INBOTS

Inclusive Robotics for a better Society

INBOTS WP5. D2.1

Andrea Bertolini, Assistant Professor of Private Law of the Scuola Superiore Sant'Anna
Nicoleta Cherciu, Research Fellow in Private Law of the Scuola Superiore Sant'Anna
Francesca Episcopo, Affiliated Research Fellow to the Scuola Superiore Sant'Anna

Individual contributions:

- 4.3. Standardize ethics (Michael Funk, Researcher, University of Vienna, Department of Philosophy)
- 5.2. A critical review of the measures recently proposed by the European Parliament for the governance of robotics (Maria Amparo Grau Ruiz, Professor of Financial and Tax Law, Universitad Complutense Madrid)
- 5.3. Brief reflections on other governance structures to provide alternative (Maria Amparo Grau Ruiz, Professor of Financial and Tax Law, Universitad Complutense Madrid)

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

INBOTS

Inclusive Robotics for a better Society

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

Table of contents

1. The need for definitions. and the choice between technology-neutral and technology-specific regulation. 4

1.1. Definitions 4

1.2. Opting between technology-neutral and technology-specific regulation 6

2. Robots are products 8

2.1. Setting the scene for future regulation. Robots are products 8

2.2. Robots as legal subjects from in functional perspective only 9

3. Liability and the Risk-Management Approach 13

3.1. Liability. General considerations 13

3.2. Strict and/or semi-strict liability rules 14

3.3. The Risk-Management Approach 17

4. Product safety regulation 20

4.1. Extant EU product safety legal framework 20

4.2. The need to update product safety regulation 24

4.3. Standardize ethics 26

4.4. A procedural approach towards transparency and disclosure 32

5. The need to update regulation & the creation of an Agency 34

5.1. Extant initiatives and their suitability 34

5.2. A critical review of the measures recently proposed by the European Parliament for the governance of robotics 37

5.3. Brief reflections on other governance structures to provide alternatives 41

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

1. The need for definitions. and the choice between technology-neutral and technology-specific regulation.

1.1. Definitions

The field of interactive robotics both at EU and international level is as prolific and broad one. Interactive robots are commonly defined as robots conceived to perform their intended tasks by interacting physically and cognitively with different environments and different situations, and most importantly with humans[1].

However, what exactly falls within this category is uncertain. Many applications and types of robotic solutions are categorized under the umbrella term of interactive robotics, which is usually further divided into diverse sub-fields such as Humanoids, or service robotics (HUM), and Wearable Robots (WRs).

As for the first category, it may include on the one hand, applications such as collaborative robots used for a myriad of tasks, including for domestic/household ones, for entertainment purposes, for the assistance of the elderly or people with disabilities, usually and collectively referred to as "companion robots". On the other hand, these collaborative robots may include applications such as industrial robots implemented in the so-called "smart-factories", where robots – with different degrees of autonomy and freedom, – and workers cooperate. Furthermore, the category of collaborative robots also includes surgical robots which are operated by a surgeon, to perform different medical procedures.

With respect to the second category, WRs – such as exoskeletons and robotic prostheses – are robots that are physically connected to the user's body and interact physically with the latter by exercising mechanical power on and exchanging forces with him,. WRs may be used in different fields such as healthcare, manufacturing and in the consumer industry.

Furthermore, the functioning of all these hardware applications usually requires the massive processing of information, which in turn creates the robots' capability to react in real time, to be autonomous, navigate and exhibit cognitive perception. All these functions and capabilities are achieved by implementing and embedding novel artificial intelligence ("AI") technologies into the hardware components.

Thus, interactive robots may be currently broadly defined in a functional perspective, as indicated above. Under such definition many classes of applications may fall, each with varying degrees of automation, autonomy, and with different use cases and purposes, while each of them gives rise in turn to different risks, to different types of harms, and to different legal and ethical issues.

[1] INBOTS White Paper on Interactive Robotics Regulatory Framework & Risk Management Framework, available at http://inbots.eu/wp-content/uploads/2019/07/Attachment_0-4.pdf.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

Nevertheless, by relying on the functional definition provided above and on the exemplified classes of applications and their components, interactive robots may be deemed as a sub-category of the general notions of robots and AI.

At EU level, current attempts to define robots and AI are underway, and from a policy-making perspective, a tendency towards regulating robotics and AI unitarily seems to be emerging , also with respect to the liability and safety framework. This approach is apparent from the 2020 EU Parliament's recommendations to the Commission on a civil liability regime for artificial intelligence[2] and its 2020 recommendations to the Commission on a framework of ethical aspects of artificial intelligence, robotics and related technologies[3], including the therein provided text of the legislative proposal requested.

Under the two aforementioned policy-making instruments, the followings definitions are provided:

- "'artificial intelligence' means a system that is either software-based or embedded in hardware devices, and that displays intelligent behavior by, inter alia, collecting, processing, analyzing, and interpreting its environment, and by taking action, with some degree of autonomy, to achieve specific goals";

- "'robotics' means technologies that enable automatically controlled, reprogrammable, multi-purpose machines to perform actions in the physical world traditionally performed or initiated by human beings, including by way of artificial intelligence or related technologies";

- "'related technologies' means technologies that enable software to control with a partial or full degree of autonomy a physical or virtual process, technologies capable of detecting biometric, genetic or other data, and technologies that copy or otherwise make use of human traits"[4].

However, as discussed in depth in § 1.2, unitarily regulating advanced technologies under one piece of legislation by resorting to broad, vague and subjective notions such as robotics and AI severely challenges this aim of it being future-proof since its scope of application may become outdated fast due to technological advancements. Such outcome may be anticipated by the continuous change and evolution of the meaning of AI reflected in the so-called "AI effect" according to which "[a]s soon as it works, no one calls it AI anymore" (Vardi 2012)[5]. Thus, the notion of AI might be considered a moving target, itself changing with technological advancement in a way that what might be deemed one of its applications at a given moment in time, might not be any longer classified as such later on (Bertolini 2020). For example, search engines and recommender systems are

[2] European Parliament (2020). Report with with recommendations to the Commission on a civil liability regime for artificial intelligence (2020/2014(INL)). Plenary sitting, European Parliament. available at https://www.europarl.europa.eu/doceo/document/A-9-2020-0178_EN.html

[3] European Parliament (2020). Report with recommendations to the Commission on a framework of ethical aspects of artificial intelligence, robotics and related technologies (2020/2012(INL)). Brussels, European Parliament. available at PR_INL (europa.eu).

[4] See in this respect art. 4 "Definitions" under ibid. and art. 3 "Definitions" under European Parliament (2020). Report with with recommendations to the Commission on a civil liability regime for artificial intelligence (2020/2014(INL)). Plenary sitting, European Parliament.

[5] Bertolini, A. (2020). Artificial Intelligence and Civil Liability. Bruxelles, European Parliament - Committee on Legal Affairs: 1-132.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

created by using different AI techniques (Russell and Norvig 2010), but they are not being addressed neither by the general public, nor by the legislator as AI, as further evidenced by the fact that such systems are to be regulated under a separate body of legislative initiatives, such as the new Digital Services Act (European Parliament 2020).

1.2. Opting between technology-neutral and technology-specific regulation

These definitions, if adopted, will constitute the basis for regulating all new advanced technologies.

Indeed, any normative intervention needs to specify the object it that falls under its scope of application.

However, the choice to adopt broad and all-encompassing definitions, in the attempt to simplify and regulate uniformly, however, raises concerns due to the extreme vagueness of such a notion of AI and advanced technologies, and the evident technical differences, societal concerns and incentive structures that characterizes them. This, in turn, could cause regulation to be both over-,and under-inclusive. From a policy making and legal responsibility perspective, clearly identifying the object to be regulated and the scope of regulation is not simply of paramount importance but strictly necessary, since, absent a clear definition, the robotics and AI ("R&AI") field will be subject to uncertainties, which may in turn cause litigation and hamper the very development of desirable technologies, instead of fostering innovation (Bertolini 2020).

Indeed, a toothbrush, an expert system to be used in the medical practice, or instead, in capital markets, a driverless vehicle, a collaborative robot, and a drone, are all extremely diverse technologies.

At the same time, since they also give rise to very diverse societal concerns, even with respect to the sole issue of liability, they are regulated by separate bodies of norms. The consumer sales field is, in fact, addressed by a separate body of norms from medical malpractice, intermediaries' or employer's responsibility, as well as drone's operators' and car owners' and drivers' liability.

Moreover, the fact that in such domains AI-based applications could be used, that, however, are so different from one another also in a purely technological perspective, does not provide sufficient an argument to justify regulating them through a simple set of norms.

Such a regulatory technique is based on the long-standing EU principle of technology neutrality, characterizing he EU data protection frameworkas defined by the General Data Protection Regulation ("GDPR")[6] – which sets forth the principle of privacy-by-design, and which states that "the protection of natural persons should be technologically neutral and should not depend on the techniques used"[7]

[6] See Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) (Text with EEA relevance), OJ L 119, 4.5.2016, p. 1–88.

[7] See in this respect Recital 15, Recital 78 and Article 25 of the Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) OJ L 119, 4.5.2016, 1–88.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

– and the product liability framework – which sets forth liability rules applicable to all tangible, mass-produced and movable goods[8].

Technology neutrality is however often a myth. Indeed, the breadth of the GDPR causes commentators to challenge its adequacy for some very relevant specific fields, such as AI itself[9]. Similarly, merely observing the litigation occurred pursuant to the Product Liability Directive ("PLD")[10], it is possible to identify clear clusters for only some very specific product categories, typically technologically simple (raw materials) or infringing upon the most fundamental rights of the individual, such as health and bodily integrity (pharmaceutical and medical devices)[11].

The final aim for employing the technology neutrality technique is, among others[12], that of providing future-proof regulation (Koops 2006). Such considerations radically challenge the most relevant theoretical justification for a technology-neutral approach, namely its future=proof nature. Indeed, every regulation conceived as such will self-select those applications where the incentive structure it provides might be deemed adequate. Yet, such choice, instead of being operated ex ante through careful ascertainment by the policymaker, will be made *ex post* and possibly will be determined by unintended failures of the market or legal system, or both.

Even more, the technology neutrality paradigm justifying the adoption of broad definitions is based on certain premises, which are not free from criticism. One of such premises is that legislation should have an effect of equivalence, in the sense that legal rules should provide for a similar treatment across different technologies so as "to avoid limiting a right only to its exercise in extant technology or discriminating against older technology simply because it existed when the law was enacted"[13].

This may be true for classes of applications having the same characteristics, similar use cases and area of application and comparable risks. Such rationale should not be construed as regulating uniformly the entire soon to become digitalized worldwide economy. To a certain extent, even the

---

[8] See Council Directive 85/374/EEC of 25 July 1985 on the approximation of the laws, regulations and administrative provisions of the Member States concerning liability for defective products, OJ L 210, 7.8.1985, p. 29–33.

9 Finck, M. (2019). Blockchain and the General Data Protection Regulation. Can distributed ledgers be squared with European data protection law? The Panel for the Future of Science and Technology (STOA). Brussels, European Parliament. and Espinoza, J. (2021). "EU must overhaul flagship data protection laws, says a 'father' of policy." https://www.ft.com/content/b0b44dbe-1e40-4624-bdb1-e87bc8016106.

[10] See Council Directive 85/374/EEC of 25 July 1985 on the approximation of the laws, regulations and administrative provisions of the Member States concerning liability for defective products, OJ L 210, 7.8.1985, 29–33

[11] See in this respect Ernst&Young, Technopolis and VVA (2018). Evaluation of Council Directive 85/374/EEC on the approximation of laws, regulations and administrative provisions of the Member States concerning liability for defective products. Brussels, European Commission., pp. 19-21.

[12] For further reference please see Bertolini, A. (2020). Artificial Intelligence and Civil Liability. Bruxelles, European Parliament - Committee on Legal Affairs**:** 1-132.

[13] Greenberg, B. A. (2016). "Rethinking Technology Neutrality." Minnesota Law Review 100:1495., 1513. Also see(1999). Communication from the Commission to the European Parliament, the Council, the Economic and Social Committee and the Committee of the Regions, Towards a new Framework for Electronic Communications Infrastructure and Associated Services: the 1999 Communications Review COM (1999) 539 final, 10 November 1999 Brussels, Commission of the European Communities., 14: "Technological neutrality means that legislation should define the objectives to be achieved and should neither impose, nor discriminate in favour of, the use of a particular type of technology to achieve those objectives".

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

European Parliament recognizes that different rules should be adopted for different applications, as it can be seen from its 2020 Report with recommendations to the Commission on a civil liability regime for artificial intelligence whereby the European Parliament distinguishes between high- and low- risk technologies with respect to liability rules (European Parliament 2020). Moreover, legal rules will indirectly affect technology, and thus "discriminate" between certain features of advanced applications, since liability and safety rules shape and incentivize both the development and the adoption of certain technologies over others and impact the kind of technology that ultimately prevails and their adoption (Bertolini 2020). For example, while in the case of surgical robots it is justified to hold liable both the producer, and the medical doctor as professional operator, for fully autonomous driverless cars holding the operator/driver liable, may not only be unjustified, but also impede the adoption of and trust in autonomous driving solutions since fear for the human user to be held responsible may act as a deterrent for the early uptake of such technology (Bertolini and Riccaboni 2020).

Since legal norms do not function in a vacuum, but directly affect the market and its innovation, it is evident that liability and safety rules impact the design of new technologies and discriminate between technologies by easing or impeding their diffusion, which in turn renders ineffective the envisioned and desired functions of technology-neutral regulation.

2. Robots are products

2.1. Setting the scene for future regulation. Robots are products

The robots' features of being autonomous and "intelligent" gave rise to many debates with respect the need of constructing new liability rules (Calo 2015), yet such a necessity may be argued on both ontological and functional grounds. The same applies for AI, which is more often than not, embedded into robotics applications, with respect to which, in many cases, the "intelligent" feature is seen as the AI's ability to perform functions that are generally associated with human intelligence, such as reasoning, learning and self-improvement[14].

Ontological arguments justify the need for a new liability framework on the basis of the intrinsic characteristics of the machine (e.g. its autonomy, and/or ability to learn) that would require us to consider it as a subject and not as a mere object of the law. Functional arguments formulate a policy claim for legal reform whenever the existing incentive structure is deemed inadequate, eventually failing to ensure the victim's compensation, due to a number of possible considerations, not merely related to the technical features of the application itself.

Ontological grounds are to be radically dismissed. Robots could be deemed as subjects of law, and thence responsible only if they possessed moral agency defined as the ability to perceive one own's existence, possess individual preferences, and be able to coordinate actions in order to achieve the intended result. Such a strong autonomy (Gutman, Rathgeber et al. 2012) is not displayed by any

[14] See Smart Dubai. (2019). "Artificial Intelligence Principles& Ethics." from https://www.smartdubai.ae/initiatives/ai-principles-ethics. Bertolini, A. (2020). Artificial Intelligence and Civil Liability. Bruxelles, European Parliament - Committee on Legal Affairs**:** 1-132.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

existing or reasonable foreseeable application, nor is to be deemed a desirable feature to pursue in the design of advanced technologies.

Short of that, machines only possess weak autonomy (Gutman, Rathgeber et al. 2012), defined as the ability to act even under great degrees of autonomy and absent human constant supervision and understanding of the inner mechanisms and processes delivering a specific outcome, yet pursuing and performing a task for which it was specifically created and/or used.

Moreover, even if a robot's "behaviour" may deviate from its standard settings when build with self-learning capabilities, such outcomes can be mitigated and/or eliminated through design, testing and security measures, as per the current product safety and liability legislation, which in turn excludes in any case the so-called "responsibility gap".

Therefore, robots cannot be considered moral agents. Instead, robots are mere objects, that is «artefacts crafted by human design and labor, for the purpose of serving identifiable human needs» (Bryson and Kime 2011, Bertolini 2013). This conclusion excludes any necessity of reforming liability rules based on ontological reasons.

If ontological claims need thence to be altogether disregarded as implausible, the same may not be said for functional ones. A number of policy considerations may eventually be formulated with respect to a multitude of advanced technological applications.

2.2. Robots as legal subjects from in functional perspective only

So conceived, a proposal, even the European Parliament's, to attribute – in certain cases only – "electronic personhood"[15] to more advanced application may be read in a functional perspective.

The debates carried out by scholars, policymakers and the media alike, all based on ontological and sometimes sentimental reasons[16], however, created a background of controversy, whereby the idea of making robots legal subjects was equated with granting them human rights.

Against this background, the European Parliament's recommendation – presented in the 2017 European Parliament Resolution on Civil Law Rules on Robotics ("CLRR")[17] – on creating an "electronic personhood" – was subject to outright dismissal. Under the CLRR, the European Parliament stated that:

"creating a specific legal status for robots in the long run, so that at least the most sophisticated autonomous robots could be established as having the status of electronic persons responsible for

[15] See in this respect European Parliament (2017). European Parliament resolution of 16 February 2017 with recommendations to the Commission on Civil Law Rules on Robotics. 2015/2103(INL), European Parliament. whereby the European Parliament called on the Commission "when carrying out an impact assessment of its future legislative instrument, to explore, analyze and consider the implications of all possible legal solutions".

[16] See in this respect the statement of Stephen Hawking that "the development of full artificial intelligence could spell the end of the human race", available at https://www.bbc.com/news/technology-30290540 (last accessed February 2021).

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

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

making good any damage they may cause, and possibly applying electronic personality to cases where robots make autonomous decisions or otherwise interact with third parties independently"[18].

Many criticized this recommendation, including the European Economic and Social Committee[19] and other intellectuals from the legal and scientific community[20]. Some even argued, among other things, that the attribution of legal personhood would either impair the preventive remedial effect of liability, or would result in granting robots human rights.

Similarly, t the Expert Group on Liability and New Technologies (the "EG") appointed in 2019 by the European Commission, denied the necessity to create and adopt a new form of legal personhood for robots and AI, stating among others, that: (i) it would first raise certain ethical issues by granting to new technologies rights; (ii) it would be superfluous, as the laws directed at the individuals manufacturing, operating and/or owning the machine would be better suited for tackling the issue of liability for damages arising from the use of robots and AI; (iii) it would disrupt and burden the legal system, as it would imply creating an entire new legislative framework for allowing electronic agents to own assets, which is a prerequisite for any liability system[21].

The above referred criticism may be grouped into two categories, namely those rejecting the idea of electronic personhood based on ontological grounds (e.g. this is the case for the grounds provided in the open letter[22] and partially in the EG's report when referring to "ethical issues") and on functional grounds (e.g. the futility and possibly disruptive effect on extant liability and legislative frameworks).

With respect to the first category (ontological), it was already clarified that machines do not satisfy the philosophical and legal arguments to be deemed agents, and thence, subjects of the law and bearers of rights. Such considerations should suffice to overcome all concerns of ambiguity and exclude the possibility to equate machines to human beings or even simply animals.

Indeed, the concept of legal personhood is based entirely on functional grounds, and the legal entity was never, and it is not considered to be a moral, but an economic agent, indispensable for the functioning of the modern economy.

The attribution of legal personhood, in the majority of legal systems, with very little variations, may be justified through reference to the following purposes and or characteristics, that legislators

[18] European Parliament (2017). European Parliament resolution of 16 February 2017 with recommendations to the Commission on Civil Law Rules on Robotics (2015/2103(INL)., paragraph 59.

[19] Muller, C. (2017). Opinion of the European Economic and Social Committee on 'Artificial intelligence - The consequences of artificial intelligence on the (digital) single market, production, consumption, employment and society' (own-initiative opinion) INT/806 Artificial intelligence. European Economic and Social Committee., paragraph 1.12.

[20] The open letter may be found at http://www.robotics-openletter.eu/ (last accessed February 2021).

[21] Expert Group on Liability and New Technologies (2019). Report on Liability for Artificial Intelligence and other emerging digital technologies. Brussels, European Commission., p. 38. Also see Bertolini, A. (2020). Artificial Intelligence and Civil Liability. Bruxelles, European Parliament - Committee on Legal Affairs**:** 1-132.

[22] The open letter may be found at http://www.robotics-openletter.eu/ (last accessed February 2021).

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

pursued by enacting this notion, namely: (i) coordination and simplification; (ii) separation of assets and limited liability; (iii) the possibility to conceive different taxation regimes[23].

Legal personhood is, in fact, a coordination vehicle that serves as a nexus of contracts by entering into and performing contracts with consumers, employees, services and products providers, and even with state authorities[24]. Although such contractual economic relationships could be replicated via multiple webs of contracts, the firm appears as an alternative and more efficient method to manage production due to reduced transaction costs[25], which also benefits the companies' creditors since the firm acts as a one-stop-shop for liability claims.

Assets' separation allows the company to own assets, separately from its founders and investors, and to pledge them for the benefit of its creditors who will have the possibility of recovering damages against said assets, without having to compete with the founders' or investors' personal creditors. Moreover, the assets brought into the company as capital contribution cannot be withdrawn by the individuals who subscribed such assets. These two essential features are also known as "affirmative asset partitioning" and "liquidation protection"[26].

With respect to the limited liability of the company to the assets it owns, this legal construct functions as an incentive for economic development and for fertile investments. This limited liability is what allows investors to infuse capital in financially risky activities (e.g. investing money in a small spin-off company creating new and revolutionary robotic products), which in turn allows society to achieve the subsequent economic welfare.

Although the firm's limited liability may be regarded as a negative aspect, it is worth mentioning that, first, liability caps are a usual and important aspect of the modern economy, and are provided in many different pieces of legislation[27], including under the Product Liability Directive (see art. 16 PLD)[28] applicable to robots and the 2020 European Parliament's Report with recommendations to the Commission on a civil liability regime for artificial intelligence (see art. 5 under the Annex to the motion for a resolution)[29]. Therefore, the person entitled to damages arising from the use of robotics and related technologies, would first and foremost be limited to recovering damages by the liability

---

[23] Bertolini, A. (2020). Artificial Intelligence and Civil Liability. Bruxelles, European Parliament - Committee on Legal Affairs**:** 1-132.

[24] Kraakman, R., P. Devies, H. Hansmann, G. Hertig, K. Hopt, H. Kanda and E. Rock (2006). The Anatomy of Corporate Law: A Comparative and Functional Approach. and Bertolini, A. (2020). Artificial Intelligence and Civil Liability. Bruxelles, European Parliament - Committee on Legal Affairs**:** 1-132.

[25] Coase, R. H. (1937). "The Nature of the Firm." Economica **4**(16)., p. 388.

[26] Kraakman, R., P. Devies, H. Hansmann, G. Hertig, K. Hopt, H. Kanda and E. Rock (2006). The Anatomy of Corporate Law: A Comparative and Functional Approach., 424.

[27] For example, this is the case of aircraft liability as stated under Article 11 of the 1952 Rome Convention on Damage Caused by Foreign Aircraft To Third Parties on the Surface, and of the liability for nuclear installations, as stated under Article 7 of the 1960 Paris Convention on Third Party Liability in the Field of Nuclear Energy.

[28] See Council Directive 85/374/EEC of 25 July 1985 on the approximation of the laws, regulations and administrative provisions of the Member States concerning liability for defective products, OJ L 210, 7.8.1985, 29–33

[29] European Parliament (2020). Report with with recommendations to the Commission on a civil liability regime for artificial intelligence (2020/2014(INL)). Plenary sitting, European Parliament.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

caps provided for under specific legislation, rather than due to the electronic legal personhood of the machine itself.

Second, the limited liability of legal persons can be surmounted by piercing the corporate veil[30] or by regulating new mechanism, such as that of posting a bond "equal to the highest reasonable estimate of the probable extent of its tort liability"[31].

This being said, it is clear that when addressing the suitability of creating legal personhood, the arguments against and for such a solution should be focused on these economic and market specific grounds.

Similar, if not identical considerations could thence justify a similar solution for certain robotic applications despite it being always possible to identify a relevant human behind it that could be deemed – more or less – directly responsible for the harmful outcome.

Electronic personhood allows for the identification of a single entry point of all litigation providing another alternative to the one-stop-shop solution represented by the *ex ante* identification of one responsible party among those possibly involved in the development and use of the application (e.g. producer, service provider, personal and professional user). On legal personhood, it has been rightfully argued that it is a more effective framework of sanctioning and remedying acts and omissions of both a criminal and civil nature. Thus, it has been said that liability against the legal entity has a stronger preventive effect by sanctioning the real decision makers, and brings about the benefit identifying rapidly the responsible party, since absent such personality, it would be more difficult to determine who committed the offence, given the often complex management structure and multiple layers of decisions found in firms[32].

To identify the responsible party in alternative causation scenarios, typical of advanced technologies – where multiple parties are involved in manufacturing the same product, by embedding different stand-alone components into a final machine –, often represents an obstacle for the victim attempting to identify party that should be required to compensate for the damage suffered.

In this sense, electronic personhood would simplify *prima facie* litigation and thence victim's compensation by creating a single-entry point for litigation whereby the victim could request compensation from one single party, that is the electronic legal person, and thus reduce litigation time and costs for the victim, and ease its burden of proof. The entities or individuals acting as subscribers or founders of the electronic person would participate and share all liabilities pursuant to their agreement included in the articles of association and/or incorporation.

---

[30] See in this respect the USA case-law Walkovsky v Carlton, where it is stated that "the courts will disregard the corporate form, or, to use accepted terminology, "pierce the corporate veil", whenever necessary "to prevent fraud or to achieve equity" and "whenever anyone uses control of the corporation to further his own rather than the corporation's business, he will be liable for the corporation's acts "upon the principle of *respondeat superior* applicable even where the agent is a natural person".

[31] Posner, R. A. (2007). Economic Analysis of Law, Wolters Kluwer Law & Business., 438 and Bertolini, A. (2020). Artificial Intelligence and Civil Liability. Bruxelles, European Parliament - Committee on Legal Affairs**:** 1-132.

[32] Beck, S. (2014). Corporate Criminal Liability. The Oxford Handbook of Criminal Law. M. D. Dubber and T. Hörnle. United States of America, Oxford University Press. Bertolini, A. (2020). Artificial Intelligence and Civil Liability. Bruxelles, European Parliament - Committee on Legal Affairs**:** 1-132.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

Moreover, electronic personhood could, if accurately tailored, represent a suitable mechanism for greater transparency and accountability about the different stakeholders involved in the device production or service operation, through registration and disclosures duties typically imposed upon legal entities[33].

Additionally, electronic personhood could: (i) ease the exploitation and benefits' allocation of profits arising from intellectual property rights, including those on works created by the systems themselves[34], and (ii) under the, possibly indispensable shield of limited corporate liability, it could allow for the fast development of innovative solutions financed and funded by investors which absent such a cap would not undertake unlimited liability risks[35].

3. Liability and the Risk-Management Approach

3.1. Liability. General considerations

Liability rules ensure that the wrongdoer is responsible for the harm caused, under civil, administrative and criminal law. Under criminal liability, the wrongdoer, which is prosecuted by state agents, may be sanctioned with fines, imprisonment and/or accessory measures, when its acts and omissions represent a crime/offence, that is when the latter fulfil all the objective and subjective elements required by the norm. Under administrative law, the sanctions are financial in nature and are imposed by agents of the public administration with the aim of sanctioning and/or compensating for the harm caused.

Civil liability rules are aimed at determining the person who should bear the negative consequences of a certain act or omission, usually by burdening the latter with the obligation to compensate the damages arising therefrom, while requiring the victim with proving the liability constitutive elements (e.g. causation, the damage, the wrongdoer's fault, etc.).

Liability mechanisms have two functions, namely: (i) *ex ante* deterrence by forcing the wrongdoer to incentivize the cost he creates, and (ii) *ex post* victim's compensation based on the corrective justice theory whereby the breach of legal duties gives rise to so-called second order duties to compensate[36].

These rationales or combinations thereof give rise to different variations of liability systems or mechanisms, namely: (i) liability based on fault specific to the classical tort law systems which aims both at sanctioning, deterring and compensating for the wrongful act; (iii) strict or semi-strict liability rules grounded on the idea that due to the particular position the wrongdoer holds towards the

[33] See in this respect Chapter III under Directive (EU) 2017/1132 of the European Parliament and of the Council of 14 June 2017 relating to certain aspects of company law (Text with EEA relevance.) OJ L 169, 30.6.2017, 46–127 and Bertolini, A. (2020). Artificial Intelligence and Civil Liability. Bruxelles, European Parliament - Committee on Legal Affairs**:** 1-132.

[34] See European Parliament (2020). Report on intellectual property rights for the development of artificial intelligence technologies (2020/2015(INI)), Plenary sitting.

[35] Similarly see Bertolini, A. (2020). Artificial Intelligence and Civil Liability. Bruxelles, European Parliament - Committee on Legal Affairs**:** 1-132.

[36] Coleman, J., S. Hershovitz and G. Mendlow (Winter 2015). Theories of the Common Law of Torts. The Stanford Encyclopedia of Philosophy. E. Zalta. https://plato.stanford.edu/archives/win2015/entries/tort-theories/.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

source of the damage, fault is irrelevant and only objective external elements could exclude liability, such as acts of God and force majeure. Law and economics theories at times suggest one rationale ought to prevail over the other, and in this domain, eventually, that liability should be borne by the party that is best positioned to identify and minimize the risks arising from a certain activity and to compensate for the damages arising therefrom by internalizing the costs associated therewith[37].

3.2. Strict and/or semi-strict liability rules

As previously mentioned, robots are products, thence the EU product liability directive applies[38].

The PLD is the cornerstone piece of legislation governing the compensation of damages arising from the use of a defective product[39], and for the subject matter here considered, robots. Pursuant to art. 1 PLD), liability is incumbent upon the producer – "the manufacturer of a finished product, the producer of any raw material or the manufacturer of a component part and any person who, by putting his name, trade mark or other distinguishing feature on the product presents himself as its producer", or the importer of a product within the European Union, and the seller of the product – in case the producer cannot be identified (art. 1 and art. 3 PLD).

As per art. 6 PLD, a defective product is one which "does not offer the safety that a person is entitled to expect, considering all circumstances", such as the presentation of the product, its reasonably expected use, and the time in which it was put into circulation.

The PLD offers redress to consumers for damages caused by death or by personal injuries and/or damage to, or destruction of, any item of property other than the defective product itself, provided that the item of property (i) is of a type ordinarily intended for private use or consumption, and (ii) was used by the injured person mainly for his own private use or consumption (art. 9 PLD). In order to obtain compensation, the victim must prove the damage, the defect and that there is a causal nexus between the defect and the damage which compensation is sought for.

Despite often being described as a form of strict liability, the one set by the PLD is actually a form of semi-strict liability, since manufacturers may escape liability by proving one of the defences put forth by art. 7 PLD. Most prominently, art. 7 (e) PLD, introduces the so-called development risk defence, whereby the duty to compensate is excluded, despite the product being indeed defective, so long as "the state of scientific and technical knowledge at the time when he put the product into circulation was not such as to enable the existence of the defect to be discovered". Pursuant to scholars and case-law[40] this could equate the application of a rule of presumed fault, or a simple

[37] Bertolini, A. (2016). "Insurance and Risk Management for Robotic Devices: Identifying the Problems." Global Jurist **16**(3): 291-314.

[38] See Council Directive 85/374/EEC of 25 July 1985 on the approximation of the laws, regulations and administrative provisions of the Member States concerning liability for defective products, OJ L 210, 7.8.1985, p. 29–33.

[39] As per art. 3 PLD, "'product' means all movables, with the exception of primary agricultural products and game, even though incorporated into another movable or into an immovable. 'Primary agricultural products' means the products of the soil, of stock-farming and of fisheries, excluding products which have undergone initial processing. 'Product' includes electricity".

[40] Timan, T., R. Snijders, M. Kirova, S. Suardi, M. v. Lieshout, M. Chen, P. Costenco, E. Palmerini, A. Bertolini, A. Tejada, S. v. Montfort, M. Bolchi, S. Alberti, R. Brouwer, K. Karanilokova, F. Episcopo and S. Jansen (2019). Study on safety of non-embedded software. Service, data access, and legal issues of advanced

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

reversal of the burden of proof with respect to the element of negligence. Indeed, a judgement of negligence is based upon a two-prong assessment, namely: (i) determining whether the relevant knowledge and expertise is possessed by even one single man, and (ii) whether that knowledge could have been demanded of the specific agent involved[41].

Art. 7 (e) PLD excludes negligence on the basis of (i) above, and therefore might be considered as equivalent a solution to a rule of responsibility for fault, despite with an altered distribution of the burden of proof. Such a conclusion, however, clashes with the statement contained in the second recital of the PLD, whereby "liability without fault on the part of the producer is the sole means of adequately solving the problem, peculiar to our age of increasing technicality of a fair apportionment of the risks inherent in modern technological production".

Furthermore, producer's liability may also be reduced when victim's contributory negligence is demonstrated.

The PLD is intended to set a regime of maximum harmonization for product liability claims. However, due to the fact that Member States were able to implement it with certain variations deriving from their national different systems of liability, a significant margin of fragmentation was created[42].

Moreover, 30 years have passed since the PLD's enactment, which raises the question whether the PLD is still fit for purpose to tackle emerging challenges associated with ever more complex products. In this respect, the latest PLD official evaluation (European Commission 2018) and adjacent study (Ernst&Young, Technopolis et al. 2018) provided an empirical assessment of the PLD's application across Member States.

Furthermore, the European Commission appointed an Expert Group to evaluate the applicability of the PLD to traditional products and new technologies (Expert Group on Liability and New Technologies, 2019).

The aforementioned reports and studies show, that although the PLD may be deemed adequate overall to face the challenges brought about by existing products, certain criticalities become at the same time apparent from the empirical data and analysis provided therein, which hinder the PLD's suitability with respect to advanced and complex technologies.

Specifically, the PLD has in practice a limited scope of application, albeit the technology neutrality principle on which it is based. This is as evidenced by the limited number of cases litigated as indicated in the PLD's evaluation report, which instead are solved in out-of-courts settlements[43].

---

robots, autonomous, connected, and AI-based vehicles and systems: final study report regarding CAD/CCAM and industrial robots. Brussel, European Commission.

[41] Padovani, T. (2002). Diritto Penale. Milano, Giuffrè.

[42] For an overview of the directive and its implementation among Member States, see Timan, T., R. Snijders, M. Kirova, S. Suardi, M. v. Lieshout, M. Chen, P. Costenco, E. Palmerini, A. Bertolini, A. Tejada, S. v. Montfort, M. Bolchi, S. Alberti, R. Brouwer, K. Karanilokova, F. Episcopo and S. Jansen (2019). Study on safety of non-embedded software. Service, data access, and legal issues of advanced robots, autonomous, connected, and AI-based vehicles and systems: final study report regarding CAD/CCAM and industrial robots. Brussel, European Commission. and Machnikowski, P., Ed. (2016). European Product Liability. An Analysis of the State of the Art in the Era of New Technologies. Cambridge, Intersentia.

[43] See in this respect Ernst&Young, Technopolis and VVA (2018). Evaluation of Council Directive 85/374/EEC on the approximation of laws, regulations and administrative provisions of the Member States concerning

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

Furthermore, the report shows that for a period of 16 years spanning from 2000 to 2016, only 798 claims were formulated based on product liability rules[44]. And that, in 20% of these cases, consumers were granted redress by the courts based on national tort or contract liability rules, and not product liability rules implementing the PLD. These figures and findings may show that consumers either struggle to find protection under the PLD, or do not seek it altogether, and prefer to resort to alternatives such as tort or contract law, and even national non-harmonized rules and rules on product conformity[45].

Moreover, the PLD's narrow practical application is also demonstrated by the very specific domains and products where the PLD is being applied. Thus, the PLD's evaluation report shows that the highest number of cases litigated involves either raw materials, that is technology simple products, or pharmaceuticals and vehicles, that is products which involve both considerable harms with respect to the type of damages caused and/or sophisticated parties[46].

Also, it is still unclear whether software could be included in the notion of product. As indicated by the EG, "it is also questionable whether software is covered by the legal concept of product or product component" and "it is particularly discussed whether the answer should be different for embedded and non-embedded software, including over-the-air software updates or other data feeds"[47]. This creates major uncertainty on the PLD's applicability, especially with respect to technologically advanced products which display both hardware and software components.

Additionally, the victim's burden of proving the causal nexus between the defect and the damage substantially encumbrances the claimant, especially in cases of advanced robotics where determining that the product is defective, and that the harm is the consequence of a defect in the functioning of the device requires extensive data about the design and functioning of the product and technical expertise, and thus substantial litigation costs[48].

Furthermore, the PLD defences, in particular the development risk defence which allows producers to escape liability as previously mentioned, may no longer be suitable for the current status quo of technological development, where the producers control the functioning of the product even after being placed onto the market via over-the-air updates and maintenance service[49], and are thus able either to prevent and/or mitigate potential harms after the product was put into circulation.

---

liability for defective products. Brussels, European Commission., p. 18 whereby it is stated that "all categories of stakeholders indicated that extra-judicial arrangements represent a common way to settle cases, and that most cases are settled out of court".

[44] Ibid., p. 21.

[45] See in this respect Directive 1999/44/EC of the European Parliament and of the Council of 25 May 1999 on certain aspects of the sale of consumer goods and associated guarantees.

[46] See in this respect Ernst&Young, Technopolis and VVA (2018). Evaluation of Council Directive 85/374/EEC on the approximation of laws, regulations and administrative provisions of the Member States concerning liability for defective products. Brussels, European Commission., p. 20.

[47] Expert Group on Liability and New Technologies (2019). Report on Liability for Artificial Intelligence and other emerging digital technologies. Brussels, European Commission., p.28.

[48] See in this respect ibid. and Bertolini, A. (2020). Artificial Intelligence and Civil Liability. Bruxelles, European Parliament - Committee on Legal Affairs**:** 1-132.

[49] Similarly see the Twigg-Flesner, C. (2021). Guiding Principles for Updating the Product Liability Directive for the Digital Age. ELI Innovation Paper Series, European Law Institute.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

Thus, it can no longer be upheld that the current liability regime under the PLD fulfils its sought effectiveness and intended functions, including that of leading to a high standard of product safety through its deterrent effect.

Due to such criticalities and based on settled law and economic theories and research, an alternative model of liability, namely the Risk-Management-Approach (the "RMA"), as further developed below, shall be adopted, and implemented.

3.3. The Risk-Management Approach

The RMA attributes liability to the party that is best positioned to (i) identify a risk, (ii) control and minimize it through its choices, and (iii) manage it – ideally pooling and distributing it among all other parties – eventually through insurance, and/or no-fault compensation funds[50].

To do so, it resorts to strict – if not absolute – liability rules[51] and it differs from a fault-based liability rule because it does burden the party responsible for having departed from an intended conduct, but the best party to tackle the risks which may materialize by causing damages, even if such party might not be blameworthy[52].

To a certain extent, similar solutions are being taken into account at EU policy-making level[53].

The RMA departs from the traditional paradigm that liability can achieve both functions of ensuring compensation and of providing the accurate incentives for enhancing product safety, as showed under law and economics theories[54], and thus, it decouples product liability from product safety rules. Under the RMA, product safety is best achieved through narrow tailored *ex ante* regulation detailing relevant safety requirements, while the only function of liability is that of ensuring the victims' compensation[55].

In order to allow an efficient redress and compensation right for the victim and by placing the responsibility and liability on the best party to identify and manage the risks, the RMA creates a one-stop-shop for liability claims.

Such an approach is not novel to EU legislation. Indeed, the Consumer Sales and Guarantees Directive establishes the responsibility of the final seller towards the consumer whenever they lack

---

[50] Bertolini, A. (2020). Artificial Intelligence and Civil Liability. Bruxelles, European Parliament - Committee on Legal Affairs**:** 1-132., p. 99

[51] Ibid., p. 99.

[52] Ibid.

[53] See Expert Group on Liability and New Technologies (2019). Report on Liability for Artificial Intelligence and other emerging digital technologies. Brussels, European Commission. and European Parliament (2020). Report with with recommendations to the Commission on a civil liability regime for artificial intelligence (2020/2014(INL)). Plenary sitting, European Parliament.

[54] Posner, R. A. (2007). Economic Analysis of Law, Wolters Kluwer Law & Business. and Polinsky, M. A. and S. Shavell (2009-2010). "The uneasy case for product liability." Harvard Law Review **123**: 1437-1492.

[55] Bertolini, A. (2020). Artificial Intelligence and Civil Liability. Bruxelles, European Parliament - Committee on Legal Affairs**:** 1-132.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

conformity with the contract of sale, despite the former not being responsible for the lack of conformity displayed by the product itself[56].

Similarly, RMA eliminates the burden to pinpoint liability by identifying the one responsible party among the numerous potential contributors, including the producer, the owner, the user, and all connected service providers, cooperating in the functioning of the product or provision of the service, based upon the advanced technology. Such approach simplifies the structure responsibility in first instance litigation, leaving unaltered the possibility to pursue different agents involved through secondary litigation, acting in regress through other contractual tools. The effect pursued, of increased simplification for the claimant, counterbalances the the opacity and complexity of advanced technologies[57] which make it, if not impossible, at least considerably difficult for the victim to identify: (i) the cause of the defect, and thence (ii) the responsible party, and (iii) even the causal nexus between the defect and the damage, having in the end to carry out a *probatio diabolica*.

Instead, the RMA clearly identifies the party who, *prima facie,* should be called in to compensate for the damage suffered[58], thus easing access to justice.

The party required to compensate for the caused damages shall not necessarily be the party who bears the economic consequences of the harmful activity. Thus, additional functional mechanisms should be attached to a party's strict – or even absolute – liability, such as the enactment of an adjacent right to sue in recourse the other parties contributing to the marketing of the defective and harmful product, via contractual and/or non-contractual claims. Such a right to sue in recourse allows the primary responsible party to efficiently internalize and then distribute the costs along the entire value chain, to the party that is specifically in control of the one risk that materialized[59].

Additionally, under RMA first- or third-party insurance is being considered. The duty to purchase and maintain such insurance coverage would rest upon the party having the obligation under the law to compensate for the damages caused by a defective product. In this manner, the damages would be paid by the insurer, and would thus allow the liable party to transform the *ex post* costs of liability into *ex ante* predictable and manageable insurance costs, such as premiums. These cost externalities can be subsequently transferred via price mechanisms to the users (e.g. such as a products price increase), benefiting from a certain service, product or activity[60].

Nevertheless, creating and adopting compulsory insurance mechanism for this purpose and at the current state of the art, may also prove problematic. Certain factual limitations may affect, yet, the creation of a mature and efficient insurance market for advanced technologies. First, legal risks might be hard to define, especially in cases of multiple potential tortfeasors and of absence of clear-cut liability rules such as those referenced above with respect to the PLD (e.g. multiple alternative

[56] See in this respect art. 3 of Directive 1999/44/EC of the European Parliament and of the Council of 25 May 1999 on certain aspects of the sale of consumer goods and associated guarantees OJ L 171, 7.7.1999, p. 12–16. Moreover, art. 4 also provides the final seller's right of redress against the producer.

[57] See in this respect Expert Group on Liability and New Technologies (2019). Report on Liability for Artificial Intelligence and other emerging digital technologies. Brussels, European Commission., p. 28.

[58] Bertolini, A. (2020). Artificial Intelligence and Civil Liability. Bruxelles, European Parliament - Committee on Legal Affairs**:** 1-132.

[59] Ibid.

[60] Bertolini, A. (2016). "Insurance and Risk Management for Robotic Devices: Identifying the Problems." Global Jurist **16**(3): 291-314.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

causation scenarios, producer's defenses)[61]. Second, the absence of statistically relevant risk data may curtail the creation of a mature and stable insurance market[62]. Third, the difficulty in identifying *ex ante* unknown risks, and in assessing their probability, due to advanced technologies' capacity to self-modify their functioning over time[63], may preclude the creation of suitable insurance packages. Forth, the limited market for specific classes of applications and the lack of limited benchmarking data may cause concerns and limitations to the uptake of said insurance markets[64].

These functional and practical caveats render the conclusion that compulsory insurance may not be the best suitable solution for all classes of applications and that a technology specific approach should be pursued.

On the victims' side, the RMA also envisions the creation of no-fault compensation funds, publicly or privately operated or managed, and possibly funded through taxation or through compulsory contributions. Such funds would be necessary to allow and ease a fast and efficient compensation, in cases where the damage was caused by an unidentified or uninsured technology and/or tortfeasor (e.g. such as in the case of damages caused by hacking techniques or due to cybercrime)[65].

Indeed, the RMA avoids the technology neutrality and one-size-fits-all approach paradigms, and it employs a technology specific approach. Although, robotics and AI are seen by the general public and policy makers alike as a single entity, and thus, these two notions are used as umbrella terms to unitarily denominate, and eventually regulate, vast and diverse classes of applications, this trend is highly problematic because there are many devices that might be deemed robotics- and AI-based, so diverse from one another as driverless cars, a smart-toothbrush, a robot-companion, a non-embedded expert system for medical diagnosis, all of which may give rise to different types of harms and risks[66]. This sheer diversity severally challenges the aim of drafting rules suitable for the intended purpose that should be neither over- nor under- inclusive, since these diverse classes of applications differ profoundly among one another on technical grounds and with respect to the ethical and legal implications they give rise to.

---

[61] Bertolini, A. (2020). Artificial Intelligence and Civil Liability. Bruxelles, European Parliament - Committee on Legal Affairs**:** 1-132.

[62] With respect to the lack of performance date for autonomous vehicles, please see Kalra, N. and S. M. Paddock (2016). "Driving to Safety. How Many Miles of Driving Would It Take to Demonstrate Autonomous Vehicle Reliability?" Transportation Research Part A: Policy and Practice **94**: 182-193.

[63] Expert Group on Liability and New Technologies (2019). Report on Liability for Artificial Intelligence and other emerging digital technologies. Brussels, European Commission.

[64] Bonsignorio, F., E. Messina and A. P. del Pobil (2014). "Fostering Progress in Performance Evaluation and Benchmarking of Robotic and Automation Systems." IEEE Robotics & Automation Magazine **21**(1): 22-25. Bertolini, A. (2015). "Robotic prostheses as products enhancing the rights of people with disabilities. Reconsidering the structure of liability rules." International Review of Law, Computers & Technology **29**(2-3): 116-136. and Bertolini, A. (2020). Artificial Intelligence and Civil Liability. Bruxelles, European Parliament - Committee on Legal Affairs**:** 1-132.

[65] Also see Expert Group on Liability and New Technologies (2019). Report on Liability for Artificial Intelligence and other emerging digital technologies. Brussels, European Commission., p. 62.

[66] Bertolini, A. (2020). Artificial Intelligence and Civil Liability. Bruxelles, European Parliament - Committee on Legal Affairs**:** 1-132.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

The incentive structure we might want to conceive for the use of advanced technologies in the medical profession most likely differs from that we could design for intermediaries operating in the capital markets.

As of today, all legal systems address those domains separately, and provide different incentive structures. Avoiding defensive medicines practices has induced legislators to intervene and – at least attempt to – shield medical professionals from liability, eventually primarily burdening the hospitals within which they operate[67].

To the contrary, intermediaries are typically and frequently strictly liable, due to the relevant role they play on financial markets and the control they might exert over a number of operations. If both were applied the identical standard of liability, by qualifying each one as an "operator"[68], very different outcomes could be observed, not equally desirable. In the medical profession, the uptake of said – otherwise beneficial technology – could be discouraged, since a practitioner using them would be held to a comparably much more severe standard of liability, as opposed his colleagues not using less sophisticated applications[69].

4. Product safety regulation

4.1. Extant EU product safety legal framework

Product safety rules aim at balancing opposing interests, namely that of promoting and ensuring the development of safe products and easing their distribution into the market (Palmerini, Bertolini et al. 2016). At Union level, product safety and product liability provisions are two complementary mechanisms to pursue the same policy goal of a functioning single market for goods that ensures high levels of safety, i.e. minimise the risk of harm to users and provides for compensation for damages resulting from defective goods[70].

The European product safety framework consists of both general and specific applicable rules which set forth mandatory and statutory safety requirements under EU or national law. This framework is

[67] See in this respect Legge 8 marzo 2017, n. 24, Disposizioni in materia di sicurezza delle cure e della persona assistita, nonche' in materia di responsabilita' professionale degli esercenti le professioni sanitarie (17G00041).

[68] See in this respect art. 4 of the European Parliament (2020). Report with with recommendations to the Commission on a civil liability regime for artificial intelligence (2020/2014(INL)). Plenary sitting, European Parliament. and art. 3 of the same, whereby "operator" is defined as "both the frontend and the backend operator as long as the latter's liability is not already covered by Directive 85/374/EEC", while "frontend operator" is defined as any natural or legal person who exercises a degree of control over a risk connected with the operation and functioning of the AI-system and benefits from its operation" and "backend operator" is defined as "any natural or legal person who, on a continuous basis, defines the features of the technology and provides data and an essential backend support service and therefore also exercises a degree of control over the risk connected with the operation and functioning of the AI-system".

[69] Additionally, see WP2.

[70] European Commission (2020). Report from the Commission to the European Parliament, the Council and the European Economic and Social Committee. Report on the safety and liability implications of Artificial Intelligence, the Internet of Things and robotics. COM(2020) 64 final. Brussels, European Commission.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

the result of the adoption of the so called "New Approach"[71] resolutions, and the subsequent "New Legislative Framework Approach"[72] regulations.

At EU level, the safety general requirements are determined by the General Product Safety Directive ("GPSD")[73]. As per the GPSD, producers of products, and more specifically of robots[74], are obliged to: i) ensure that products placed on the marker are safe; ii) inform consumers of any risks associated with the products supplied; and iii) take corrective action when that products prove to be unsafe.

While enumerating all the legislative acts setting forth specific security requirements would fall outside the scope of this research, the below directives are of specific relevance to the field of interactive robotics, namely:

- The Machinery Directive[75], applicable to certain classes of robotics applications which may be considered "machinery" or "partly completed machinery";

[71] Council Resolution of 7 May 1985 on a new approach to technical harmonization and standards, OJ C 136, 4.6.1985, 1–9.

[72] Regulation (EC) No 765/2008 of the European Parliament and of the Council of 9 July 2008 setting out the requirements for accreditation and market surveillance relating to the marketing of products and repealing Regulation (EEC) No 339/93, OJ L 218, 13.8.2008, 30–47; Decision No 768/2008/EC of the European Parliament and of the Council of 9 July 2008 on a common framework for the marketing of products, and repealing Council Decision 93/465/EEC, OJ L 218, 13.8.2008, 82–128; Regulation (EC) No 764/2008 of the European Parliament and of the Council of 9 July 2008 laying down procedures relating to the application of certain national technical rules to products lawfully marketed in another Member State and repealing Decision No 3052/95/EC, OJ L 218, 13.8.2008, 21–29.
For an overview of this approach, see https://ec.europa.eu/growth/single-market/goods/new-legislative-framework_en (last accessed February 2021), and, more in detail, European Commission (2016). The 'Blue Guide' on the implementation of EU products rules 2016. For a description and an assessment of the product safety framework in the field of industrial robots, see Timan, T., R. Snijders, M. Kirova, S. Suardi, M. v. Lieshout, M. Chen, P. Costenco, E. Palmerini, A. Bertolini, A. Tejada, S. v. Montfort, M. Bolchi, S. Alberti, R. Brouwer, K. Karanilokova, F. Episcopo and S. Jansen (2019). Study on safety of non-embedded software. Service, data access, and legal issues of advanced robots, autonomous, connected, and AI-based vehicles and systems: final study report regarding CAD/CCAM and industrial robots. Brussel, European Commission., Annex 3, Task 3&4.

[73] Directive 2001/95/EC of the European Parliament and of the Council of 3 December 2001 on general product safety, (Text with EEA relevance), OJ L 11, 15.1.2002, p. 4–17.

[74] As per art. 2 (a) of the GPSD, ""product" shall mean any product - including in the context of providing a service - which is intended for consumers or likely, under reasonably foreseeable conditions, to be used by consumers even if not intended for them, and is supplied or made available, whether for consideration or not, in the course of a commercial activity, and whether new, used or reconditioned". Under such a definition, most robotics applications will likely fall.

[75] Directive 2006/42/EC of the European Parliament and of the Council of 17 May 2006 on machinery, and amending Directive 95/16/EC, in OJ L 157, of June 9th, 2006. The directive defines "machinery" 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", and "partly completed machinery" is defined as "an assembly which is almost machinery but which cannot in itself perform a specific application".

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

- The Personal Protective Equipment Directive, or the Regulation repealing it[76], applicable in certain cases to components of robotics applications such as exoskeletons which may be considered as "personal protective equipment";

- The Low Voltage Directive[77], applicable to robotics applications complying with the applicability requirements provided therein;

- The Electromagnetic Compatibility Directive[78], virtually applicable to all robots which involve electricity and which may be deemed as "apparatus" or as "fixed installation", as defined therein.

- The Regulation on Medical Devices[79], repealing Directive 93/42/EEC on Medical Devices[80], and which includes a broadened definitions of regulated devices, now including new devices which can be related to the use of interactive robots, e.g. medical purpose devices, cleaning products, and liposuction equipment[81].

Albeit certain necessary differences are included in these sectorial safety directives, they all provide more or less for the same safety framework.

Thus, under this safety framework, a product is presumed to be safe it meets all mandatory and statutory safety requirements under EU or national law. When such safety requirements are not in place, products are deemed safe when they conform with national standards, Commission recommendations, codes of practice and best practice, state of the art and technology and /or reasonable consumer safety expectations.

The specific implementation of these essential safety requirements is not provided under the directives which allow the manufacturers to choose the best practical and technical manner for complying with and implementing such requirements. Instead, other relevant and technical

[76] Regulation (EU) 2016/425 of the European Parliament and of the Council of 9 March 2016 on personal protective equipment and repealing Council Directive 89/686/EEC (Text with EEA relevance), OJ L 81, 31.3.2016, p. 51–98. Personal Protective Equipment is defined as "equipment designed and manufactured to be worn or held by a person for protection against one or more risks to that person's health or safety".

[77] Directive 2014/35/EU of the European Parliament and of the Council of 26 February 2014 on the harmonization of the laws of the Member States relating to the making available on the market of electrical equipment designed for use within certain voltage limits Text with EEA relevance, in OJ L 96, 29.3.2014..

[78] Directive 2014/30/EU of the European Parliament and of the Council of 26 February 2014 on the harmonization of the laws of the Member States relating to electromagnetic compatibility (recast) Text with EEA relevance. See OJ L 96, 29.3.2014. The directive applies to any "apparatus", i.e. any finished appliance or combination thereof made available on the market as a single functional unit, intended for the end-user and liable to generate electromagnetic disturbance, or the performance of which is liable to be affected by such disturbance.

[79] Regulation (EU) 2017/745 of the European Parliament and of the Council of 5 April 2017 on medical devices, amending Directive 2001/83/EC, Regulation (EC) No 178/2002 and Regulation (EC) No 1223/2009 and repealing Council Directives 90/385/EEC and 93/42/EEC (Text with EEA relevance.), OJ L 117, 5.5.2017, p. 1–175.

[80] Council Directive 93/42/EEC of 14 June 1993 concerning medical devices, OJ L 169, 12.7.1993, p. 1–43.

[81] See in this respect INBOTS (2019). Preliminary White Paper on Standardization and Interactive Robots. http://inbots.eu/wp-content/uploads/2019/07/Attachment_0-3.pdf.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

specifications are set forth in national and international technical standards identifying the relevant best practices or state of the art techniques[82].

Applying and/or implementing standards is not mandatory, but such implementation could be a preferred solution for producers, especially with respect to harmonized technical standard (hEN), developed by a recognized European Standardization Organizations – CEN, CENELEC, or ETSI, that, if implemented, give rise to a presumption of the products' conformity with the essential safety requirements[83]. Instead, applying different technical solutions – not included in hEN – comes with the additional obligation on the side of the manufacturer of proving that said employed solutions comply with the mandatory product safety essential requirements.

Moreover, hEN standards, as opposed to standards which are adopted by other international organizations (e.g. ISO) ease the procedure for placing a product onto the market since, their implementation provides for a simplified and flexible procedure of conformity whereby in certain cases the producer is required to only issue a declaration of conformity.

Other products, usually depending on the risk their use carries, can be subject to mandatory certification – and not only to procedures of self-declarations of conformity – whereby compliance with the essential safety requirements is assessed by national authorized bodies. This procedure, when successfully completed, results in the issuance of a safety certificate.

In all cases, be it with respect to self-assessment of conformity or to certification procedures, the products placed onto the EU market must bear the "conformity mark" CE which indicates to consumers that products placed on the marked comply with the minimum mandatory and statutory safety requirements.

Nevertheless, obtaining certification does not imply an exemption from liability[84]. Moreover, certain safety-related measures can be adopted by surveillance authorities even after the product has been placed onto the market, if it later on is ascertained as unsafe, such as recalling the products from the market[85].

---

[82] It should be noted though that the ECJ stated in the James Elliott ruling that hENs are part of EU law, thus falling its own jurisdiction under Article 267 of the Treaty on the Functioning of the European Union. It is to be seen whether this means that the ECJ has the competence to merely interpret hEN and/or even to review them. See C-613/14, James Elliott Construction Limited v Irish Asphalt Limited, Judgment of the Court (Third Chamber) of 27 October 2016, ECLI:EU:C:2016:821.

[83] See in this respect art. 8 (1) of the Regulation (EU) 2017/745 of the European Parliament and of the Council of 5 April 2017 on medical devices, amending Directive 2001/83/EC, Regulation (EC) No 178/2002 and Regulation (EC) No 1223/2009 and repealing Council Directives 90/385/EEC and 93/42/EEC (Text with EEA relevance.), OJ L 117, 5.5.2017, p. 1–175 and art. 7 (1) of the Directive 2006/42/EC of the European Parliament and of the Council of 17 May 2006 on machinery, and amending Directive 95/16/EC, in OJ L 157, of June 9th, 2006.

[84] See in this respect recital 36 of the Directive 2001/95/EC of the European Parliament and of the Council of 3 December 2001 on general product safety (Text with EEA relevance), OJ L 11, 15.1.2002, p. 4–17, pursuant to which "this Directive should not affect victims' rights within the meaning of Council Directive 85/374/EEC of 25 July 1985 on the approximation of the laws, regulations and administrative provisions of the Member States concerning liability for defective products".

[85] See in this respect art. 3 (4) of Directive 2001/95/EC of the European Parliament and of the Council of 3 December 2001 on general product safety (Text with EEA relevance), OJ L 11, 15.1.2002, p. 4–17 and art. 95 of Regulation (EU) 2017/745 of the European Parliament and of the Council of 5 April 2017 on

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

4.2. The need to update product safety regulation

Despite being a well-established safety framework, the current legislation fails to address certain emerging risks specific to advanced technologies. Although some of these risks could not have been identified earlier, the EU should work towards setting forth a revised and refitted safety framework covering new, yet evident risks, such a mental, dignity and deception risks arising from the use of new technologies.

Such risks are apparent especially with respect to collaborative robotics.

Namely, the robots' ability to simulate empathy and intelligence effectively, and thus, deceive, is not only a criterion for assessing intelligence in machines[86], but also a feature that is sought by their producers and which is exhibited by many existing applications.

However, this technical capacity can give rise to many risks, especially when used by private users independently. In this respect, studies have shown that, especially in the case of vulnerable categories of people, namely children and the elderly, there is a blurred line between a robot companion developed to respond to the needs of the human and one that manipulates emotions[87].

With respect to children, it has been argued that the robots' autonomy and goal-oriented movement, their morphology and responsiveness to the environment, or simply the children's inference of intentionality of the action from its mere occurrence (Piaget 1951, Cameron, Fernando et al. 2017) (Bertolini 2018), and considered the status of their still fragile mental development, it may cause the misqualification of machines as living and sentient beings, equivalent to a pet.

Similarly, endearing robots' features may exploit human reward-mechanism and trigger the child into believing that a meaningful relationship may be developed with said robots (Kanda, Sato et al. 2007), ultimately deceiving the child and affecting his or her psychological development.

That way users might be induced to developing emotional attachments that are typical of a relationship – which however may only be established with a living being and primarily a human[88] – when, instead, they are part to a mere interaction with a technological product.

Comparable cognitive and psychological risks arise also with respect to the use of robots for elderly companionship and care, coupled with the risks of deception and impairments of their dignity. This is the case of robots eliciting emotional attachment and/or those which provide feedback in the form of sound, movement and light the more they are being engaged with, and thus exploit the same reward mechanism as in the case with children (Bertolini 2018)[89].

---

medical devices, amending Directive 2001/83/EC, Regulation (EC) No 178/2002 and Regulation (EC) No 1223/2009 and repealing Council Directives 90/385/EEC and 93/42/EEC (Text with EEA relevance.), OJ L 117, 5.5.2017, p. 1–175.

[86] Turing, A. (1950). "Computing Machinery and Intelligence." Mind **49**: 433-460.

[87] Bertolini, A. (2018). "Human-Robot Interaction and Deception." Osservatorio del diritto civile e commerciale, Rivista semestrale(2): 645-659.

[88] See in this respect Donati, P. (2013). Sociologia della relazione. Italia, Il Mulino.

[89] See in this respect Cameron, D., S. Fernando, E. C. Collins, A. Millings, M. Szollosy, R. Moore, A. Sharkey and T. J. Prescott (2017). You made him be alive: Children's perceptions of animacy in humanoid robot. Biomimetic and Biohybrid Systems. 6th International Conference, Living Machines 2017, Stanford, CA,

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

The risks of users' manipulation have been recognized also at international and scientific level. In this respect, an Empathic Technology Working Group was created within the IEEE Society on Social Implications of Technology/Social Implications of Technology Standards Committee for the development of the P7014 - Standard for Ethical considerations in Emulated Empathy in Autonomous and Intelligent Systems[90], still in progress[91].

As indicated in the "Need for the Project" justification, "emotions and cognitive states are closely related to decision-making, health and general wellbeing" and by applying predictive modelling to signals of feelings or behaviour in user data, these systems have the potential to monitor, measure and interact with those users at highly intimate and personal levels. When said effects are not clearly studied and observed, and design fails to take into account the overall – including psychological and emotional – wellbeing of the user, as well as the respect of the existing legal framework, – primarily defined by fundamental rights and principles –,both predictable and unexpected harm may be caused to users, eventually leading to wider social consequences [92].

Thus, it is argued that "ethical standards for empathic technology could help mitigate the exploitation of users through the classification or manipulation of their emotions and cognitive states, and provide guidelines for the positive use of empathic systems"[93].

With respect to the scope and subject matter of the notion of ethics employed by the IEEE, in the IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems it is stated that "we understand "ethical" to go beyond moral constructs and include social fairness, environmental sustainability, and our desire for self-determination. Our analyses and recommendations in Ethically Aligned Design address values and intentions as well as implementations, both legal and technical"[94].

The "General Principles" suggested by the IEEE as a guiding foundation for ethical and values-based design, development, and implementation of autonomous and intelligent systems ("A/IS") are the following:

- *Human Rights*. A/IS shall be created and operated to respect, promote, and protect internationally recognized human rights;
- *Well-being*. A/IS creators shall adopt increased human well-being as a primary success criterion for development;
- *Data Agency*. A/IS creators shall empower individuals with the ability to access and securely share their data, to maintain people's capacity to have control over their identity;
- *Effectiveness*. A/IS creators and operators shall provide evidence of the effectiveness and fitness for purpose of A/IS;
- *Transparency*. The basis of a particular A/IS decision should always be discoverable;

USA, July 26–28, 2017, Proceedings. Living Machines 2017. M. Mangan, M. Cutkosky, A. Mura et al. Stanford University, California, Springer. **Lecture Notes in Computer Science, vol 10384:** 73-85.

[90] See myProject (ieee.org) (last accessed February 2021).

[91] See P7014 - Standard for Ethical considerations in Emulated Empathy in Autonomous and Intelligent Systems (ieee.org) (last accessed February 2021).

[92] See myProject (ieee.org) (last accessed February 2021).

[93] See myProject (ieee.org) (last accessed February 2021).

[94] IEEE (2019). Ethically Aligned Design. A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems, IEEE.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

- *Accountability*. A/IS shall be created and operated to provide an unambiguous rationale for all decisions made.
- *Awareness of Misuse*. A/IS creators shall guard against all potential misuses and risks of A/IS in operation;
- *Competence*. A/IS creators shall specify and operators shall adhere to the knowledge and skill required for safe and effective operation[95].

Some of these principles are well established concepts of mandatory law and they represent statutory obligations for developers, at least under EU law. For example, the principle of respecting human rights is not merely an ethical principle, but an obligation of primary law, as set forth in the European Charter of Fundamental Rights and in the European Convention of Human Rights. Moreover, the principles of data agency, transparency and accountability are enshrined under the GDPR as statutory obligations whose breach is sanctionable with considerable fines (see for example art. 12 and art. 22 of the GDPR). Additionally, the principle of effectiveness which uses the notion of "fitness for purpose" can also be tied to an extant legal obligation, namely under the PLD the producer is to be held liable when the product is defective, that is also when the product does not comply with its reasonably expected use.

In this respect, the choice of naming these legal obligations as ethical principles, included in voluntary standards, may be criticized. Instead, creating standards for the implementation and compliance with these obligations may – at least under certain conditions – be recommended. However, distinction between binding legal norms and ethical principles ought never be disregarded, and greater attention by policy workers should be required to keep such considerations separate.

The results of implementing this standard are yet to be gathered and observed, however, given the breadth of the notion of well-being, it should be explored on a theoretical and empirical level well-being can be standardized. Below, this report shall undertake such analysis.

4.3. Standardize ethics

4.3.1. Different forms of standardization and approaches to ethical rules

Several attempts to justify, create and apply standards of ethics and ethical reasoning have been developed in the history of philosophy. Although there are controversial debates about the concrete normative content of ethical standards and their justification through today, standardization is generally possible in ethics. Indeed, the WP5 First Deliverable focused on the conceptualization and methodology of standardizing moral reasoning, and differentiated among different ways of standardizing, which are shortly summarized below.

Indeed, morals describes the contingent concrete values and habits that shape human behaviour in social life. Those values and habits can strongly differ between certain groups of persons.[96] One example is table manners, where simple forms of standardizing morals relate in a weak sense to the

[95] Ibid.

[96] Birnbacher, D. (2007). Analytische Einführung in die Ethik. 2. Auflage. . Berlin, New York, Walter de Gruyter., p. 7-56; Frankena, W. K. (2017). Ethik. Eine analytische Einführung. Wiesbaden, Springer., p. 6-11, Pieper, A. (2017). Einführung in die Ethik. Tübingen, Francke., p. 22-35.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

sociocultural background of socialization. In contrast to ethics (1.), standardization is pragmatic and involves implicit knowledge, as we often we follow moral rules without making them explicit.

As soon as we start to think explicitly and rationally about our moral habits, we do ethics. Ethics (1.) is the science of morals (2.). In contrast to morals, ethics includes dialogical argumentation, counterarguments and has a high theoretical demand. Insofar, the standardization of ethics is challenged by theoretical demands that are linked to overarching reasons and universal moral laws. Those laws should be created on the basis of explicit knowledge and are intended to be true for every human being – not only for certain groups of persons. Two forms of ethics can be differentiated: (1.a.) descriptive ethics, where the object of investigation (morals (2.)) is described in explicit and therefore standardized phrases; and (1.b.) normative ethics, where the object of investigation is critically evaluated.[97] In this sense, ethics covers the rational justification of moral values and habits, but also the critical evaluation of concrete unmoral actions.

Another third term is ethos, codex or code of ethics (3.). Examples are the *Hippocratic Oath* and Isaac Asimovs *Robo Laws*[98], and the FEANI *Ethics and Conduct of Professional Engineers*[99]. Per definition an ethos is a strict form of normative standardisation because it summarizes at (least certain) moral values and habits in an explicitly written form.[100] A code of ethics does not need to totally regulate the whole range of possible moral behaviour. It fulfils its function when at least some rules are expressed in standardized linguistic phrases that can be passed on to new generations. What differentiates this from ethics (1.) is that an ethos might be the result of tradition and maybe ethical reasoning as well, but this is not required. A code of ethics can also be the naïve and uncritical summary of habitual heritage.

Therefore, ethical standardization may assume three forms:[101]

1. "ethics": science of morals, relating to rational reasoning, justification and theoretical approaches on the basis of explicit knowledge

1.a. descriptive ethics: standardizing the object of observation (= morals (2.)) by putting it into normalized linguistic formulations, describing how people behave and making implicit habits explicit

---

[97] Birnbacher, D. (2007). Analytische Einführung in die Ethik. 2. Auflage. . Berlin, New York, Walter de Gruyter., S. 1-63, Frankena, W. K. (2017). Ethik. Eine analytische Einführung. Wiesbaden, Springer., p. 4-5, Pieper, A. (2017). Einführung in die Ethik. Tübingen, Francke., p. 15-50.

[98] Jordan, J. (2017). Roboter. Wiesbaden, Berlin University Press., p. 51-54, Laßmann, G. (2017). Asimovs Robotergesetze. Was leisten sie wirklich? Hannover, Heise., Chpater 2, Chapter 2.2 and Chapter 2.3, Asimov, I. (2016). Ich der Roboter. Erzählungen. München, Wilhelm Heyne.

[99] FEANI (2006). FEANI position paper on Code of Conduct: Ethics and Conduct of Professional Engineers, FEANI General Assembly., available at https://www.feani.org/sites/default/files/PDF_Documents/Position_papers/Position_Paper_Code_of_Conduct_Ethics_approved_GA_2006.pdf

[100] Honnefelder, L. (2006). Sittlichkeit / Ethos. Handbuch Ethik. M. Düwell, C. Hübenthal and M. Werner. Stuttgart/Weimar, J.B. Metzler**:** 508–513., p. 21-26.

[101] Funk, M. (2020). What Is Robot Ethics? …And Can It Be Standardized? Culturally Sustainable Social Robotics. Proceedings of Robophilosophy. M. Nørskov, J. Seibt and O. S. Quick. Amsterdam, IOS Press**:** 469-480., p. 473-477. See also Funk, M. (2021). Roboter- und KI-Ethik. Eine methodische Einführung – Grundlagen der Technikethik Band 1. Wiesbaden, Springer Vieweg., Chapter 3, Chapter 4, Chapter 5.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

1.b. normative ethics: standardizing the object of observation (= morals (2.)) by critically reflecting its content and methodically creating universal ethical rules, evaluating how people behave, and arguing how they *should* behave and why

2. "morals": standardizing contingent moral values and habits by tradition, education and socialization, primarily based on implicit knowledge, including technical practice

3. "ethos"/"codex"/"code of ethics": standardizing moral rules – therefore making them explicit – by putting them into a normalized linguistic form, it´s close to but also more than morals since a code of ethics relates to a linguistic standardization on the basis of explicit knowledge, and it´s neither descriptive ethics nor normative ethics, but maybe – not necessarily –the output of normative ethical assessment.

Having these different forms of standardization in mind, the first WP5 Deliverable how the concrete content, values and habits which are subject to ethical standardization may be justified, distinguishing between two different approaches, namely, the deontological approach (I.) and in utilitarism (II.).

In the deontological tradition, the general starting point of ethical reasoning is primarily located in the motivation of an action. Consequently, the good will received the status of a key term. Because of the initial reasoning before an action is performed, the regulation operates top-down: from abstract duties to concrete performances in real life. Therefore, a main principle is formulated with the demand that it be universally valid. Every maxim – the subjective norm of action – is fundamentally deduced from the categorical imperative. Standardization in deontological ethics means the application of a universally true abstract principle (which is the categorical imperative) to the reasons of actions (top-down approach, 1.b.I.). Kant himself created several formulations. One common English translation reads: "Act only according to that maxim whereby you can, at the same time, will that it should become a universal law."[102]

On the contrary, utilitarian ethics are characterized by a certain attention to the consequences of actions, belonging to the category of consequentialism. Insofar the abstract general principle of maximizing utility is applied, utilitaristic ethics proceed – just like the Kantian approach – top down. The crucial theoretical criterion is therefore the maximizing of wellbeing, benefit or happiness for a maximum of people.[103] On the other hand, the strict consideration of consequences for the ethical evaluation of an action put a high methodical priority on bottom-up procedures. The general principle and abstract reasoning top down becomes secondary. Pragmatism, empirical issues and the evaluation of the concrete factual action receive a primary bottom-up status – proceeding from the empirical anticipation of consequences of an action to the ethical norm that guides the moral action (1.b.II.). In conclusion it can be summarized that the deontological approach (1.b.I.) belongs to the

[102] Kant, I. (1993 [1785]). Grounding for the Metaphysics of Morals, Hackett [Akademie Ausgabe of Kant's works. Fourth volume. 4:421]., p. 30. See also Kant, I. (1974). Kritik der praktischen Vernunft. Grundlegung zur Metaphysik der Sitten. Band VII Werkausgabe. Herausgegeben von Wilhelm Weischedel. Frankfurt a.M., Suhrkamp., p. 51.

[103] Birnbacher, D. (2007). Analytische Einführung in die Ethik. 2. Auflage. . Berlin, New York, Walter de Gruyter., p. 173-240, Frankena, W. K. (2017). Ethik. Eine analytische Einführung. Wiesbaden, Springer., p. 35-55, Höffe, O. (2013). Ethik. Eine Einführung. München, C.H. Beck.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

formally and methodically strong top-down standardization, whereas utilitarism (1.b.II.) remains formally and methodically strong, but its normalization of values follows a bottom-up methodology.

4.3.2.Is Ethical Standardization possible? – Syntactic Formulations, Methodical Operations and Successful Repetitions

Standardization of ethics is not only possible but in a much more fundamental sense a key characteristic of ethics. However, the concrete kind of standardization, its epistemic foundations, practices and criteria differ heavily. Therefore, in the previous sections three notions of ethics have been differentiated: "ethics" (1. = moral philosophy, science of morals), "morals" (2. = culturally embedded moral lifestyles) and "ethos"/"codex"/"code of ethics" (3.). General differences and similarities of its standardizations can be summarized with respect to three criteria: formal language (= standardized syntax), method (= standardized scientific operations) and practice (= standardization of lifestyles by socially shared repeatedly successful ordinary actions):[104]

1. "ethics": formally strong & methodically strong standardization

1.a. descriptive ethics: formally strong, methodically weak & pragmatically medium

1.b. normative ethics: formally medium, methodically strong & pragmatically medium

Two examples:

1.a.I. Deontological approach: standardization by applying a universally true abstract principle to the reasons of actions (top-down standardization)

1.a.II. Utilitarianism: standardization by applying utility to the pragmatic consequences of actions (bottom-up standardization)

2. "morals": formally weak, methodically weak & pragmatically strong

3. "ethos"/"codex"/"code of ethics": formally strong, methodically weak & pragmatically weak

In a nutshell it can be summarized that ethics (1.) is shaped by a plurality of methods without any meta-method as overarching standard. However, each specific operational method follows the epistemic requirements of creating a specific standard that fulfils the requirements of scientific works. Scepticism and self-critique are essential elements of this more or less dynamic form of standardization – that can also be compared with Thomas Kuhn´s concept of paradigm changes (paradigm = standard of a certain scientific approach). Moral standards related to culturally embedded ordinary life and are very diverse. They heavily depend on pragmatically performed and repeatedly successful actions in social situations. Code of Ethics is per definition a linguistically formulated, syntactical standard. However, it is not to be confused with ethical standardization nor with moral standards. The triangle of these three different forms of standardisation fulfils a heuristic function in order to identify and locate concrete problems when it comes to overarching requirements

[104] Funk, M. (2020). What Is Robot Ethics? …And Can It Be Standardized? Culturally Sustainable Social Robotics. Proceedings of Robophilosophy. M. Nørskov, J. Seibt and O. S. Quick. Amsterdam, IOS Press: 469-480., p. 477-47. See also Funk, M. (2021). Roboter- und KI-Ethik. Eine methodische Einführung – Grundlagen der Technikethik Band 1. Wiesbaden, Springer Vieweg., Chapter 5.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

of regulating the use and development of robots and AI.[105] It also serves for a more differentiated understanding of Robot- and AI-Ethics.[106]

4.3.3. Ethical Standards and Ethics Guidelines – Regulating Humans and Machines

Current attempts to standardize ethics in the field of robotics and AI are primarily undertaken with respect to codes of ethics (3.) that are linguistically standardized in specific ethics guidelines. Examples include the *Ethics Guidelines for Trustworthy Artificial Intelligence (AI)* developed by the High-Level Expert Group on Artificial Intelligence of the European Commission[107] or the guidelines concerning *Ethically Aligned Design* by the Institute of Electrical and Electronics Engineers[108]. Codes like these address human agents in the field of design and development. Others like Asimov's 3/4 Rules have been written for artificial agents. However, attempts to formulate ethical or legal rules for machines bear a certain side-effect – which at the end might become the primary benefit – since they are formulated in human ordinary language and therefore force us to rethink our own behaviour. The primary reason is that ethical rules cannot be translated into machine code without losing meaning.[109] Due to its formal characteristics an ethos is close to legal standards and technical norms. However, they share the problems and drawbacks of application within very specific situations. Furthermore, every abstract regulation requires practical knowledge concerning its application.

Without morals (2.) any ethos (3.) becomes useless and senseless. And this exactly the critical point of top-down regulation: not when it comes to ethical reasoning (1.) but when it comes to the implementation in ordinary life it heavily depends on concrete human lifestyles, culturally embedded values, expectations etc. The primary obstacle of ethics guidelines concerning Robotics and AI seems to be the one sided perspective that lacks a bottom-up perspective (= developing ethics guidelines out of ordinary life). However, for instance on the level of EU regulation several attempts are initiated to involve stakeholders and develop best-practices in order to bridge this gap.[110] Its success depends more on practical approval in ordinary life – and if at all, then only in a very minor sense on its formal logical consistency. Regulation of human-robotics-interactions remains a provisional, prima facie challenge. Rules and practices are a matter of constant critical, reflective loops (ethics, 3.).

---

[105] Funk, M. (2020). What Is Robot Ethics? …And Can It Be Standardized? Culturally Sustainable Social Robotics. Proceedings of Robophilosophy. M. Nørskov, J. Seibt and O. S. Quick. Amsterdam, IOS Press: 469-480., p. 469-471, and p. 477-478. See also Funk, M. (2021). Roboter- und KI-Ethik. Eine methodische Einführung – Grundlagen der Technikethik Band 1. Wiesbaden, Springer Vieweg., Chapter 5, Figure 5.1

[106] Funk, M. (2020). What Is Robot Ethics? …And Can It Be Standardized? Culturally Sustainable Social Robotics. Proceedings of Robophilosophy. M. Nørskov, J. Seibt and O. S. Quick. Amsterdam, IOS Press: 469-480., p. 472-473, Figure 1 on p. 473. See also Funk, M. (2021). Roboter- und KI-Ethik. Eine methodische Einführung – Grundlagen der Technikethik Band 1. Wiesbaden, Springer Vieweg., Chapter 2, Chapter 6.

[107] High-Level Expert Group on AI (2020). The Assessment List for Trustworthy Artificial Intelligence (ALTAI) for self assessment. Brussels, European Commission.

[108] IEEE (2019). Ethically Aligned Design. A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems, IEEE.

[109] European Parliament (2017). European Parliament resolution of 16 February 2017 with recommendations to the Commission on Civil Law Rules on Robotics. 2015/2103(INL), European Parliament., p. 6-7.

[110] High-Level Expert Group on AI (2020). The Assessment List for Trustworthy Artificial Intelligence (ALTAI) for self assessment. Brussels, European Commission.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

Therefore, it is not only necessary to formulate new codes of ethics, but also to foster the developments of ethical terms, methods, reasoning and arguments. An ethos (3.) cannot replace (self-)critical rationality (1.) nor wise and practically skilled lifestyle (2.). Each issues is a challenge on its own right. Any attempt of successful standardization needs to take into attention the complex interrelations within this triangle.

With respect to non-human agency, there is a certain anomaly that could become the object of further investigations and initiative as well. Robot Ethics (RE) includes two levels and four meanings:[111]

LEVEL I (human agency):

RE 1 = applied ethics, ethics of technology (= ethics (1.), humans perform ethics with robots and AI a specific object)

LEVEL II (artificial agency):

RE 2 = robots/AI performing moral lifestyles (= morals (2.))

RE 3 = robots/AI performing rational ethical reasoning/reflection (= ethics (1.))

RE 4 = robots/AI functionally following normative rules, code of ethics (= ethos (3.))

The unusual conceptual option is located on LEVEL II, RE 3: machines might be meaningfully addressed as artificial ethical agents without being artificial moral agents at the same time. On LEVEL I, in contrast, human moral agency is usually the necessary, but not yet sufficient, precondition of ethical agency. This point changes the structure of the triangle of ethics (1.), morals (2.) and ethos (3.) when it comes to standardization of non-human agency.[112] However, only very few investigations have been done on this issue, since most initiatives address the most relevant needs of regulation in the field of human agency. The most important take away in this regard is the message that humans are not mere means that fulfil certain aims – they owe always moral values – whereas machines might follow standardized ethically assessed guidelines in a mere functional way. The open question is then: Is ethical standardization at all possible, when it comes to mere robotic means-end oriented processes without any moral lifestyles? The value of an answer might

[111] Funk, M. (2020). What Is Robot Ethics? …And Can It Be Standardized? Culturally Sustainable Social Robotics. Proceedings of Robophilosophy. M. Nørskov, J. Seibt and O. S. Quick. Amsterdam, IOS Press**:** 469-480., pp. 472-473, Figure 1 on p. 473. See also Funk, M. (2021). Roboter- und KI-Ethik. Eine methodische Einführung – Grundlagen der Technikethik Band 1. Wiesbaden, Springer Vieweg., Chapter 2, Chapter 6; classical sources within the current debate include Veruggio, G. and K. Abney (2012). Roboethics: The Applied Ethics for a new Science. Robot Ethics. The Ethical and Social Implications of Robotics. P. Lin, K. Abney and G. Bekey. Cambridge MA, London, The MIT Press**:** 347-363., Veruggio, G. (2006). EURON ROBOETHICS ROADMAP. Release 1.1. EURON Roboethics Atelier. Genoa., Wallach, W. and C. Allen (2009). Moral Machines. Teaching Robots Right from Wrong. Oxford, Oxford University Press., Allen, C., I. Smit and W. Wallach (2005). "Artificial Morality: Top-down, Bottom-up, and Hybrid Approaches." Ethics and Information Technology **7**(3): 149-155., Anderson, M. and S. Anderson (2011). Machine Ethics, Cambridge Univ. Press.

[112] Funk, M. (2020). What Is Robot Ethics? …And Can It Be Standardized? Culturally Sustainable Social Robotics. Proceedings of Robophilosophy. M. Nørskov, J. Seibt and O. S. Quick. Amsterdam, IOS Press**:** 469-480., p. 476.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

not be found it perfect rules for robots (RE 4) but in a self-critical change of perspective that with benefits for human agents (RE 1).

4.4. A procedural approach towards transparency and disclosure

At the same time, hard-law will not always be the preferable tool to govern emerging technology. An all-encompassing law on all ethical related risks and harms cannot be adopted, neither desirable to do so, for it will most likely be over-inclusive and uncertain with respect to its field and scope of application.

Thus, certain aspects can be better regulated through soft-law instruments and policy governance mechanisms, also known as "smart regulation"[113].

Ethical values, when not confused with or equated to human rights, could be accurately implemented in the research and development of advanced technologies through a procedural framework, combining mandatory law with soft-law and co-regulation.

In this respect, legislation should be adopted to mandate industries to undergo an ethical assessment. This is an idea towards which the European Commission is heading, as it may be seen from its appointing the High-Level Expert Group which drafted and adopted the Assessment List for Trustworthy AI[114] – a self-evaluation and assessment tool for organizations to use in order to determine their systems' compliance with the Ethics Guidelines for Trustworthy AI[115] –.

However, at the current stage, the use of this tool is not mandatory and thus its implementation remains at firms' discretion. Given that using such a tool requires costs on the firm's side and the involvement of the company's human capital and resources[116], making the ethical assessment mandatory may be a preferable and viable alternative.

Moreover, mandatory law should also provide for the disclosure of the content and results of this assessment, which ought to include the measures undertaken to avoid the occurrence of any identified ethical risks.

---

[113] See Gunningham, N., D. Sinclair, P. Grabosky and P. N. Grabosky (1998). Smart Regulation: Designing Environmental Policy, Clarendon Press. and European Commission (2010). Communication from the Commission. Smart Regulation in the European Union. COM(2010) 543 final Brussels, European Commission.

[114] High-Level Expert Group on AI (2020). The Assessment List for Trustworthy Artificial Intelligence (ALTAI) for self assessment. Brussels, European Commission.

[115] High-Level Expert Group on Artificial Intelligence (2019). Ethics Guidelines for Trustworthy AI. Brussels, European Commission.

[116] As per the ALTAI, "this Assessment List for Trustworthy AI (ALTAI) is best completed involving a multidisciplinary team of people. These could be from within and/or outside your organisation with specific competences or expertise on each of the 7 requirements and related questions". See High-Level Expert Group on AI (2020). The Assessment List for Trustworthy Artificial Intelligence (ALTAI) for self assessment. Brussels, European Commission.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

Disclosure in the form of reports represents an essential part in this process and towards achieving better product safety. Such reports ought to be provided in a clear and easily understandable manner, as discussed below.

However, this requires that: (i) consumers are aware of the risks they face[117], and (ii) perceive them accurately[118].

However, information asymmetry annihilates the aforementioned criteria and leads to a market failure and to insufficient levels of safety investments[119].

Market forces will also induce the internalization of risks associated with a broader spectrum of values (including respect for individual's rights, the environment, equality and inclusivity, to name a few) and interests. Users are in fact sensitive to such considerations and their response will affect product's price, and industry profits, and subsequently their compliance and careful design.

By mandating a concise, synthetic and understandable by the general public ethical risk assessment disclosure, information asymmetry could be reduced, i and thence individuals would be able to make informed decisions about the risks they want to be subject to, when using certain applications.

Absent such mandatory disclosure, the benefits of an ethical assessment will be diminished. This may be observed from the application of the GDPR, which although requires for certain cases the drawing up of a data protection impact assessment, its publication by the data controller is voluntary. This impedes individuals to consider any risks when giving consent for their data processing since they do not know them. This has resulted in the creation of a situation also known as the "privacy paradox" whereby individuals' behavior does not always match reported values. Studies show that there is "a disparity between people's expressed values around data privacy and the way they interact with services that require data sharing, for example only requiring very small incentives to share personal details, independently of reported privacy concerns"[120]. It may be the case that absent an accurate knowledge of the potential risks arising from the personal data processing, individuals will never have sufficient incentives to claim better protection, nor will they have sufficient knowledge for their consent to be informed.

Even if such solutions may never replace an adequately conceived legal framework, still to be designed and implemented, it would provide additional guidance and incentives, through market mechanisms, and societal awareness.

[117] Viscusi, W. K. (1985). "Market Incentives for Safety." Harvard Business Review **63**., p. 134.
[118] Sarumida, H. (1996). "Comparative Institutional Analysis of Product Safety Systems in the United States and Japan: Alternative Approaches to Create Incentives for Product Safety." Cornell International Law Journal **79**., p. 132 -133.
[119] Viscusi, W. K. (1995). Fatal Tradeoffs. Public and Private Responsibilities for Risk, Oxford University Press., p. 110
[120] UK Government Office for Science (2020). Evidence and scenarios for global data systems. The Future of Citizen Data Systems. United Kingdom., UK, p. 4.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

5. The need to update regulation & the creation of an Agency

5.1. Extant initiatives and their suitability

As anticipated above, the emergence and diffusion of new advanced technologies pose different and diverse challenges, both for society and for policymakers alike, raising inevitably the need for adopting and updating regulation.

To this end, the European Parliament and the European Commission have started several initiatives meant to align the liability and safety legal framework to the peculiarities these new technologies may bring about. In this respect, several initiatives are notable, namely:

- The 2020 European Parliament's recommendations to the Commission on a civil liability regime for artificial intelligence[121] ("CLRAI");
- The 2020 European Parliament's recommendations to the Commission on a framework of ethical aspects of artificial intelligence, robotics and related technologies[122] ("Ethical framework");
- The Review of the Directive on General Product Safety (2001/95/EC)[123].

Among others, these initiatives advance two important and game changing solutions, which shall be analysed below.

First, the CLRAI departs from the framework of strict liability framework, by setting forth two types of liabilities, namely (i) a s strict liability system for high-risk[124] AI-systems on their operator[125], and (ii) a fault-based liability for other AI-systems that do not constitute a high-risk AI-system, on their operator as well.

[121] European Parliament (2020). Report with with recommendations to the Commission on a civil liability regime for artificial intelligence (2020/2014(INL)). Plenary sitting, European Parliament.

[122] European Parliament (2020). Report with recommendations to the Commission on a framework of ethical aspects of artificial intelligence, robotics and related technologies (2020/2012(INL)). Brussels, European Parliament.

[123] See Have your say (europa.eu) (last accessed February 2021).

[124] As per art. 3 letter c) of the CLRAI "'high risk' means a significant potential in an autonomously operating AI-system to cause harm or damage to one or more persons in a manner that is random and goes beyond what can reasonably be expected; the significance of the potential depends on the interplay between the severity of possible harm or damage, the degree of autonomy of decision-making, the likelihood that the risk materializes and the manner and the context in which the AI-system is being used".

[125] As per art. 3 letter d)-f) of the CLRAI, "'operator' means both the frontend and the backend operator as long as the latter's liability is not already covered by Directive 85/374/EEC" and "'frontend operator' means any natural or legal person who exercises a degree of control over a risk connected with the operation and functioning of the AI-system and benefits from its operation", while "'backend operator' means any natural or legal person who, on a continuous basis, defines the features of the technology and provides data and an essential backend support service and therefore also exercises a degree of control over the risk connected with the operation and functioning of the AI-system."

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

This solution has been already criticized and rightfully so. Respectively, it has been argued that this liability solution may be doomed to fail, based on the following criticalities[126]:

- The distinction between high- and low-risk technologies is based on vague, subjective and broad terms, such as "significant potential to cause harm" and "autonomy that goes beyond what can be reasonably expected", which, in turn, will cause uncertainty with respect to the scope of application of this regulation, making it thus, *ab initio* uncertain and unpredictable, contrary to all fundamental principles of lawful and good regulation;
- Consequently, since the extent and cases where liability may be incurred are uncertain, the intended aim of these liability norms to incentivize safety will not be achieved. As per the CLRAI "the concept of 'liability' plays an important double role in our daily life: on the one hand, it ensures that a person who has suffered harm or damage is entitled to claim compensation from the party held liable for that harm or damage, and on the other hand, it provides the economic incentives for persons to avoid causing harm or damage in the first place". However, it is very likely that under such a regulation the operators will be incapable of determining *ex ante* the correct threshold of safety investment demanded to escape liability in cases of design defect;
- Moreover, the classes of applications to represent high-risk technologies will be outlined in an annex to the regulation, to be constantly updated. However, the distinction between high- and low-risk technologies and the creation of such a list can only be implemented, presuming that such implementation will not take place randomly, based on statistical inferences (see again the notion of "significant *potential* to cause harm"). However, this envisioned scenario fails to take into account a fundamental constraint of emerging technologies, frequently recalled by scholars and practitioners alike, namely the unavailability of data with respect to the risks brought about by said technologies[127].

This preference on unjustly distinguishing between high- and low-risk technologies, not only at EU but also at international level, has also brought about fervent criticism from relevant stakeholders and society representatives. For example, in 2020 Access Now has resigned from the Partnership on AI stating that "as a human rights organization, we support human rights impact assessments and red lines around use of these technologies, rather than an ethics, risk-based, or sandboxing approach"[128].

Second, the Ethical framework borrows the same rationale of the CLRAI by distinguishing between high- and low-risk technologies. One of the consequences of such distinction is that only high-risk AI-based and/or robotics-based technologies will be subject to certification and certain essential safety and human rights requirements such as transparency, non-discrimination and accountability are provided in a firm and mandatory manner only for high-risk technologies[129]. This, of course, is a

[126] See Bertolini, A. (2020). Artificial Intelligence and Civil Liability. Bruxelles, European Parliament - Committee on Legal Affairs**:** 1-132.
[127] Ibid.
[128] See Access Now resigns from the Partnership on AI - Access Now (last accessed February 2021).
[129] European Parliament (2020). Report with recommendations to the Commission on a framework of ethical aspects of artificial intelligence, robotics and related technologies (2020/2012(INL)). Brussels, European Parliament.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

commendable solution, given that certain essential requirements should be applicable to all products placed onto the EU market, in accordance with the extant safety and liability framework. What is more striking is that, although the European Parliament insists on adopting a technology neutral legal framework for advanced technologies, the very premise of the aforementioned recommendations is based on an attempt to differentiate between certain types of applications and it does so in a vague and unclear manner.

Nonetheless, the newly suggested Ethical framework puts forward an applaudable solution, namely that of creating a "Supervisory Authority", which shall be "responsible for ensuring, assessing and monitoring the compliance of the development, deployment and use of high-risk artificial intelligence, robotics and related technologies with the Union's regulatory framework for AI; and for allowing discussions and exchanges of views in close cooperation with relevant stakeholders and civil society; whereas national supervisory authorities should cooperate with each other"[130]. Such an authority shall also be responsible for granting certificates assessing the machines' compliance with this Ethical framework.

This solution has been long proposed and it is clear that there is a need for such an agency[131]. However, there are also many challenges ahead that should not be overlooked. Namely, it is unclear how such an authority will cooperate with other national authorities and how their functions will be split and/or merged in certain cases. For example, the certification of a collaborative robot would most likely require: (i) personal data protection assessment and compliance with the GDPR which is currently within the ambit of the national data protection authorities, (ii) hardware safety requirements assessments currently under the scope of national certification bodies, and then an Ethical framework assessment which will be placed within the ambit of the new Supervisory Authority's tasks.

The accurate creation of such an authority shall also consider other criticalities and requires a multi- and interdisciplinary approach, where the failures of the current structure are taken into account. Just to name a couple, with respect to the national data protection authorities it has been showed that they are understaffed and overwhelmed by the myriad of tasks placed on them by the current data protection legislation[132], whereas reports mandated by the European Commission have showed that safety assessments are usually more accurate when carried out internally by the producers through the self-assessment and self-declaration of conformity procedures, than when such assessments are being carried out by national bodies[133].

[130] Ibid.

[131] Palmerini, E., A. Bertolini, F. Battaglia, B.-J. Koops, A. Carnevale and P. Salvini (2016). "RoboLaw: Towards a European framework for robotics regulation." Robotics and Autonomous Systems **86**: 78-85.

[132] Centre for Public Reform (2010). Final Report: Comparative Study on Different Approaches to Privacy Challenges, in Particular in the Light of Technological Developments. Brussels, European Commission,.

[133] European Commission (2018). Commission Staff Working Document. Evaluation of the Machinery Directive. SWD(2018) 160 final. Brussels, European Commission.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

5.2. A critical review of the measures recently proposed by the European Parliament for the governance of robotics

In the Resolution of 20th of October 2020, with Recommendations to the European Commission on a framework of ethical aspects of artificial intelligence, robotics and related technologies[134], some hints are offered for a feasible governance structure to be designed through the creation of supervisory authorities. This governance structure would apply not only to robotics, but also to artificial intelligence and other related technologies. Certain specific features regarding this structure are highlighted or criticized, in the following paragraphs.

The governance structure basically comprises both the EU and the national levels, having in mind the need of cooperation within other existing authorities and among them. At the same time, the importance of exchanging views with the stakeholders and the civil society is adequately stressed[135]. There is a clear need of a "pluralistic forum for reflection and exchange of views so as to achieve comprehensible and accurate conclusions for the purpose of guiding how governance is regulated".

This proposal, on the one hand, reasonably relies on a common framework for the governance of technologies, implemented by national supervisory authorities in each Member State and coordinated by the Commission (and/or any relevant institutions or bodies of the Union) to "ensure a coherent Union approach and prevent a fragmentation of the single market". Given the economic impact of the expected and current technological change within the EU, this approach makes perfect sense. Accordingly, the coordination at Union level should be structured to ensure a "harmonised approach" across the Member States on the "mandates and actions" of the national supervisory authorities, liaising with them, sharing "best practices", and contributing to cooperation in "research and development". The cooperation will be also necessary for the establishment of binding guidelines on the methodology to be followed for compliance assessments. Finally, the Commission will draw up and update, by means of delegated acts, a common list of high-risk technologies identified in cooperation with the supervisory authorities. On the other hand, the creation of a centre of expertise by the Commission is also envisaged, bringing together academia, research, industry, and individual experts at Union level to foster exchange of knowledge and technical expertise. This openly shows an inclusive way to proceed within the European Union, that hopefully will be reflected when promoting the Union's approach through international cooperation in the attempt to ensure a consistent reply worldwide to the opportunities and risks inherent in these technologies.

---

[134] See European Parliament (2020). Report with recommendations to the Commission on a framework of ethical aspects of artificial intelligence, robotics and related technologies (2020/2012(INL)). Brussels, European Parliament. In the Annex, the report contains the Proposal for a Regulation of the European Parliament and of the Council on ethical principles for the development, deployment and use of artificial intelligence, robotics and related technologies. Article 14 on Risk assessment, Article 15 on Compliance assessment, Article 17 Governance standards and implementation guidance, Article 18 on Supervisory authorities and Article 20 on Coordination at Union level deserve a careful reading.

[135] The wording is very expressive of the desired wide scope: "exchanges by providing assistance to researchers, developers, and other relevant stakeholders, as well as to less digitally-mature companies, in particular small and medium-sized enterprises or start-ups; in particular regarding awareness-raising and support for development, deployment, training and talent acquisition to ensure efficient technology transfer and access to technologies, projects, results and networks" (para. 126) of ibid.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

Member States should appoint an "independent administrative authority" to act as a supervisory authority, although there should be a "margin for implementation" by Member States, including "how the mandate of their respective national supervisory authority is to be carried out". These authorities will have the "responsibility of the good governance of these technologies under the coordination of the Commission" and an important role to play in "promoting the trust and safety of Union citizens, as well as in enabling a democratic, pluralistic and equitable society". National supervisory authorities should ensure the gathering of a maximum number of stakeholders such as industry, businesses, social partners, researchers, consumers and civil society organisations, to "facilitate cooperation with and collaboration between stakeholders", in particular from academia and individual experts.

In that sense, they should provide professional administrative guidance and support to developers, deployers and users. Still, the main purpose of designating these national supervisory authorities is to "ensure, assess and monitor the compliance" with the Union's regulatory framework for the development, deployment and use of robotics. Occasionally, in the proposal, compliance with ethical principles is also mentioned. This may cause some problems, due to their different nature with respect to legal obligations. The latter are binding, and cases of non-compliance may result in a sanction. Thus, to reinforce the efficacy of the ethical principles, they should be clearly inserted in the applicable legal rules. In other words, the role of the Supervisory Authority in each Member State should be clarified: Is it to only ensure that ethical principles are applied to artificial intelligence, robotics and related technologies? Or to take into account legal rules as well? This double perspective happens to be proposed when pursuing the introduction of a European certificate of compliance (referred to both ethical principles and legal obligations as laid down in the proposal for a Regulation requested and relevant Union law)[136]. This is relevant as in the context of aspirational Law the use of standards frequently increases (as opposite to more stringent legal obligations) and positive liberties are developed (instead of focusing regulation on negative liberties).

High-risk technologies should respect the principles of safety, transparency, accountability, non-bias or non-discrimination, social responsibility and gender equality, right to redress, environmental sustainability, privacy and good governance. The national supervisory authority should carry out "an impartial, objective and external risk assessment" in accordance with the criteria provided for in the Regulation and in the list set out in its annex, taking into account the views and any self-assessment made by the developer or deployer. This authority should inform other authorities carrying out risk assessments in accordance with any sector-specific legislation when these technologies are assessed as high-risk.

---

[136] Article 16 on "European certificate of ethical compliance" states the following:

"1.Where there has been a positive assessment of compliance of high-risk artificial intelligence, robotics and related technologies, including software, algorithms and data used or produced by such technologies, carried out in line with Article 15, the respective national supervisory authority shall issue a European certificate of ethical compliance.

2. Any developer, deployer or user of artificial intelligence, robotics and related technologies, including software, algorithms and data used or produced by such technologies, that are not considered as high-risk and that are therefore not subject to the obligations laid down in Articles 6 to 12 and to the risk assessment and compliance assessment provided for in Articles 14 and 15, may also seek to certify the compliance with the obligations laid down in this Regulation, or part of them where so justified by the nature of the technology in question as decided by the national supervisory authorities. A certificate shall only be issued if an assessment of compliance has been carried out by the relevant national supervisory authority and that assessment is positive."

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

Special attention must be paid to the cases where information should be provided twice: "to the national supervisory authorities and national consumer protection authorities[137]", as this may generate overlaps and increase administrative compliance costs. Instead, regular channels for an automatic flow of the information could be established[138]. In the same vein, attention is also needed in cases where accessibility must be granted to various public authorities: "such as national supervisory authorities and market surveillance authorities" (which may have additional prerogatives). It makes no sense to complicate the different ways to meet all the transparency requirements with excessive burdens related to redundant documentation[139]. If the information to be communicated is the same ("the algorithms and data sets used or produced by artificial intelligence, robotics, and related technologies")[140], it would be more efficient to deal with a unique competent authority ("sole point of entry") that would store and share the data with the rest of authorities. This would simplify the possibility of receiving multiple requests by one single person or company.

One could argue that, as data sets should be auditable by national supervisory authorities to ensure their conformity to principles, they should be positioned as the competent authorities (although this task could be probably well carried out in cooperation with or by the data protection authorities). However, going further by allowing that "national supervisory authorities could be addressed by consumers with requests for redress" seems to a certain extent questionable. Obviously, being conceived as the first point of contact in cases of suspected breaches of the Union's regulatory framework, they could receive the requests, but it might be more efficient to transmit it to the consumer protection authorities (or the substantive competent authorities) for them to solve the case, with the possibility of adding an expert report on the technological issues at stake. It is noteworthy that the protection of reporting persons is explicitly considered[141].

The scheme devised so far by the European Parliament counts on the assistance by the EU body/agency to the national supervisory authorities concerning their role as first points of contact in cases of suspected breaches of the legal obligations and ethical principles (including non-discrimination) to carry out compliance assessments, namely by supporting the consultation of other competent authorities in the Union, (e.g. Consumer Protection Cooperation Network and national consumer protection bodies, civil society organizations and social partners located in other Member States). This may prove complicated to implement as it would entail incurring in more steps, through intermediaries, to reach a solution. Hence the procedures could be unnecessarily lengthy and better

[137] The right to information of consumers demands transparency "regarding interaction with artificial intelligence systems, including automation processes, and regarding their mode of functioning, capabilities, for example how information is filtered and presented, accuracy and limitations".

[138] See Para. 17 on "Safety features, transparency and accountability" of European Parliament (2020). Report with recommendations to the Commission on a framework of ethical aspects of artificial intelligence, robotics and related technologies (2020/2012(INL)). Brussels, European Parliament.

[139] Article 8 on Safety, transparency and accountability states that technologies, shall be developed, deployed and used in transparent and traceable manner so that their elements, processes and phases are documented to the highest possible and applicable standards.

[140] See Para. 19 of European Parliament (2020). Report with recommendations to the Commission on a framework of ethical aspects of artificial intelligence, robotics and related technologies (2020/2012(INL)). Brussels, European Parliament.

[141] Article 19 precisely affirms that the Directive (EU) 2019/1937 of the European Parliament and of the Council shall apply to the reporting of breaches of this Regulation and the protection of persons reporting such breaches.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

advantage could be taken of the accumulated expertise of this Network with existing expertise in the field of consumers' protection. Due to the rhythm of implementation and work of these technologies, procedural speed should be duly taken into consideration in the governance structure as a key factor. Moreover, the specificity of this sort of technologies should not be an excuse to allow them to deviate from a common integrative perspective that already serves to solve problems in many different sectors. The search of consistent solutions can be easily made through widely accepted ways, through formulas with a longstanding tradition (e.g. mediation).

It seems to be an added value of having designated national supervisory authorities in each Member State, but their role and mandate should be clarified in a relatively homogeneous manner, in order to guarantee coherent cross-border action. There is a risk to duplicate tasks with cross-sectoral authorities, despite it is acknowledged that the national supervisory authorities should cooperate with the authorities responsible for implementing sectorial legislation.

In addition, in a context of economic recovery from the pandemic, the costs of the structure to set become crucial. There is an express call for "sufficient funding" by each Member State of their designated national supervisory authorities and the reinforcement of national market surveillance authorities. Their size should be carefully considered in line with their respective missions.

Each national supervisory authority will make an assessment of compliance, upon request by any developer, deployer or user of technologies not considered as high-risk and issue a "European certificate of ethical compliance" following common criteria and an application process developed in the context of coordination at Union level. In fact, this task may be overwhelming in practice and may face some problems that other existing certificates have experienced in the past[142]. Undoubtedly, the topics to be covered by the certificate are ambitious. The multi-purpose certificate could be a mix of previous well-established certificates, that are specific for each line of action. The combination of scores under each criterion may be difficult when assessing them altogether in a synthetized manner. At the end of the day, it may result in a lack of clarity to send proper signals to the market. Obviously, the level of performance will not always be the same in all fronts (e.g. environmental performance or gender) and it can be problematic to find a balance. Additionally, these certificates should point out the room for improvement in the respective areas that it comprises to any person using them, so the details provided with regard to each topic are pretty valuable.

To summarize, despite the big efforts made by the European Parliament to provide an institutional oversight for sufficient protection in this new field, there is still a large dose of legal uncertainty due to the broad material scope covered in the proposed regulation[143] and the array of supervisory bodies involved. Regarding the former, there is a risk that too general vague recommendations come to nothing[144]. With regard to the latter, some references should be streamlined depending on the

[142] Grau Ruiz, M. A. (2019) "Some Lessons Learnt From Environmental Labelling Information Schemes: Could Certification Of Inclusive Robotics Follow A Similar Path?".

[143] Article 17 on "Governance standards and implementation guidance", provides that: "Artificial intelligence, robotics and related technologies developed, deployed or used in the Union shall comply with relevant governance standards established in accordance with Union law, principles and values by the national supervisory authorities referred to in Article 18 in accordance with Union law, principles and values, under the coordination of the Commission and/or any relevant institutions, bodies, offices and agencies of the Union that may be designated for this purpose and in consultation with relevant stakeholders".

[144] Article 11 on "Environmental sustainability" states that: "Any high-risk artificial intelligence, robotics and related technologies, including software, algorithms and data used or produced by such technologies, shall

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

essence of the case and for the sake of clarity (e.g. "by the national supervisory authorities or, where applicable, other national or European sectorial supervisory bodies").

5.3. Brief reflections on other governance structures to provide alternatives

The institutional architecture will compromise the success of the implementation of the proposed Regulation. To a certain extent, it may be seen as its Achilles' heel. Therefore, special care has to be put in its definition. In the text of the Resolution, in order to find the most appropriate solution for the coordination at EU level, some examples are given: "ENISA, the EDPS and the European Ombudsman" (para. 128). However, for a better design of the system as a whole, some governance structures should be analysed.

Today, more than ever before in world history, there is a clear need to detect alternatives, debate conflicting opinions and generate consensus in order to opt for sufficiently balanced solutions, the more global the better. Consequently, a sensible approach to this issue requires a continuous dialogue of an interdisciplinary and intercultural nature.

Intelligent robotization does not only imply incorporating artificial intelligence techniques into robots to enable them to interact more closely with each other and with humans - as this term is often abused in modern times. The intelligence to which we must appeal in the process currently underway, in which we are forcibly immersed, is truly human intelligence. A firm commitment to understand and comprehend the magnitude of this phenomenon and to make meaningful proposals to solve the problems it may cause requires, to a large extent, knowing and sharing experiences to foster the skills that may be useful in this regard.

Once this changing situation has been assimilated, it is worth considering what governance schemes are desirable and feasible to ensure people's well-being. In order to protect human, social and political rights, it is absolutely necessary to establish a system of oversight or control, especially when making decisions or carrying out certain sensitive activities. How to arbitrate governance in the most effective way is perfectly debatable, as it may rely on self-monitoring, create a novel system or try to combine other existing ones that may be relevant, having some more or less direct connection. In any case, it is beyond doubt that supervision must ultimately be typically human.

The emphasis is often placed on sectors where robotics can have the greatest impact, such as manufacturing, medicine, agriculture, logistics and transport and consumption (European Commission, Strategic Research Agenda for Robotics in Europe 2020-2021), relying on pre-existing supervision mechanisms in each of them. This seems to have been the case to date. However, an attempt to look introspectively at the robotics sector itself may prove enormously useful. At least as far as the relative invisibility of its sectoral analysis is concerned, as it is easy to observe that it is frequently overlooked.

---

be assessed as to their environmental sustainability, ensuring that measures are put in place to mitigate and remedy their general impact as regards natural resources, energy consumption, waste production, the carbon footprint, climate change emergency and environmental degradation in order to ensure compliance with the applicable Union or national law, as well as any other international environmental commitments the Union has undertaken".

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

At this point, it is obvious that the results of robotization processes must be monitored (specifically their negative and positive impacts); but, simultaneously, this does not exempt from monitoring the processes themselves. Therefore, beyond the rate of implementation of robots and their proper functioning, we should be concerned not only that they do not harm others (as was postulated in Roman law: *neminem laedere* or *alterum non laedere*), but also that the opportunity to use them for the common good is seized, without wasting their potential in this sense.

Indeed, there is an important restriction at the outset: the system of governance to be designed must certainly be effective and sufficiently provide enough guarantees, but it must not hold back the advantages that innovation in robotics can bring with it.

In the first stage, it is logical that the law should be fundamentally reactive with respect to the possible drawbacks that are discovered. Thus, the interests of the parties involved in certain factual scenarios are usually initially served by rules on liability and insurance, and even on data protection in the event of possible breaches of the legal system, offering varying levels of protection depending on the circumstances of the specific case. Often, the idea of remedying, or failing that, palliating or compensating the damage, by procuring the necessary funds, is the main solution on which the legislator is inspired, anticipating a quantitative estimate for this purpose. The way of collecting the necessary monetary amounts varies according to the risk and the expected impact.

At a later and more mature stage, it may be possible to go a step further and guide the law proactively. This means articulating incentive measures to encourage behavior that society values as useful because it brings not only economic benefits, but also environmental, good governance and social benefits. Any future regulation of robotics will have to move towards this point, heading in the direction traced by the adoption of the different regulatory measures and, where appropriate, the margin left for self-regulation.

In addition - following the maritime simile - once docked in that safe harbor, for any subsequent voyage, it will be convenient to have a kind of GPS that measures the drift due to adverse winds and currents in order to correct the course nimbly. This adaptative quality will have to be incorporated into the governance system. Its good regulation will result in greater legal certainty and, therefore, in greater social acceptance of robotics.

Today, it is well known that some sectors critical to economic development have specific governance systems. For example, in the banking sector, compliance with special rules on accounting, security and economic stability has traditionally been overseen by the National Banks or the European Central Bank. On the other hand, the situation of generalized vulnerability that could cause their bankruptcy to individuals and companies means that banks are periodically subjected to stress tests, simulating different economic scenarios to find out the hypothetical situation in which they would find themselves. The injection of capital ends up being the most commonly used tool in practice to try to solve problems that are to some extent a priori unthinkable or unavoidable.

The foreseeable growing dependence on robotics for the exercise of diverse economic activities and for the normal development of people's lives in the heart of a technologically more advanced society should make us think whether, in this situation of possible vulnerability, it would be possible to arbitrate some system of governance similar to the banking system, reinforcing it and stockpiling minimally sufficient funds in the event of possible disasters. In this context, some proposals on the registration of robots, which would entail the payment of the respective insurance policies and the creation of funds to deal with future contingencies, are not surprising.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

Or perhaps we could consider other governance systems in force in the European Union that, in fact, already have a more or less direct influence in related fields. In particular, it would be of interest to analyze the governance systems applicable in the world of data, as these are closely linked to the operation of interactive robots in which artificial intelligence operates. Likewise, given the impact that robotization has on competitiveness, insofar as the inappropriate use of public funding can alter competition in markets, it would be appropriate to pay attention to governance systems in this area, given their effects on interpersonal and inter-territorial inequality.

To begin with, the design of a governance system may involve private or public institutions (or both), with centralized or decentralized schemes, taking into account the number of levels when assessing the costs (economic, time, complexity of operations, etc.) and the advantages they entail. It would be advisable to optimize the creation of institutions, limiting them to those strictly necessary according to their objectives and the appropriate distribution of competencies (according to territory or subject matter, avoiding overlaps that give rise to conflicts). In addition, advantage should be taken of the synergies that may arise when considering cooperative approaches with other organizations.

The ambition of the governance system should be to reach as many countries and companies as possible, without losing effectiveness. It is also important to point out the tasks assigned to the agency in question in the governance framework (e.g. regarding control, decision making and/or implementation).

European standards that have already established some governance systems of potential interest to the robotics (artificial intelligence and big data) sector include the following:

- Regulation (EC) No. 1060/2009 of the European Parliament and of the Council of 16 September 2009 on credit rating agencies.
- Regulation (EU) No. 1095/2010 of the European Parliament and of the Council of 24 November 2010 establishing a European Supervisory Authority (European Securities and Markets Authority), amending Decision 716/2009/EC and repealing Commission Decision 2009/77/EC.
- Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of individuals with regard to the processing of personal data and on the free movement of such data and repealing Directive 95/46/EC.
- Regulation (EU) 2017/2394 of the European Parliament and of the Council of 12 December 2017 on cooperation between national authorities responsible for the enforcement of consumer protection laws and repealing Regulation (EC) No. 2006/2004.

In view of its content, depending on the degree of commitment of the Member States, either a European agency could be established, or to have a network of state authorities coordinated by the EU.

In the first case, decision-making could be vested in a Committee of European decision-makers, which would include the national authorities (with voting rights) together with a representative of the European Commission (without voting rights). It should also be clarified which decisions would require approval by simple or qualified majority.

In the second case, a network of cooperation between national authorities coordinated at the European level could be established, which could be supported by research and implementation in the national structures already available. The European Commission would bring to the attention of

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

the network the cases in which it considers its action relevant. In order to determine the competent authority, the main headquarters of the company concerned could be taken into account (with the risk that sites with lower standards would be frequented) or the authorities involved could choose the national authority best positioned to coordinate the case, prepare a draft decision and discuss it with the others in the search for consensus, providing for mechanisms for unblocking, if necessary, by a European Committee. The territorial scope of the effects of the decision taken should be specified.

Other hybrid governance models can be devised as a result of combinations in the distribution of tasks between European and national authorities. In any case, the organizational changes involved in the system to be applied must be weighed up, avoiding duplication of structures in each State and unnecessary costs in times of economic recovery after the crisis caused by Covid-19. Governance must respond to the criterion of economically efficient centralization without undermining the necessary legal decentralization in accordance with the principle of subsidiarity that must guide the degree of intervention of the European institutions.

In order to adequately finance the functioning of the governance system, the possibility of requiring the corresponding registration or supervision fees, or both, should be emphasized, and should be proportional to the cost incurred by the competent authority.

Aiming at the compliance of a regulatory framework under development is commendable, but one should not fall into temptation of passing too many laws and creating too many organisms (following the trend as regards proliferation of independent authorities in many ambits), as inflation might devaluate them in the long run and cause social mistrust. At present, there is a tangle of soft law provisions in which sometimes not entirely harmonious orientations can be appreciated without a clear pattern, so it would be desirable to have as soon as possible some precise legal rules of harmonization and/or coordination -preventive, not only remedial-, which would guarantee the dignity and solidarity of a Community of Law formed on the basis of Social States, historically born to meet the basic needs of individuals of the human species.

6. Bibliography

Allen, C., I. Smit and W. Wallach (2005). "Artificial Morality: Top-down, Bottom-up, and Hybrid Approaches." Ethics and Information Technology **7**(3): 149-155.
Anderson, M. and S. Anderson (2011). Machine Ethics, Cambridge Univ. Press.
Asimov, I. (2016). Ich der Roboter. Erzählungen. München, Wilhelm Heyne.
Beck, S. (2014). Corporate Criminal Liability. The Oxford Handbook of Criminal Law. M. D. Dubber and T. Hörnle. United States of America, Oxford University Press.
Bertolini, A. (2013). "Robots as Products: The Case for a Realistic Analysis of Robotic Applications and Liability Rules." Law Innovation and Technology **5**(2): 214–247.
Bertolini, A. (2015). "Robotic prostheses as products enhancing the rights of people with disabilities. Reconsidering the structure of liability rules." International Review of Law, Computers & Technology **29**(2-3): 116-136.
Bertolini, A. (2016). "Insurance and Risk Management for Robotic Devices: Identifying the Problems." Global Jurist **16**(3): 291-314.
Bertolini, A. (2018). "Human-Robot Interaction and Deception." Osservatorio del diritto civile e commerciale, Rivista semestrale(2): 645-659.
Bertolini, A. (2020). Artificial Intelligence and Civil Liability. Bruxelles, European Parliament - Committee on Legal Affairs**:** 1-132.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

Bertolini, A. and M. Riccaboni (2020). "Grounding the case for a European approach to the regulation of automated driving: the technology-selection effect of liability rules." European Journal of Law and Economics.
Birnbacher, D. (2007). Analytische Einführung in die Ethik. Berlin, New York, Walter de Gruyter.
Birnbacher, D. (2007). Analytische Einführung in die Ethik. 2. Auflage. . Berlin, New York, Walter de Gruyter.
Bonsignorio, F., E. Messina and A. P. del Pobil (2014). "Fostering Progress in Performance Evaluation and Benchmarking of Robotic and Automation Systems." IEEE Robotics & Automation Magazine **21**(1): 22-25.
Bryson, J. J. and P. P. Kime (2011). Just an artifact : why machines are perceived as moral agents. Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence: Barcelona, Catalonia, Spain, 16–22 July 2011. T. Walsh. Menlo Park, CA, USA AAAI Press. **II:** 1641-1646.
Calo, R. (2015). "Robotics and the Lessons of Cyberlaw." California Law Review **103**(3): 513-563.
Cameron, D., S. Fernando, E. C. Collins, A. Millings, M. Szollosy, R. Moore, A. Sharkey and T. J. Prescott (2017). You made him be alive: Children's perceptions of animacy in humanoid robot. Biomimetic and Biohybrid Systems. 6th International Conference, Living Machines 2017, Stanford, CA, USA, July 26–28, 2017, Proceedings. Living Machines 2017. M. Mangan, M. Cutkosky, A. Mura et al. Stanford University, California, Springer. **Lecture Notes in Computer Science, vol 10384:** 73-85.
Centre for Public Reform (2010). Final Report: Comparative Study on Different Approaches to Privacy Challenges, in Particular in the Light of Technological Developments. Brussels, European Commission,.
Coase, R. H. (1937). "The Nature of the Firm." Economica **4**(16).
Coleman, J., S. Hershovitz and G. Mendlow (Winter 2015). Theories of the Common Law of Torts. The Stanford Encyclopedia of Philosophy. E. Zalta. https://plato.stanford.edu/archives/win2015/entries/tort-theories/.
Donati, P. (2013). Sociologia della relazione. Italia, Il Mulino.
Ernst&Young, Technopolis and VVA (2018). Evaluation of Council Directive 85/374/EEC on the approximation of laws, regulations and administrative provisions of the Member States concerning liability for defective products. Brussels, European Commission.
Espinoza, J. (2021). "EU must overhaul flagship data protection laws, says a 'father' of policy." https://www.ft.com/content/b0b44dbe-1e40-4624-bdb1-e87bc8016106.
European Commission (2010). Communication from the Commission. Smart Regulation in the European Union. COM(2010) 543 final Brussels, European Commission.
European Commission (2018). Commission Staff Working Document. Evaluation of Council Directive 85/374/EEC of 25 July 1985 on the approximation of the laws, regulations and administrative provisions of the Member States concerning liability for defective products. SWD(2018) 157 final. Brussels, European Commission.
European Commission (2018). Commission Staff Working Document. Evaluation of the Machinery Directive. SWD(2018) 160 final. Brussels, European Commission.
European Commission (2020). Report from the Commission to the European Parliament, the Council and the European Economic and Social Committee. Report on the safety and liability implications of Artificial Intelligence, the Internet of Things and robotics. COM(2020) 64 final. Brussels, European Commission.
European Parliament (2017). European Parliament resolution of 16 February 2017 with recommendations to the Commission on Civil Law Rules on Robotics. 2015/2103(INL), European Parliament.
European Parliament (2020). Digital Services Act: Improving the functioning of the Single Market European Parliament resolution of 20 October 2020 with recommendations to the Commission on

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

the Digital Services Act: Improving the functioning of the Single Market (2020/2018(INL)). TEXTS ADOPTED. Provisional edition, European Parliament.
European Parliament (2020). Report on intellectual property rights for the development of artificial intelligence technologies (2020/2015(INI)), Plenary sitting.
European Parliament (2020). Report with recommendations to the Commission on a framework of ethical aspects of artificial intelligence, robotics and related technologies (2020/2012(INL)). Brussels, European Parliament.
European Parliament (2020). Report with with recommendations to the Commission on a civil liability regime for artificial intelligence (2020/2014(INL)). Plenary sitting, European Parliament.
Expert Group on Liability and New Technologies (2019). Report on Liability for Artificial Intelligence and other emerging digital technologies. Brussels, European Commission.
FEANI (2006). FEANI position paper on Code of Conduct: Ethics and Conduct of Professional Engineers, FEANI General Assembly.
Finck, M. (2019). Blockchain and the General Data Protection Regulation. Can distributed ledgers be squared with European data protection law? The Panel for the Future of Science and Technology (STOA). Brussels, European Parliament.
Frankena, W. K. (2017). Ethik. Eine analytische Einführung. Wiesbaden, Springer.
Funk, M. (2020). What Is Robot Ethics? …And Can It Be Standardized? Culturally Sustainable Social Robotics. Proceedings of Robophilosophy. M. Nørskov, J. Seibt and O. S. Quick. Amsterdam, IOS Press**:** 469-480.
Funk, M. (2021). Roboter- und KI-Ethik. Eine methodische Einführung – Grundlagen der Technikethik Band 1. Wiesbaden, Springer Vieweg.
Grau Ruiz, M. A. (2019) "Some Lessons Learnt From Environmental Labelling Information Schemes: Could Certification Of Inclusive Robotics Follow A Similar Path?".
Gunningham, N., D. Sinclair, P. Grabosky and P. N. Grabosky (1998). Smart Regulation: Designing Environmental Policy, Clarendon Press.
Gutman, M., B. Rathgeber and T. Syed (2012). Action and Autonomy: A hidden Dilemma in Artificial Autonomous Systems. Robo- and Informationethics. Some Fundamentals. M. Decker and M. Gutman. Zürich, Lit**:** 231-256.
High-Level Expert Group on AI (2020). The Assessment List for Trustworthy Artificial Intelligence (ALTAI) for self assessment. Brussels, European Commission.
High-Level Expert Group on Artificial Intelligence (2019). Ethics Guidelines for Trustworthy AI. Brussels, European Commission.
Höffe, O. (2013). Ethik. Eine Einführung. München, C.H. Beck.
Honnefelder, L. (2006). Sittlichkeit / Ethos. Handbuch Ethik. M. Düwell, C. Hübenthal and M. Werner. Stuttgart/Weimar, J.B. Metzler**:** 508–513.
IEEE (2019). Ethically Aligned Design. A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems, IEEE.
INBOTS (2019). Preliminary White Paper on Standardization and Interactive Robots. http://inbots.eu/wp-content/uploads/2019/07/Attachment_0-3.pdf.
Jordan, J. (2017). Roboter. Wiesbaden, Berlin University Press.
Kalra, N. and S. M. Paddock (2016). "Driving to Safety. How Many Miles of Driving Would It Take to Demonstrate Autonomous Vehicle Reliability?" Transportation Research Part A: Policy and Practice **94**: 182-193.
Kanda, T., R. Sato, N. Saiwaki and H. Ishiguro (2007). "A two-month field trial in an elementary school for long-term human robot interaction." IEEE Trasactions on Robotics **23**(5): 962-971.
Kant, I. (1974). Kritik der praktischen Vernunft. Grundlegung zur Metaphysik der Sitten. Band VII Werkausgabe. Herausgegeben von Wilhelm Weischedel. Frankfurt a.M., Suhrkamp.
Kant, I. (1993 [1785]). Grounding for the Metaphysics of Morals, Hackett [Akademie Ausgabe of Kant's works. Fourth volume. 4:421].

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

Koops, B.-J. (2006). Should ICT Regulation Be Technology-Neutral? Starting Points for ICT Regulation. Deconstructing Prevalent Policy One-Liners. B.-J. Koops, M. Lips, C. Prins and M. Schellekens. The Hague, T.M.C. Asser Press. **9:** 77-108.
Kraakman, R., P. Devies, H. Hansmann, G. Hertig, K. Hopt, H. Kanda and E. Rock (2006). The Anatomy of Corporate Law: A Comparative and Functional Approach.
Laßmann, G. (2017). Asimovs Robotergesetze. Was leisten sie wirklich? Hannover, Heise.
Machnikowski, P., Ed. (2016). European Product Liability. An Analysis of the State of the Art in the Era of New Technologies. Cambridge, Intersentia.
Maio, G. (2012). Mittelpunkt Mensch: Ethik in der Medizin. Ein Lehrbuch. Stuttgart, Schattauer.
Muller, C. (2017). Opinion of the European Economic and Social Committee on 'Artificial intelligence - The consequences of artificial intelligence on the (digital) single market, production, consumption, employment and society' (own-initiative opinion) INT/806 Artificial intelligence. European Economic and Social Committee.
Padovani, T. (2002). Diritto Penale. Milano, Giuffrè.
Palmerini, E., A. Bertolini, F. Battaglia, B.-J. Koops, A. Carnevale and P. Salvini (2016). "RoboLaw: Towards a European framework for robotics regulation." Robotics and Autonomous Systems **86**: 78-85.
Piaget, J. (1951). The child's conception of the world. Lanham, Maryland, Rowman & Littlefield.
Pieper, A. (2017). Einführung in die Ethik. Tübingen, Francke.
Polinsky, M. A. and S. Shavell (2009-2010). "The uneasy case for product liability." Harvard Law Review **123**: 1437-1492.
Posner, R. A. (2007). Economic Analysis of Law, Wolters Kluwer Law & Business.
Russell, S. J. and P. Norvig (2010). Artificial Intelligence. A Modern Approach, Prentice Hall.
Sarumida, H. (1996). "Comparative Institutional Analysis of Product Safety Systems in the United States and Japan: Alternative Approaches to Create Incentives for Product Safety." Cornell International Law Journal **79**.
Smart Dubai. (2019). "Artificial Intelligence Principles& Ethics." from https://www.smartdubai.ae/initiatives/ai-principles-ethics.
Timan, T., R. Snijders, M. Kirova, S. Suardi, M. v. Lieshout, M. Chen, P. Costenco, E. Palmerini, A. Bertolini, A. Tejada, S. v. Montfort, M. Bolchi, S. Alberti, R. Brouwer, K. Karanilokova, F. Episcopo and S. Jansen (2019). Study on safety of non-embedded software. Service, data access, and legal issues of advanced robots, autonomous, connected, and AI-based vehicles and systems: final study report regarding CAD/CCAM and industrial robots. Brussel, European Commission.
Turing, A. (1950). "Computing Machinery and Intelligence." Mind **49**: 433-460.
Twigg-Flesner, C. (2021). Guiding Principles for Updating the Product Liability Directive for the Digital Age. ELI Innovation Paper Series, European Law Institute.
UK Government Office for Science (2020). Evidence and scenarios for global data systems. The Future of Citizen Data Systems. United Kingdom.
Vardi, M. Y. (2012). "Artificial Intelligence: Past and Future." Communications of the ACM **55**(1): 5.
Veruggio, G. (2006). EURON ROBOETHICS ROADMAP. Release 1.1. EURON Roboethics Atelier. Genoa.
Veruggio, G. and K. Abney (2012). Roboethics: The Applied Ethics for a new Science. Robot Ethics. The Ethical and Social Implications of Robotics. P. Lin, K. Abney and G. Bekey. Cambridge MA, London, The MIT Press**:** 347-363.
Viscusi, W. K. (1985). "Market Incentives for Safety." Harvard Business Review **63**.
Viscusi, W. K. (1995). Fatal Tradeoffs. Public and Private Responsibilities for Risk, Oxford University Press.
Wallach, W. and C. Allen (2009). Moral Machines. Teaching Robots Right from Wrong. Oxford, Oxford University Press.
Wiesing, U. (2008). Der hippokratische Eid. thik in der Medizin. Ein Studienbuch. U. Wiesing. Stuttgart, Reclam**:** 36–42.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073