2018.654
- Robinson, N. L., Cottier, T. V., and Kavanagh, D. J. (2019). Psychosocial health interventions by social robots: systematic review of randomized controlled trials. *Journal of medical Internet research* 21 (5), e13203. doi:10.2196/13203
- Romero, D., Stahre, J., Wuest, T., Noran, O., Bernus, P., Fast-Berglund, Å., and Gorecky, D. (2016). “Towards an operator 4.0 typology: a human-centric perspective on the fourth industrial revolution technologies,” in International conference on computers & industrial engineering (CIE46), 1–11.
- Rouaix, N., Retru-Chavastel, L., Rigaud, A. S., Monnet, C., Lenoir, H., and Pino, M. (2017). Affective and engagement issues in the conception and assessment of a robot-assisted psychomotor therapy for persons with dementia. *Frontiers in psychology* 8, 950. doi:10.3389/fpsyg.2017.00950
- Salay, R., Queiroz, R., and Czarnecki, K. (2017). *An analysis of ISO 26262: Using machine learning safely in automotive software*. CoRR abs/1709.02435. <http://arxiv.org/abs/1709.02435>.
- Salem, M., Lakatos, G., Amirabdollahian, F., and Dautenhahn, K. (2015). “Towards safe and trustworthy social robots: ethical challenges and practical issues,” in International conference on social robotics, Paris, France, October 26–30, 2015 (Springer), 584–593. doi:10.1007/978-3-319-25554-5_58
- Schaefer, K. E., Chen, J. Y. C., Szalma, J. L., and Hancock, P. A. (2016). A Meta-Analysis of Factors Influencing the Development of Trust in Automation. *Hum Factors* 58, 377–400. doi:10.1177/0018720816634228
- Stacey, N., Ellwood, P., Bradbrook, S., Reynolds, J., and Williams, H. (2017). *Key trends and drivers of change in information and communication technologies and work location* Key trends and drivers of change in information and communication technologies and work location. European Agency for Safety and Health at Work.
- Strathearn, C., and Ma, M. (2018). Biomimetic pupils for augmenting eye emulation in humanoid robots. *Artif Life Robotics* 23 (4), 540–546. doi:10.1007/s10015-018-0482-6
- Robinette, P., Li, W., Allen, R., Howard, A. M., and Wagner, A. R. (2016). “Overtrust of robots in emergency evacuation scenarios,” in 2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI), Christchurch, New Zealand, 7–10 March 2016 (IEEE), 101–108. doi:10.1109/HRI.2016.7451740
- Terzioglu, Y., Mutlu, B., and Sahin, E. (2020). “Designing social cues for collaborative robots: The role of gaze and breathing in human-robot collaboration,” in ACM/IEEE International Conference on Human-Robot Interaction, March 2020 (ACM), 343–357. doi:10.1145/3319502.3374829
- Van Damme, M., Vanderborght, B., Verrelst, B., Van Ham, R., Daerden, F., and Lefeber, D. (2009). Proxy-based sliding mode control of a planar pneumatic manipulator. *The International Journal of Robotics Research* 28 (2), 266–284. doi:10.1177/0278364908095842
- Vrousalis, N. (2013). Exploitation, vulnerability, and social domination. *Philos Public Aff* 41 (2), 131–157. doi:10.1111/papa.12013
- Welke, B., Jakubowitz, E., Seehaus, F., Daniilidis, K., Timpner, M., Tremer, N., Hurschler, C., and Schwarze, M. (2019). The prevalence of osteoarthritis: Higher risk after transfemoral amputation?—A database analysis with 1,569 amputees and matched controls. *PLoS ONE* 14, e0210868. doi:10.1371/journal.pone.0210868
- Winfield, A. (2019). Ethical standards in robotics and AI. *Nat Electron* 2 (2), 46–48. doi:10.1038/s41928-019-0213-6
- Wisskirchen, G., Thibault, B., Bormann, B. U., Muntz, A., Niehaus, G., Soler, G. J., and Von Brauchitsch, B. (2017). *Artificial Intelligence and Robotics and Their Impact on the Workplace*. IBA Global Employment Institute, 120.
- Yang, G. Z., Cambias, J., Cleary, K., Daimler, E., Drake, J., Dupont, P. E., and Taylor, R. H. (2017). Medical robotics—Regulatory, ethical, and legal considerations for increasing levels of autonomy. *Science Robotics* 2 (4), 8638. doi:10.1126/scirobotics.aam8638

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Martinetti, Chemweno, Nizamis and Fosch-Villaronga. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.