Discussion Paper Series – CRC TR 224

Discussion Paper No. 371
Project B 05

The Law and Economics of AI Liability

Miriam Buiten [1]
Alexandre de Streel [2]
Martin Peitz [3]

October 2022

[1] University of St. Gallen
[2] University of Namur and NADI
[3] University of Mannheim, Email: martin.peitz@gmail.com

Funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through CRC TR 224 is gratefully acknowledged.

The Law and Economics of AI Liability

October 2022

Miriam Buiten*

Alexandre de Streel†

Martin Peitz‡

Abstract

When AI systems possess the characteristics of autonomy and unpredictability, they present challenges for the existing liability framework. (Semi)-autonomous AI systems shift control over these systems away from users and towards producers, while errors of AI systems may be difficult to foresee. Policymakers are faced with the questions when existing civil liability rules do not adequately cover risks arising in the context of AI systems, and how they then should be adapted. This paper addresses these two questions for EU non-contractual liability rules. It considers how liability rules affect the incentives of producers, users, and bystanders that may be harmed by AI. The paper provides concrete recommendations for updating the EU Product Liability Directive and for an EU liability framework for owners and users of AI.

Keywords: Artificial Intelligence, Liability, EU law

JEL Classification: K13, O33

* Miriam Buiten, University of St.Gallen, Law School, St.Gallen, Switzerland; CERRE, Brussels, Belgium.
† Alexandre de Streel, University of Namur and NADI, Belgium; CERRE, Brussels, Belgium.
‡ Martin Peitz, University of Mannheim; Department of Economics and MaCCI, Mannheim, Germany.

This paper is based on a report prepared in March 2021 by the authors for the Centre on Regulation in Europe (CERRE <http://www.cerre.eu/>). Martin Peitz gratefully acknowledges support from Deutsche Forschungsgemeinschaft (DFG) through CRC TR 224 (project B05).

1. Introduction

Artificial Intelligence, equipped with learning abilities, can improve human decision-making by approaching problems in fundamentally different ways. In doing so, AI systems promise to improve societal well-being and increase efficiency in numerous fields, including healthcare, transport, and consumer products.

As a result of the different approaches to problem-solving, a possible downside of AI is that in the event that errors occur in AI systems, these errors may be less predictable to humans. While AI may lead to many benefits, it seems unlikely that Al producers will be able to completely prevent AI-inflicted injuries. As AI systems advance, they will increasingly be used to support and – possibly, in some cases - replace human decision-making. It may then be unclear under which circumstances a human supervisor should intervene and override the decision taken by the AI system.

When AI systems possess the characteristics of unpredictability and autonomy, they present challenges for the existing liability framework. Liability rules have to address who bears responsibility for accidents in which AI systems are involved. The arrival of autonomous or semi-autonomous AI systems tends to shift control over these systems away from owners and users ("operators") and towards producers. Given that producers of autonomous AI systems can exercise more control over the performance of these systems than producers of mechanical products, it may be justified to shift more responsibility towards producers of AI systems.

Two questions about the liability of Artificial Intelligence (AI) deserve attention from policymakers: 1) Do existing civil liability rules adequately cover risks arising in the context of AI systems? 2) How would modified liability rules for producers, owners, and users of AI play out? The paper addresses these two questions for EU non-contractual liability rules while acknowledging the interaction of such rules with other regulatory instruments. The normative framework of the paper is based on an economic analysis of the law, considering how liability rules affect the incentives of producers, users, and other parties that may be harmed.

The paper proceeds as follows. Section 2 lays out the broader regulatory framework. Section 3 identifies the gaps in liability flowing from the unique characteristics of AI. Section 4 analyses the efficient liability rules for AI, discussing what parties should be held liable and what standard of liability is efficient. Section 5 discusses in more detail what changes are appropriate to the EU Product Liability Directive and evaluates the need for an EU liability regime for AI operators. Section 6 concludes.

2. Existing EU regulatory framework for safety and liability

2.1. Safety Regulation

The rules for non-contractual liability are part of a broader legal framework, which includes contractual liability rules, general safety rules and sector-specific liability and safety rules that apply in high-risk sectors.

Within the EU, the product safety rules, which are mostly ex ante, are divided into two levels of legislation. Specific rules regulate certain sectors or products, and in absence of such specific requirements, the general rules set out in the General Product Safety Directive (henceforth, "GPSD")[1] apply. The directive ensures that only safe consumer products are placed on the market by manufacturers.[2] The GPSD obliges producers as well as distributors to provide safe products to consumers; to take all possible steps to identify any hazards of their products and to inform consumers of the existence of such risks; and if necessary to withdraw dangerous products.[3] After reviewing the GPSD, the Commission decided to move from a Directive to a Regulation and therefore published a proposal for a General Product Safety Regulation (henceforth, "GPSR")[4]. The proposed

[1] Directive 2001/95 of the European Parliament and of the Council of 3 December 2001 on general product safety [2002] OJ L 11 (General Product Safety Directive).
[2] General Product Safety Directive, art 3
[3] *Ibid*.
[4] Commission, 'Proposal for a Regulation of the European Parliament and of the Council on general product safety, amending Regulation (EU) No 1025/2012 of the European Parliament and of the Council, and repealing Council Directive 87/357/EEC and Directive 2001/95/EC of the European Parliament and of the Council' COM (2021) 346 final

GPSR aims to regulate safety risks emerging from new technologies.[5] Additionally, the Market Surveillance Regulation sets out requirements for accreditation and market surveillance.[6]

For high-risk sectors, the EU product safety framework complements the horizontal rules with sector-specific rules. For instance, in healthcare the Medical Devices Regulation[7] and the In Vitro Diagnostic Medical Devices Regulation[8] apply; in the domain of transportation the General Vehicles Safety Regulation[9], the Approval and Market Surveillance of Vehicles Regulation[10] and the Motor vehicles Insurance Directive (MID) apply[11]; and for machinery the Machinery Directive[12] applies and a CE Declaration is required.

2.2. Liability rules

2.2.1. Standards of liability

The fault-based liability regime is the standard liability applicable in Member States.[13] Under such a regime, claimants need to prove three cumulative conditions to get damages: that the injurer was at fault, that they suffered damage, and that there is a link of causality between the fault and the damage. As the proof of each of those conditions may be difficult, some other – more victim-friendly – liability regimes have been established in Member States for specific situations where the legislator (or the case-law) has estimated that the victims need to be better protected.

A first victim-friendly regime consists of starting from the standard fault-based regime but changing the burden of proof of some of the three conditions. A rebuttable presumption for the fault requirement and/or for the causal link can facilitate victims in obtaining compensation, and/or can help reduce information asymmetry between the victim and the wrongdoer. A presumption regime may be linked to a diverse set of factual situations generating different types of risks and damages, such as (i) the responsibility of the owner/possessor of the building in case of damages causes by his/her building (unless he/she proves that he/she observed appropriate care to avoid the damage); (ii) the responsibility of a person carrying out a dangerous activity (unless he/she proves that all appropriate measures to avoid the damage have been taken); (iii) the responsibility of the employer/the principal for the actions executed on his behalf or interest by his employees/agents (unless he proves that he used appropriate care in the selection and the management of the agent/employee) or (iv) the

[5] For an overview see Nikolina Šajn, 'General product safety regulation - European Parliamentary Research Service Briefing' (PE 698.028, September 2021) <https://www.europarl.europa.eu/RegData/etudes/BRIE/2021/698028/EPRS_BRI(2021)698028_EN.pdf> accessed on 07 August 2022

[6] Regulation (EU) 2019/1020 of the European Parliament and of the Council of 20 June 2019 on market surveillance and compliance of products and amending Directive 2004/42/EC and Regulations (EC) No 765/2008 and (EU) No 305/2011 [2019] OJ L 169/1. This Regulation will be applicable starting July 2021 and replaces Regulation 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 [2008] OJ L 218/39

[7] 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 [2017] OJ L 117/1

[8] Regulation (EU) 2017/745 of the European Parliament and of the Council of 5 April 2017 on in vitro diagnostic medical devices and repealing Directive 98/79/EC and Commission Decision 2010/227/EU [2017] OJ L 117/176

[9] Regulation (EU) 2019/2144 of the European Parliament and of the Council on type-approval requirements for motor vehicles and their trailers, and systems, components and separate technical units intended for such vehicles, as regards their general safety and the protection of vehicle occupants and vulnerable road users, amending Regulation (EU) 2018/858 of the European Parliament and of the Council and repealing Regulations (EC) No 78/2009, (EC) No 79/2009 and (EC) No 661/2009 of the European Parliament and of the Council and Commission Regulations (EC) No 631/2009, (EU) No 406/2010, (EU) No 672/2010, (EU) No 1003/2010, (EU) No 1005/2010, (EU) No 1008/2010, (EU) No 1009/2010, (EU) No 19/2011, (EU) No 109/2011, (EU) No 458/2011, (EU) No 65/2012, (EU) No 130/2012, (EU) No 347/2012, (EU) No 351/2012, (EU) No 1230/2012 and (EU) 2015/166 [2019] OJ L 325/1, art 3(22): "a motor vehicle that has been designed and constructed to move autonomously without any driver supervision".

[10] Regulation 2018/858 of the European Parliament and of the Council of 30 May 2018 on the approval and market surveillance of motor vehicles and their trailers, and of systems, components and separate technical units intended for such vehicles [2018] OJ L 151/1. This Regulation took force on 1 September 2020 and repeals Directive 2007/46 of the European Parliament and of the Council of 5 September 2007 establishing a framework for the approval of motor vehicles and their trailers, and of systems, components and separate technical units intended for such vehicles (Framework Directive) [2007] OJ L 263/1

[11] Directive 2009/103 of the European Parliament and of the Council of 16 September 2009 relating to insurance against civil liability in respect of the use of motor vehicles, and the enforcement of the obligation to insure against such liability [2009] OJ L 263/11

[12] Directive 2006/42/EC of the European Parliament and of the Council of 17 May 2006 on machinery, and amending Directive 95/16/EC (recast) [2006] OJ L 157/23

[13] Konrad Zweigert. K., & Hein Kötz, *Einführung in die Rechtsvergleichung auf dem Gebiete des Privatrechts* (3rd edn, Mohr Siebeck 1996) 650.

responsibility of parents/tutors/guardians/teachers for damages caused by a minor, pupil, student/apprentice, or mentally impaired person (unless they can prove that they were not able to prevent the damages from happening). [14]

A second more victim-friendly regime consists of facilitating the standard of proof of the victim by changing the conditions which needed to be proven to get damages. Under a strict liability regime, victims need to prove the *default* or the risks taken by the wrongdoers, which is easier to prove than the fault - or the negligence - of the wrongdoer.

A strict liability standard can be justified when the likelihood of damage is linked to the unpredictable behaviour of specific risk groups.[15] In these cases – as can be seen with liability rules for autonomous third parties such as animals or children – liability is attributed to the individual that is responsible to supervise said third party, as they are deemed able to adopt measures to prevent or at least reduce the risk. Strict liability may also be justified because the risk of damages is linked to dangerous activities: some jurisdictions may attribute liability to the person that carries out such dangerous activity (e.g. the operator of a nuclear power plant or an aircraft, or the driver of a car) or are ultimately responsible for the dangerous activity to happen (e.g. the owner of a vehicle). The rationale typically is that this person has created a risk, which materialises in damage and at the same time also derives an economic benefit from this activity. Those strict liability regimes may apply to a diverse set of factual situations generating different types of risks and damages, such as (i) the liability of the owners of animals for the damages caused by the animals under their custody, (ii) the strict liability of the person responsible for carrying out an unspecified or specified dangerous activity (for example the operation of nuclear power plants, aircraft, or motor vehicles or (iii) other cases linked to a legal or factual relationship between two persons or a person and an object, such as when the damages are caused by someone executing a task in the interest of someone else (employee/employer) or by an object that is under his/her custody.

Some forms of strict liability may go even a step further by linking liability simply to the materialisation of risk and/or making the discharge of liability either impossible or possible only under the proof that the damaging event was caused by an exceptional/unforeseen circumstance that could not be avoided. In effect, those stricter regimes establish non rebuttable presumptions of a causality link to facilitate the compensation of the victim of damages in situations where the legislator considers it too burdensome or unbalanced to require the victim to prove such causality link.

As the strict liability regime tilts the balance in favour of the victim at the expense of the person responsible, they are in general accompanied with limiting principles, esp. regarding the type of damages which can be compensated or the maximum of damages which can be granted. Thus, relying on a strict liability regime (when it exists), the victim may be compensated more easily than under the fault-based liability regime but a more limited manner. If the victim seeks compensation for more damages than the ones covered by strict liability, the victim needs to also launch a complementary action against the person responsible under fault-based liability. It is also interesting to note that strict liability regimes may also be coupled with mandatory insurance requirements which reduces the risks for the victims of not being compensated.

2.2.2. Product Liability Directive

The EU Product Liability Directive (henceforth, "PLD"), adopted in 1985, established a strict liability regime where producers are liable for their defective products regardless of whether the defect is their fault.[16] The PLD is a technology-neutral instrument that fully harmonises product liability rules throughout the EU. It applies to any product sold in the EEA with a three-year limit for the recovery of damages. The PLD applies to "movables" (Art.

[14] Commission, 'Report of 7 May 2018 on the Application of the Council Directive on the approximation of the laws, regulations, and administrative provisions of the Member States concerning liability for defective products (85/374/EEC)' COM (2018) 246 final, 5-6

[15] Commission, 'Liability for emerging digital technologies – Accompanying the document Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions Artificial intelligence for Europe' (Staff Working Document) COM (2018) 237 final, 8

[16] Council Directive 85/374/EEC on the approximation of the laws, regulations and administrative provisions of the Member States concerning liability for defective products [1985] OJ L 210/29 (Product Liability Directive). Given that liability is restricted to defective products, some argue that this is in fact a fault-based liability regime. See e.g. Herbert Zech, 'Liability for AI: public policy considerations' (2021) 22 ERA Forum 147

2), which are interpreted as tangible goods.[17] The Court of Justice of the EU has indicated that the PLD applies to products used while providing any service.[18]

The PLD assigns liability to the "producer" (Art. 1 PLD). It defines "producer" as the manufacturer of a finished product, the producer of any raw material or the manufacturer of a part, and any person who, by putting his name, trademark or other distinguishing feature on the product presents himself as to its producer. Any person who imports a product for sale, hire, leasing or any form of distribution in the course of his business is responsible in the same way as the producer (Art. 3(2) PLD).[19]

Producer liability under the PLD arises in case of a defect. The PLD defines a "defective" product as a product that does not provide the safety the consumer is entitled to expect, considering all circumstances. This includes, for instance, the presentation of the product, the use to which it could reasonably be expected that the product would be put, the time when the product was put into circulation. The notion of defect focuses on consumers' safety expectations to physical harm, excluding possible privacy harm, cybersecurity flaws, or other risks that may arise for IoT products.[20] A defect shall be assessed considering "the time when the product was put into circulation" (Art. 6(1)(b) PLD). This concept has raised interpretative questions.[21]

The PLD defines damage as death, personal injury, or damage to the product or other property with a ceiling.[22] Thus, the PLD limits the type of damages which can be compensated as it does not cover immaterial damages. It also allows Member States to cap the number of damages.[23]

There is room for different national approaches, for example on systems to settle claims for damages, or on how to bring proof of damage. Member States may also introduce or maintain other national instruments for the liability of producers based on fault.[24]

3. Gaps in liability rules for AI-based products

3.1. Risks associated with AI

Artificial Intelligence is an umbrella term for various technologies that rely on algorithms, which have different features and are designed for different fields of application. AI has been defined in terms of its perceived intelligence[25], its ability to act autonomously[26] or its characteristic of evolving in an unforeseeable way.[27] From a technical perspective, it is clear that not all algorithms constitute AI, but there is no consensus over what subset of algorithms, such as machine learning algorithms and neural networks, can be considered AI. The Proposed AI Act defines AI as software developed with machine learning, logic-based and statistical approaches that can generate output influencing the environments with which they interact.[28]

[17] See e.g. BEUC (The European Consumer Organisation), 'PRODUCT LIABILITY 2.0 How to make EU rules fit for consumers in the digital age' (BEUC-X-2020-024, 07 May 2020) <www.beuc.eu/publications/beuc-x-2020-024_product_liability_position_paper.pdf> accessed on 07 August 2022, 12.
[18] Case C-203/99 *Veedfald*, [2001] ECR I-3569.
[19] BEUC (n 17) 18-19
[20] BEUC (n 17) 13
[21] See e.g. Case C-127/04 *Declan O'Byrne v. Sanofi Pasteur MSD* [2006] ECR I-01313
[22] See BEUC (n 17) 15.
[23] Product Liability Directive, art 16
[24] COM (2018) 246 final 4
[25] John J. McCarthy, Marvin L. Minsky, and Nathaniel Rochester, *Artificial Intelligence* (Research Laboratory of Electronics at the Massachusetts Institute of Technology 1959)
[26] Matthew U. Scherer, 'Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, And Strategies' (2016) 29 Harvard Journal of Law & Technology 353, 363
[27] Urs Gasser and Virgilio A.F. Almeida, 'A Layered Model for AI Governance' (2017) 21 IEEE Internet Computing 58
[28] Commission, 'Proposal for a Regulation of the European Parliament and of the Council laying down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain Union Legislative Acts' COM (2021) 206 final ("Proposed AI Act"), art 3 (1): In particular, Annex 1 names three specific techniques that are to fall under the term AI: 'machine learning approaches', including

The EU AI High-Level Expert Group Report recognises complexity, opacity, openness, autonomy, predictability, data-drivenness, and vulnerability as unique characteristics of AI.[29] In order to identify the appropriate liability regime for AI, it is necessary to assess how these characteristics affect risks associated with AI-based technologies.

First, AI can be *complex* because of the involvement of multiple stakeholders and the interdependence of AI components. The various parts of digital goods, such as hardware and digital contents, may be sold separately and produced by multiple parties. This can make it difficult to trace back the source of a malfunction, or to attribute liability for the malfunction to a single manufacturer. Injured parties may be confronted with hardware manufacturers, software designers, software developers, facility owners, and so on.[30] Consumers may have difficulty proving why their product does not work.[31] Nevertheless, the involvement of multiple stakeholders is neither new nor limited to AI products. Multiple stakeholders are involved in the production of many tangible products, such as cars, which are effectively regulated by the EU's existing liability regime.

Risk of harm related to AI can also increase if interdependent components of AI systems from different manufacturers raise compatibility issues.[32] AI systems could become more vulnerable as multiple systems, such as sensors, hardware and data,[33] become interconnected, and therefore, the risk of unanticipated or cascading problems grows In a litigation setting, the question may arise which party was responsible for ensuring safety and compatibility in the specific case.

Second, data-driven, probabilistic AI systems can be *opaque*. Such a lack of transparency[34] can make it difficult to identify causality, as it may be unclear exactly how input resulted in output.[35] Injured parties may not be able to identify that they have been harmed or trace back the source of the harm.[36] Opaque systems can therefore make it more difficult to hold decision-makers accountable or liable for the outcomes of these systems.[37]

Third, as AI systems become more *autonomous*, it also becomes difficult to trace back outcomes to human decisions, and to attribute liability to a specific actor.[38] Autonomy shifts control away from users, and possibly from manufacturers. AI systems that can act independently may be unpredictable.[39] If manufacturers cannot foresee how an AI application will decide or act once placed on the market, it may be difficult to hold them liable under producer liability.[40] A system's level of autonomy should not be confused with its level of automation. A system could be automated but its output completely pre-programmed. What is new and salient about AI-based

supervised, unsupervised and reinforcement learning, using a wide variety of methods including deep learning; 'logic- and knowledge-based approaches', including knowledge representation, inductive (logic) programming, knowledge bases, inference and deductive engines, (symbolic) reasoning and expert systems; and, 'statistical approaches' that include Bayesian estimation, search and optimisation methods.

[29] Expert Group on liability and New Technologies, 'Liability for artificial intelligence and other emerging digital technologies' (Report for the European Commission, 27 November 2019), 7.

[30] Ebers (n 84), 44. See also Christiane Wendelhorst, Sale of goods and supply of digital content – two worlds apart?: Why the law on sale of goods needs to respond better to the challenges of the digital age (*Study for the Juri Committee by the Directorate General for Internal Policies, Policy Department C*, 2016) <https://www.europarl.europa.eu/cmsdata/98774/pe%20556%20928%20EN_final.pdf> accessed on 5 August 2022, 7.; BEUC (n 17) 5; Yaniv Benhamou and Justine Ferland, 'Artificial Intelligence & Damages: Assessing Liability And Calculating The Damages' in Giuseppina (Pina) D'Agostino, Carole Piovesan and Aviv Gaon (eds.), *Leading Legal Disruption: Artificial Intelligence and a Toolkit for Lawyers and the Law* (Thomson Reuters Canada 2020), 6.

[31] Ebers (n 84) 44.

[32] A possible response by manufacturers is to exclude warranties when third-party software is used. However, to encourage competition in the market, it may be beneficial to find a legal solution that encourages manufacturers to allow third-party developers access to their products.

[33] Benhamou and Ferland (n 30) 6, referring to COM (2018) 237 final.

[34] Commission, 'On Artificial Intelligence – A. European approach to excellence and trust' (White Paper) COM (2020) 65 final, 17. This is also referred to as the "black box-effect" of AI (see COM (2020) 65 final 1)

[35] Ebers (n 84) 48.

[36] Expert Group on liability and New Technologies (n 29) 33.

[37] Lilian Edwards and Michael Veale, 'Slave to the algorithm: Why a right to an explanation is probably not the remedy you are looking for' (2017) 16 Duke Law & Technology Review 18.

[38] EP Resolution, 6, number 7.

[39] Scherer (n 26) 363-64; Peter M. Asaro, 'The Liability Problem for Autonomous Artificial Agents' (2016) AAAI Spring Symposia 2; Harry Surden and Mary-Anne Williams, 'Technological opacity, predictability, and self-driving cars' (2016) 38 Cardozo Law Review 121

[40] Hannah R. Sullivan and Scott J. Schweikart, 'Are current tort liability doctrines adequate for addressing injury caused by AI?' (2019) 21. AMA journal of ethics 160; Asaro (n 39) 2.

technologies is that a given input may not produce a predetermined output. The more autonomy AI systems achieve, the more tenuous it becomes to attribute legal responsibility for their actions to human beings.[41]

For safety purposes, autonomy is not necessarily a continuum: fully autonomous AI may be safer than human action, but an intermediate solution where AI complements human decision-making is only safer than human action if humans supervise it appropriately.[42] Therefore, fully autonomous AI presents different challenges for liability than AI that still requires some level of human supervision. If a task is fully delegated to AI, humans need to be able to rely on it functioning on its own. If AI supports decision-making but humans still need to monitor the system, the human-machine interface needs to work well.[43] The question arises in which circumstances users are allowed to rely on the AI system and at what point they should override its decisions. Users may over-rely on the system's autonomy, leading to accidents.[44] The liability question may be easier when there is no human-machine interaction, and the AI system functions fully autonomously.

Overall, AI systems are not inherently riskier than non-AI-based technologies. Rather, we can expect AI to be generally safer than non-AI counterparts. This means that there are opportunity costs of not relying on AI: if an AI system is not employed, a less safe human alternative will take its place. But this does not mean that AI does not affect risk or present challenges for liability: because of its characteristics, AI can be unpredictable and its decisions can be difficult to attribute to a human.

3.2. Gaps in liability associated with characteristics of AI

We identify three possible gaps in the existing liability regime with respect to AI: establishing fault, proving causality and dividing responsibility among different parties.

First, it is unclear what constitutes *fault* if AI actions cannot be reasonably anticipated.[45] As a result of the unpredictability of AI systems and the lack of control on the side of users, complex automated systems pose unique problems to fault-based liability regimes. Fault on the side of users would need to be established in terms of a failure to maintain the automated system or to oversee its functioning.[46] Starting from the premise that AI systems are primarily tools, fault-based liability can continue to hold their users to a duty of reasonable care while using it.[47] However, it is not clear whether the decision of a user to put an automated system into operation could be considered negligent if the system causes harm, at least not in all Member States.

Fault-based liability may run into problems particularly for decision-assistance AI, which is designed to interfere with human decision-making. If AI systems are to improve upon human decision-making, and we often lack

---

[41] Mark A. Chinen, 'The co-evolution of autonomous machines and legal responsibility' (2016) 20 Virginia Journal of Law & Technology 338
[42] The tragic crashes of Boeing 737 Max airplanes in 2018 and 2019 are another example of the risk involved in the interaction between humans and technology. The Boeing 737 Max relied on an automated software tool (the Manoeuvring Characteristics Augmentation System, (MCAS)" that was meant to work discreetly in the background. The system was therefore not mentioned in the pilot manual. Investigations identified the MCAS software as the proximate cause of the accidents but illustrated that failing pilot training and regulatory oversight also played roles (Chris Palmer, 'The Boeing 737 Max Saga: Automating Failure' (2020) 6 Engineering 2; Jack Nicas, Natalie Kitroeff, David Gelles, and James Glanz, 'Boeing built deadly assumptions into 737 Max, blind to a late design change' (*The NY Times*, 1 June 2019) <https://www.nytimes.com/2019/06/01/business/boeing-737-max-crash.html> accessed on 04 August 2022). The example shows that technology aimed at reducing a primary risk can create or exacerbate a distinctive type of secondary risk, arising from the interaction between the product and the user's experience with it (W. Bradley Wendel, 'Technological solutions to human error and how they can kill you: Understanding the boeing 737 max products liability litigation' (2019) 84 Journal of Air Law and Commerce 379, 431-32; see also Charles Perrow, *Normal accidents: Living with high risk technologies* (rev edn, Princeton University Press 2011)).
[43] Daniel Schönberger, 'Artificial intelligence in healthcare: a critical analysis of the legal and ethical implications' (2019) 27 International Journal of Law and Information Technology 171, 194.
[44] Alberto Galasso and Hong Luo, 'Punishing Robots: issues in the economics of tort liability and innovation in artificial intelligence' in Ajay Agrawal, Joshua Gans, and Avi Goldfarb (eds), *The Economics of Artificial Intelligence: An Agenda* (University of Chicago Press 2018), 5.
[45] William D. Smart, Cindy M. Grimm, and Woodrow Hartzog, 'An education theory of fault for autonomous systems' (2021) 2 Notre Dame Journal on Emerging Technologies 34; Jin Yoshikawa, 'Sharing the costs of artificial intelligence: Universal no-fault social insurance for personal injuries' (2019) 21 Vanderbilt Journal of Entertainment & Technology Law 1155 referring to Curtis E. A. Karnow, 'The application of traditional tort theory to embodied machine intelligence' in Ryan Calo, Michael A. Froomkin, and Ian Kerr (eds), *Robot Law* (Edward Elgar Publishing 2016), 52. See also Andrew D. Selbst, 'Negligence and AI's Human Users' (2020) 100 Boston University Law Review 1315, 1331 et seq.
[46] See also Ignacio N. Cofone, 'Servers and Waiters: What Matters in the Law of AI' (2018) 21 Stanford Technology Law Review 167
[47] See for a US perspective Selbst (n 45) 1320.

understanding of how it does this, it is questionable whether humans be considered negligent for relying on the AI system, when this leads to harm.

Second, proving *causality* can also be problematic if there is no traceable and predictable line between AI design and harm. Developers do not control automated systems quite the same way that, for instance, car manufacturers control how airbags deploy.[48]

Both of these aspects raise questions about the division of responsibility between manufacturers and users. It may be unclear if the harm is the result of a product defect or improper use if it is at least partly attributed to the general unpredictability of the system.[49] Autonomous systems will likely shift responsibility towards manufacturers.[50] The question arises what the limit of producer liability is for AI systems with a high level of autonomy – for instance, if any harmful action constitutes a defect, or if we accept that well-functioning AI systems may nevertheless cause harm from time to time. It is not necessarily clear what liability should continue to fall on owners and users. Moreover, problems may arise when dividing responsibility among manufacturers and other stakeholders involved in the functionality of the product, such as data providers.

4. Efficient liability rules for AI

Imposing tort liability on those engaged in activities that may cause harm operates as a mechanism for internalising harmful externalities. One objective of tort law is to create incentives for potential wrongdoers to invest in safety at an efficient level by making them pay damages.[51] Taking precaution generally involves the loss of money, time, or convenience.[52] Therefore, zero risks are typically not the socially optimal level of risk since a reduction of risk typically comes at an increasing marginal cost.[53] Depending on the particular application, precaution may take different forms: additional testing of AI-based solutions, possibly by outside experts operating as certifiers; a commitment to human supervision; and a careful design of the HMI interface to reduce human decision-making errors.

4.1. Liability standards: Fault-based or strict liability

When choosing between a fault-based or a strict liability regime for AI, information costs, the role of the injured party, the value of the (risky) activity and the type of risk are relevant considerations.

4.1.1. Information costs and incentives of the victims

Under a fault-based regime, the owner of e.g. a drone is held liable if the owner failed to take the safety precautions demanded by the standard of care. The owner is induced to take efficient precautions if lawmakers and courts determine the duty of care correctly. If the standard is set too high or too low, the owner of the drone will be induced to take a suboptimal level of precautions. In the case of AI, a fault-based regime is potentially suboptimal if courts cannot accurately assign liability, because legal conditions for liability, such as fault and causation, are difficult to prove for AI applications. The efficient level of precautions may not be easy to determine for AI on a general level. They may depend on the technical possibilities to control the actions of AI when designing it. There may be a trade-off here between the safety of AI and its sophistication. That is, more sophisticated AI may offer more benefits to users but may also become increasingly complex or unpredictable, and thereby riskier. If owners and users cannot control an AI system, fault-based liability does not serve its goal of steering them towards more cautious behaviour. In other contexts, this has been a reason to introduce a type of risk-based or strict liability.

[48] Smart, Grimm, and Hartzog (n 45) 12-13.
[49] See also *Ibid* 13.
[50] Selbst (n 45) 1322.
[51] Robert B. Cooter and Thomas Ulen, *Law and Economics* (6th edn, Pearson 2012), 189-190.
[52] Haydn Belfield, José Hernández-Orallo, Seàn Ó hÉigeartaigh, Matthijs M. Maas, Alex Hagerty, and Jess Whittlestone, 'Response to the European Commission's consultation on AI' (*Report by the Centre for the Study of Existential Risk*, 15 June 2020) <https://www.cser.ac.uk/news/response-european-commissions-consultation-ai/> accessed on 05 August 2022.
[53] Also, additional measures may be less effective. Assuming that precautions reduce the likelihood of an accident or the amount of harm, but at a decreasing rate of success, the optimal expenditure on precautions is finite. The efficient level of precaution prevails when the additional cost of a precaution measure equals the resulting reduction in expected costs of harm ("marginal costs equal marginal benefits"). Ethical concerns can be raised against such a cost-benefit approach. Based on ethical concerns it is conceivable to prohibit or at least limit the use of AI for certain types of activities.

The advantage of strict liability is that the legislator or the court does not need to have information on the optimal level of precaution. A strict liability rule induces the owner of the drone to take optimal precautions because it shifts all the costs of an accident on her. Theoretically (under perfect compensation), a strict liability rule internalises the costs of harm by requiring the injurer to pay for the social costs of his/her activity, regardless of the level of care taken.

However, because the injurer bears all the costs, a strict liability fails to set incentives for victims to take the appropriate care in situations where they, too, can affect the likelihood of an accident. In the economics literature, this is coined a double moral hazard problem.

4.1.2. Level of activity and innovation

Abstracting from the victims' incentives, shifting the full costs on injurers also means that a strict liability rule not only induces the optimal level of care but the optimal activity level as well. If an activity is inherently risky, even despite efficient precautions, we may want to refrain injurers from engaging in this activity altogether (or, at least, to reduce the level of this activity).[54]

A fault-based regime does not achieve this, since an injuring party can avoid paying for the costs of her activity by taking the required level of care. This explains why most jurisdictions impose strict liability for driving a car, for instance. In certain contexts, AI applications could also cause serious harm, even if proper precautions are taken, e.g. because the AI cannot be trained on sufficiently rich data.

However, the flipside of this is that if an activity is beneficial to society, the potential wrongdoer may become too careful. Strict liability may reduce their activity below the efficient level because negative externalities (i.e. harm) are internalised while positive externalities (i.e. external benefits to society) may not all flow back to them. AI applications produce clear benefits to third parties: cars with autonomous features may be safer, AI diagnostic tools may be superior to humans in detecting diseases, and algorithms produce all types of digital services that consumers enjoy. While employing AI reduces harm as compared to the alternative, there are opportunity costs of not employing AI.[55]

Moreover, investments in AI applications and their employment may contribute to innovations in AI in other fields as well. A concern for any liability rule and, in particular, strict liability is that start-ups deploying AI may not be able to bear the associated risk and thus go bankrupt, which would shift at least part of the liability to other parties or the injured party if not fully compensated.[56] Furthermore, foreseeing these problems entrepreneurs may not put their efforts into such a start-up in the first place or may not receive funding. Mandatory insurance could, at least partly, address this issue. However, this would come at the cost of negatively affecting the injurer's incentives to efficiently reduce harm and thus prove to be rather costly.

In this context, it should be acknowledged that liability does not necessarily chill innovation: it may also encourage firms to develop risk-mitigating technologies and improve the design of their products to reduce the likelihood of harm, and in turn, increasing user trust and take-up.[57] Absent liability, there are often insufficient incentives to do so, and potential users may correctly anticipate such a problem and delay adoption. In other words, liability can be a catalyst of innovation.

[54] Gerhard Wagner, 'Robot Liability' in Sebastian Lohsse, Reiner Schulze, and Dirk Staudenmayer (eds), *Liability for Artificial Intelligence and the Internet of Things* (Nomos 2019), 30 notes with respect to liability for AI: "shielding businesses from liability for the harm that they cause, for instance, with a view to fostering innovation, also seems problematic. This is not to say that innovation is unimportant or that incentives to innovate should not be generated. It is doubtful, however, whether the liability system is the preferred tool to create such incentives. To shield certain parties from responsibility for the harm that they actually caused amounts to a subsidization of dangerous activities, leading to an oversupply of such activities."
[55] See Belfield et al (n 52).
[56] Liability rules may then act as entry barriers. See also Miriam Buiten, Alexandre de Streel and Martin Peitz, (2020) 'Rethinking liability rules for online hosting platforms', 28 International Journal of Law and Information Technology, 139, 153 with respect to liability rules for online platforms.
[57] Alberto Galasso and Hong Luo, 'When does product liability risk chill innovation? evidence from medical implants' (2018) NBER Working Paper Series, 25068 <https://www.nber.org/system/files/working_papers/w25068/w25068.pdf> accessed 04 August 2022.

4.1.3. Types of risks

To address AI liability, it is useful to elaborate on several economic environments in which third parties experience damage and the effort decision by a party affects the likelihood that damage occurs or the severity of the damage. An important distinction is whether risks are idiosyncratic across people constituting the third party or highly correlated.

In the case of idiosyncratic risk, from an individual perspective, the damage remains a random event but for the liable party, the outcome is rather predictable. For instance, a firm may invest in reducing the fraction of products that posed a risk to third parties. While an accident is then highly unpredictable for an individual, a firm faces an average number of accidents, which can be predicted rather well. Fault-based liability may be based on calculating an optimal number of accidents (based on a cost-benefit analysis) and the firm may have to contribute to a pool if the number exceeds of observed accidents is above the optimal number, with a payment that is increasing in the number of accidents.

Of course, if the fault is directly observable with little cost for the legal system, damage payments due to fault-based liability can be assessed in individual cases; then the third party receives the full damage in case fault is established. Strict liability would award damages in all cases independent of the level of care. While third parties do better under strict liability, the firm may not have to bear the burden of higher damage payments fully on its own. In particular, when a firm sells a product, it may optimally pass at least part of the increased expected cost per unit under strict relative to fault-based liability through to its customers.

In many instances, including many cases involving AI risk, individual risks of third parties are highly correlated and a failure is a rare event. For example, think of insufficient protection of personal data that are hacked despite an AI system that is supposed to detect such threats.

In theory, the application of fault-based liability would work as follows: There is an optimal level of protection implying certain damage in case of failure (damage quantified in Euro, say X, and a probability p this damage happens). Keeping the size of the damage constant, we can focus on the failure probability. The optimal failure probability p^* would be compared to the observed failure rate p^0 and damages are awarded such that, from an ex ante point of view, the firm has to pay $(p^0 - p^*)X$. Whenever a failure is observed, it is inferred that the firm's fault increased the risk and, therefore, has to pay $\frac{p^0 - p^*}{p^0}X$. For example, the investigation concluded that the optimal failure probability is 1% and the actual failure probability was 5%, then due to fault, with total harm of € 1 million to third parties, damages of € 800k should be awarded as damages. Thus, the idea of fault can in theory be applied to such probabilistic events.

We note that in terms of incentives (taking the presence in the market as given), fault-based and strict liability regimes perform equally well in theory. From an ex ante perspective when deciding about the level of care, a firm minimises the sum of expected damages and avoidance costs, which, under strict liability, is $p^0X + C(1 - p^0)$ with respect to p^0. This implies that the firm chooses the risk such that $X = C'(1 - p^0)$. Under fault-based liability, a firm minimises $(p^0 - p^*)X + C(1 - p^0)$ with respect to p^0. This implies that the firm chooses the risk such that again $X = C'(1 - p^0)$.

The above argument shows that fault-based and strict liability leads to the same (efficient) level of care. The application of fault-based liability leads to practical problems since this requires the court to be able to calculate optimal and actual risk. Strict liability does not suffer from this practical problem, as, in our example, in case of failure, simply € 1 million are awarded. Therefore, if the individual risk is highly correlated and a failure is a rare event – we may want to call such an environment a high-risk environment – practical considerations make strict liability the preferred option. A downside of strict liability is that the expected payment for an innovator is higher than under fault-based liability.

This may lead to socially insufficient innovation if the innovator does not internalise all the social benefits from innovation and therefore refrains from entering the market or scaling up activity.[58]

[58] The chosen liability regime should therefore be seen in the context of public policy towards innovation. The choice of strict instead of negligence-based liability increases the call for public support to innovations to compensate for the higher expected payments to injured parties.

Summarising, the advantage of strict liability, as compared to a fault-based regime, is that legislators and courts do not need to have information on the optimal level of precaution in designing and testing AI-based solutions. By shifting the full costs of harm on injurers, a strict liability rule induces injurers to reduce their level of activity in cases where AI applications are inherently risky, even if proper precautions are taken. A drawback of strict liability is that it may reduce the beneficial use of AI applications below the efficient level, for instance, if their superior performance reduces harm to society as compared to not employing AI. If an individual risk is highly correlated and a failure is a rare event – a high-risk environment – practical considerations make strict liability the preferred option.

4.2. Who should be liable

When multiple parties affect the risk of harm, the question arises of who should be targeted by the liability rule. From a welfare perspective, this should be the least-cost avoider, i.e., the party which can minimise harm at the lowest cost. To the extent that some harm-reducing activities are complementary, this may imply that multiple parties should be targeted.

As explained above, in many AI-based solutions, there are several parties involved in providing a product or service. While damage may be easy to show in a court, the question arises which of the involved parties should provide damages and how much the total should be. To take a look at how different liability rules work out, it is necessary to specify how failures can occur with several parties. We distinguish between two polar environments. In the first environment considered below, care is cumulative; that is the provision of care by one party is a perfect substitute for care provided by another party. In the second environment, care by all parties is essential; that is, the provision of care by one party is a perfect complement to care provided by another party. Strict liability says that the total damage has to be compensated. As is often the case, it is often unclear which of the parties is to blame. For simplicity, suppose that two parties symmetrically contribute to the risk.

4.2.1. Substitute care

First, consider the substitute case. If at least one of the two parties engages in an effort the risk is assumed to be p^*, while if none of the two exerts effort the risk is assumed to be p^0. We assume that the socially efficient decision is that one of the two parties exerts effort. If in case of an accident when it cannot be verified which party did not exert effort, the total harm is X, and one simple rule would be to equally allocate damages to the parties. With strict liability, each party then would have to pay $X/2$. Such a rule cannot achieve an efficient effort provision.[59] It would be most efficient if the least-cost provider exerts the effort.

If this party can be identified at the outset, one may assign liability to this party. However, this may be difficult to do. Alternatively, the law could specify that a certain type of party will be held liable no matter whether its effort cost is lower than that of other parties. If this party bargains efficiently with the other party, both may agree to shift liability to the least-cost provider. This would guarantee an efficient level of effort at the lowest cost. Similarly, a fault-based liability assigns damages $\frac{p^0-p^*}{p^0}X$ to one party according to a pre-specified rule. As discussed above, in high-risk environments it will be more difficult to implement such a fault-based liability rule.

4.2.2. Complement care

Second, consider the complement case. Here both parties have to exert effort to reduce the risk from p^0 to p^*. We assume that the socially efficient decision is that both parties 1 and 2 exert effort; i.e., the total cost of effort provision satisfies $C_1(1-p^*)+C_2(1-p^*)<(p^0-p^*)X$. Strict liability that allocates the total harm among the two parties according to some exogenous sharing rule does not necessarily achieve the efficient effort.[60] If

[59] When both parties are equally good at reducing risk, this can be seen as follows. The probability of harm depends on the joint cost the two parties incur, $p(C_1+C_2)$ with $p'<0$ and $p''>0$. The welfare-maximizing solution satisfies $p'(C)X+1=0$. If each party has to pay for half the damage, party 1 minimizes $p(C_1+C_2)X/2+C_1$ with respect to C_1 and party 2 minimizes $p(C_1+C_2)X/2+C_2$ with respect to C_1. Thus, parties incur costs C with $p'(C)X/2+1=0$. Hence, the overall level of care is less with this solution than in the welfare-maximizing one.

[60] When both parties are equally efficient in reducing harm given that the other party has contributed more and damage is shared equally between the two parties, no party has an incentive to invest more in reducing the probability of harm than the other party. The problem for party 1 becomes to minimize $p(\min\{C_1,C_2\})\,X/2+C_1$. For $C_1\leq C_2$, this gives $p'(C_1)X/2+1=0$. Thus, the largest effort in harm

the two parties are symmetric, effort provision by the two parties is efficient if $2C(1-p^*) < p^0X$. If each party has to bear half of the damage, a party exerts effort if $C(1-p^*) < (1/2)(p^0 - p^*)X$ provided that it expects the other party to exert effort as well. If both parties behave that way, efficient effort is provided.

However, if a party is sceptical about the other party's effort provision if will not exert effort since this does not reduce the probability of an accident. Thus, there may be a coordination failure. Coordination failures can be avoided if parties can provide proof of effort that is verifiable in court and if a party that does not provide proof will be held fully liable.

If the effort is not binary, but its level can be adjusted, both parties will exert a socially inefficient level of effort. Simply assigning the total damage to the two parties cannot lead to an efficient level of care. The overall payment must be larger than the harm that is inflicted (above the efficient level). The incremental expected payment from not exerting effort must be equal to $(p^0 - p^*)X$ for each party; from a legal perspective, this means that there may need to be punitive damages to implement the socially efficient level of care.

The feature of effort being complements may be identified in particular in high-risk environment because the effort of all parties is needed to keep risk at bay.[61] For example, self-driving cars require reliable sensors and properly functioning AI-based software. If only one of the two malfunctions, this is sufficient to significantly increase the probability of harm.

One could ask whether assigning strict liability to one pre-specified party can achieve an efficient level of care. The party that is subject to liability may contract with the other party. However, by moving part of the liability risk to this other party, this creates a free-riding problem for itself as both parties are then only subject to part of the liability risk. Thus, assigning liability to one party and efficient contracting cannot resolve the under-provision problem as long as parties only have to cover the harm that has been incurred. A similar issue arises for fault-based liability rules that only account for the incremental harm beyond the efficient level. It is thus important to acknowledge that in the presence of complementarities in which individual effort cannot be proved in court, merely compensating damages will lead to an inefficient level of care. This holds even under strict liability.

If each of the parties providing care as perfect complements is fully liable for the damage, efficient care will be provided. However, the harmed party will then receive double damages. In the spirit of fault-based liability, by assigning damages to each party based on the incremental harm above the efficient level, under some conditions, the total payment can then be kept below the money equivalent of the total damage, and still, the incentives for effort provision are efficient.[62]

To use a numerical example, suppose a lack of care by either one of the two parties implies that an accident occurs with probability $p^0 = 5\%$, while with efficient care by both parties this probability is reduced to $p^* = 3\%$. Expected incremental harm from a lack of care is 2% times damage X. When X is 1 million Euros, it is €20k. In case of an accident, each party would be required to pay $((p^0 - p^*)/p^0)X$ = €400k. Thus, the total payment would be €800k, which is less than the total damage of 1 million Euros. As discussed above, the difficulty in applying this idea in practice is the lack of information by the court about p^0 and p^*.

Overall, for many AI-based solutions, several parties are involved in providing the product or service. If care by each party is essential to avoid a failure, and courts cannot verify the source of the failure, even strict liability leads to a socially inefficient level of care when no punitive damages are allowed.

---

reduction that can be supported by the behaviour of rational parties satisfies $C_1 = C_2 = C/2$ with $p'(C/2)X/2 + 1 = 0$. By contrast, the welfare-maximizing solution satisfies $p'(C/2)X + 1 = 0$. Hence, under this liability rule both parties spend too little on harm reduction from a welfare point of view.

[61] We acknowledge that complementarity is not specific to AI (see Michael Kremer, 'The O-ring theory of economic development' (1993) 108 The Quarterly Journal of Economics 551). At the root of the Boeing 737 Max crashes lies a malfunctioning sensor and its interaction with a software. More precisely, "erroneous AOA sensor reading triggered the plane's automated Maneuvering Characteristics Augmentation System (MCAS) anti-stall software" (Benjamin Zhang, 'Boeing and Ethiopian investigators confirm a faulty sensor was triggered on the 737 Max shortly before it crashed' (*Business Insider*, 04 April 2019) <https://www.businessinsider.com/boeing-ethiopian-investigators-confirm-bad-sensor-triggered-faulty-software-before-crash-2019-4?r=US&IR=T> accessed on 05 August 2022)

[62] See Robert B. Cooter and Ariel Porat, 'Total Liability for Excessive Harm' (2007) 36 Journal of Legal Studies 63. From an economics perspective, this is a simple application of the strategic issue in the provision of Cournot complements.

4.2.3. Holding producers and operators liable

Non-contractual liability should induce producers and operators to take an efficient level of care in designing, testing, and employing AI-based solutions. Policymakers should recognise that, by shifting costs of harm, the rules on liability may influence the design choices of producers in delegating decisions to AI-based systems or humans.

While producers control the product's safety features and provide the interfaces between the product and its operator, the operator exercises control over the use of the system. The operator decides in which circumstances the system is used and is in a position to oversee it in real-world situations. It is therefore appropriate to attribute some liability to operators who choose to delegate decisions to AI systems. However, it may be difficult to attribute fault to operators when they could not have reasonably anticipated the actions of the AI system. The fault would need to be established in terms of a failure to maintain or oversee the AI system. As AI systems gain autonomy, the scope of a duty to supervise may not be clear and injured parties may be unable to prove the fault of the operator and fail to obtain compensation from them under fault-based liability. Injured parties may also have difficulty proving causality between the AI systems' actions and the harm.

5. Recommendation to adapt EU liability rules for AI

In terms of an EU liability regime for AI, it needs to be determined not only what standard is appropriate but also what level of harmonization is necessary. One possibility is to allow Member States to continue applying their national liability regimes. If the EU is to introduce new rules, options are to introduce a separate regime for all AI, a separate regime for high-risk AI applications or continuing working with current sector-specific rules.

5.1. Adapting the Product Liability Directive to AI products

5.1.1. Rationales for reviewing the Product Liability Directive

Some of the challenges to liability identified above could be resolved by the review of the Product Liability Directive (PLD), which is expected later this year. The scope of revised product liability rules helps define how responsibility is divided between manufacturers, owners and users. The PLD has a horizontal scope and the rationale for updating it is broader than the concerns identified in relation to AI. Still, this rationale is closely related to technological development.

AI systems shift the locus of control away from users towards manufacturers.[63] For technical products that do not rely on AI, the manufacturer controls the product's safety features and provides the interfaces between the product and its user, while the user exercises control over the mechanical device when employing it in real-world situations.[64] For AI systems, users will be able to exert much less control. As a result, accidents will become less dependent on the care taken by the individual user. The liability of the user is likely to increasingly recede into the background, meaning that the role of liability of the manufacturer becomes more significant for injured parties to obtain compensation.[65] In short, where producers are in a better position than consumers to control risk, an incentive-based approach would shift the relative burden of liability towards producers.[66] This incentivises producers to reduce the AI system's risk through designing and manufacturing the system.

The Commission has identified three problems with applying the provisions of the PLD in the context of IoT and autonomous connected systems. The first is the complicated product or service value chain, with interdependencies between suppliers, manufacturers and other third parties. The second is the uncertainty in relation to the legal nature of IoT devices, i.e. whether they are products, services, or products that come with the sale of a service. The third is the autonomous nature of these technologies.[67] The Expert Group Report and

[63] Wagner (n 54) 37.
[64] Gerhard Wagner, 'Robot, inc.: Personhood for autonomous systems?' (2019) 88 Fordham Law Review 591, 602.
[65] *Ibid*; Astrid Seehafer and Joel Kohler, 'Künstliche Intelligenz: Updates für das Produkthaftungsrecht?' (2020) Europäische Zeitschrift für Wirtschaftsrecht 213, 213; Carina Lutter, 'Fragen der Produkthaftung im Hinblick auf den Betrieb unbemannter Schiffe' (2017) 5 Recht der Transportwirtschaft 281, 281.
[66] Galasso and Luo (n 44) 5; Melinda F. Lohmann, 'Liability issues concerning self-driving vehicles' (2016) 7 European Journal of Risk Regulation 335, 338.
[67] Commission, 'Building a European data economy' (Communication) COM (2017) 09 final.

the White Paper on AI both concluded that some key concepts in the PLD require clarification to be apt to deal with emerging digital technologies.[68] The European Parliament has also called on the Commission "to review the Directive and consider adapting such concepts as 'product' 'damage' and 'defect' as well as adapting the rules governing the burden of proof".[69]

5.1.2. Definition of product

Currently, the PLD covers tangible products. This means that hardware components of an AI system, and software integrated into tangible AI systems are covered by the PLD[70], but for standalone software this is unclear.[71] For software, the medium can be decisive for qualifying as a product. As a result, software stored on a DVD or a Flash-drive[72] is covered by the PLD, but downloaded software may not be. Member States have applied the concept of software from the PLD differently in their national implementations.[73] Clarifying the definition of product is necessary because it determines if manufacturers of software can be held liable under the PLD.[74]

In the age of digitalisation, differentiations between tangible and intangible objects of use are more difficult to justify.[75] It is unclear why the mode in which computer programs are stored, copied, and distributed should be relevant for the application of the PLD. The main purpose of the PLD was to ensure a fair distribution of the risks associated with industrially manufactured between the injured party and the manufacturer.[76] The risks associated with downloaded software do not appear very different from their traditional counterparts supplied on CDs.[77] Once the software is introduced to a computer, it brings about material and tangible changes.[78] This is obvious where software is integrated into a machine[79] but is also easily imaginable for intangible software: one could think of an insulin therapy app used by a patient making an error, or malware corrupting all of a consumer's files. The risks involved in software, irrespective of its medium, therefore support including software in the notion of products.[80]

Such an approach would raise questions on how to delineate products from services. Items that were once consumed as products purchased by the consumer are delivered not only in the cloud but often also as

[68] COM (2020) 65 final; Commission, 'Report on the safety and liability implications of Artificial Intelligence, the Internet of Things and robotics' COM (2020) 64 final; Expert Group on Liability and New Technologies (n 29) 27-28.
[69] European Parliament, *Resolution on automated decision-making processes: ensuring consumer protection and free movement of goods and services* (2019/2915(RSP)) (EP Resolution).
[70] Jessica S. Allain, 'From Jeopardy! to jaundice: the medical liability implications of Dr. Watson and other artificial intelligence systems' (2013) 73 Louisiana Law Review 7; Susanna Navas, 'Robot Machines and Civil Liability' in Martin Ebers and Susanna Navas (eds), *Algorithms and Law* (Cambridge University Press 2020), 167 w.r.t. robots; also referring to Duncan Fairgrieve, Geraint Howells, Peter Møgelvang-Hansen, Gert Straetmans, Dimitri Verhoeven, Piotr Machnikowski, André Janssen, and Reiner Schulze, 'Product liability directive' in Piotr Machnikowski (ed), *European Product Liability: An analysis of the state of the art in the era of new technologies* (Intersentia 2016), 47. Case law and jurisprudence has largely already taken this approach. Given that the Directive covers electricity, it could be argued that a product does need to be tangible, see Ebers (n 84) 58.
[71] See further Lutter (n 65) 282. Some authors take the position that the PLD already now extends to digital content, e.g. Bernhard A. Koch, 'Product Liability 2.0 – Mere Update or New version?' in Sebastian Lohsse, Reiner Schulze, and Dirk Staudenmayer (eds), *Liability for Artificial Intelligence and the Internet of Things* (Nomos 2019), 106; Gerhard Wagner, 'Produkthaftung für autonome Systeme' (2017) 217 Archiv für die civilistische Praxis 707, 717-8; and Gerald Spindler, 'Haftung im IT-Bereich' in Lorenz Egon (ed), *Karlsruher Forum 2010: Haftung und Versicherung im IT-Bereich* (Verlag Versicherungswirtschaft 2011), 41-43.
[72] Written Question No 706/88 by Mr Gijs de Vries to the Commission: Product liability for computer programs, Official Journal [1989] OJ C 114/42.
[73] See e.g. Kristin Nemeth and Jorge Morais Carvalho, 'Time for a Change? Product Liability in the Digital Era' (2019) 8 Journal of European Consumer and Market Law 160 on the differences between the German and the Austrian implementation.
[74] See further M. Stöber, M.-C. Pieronczyk, and A. Möller, 'Die Schadensersatzhaftung für automatisierte und autonome Fahrzeuge' (2020) 90 Deutsches Autorecht 609, 612.
[75] See also e.g. *ibid* 613.
[76] Product Liability Directive, recital paras. 2 and 7.
[77] Wagner (n 64) 604.
[78] K. Alheit,'The applicability of the EU product liability directive to software' (2001) 34 Comparative and International Law Journal of Southern Africa 188.
[79] *Ibid* 201.
[80] The view that the PLD should apply independent of the mode in which computer programs are stored, copied and distributed is shared by the Expert Group on Liability and New Technologies (n 29); Christoph Schmon, 'Product Liability of Emerging Digital Technologies' (2018) 3 Zeitschrift für Internationales Wirtschaftsrecht 254, 254; Stöber, Pieronczyk, and Möller (n 74) 613; Hans Steege, 'Auswirkungen von künstlicher Intelligenz auf die Produzentenhaftung in Verkehr und Mobilität' (2021) Neue Zeitschrift für Verkehrsrecht 6, 7; Wagner (n 54) 42; Daily Wuyts, 'The product liability directive–more than two decades of defective products in Europe' (2014) 5 Journal of European Tort Law 1, 6; Rolf H. Weber, 'Liability in the Internet of Things' (2017) 3 Journal of European Consumer and Market Law 207, 210.

services.[81] For instance, where consumers would previously buy a CD, they now have a subscription to Spotify. Digital goods have blurred the distinction between products and services.[82] As cloud-computing abilities improve, more AI systems may be operated on service model as well – not just digital goods, but physical ones as well.[83] As a result, it may become increasingly difficult to draw a sharp line between products and services for IoT and AI systems.

Traditionally, the different risks associated with products and services justifies imposing strict liability for product defects, but not for services. A consumer of a service may more easily prove that the service provider was negligent, than it is for a consumer to provide evidence about the defective nature of a product.[84] The distinction between products and services may be less justified with respect to many digital goods, given that their risks may well be the same.[85] In the long term, either a common liability regime may have to be adopted for both, or clear definitional criteria will need to be developed.[86]

5.1.3. Definition of defect

The definition of "defect" is pivotal in determining producer liability for autonomously operating systems. However, some aspects are unclear.

First, producers are only liable under the PLD for a defect arising before the product was placed on the market. This reflects that producers have no control over the product from that moment onwards. This control-criterion is problematic if product safety relies on a producer's updates to the software,[87] or if AI systems are intended to continue learning once they are placed on the market.[88] A reform could consider the dynamic nature of software products, IoT devices and AI systems.[89] One option could be to extend liability to producers that fail to provide updates relevant to the safety of the product. It may be useful to clarify if consumers may expect these updates to be delivered throughout the life-cycle of the product.[90] Such an obligation to monitor the product is alien to the current PLD. If it were introduced, to prevent an over-broad and open-ended liability clear criteria are needed on how long such an obligation should reasonably exist.[91] Failure by users to install the software update should count against liability.[92] [93] For AI systems with learning capabilities, it needs to be clarified what monitoring duties the producer has with respect to these capabilities.[94]

A second clarification to the concept of "defect" is what this means in respect of autonomous AI systems. Does any instance of harm constitute a defect, or should be accepted that a well-functioning AI system could still cause harm? If *some* failure rate is accepted, *what* failure rate is acceptable? It needs to be clarified how far the concept of defect extends for deliberate, but undesirable operations of AI systems with self-learning capacities. For sophisticated AI systems, it may not be possible to draw the line between harm resulting from AI's autonomous decisions and harm resulting from a defect.[95]

---

[81] Omri Rachum-Twaig, 'Whose Robot Is It Anyway?: Liability for Artificial-Intelligence-Based Robots' (2020) University of Illinois Law Review 1141, 1157 and the text in footnote 93.
[82] BEUC (n 17) 7; Expert Group on Liability and New Technologies (n 29) 28
[83] Rachum-Twaig (n 81) 1172 names example of robots.
[84] Martin Ebers, 'Regulating AI and Robotics' in Martin Ebers and Susanna Navas (eds), *Algorithms and Law* (Cambridge University Press 2020) 58. See also Case C-495/10 *Dutreux* [2011] ECR I-14155; Commission Evaluation COM (2018) 246 final 7. See e.g. Brigid M. Carpenter and Caldwell G. Collins, 'The Shirt Off My Back: Using the Relationship Between a Product and a Service to Your Advantage' (*IADC Committee November 2021 Newsletter,* November 2012)
<https://www.bakerdonelson.com/files/Uploads/Documents/ProductLiabilityNovember2012.pdf> accessed 04 August 2022
[85] Benhamou and Ferland (n 30) 13
[86] Scott Marcus, 'Liability: When Things Go Wrong in an Increasingly Interconnected and Autonomous World: A European View' (2018) 1 IEEE Internet of Things Magazine 4.
[87] Expert Group on Liability and New Technologies (n 29) 28; BEUC (n 17) and BEUC, 'AI RIGHTS FOR CONSUMERS' (BEUC-X-2019-063, 23 October 2019) <www.beuc.eu/publications/beuc-x-2019-063_ai_rights_for_consumers.pdf> accessed 07 August 2022
[88] Commission AI Report, p. 15. See also Seehafer and Kohler (n 65) 214.
[89] See Expert Group on Liability and New Technologies (n 29) 43; Benhamou and Ferland (n 30) 13.
[90] Schmon (n 80); Seehafer and Kohler (n 65).
[91] Steege (n 80) 12..
[92] Seehafer and Kohler (n 65) 217; Steege (n 80) 12.
[93] The Commission has already pointed out that subsequent updates cannot be the sole responsibility of the manufacturer: the user would have the obligation to install safety-relevant updates (COM (2020) 64 final 15).
[94] Seehafer and Kohler (n 65) 214, referring to Tim Hey, *Die außervertragliche Haftung des Herstellers autonomer Fahrzeuge bei Unfällen im Straßenverkehr* (Springer Gabler 2019).
[95] Benhamou and Ferland (n 30**Fehler! Textmarke nicht definiert.**) 7.

There are two options: extending the concept of "defect" for fully autonomous AI applications to any harm they cause or distinguishing more clearly between different types of defects. The first option could be justified by the fact that the producer designs the learning process for the AI system, is best-placed to judge whether the product is safe enough to be put on the market, and profits from selling it.[96] Such a rule would moreover encourage producers to inform users about contexts in which the application is unable to work fully autonomously. However, from a practical perspective, we can expect producers to avoid liability by simply not marketing applications as fully autonomous. Producers would likely add extensive product manuals outlining the contexts in which users still have a duty to monitor the system. More generally, the consequences of the liability rules on the design and marketing of AI systems should be given serious thought. If products with a higher autonomy level are treated differently under product liability, this will likely affect how products are marketed and/or how they are designed.

Aside from this practical problem, it would be unreasonable to require absolute safety in the context of liability.[97] Certain situations, such as in healthcare, may require demanding absolute safety because of the high stakes involved,[98] which is reflected by regulatory safety standards. Generally, extending strict liability to AI manufacturers so that they are responsible for any AI harm shifts an undue portion of the burden on manufacturers.[99] Such a regime would force AI manufacturers to bear the negative externalities without compensation for the value of the tremendous positive externalities of AI.[100] It would also place insufficient burden on the owners who benefit from employing AI and impose risks on others by doing so (see further below). Moreover, waiting for nearly perfect AI before employing it is likely more costly than accepting a reasonable failure rate. The liability rules should reflect this.[101]

A second option would be to clarify the different types of defects and potentially differentiate the applicable liability regime. The PLD uses a single criterion to establish the defectiveness of a product. It does not distinguish between different types of defects, as is the case e.g. in American product liability law.[102] One possibility could be to limit strict liability to manufacturing defects, while a presumption of fault could be applied for defects in design and instructions to users.[103] A drawback of this approach is that in practice it would likely limit (strict) producer liability for AI systems as compared to other products. For AI systems, defects are more likely to originate in their design and instructions than in their manufacturing.[104]

A preferable option is to clarify when a design defect constitutes a defect under the PLD, i.e. determine the acceptable failure rate. This means accepting that AI systems will inevitably cause harm, even if they are free from software bugs, hardware errors, or failures of engineering precaution.[105] Rather than focusing on an individual AI system, we need to ask what error rate in a fleet of AI systems that operates by the same algorithm constitutes a defect.

Generally, the safety requirements placed on the manufacturer increase with the risks associated with the product. We may also expect AI systems to be safer than the "dumb" products they are replacing. However, we need to consider how much safer than human decision-making we require AI systems to be. It may not be possible or useful to compare the performance of an AI system with how a carefully acting human would have behaved in a specific situation. The first reason is that precisely because we require AI to do better than humans

---

[96] See also Caroline Cauffman, 'Robo-liability: The European Union in search of the best way to deal with liability for damage caused by artificial intelligence' (2018) 25 Maastricht Journal of European and Comparative Law, 527, 530.
[97] See e.g. PT Schrader, 'Haftungsfragen für Schäden beim Einsatz automatisierter Fahrzeuge im Straßenverkehr' (2016) 86 Deutsches Autorecht 242, 243; and Friedrich Graf von Westphalen, 'Produkthaftungsrechtliche Erwägungen beim Versagen Künstlicher Intelligenz (KI) unter Beachtung der Mitteilung der Kommission COM(2020) 64 final' (2020) 35 Verbraucher und Recht 248, 250.
[98] Lutter (n 65) 283.
[99] Yoshikawa (n 45) 1165 and 1171.
[100] See Section 4. See also Yoshikawa (n 45).
[101] Nidhi Kalra and David G. Groves, *The Enemy of Good: Estimating the Cost of Waiting for Nearly Perfect Automated Vehicles* (RAND Corporation 2017) <www.rand.org/pubs/research_reports/RR2150.html> accessed on 05 August 2022.
[102] See further Wuyts (n 80) 10.
[103] Navas (n 70) 168.
[104] *Ibid*; F. Patrick Hubbard, 'Sophisticated robots: balancing liability, regulation, and innovation' (2014) 66 Florida Law Review 1803, 1821-1823; Martin Ebers, Autonomes Fahren: Produkt- und Produzentenhaftung in Bernd H. Oppermann and Jutta Stender-Vorwachs (eds), *Autonomes Fahren* (C.H. Beck 2017) 111-12.
[105] Mark A. Lemley and Bryan Casey, 'Remedies for robots' (2019) 86 The University of Chicago Law Review 1311, 1327-8.

a comparison with reasonable human decision-making is pointless.[106] Second, the point of reference differs: in the case of a human being, it is the decision to act in an individual case, while for an AI system, it is whether the programming for an entire series of products could and should have been done more careful to prevent the occurrence of the damage.[107] Courts would need to identify shortcomings that could have been avoided by alternative programming.[108] Self-learning AI systems that originally function well and develop a malfunction in practical use could be considered already initially not error-free.[109] A third reason is that the pool of accidents that an autonomous system causes may be easily avoidable by humans – one can think of the ability of an autonomous car to recognise a white truck in a bright environment. Despite making errors that humans would not, AI systems may overall still make significantly fewer errors. As a result, it may be misguided to compare the standard for safety to humans.[110]

Overall, with regard to autonomous AI systems, we need to consider what design flaws for AI are unacceptable and what error rate is unacceptable. Moreover, the burden of proof (discussed below) and regulatory safety standards may help mitigate the challenges that autonomous AI systems pose for the concept of the defect.

Regulatory safety standards can also help reduce negative unintended consequences from autonomous decision-making. It is important to increase our understanding of how AI systems learn once they are placed on the market, to take the appropriate regulatory steps. A “defect” becomes more difficult to recognise or even define if AI devices continue to learn on their own once they are on the market. Such devices would be less predictable and harder to control. If AI devices are thoroughly tested after a learning process and “frozen” when placed in the market, harm from unintended actions may be less likely. If regulation precludes AI products from entering the market without “freezing” them this could reduce the need for interpreting “defect” more broadly in the product liability rules. However, such an option should always be weighed against the lost benefits of not employing these systems with learning capabilities.

5.1.4. Burden of proof

The PLD requires injured parties to prove that the product was defective and that it caused the injury. Proving a defect can be difficult for consumers for any technically complex product. National courts have therefore developed ways to facilitate the burden of proof in such situations, including by disclosure obligations for the producer, or by allocating the costs of experts' opinions.[111]

For AI products, proving a defect may be more complex if the defect is difficult to identify. If, for instance, an AI diagnosis tool delivers a wrong diagnosis, there may be no obvious malfunctioning to the user.[112] Depending on the definition of a defect, users may be asked to show that harm was the result of a flaw in the AI device, and not of its autonomous decision-making. Proving causality in the context of AI harm may be difficult, especially if some human supervision was still required.[113] The injured party may have difficulty proving that the AI system, not his negligence, caused the harm. AI developers may also try to argue that it is impossible to anticipate precisely how AI systems will act, meaning that the harm was unforeseeable.[114] While this is unlikely to succeed as a defence, such questions could arise when AI did exactly what it was intended to do (act autonomously) and nevertheless caused harm. The assessment of the causal link will often require expert advice, the cost of which may discourage injured parties from suing.[115]

[106] Jean-Sébastien Borghetti, ‘How can Artificial Intelligence be Defective?’ in Sebastian Lohsse, Reiner Schulze, and Dirk Staudenmayer (eds), *Liability for Artificial Intelligence and the Internet of Things* (Nomos 2019), 69.
[107] Seehafer and Kohler (n 65) 214 and footnote 23 therein.
[108] Wagner (n 64).
[109] See Philipp Etzkorn, ‘Bedeutung der «Entwicklungslücke» bei selbstlernenden Systemen – Rechtliche Fragen zur fortdauernden Softwareentwicklung durch maschinelles Lernen im Praxiseinsatz’ (2020) 23 Multimedia und Recht, 360, 362.
[110] Wagner (n 64) 605.
[111] Wuyts (n 80) 24 and the references therein.
[112] Borghetti (n 106) 67.
[113] Causality is governed by national rules, given that the PLD does not define a causal relationship. See on different national approaches e.g. Wuyts (n 80) 25.
[114] Weston Kowert, The foreseeability of human-artificial intelligence interactions. (2017) 96 Texas Law Review, 181, 191.
[115] Cauffman (n 96) 530; Wuyts (n 80) 24.

Reversing the burden could facilitate claims for parties injured by highly complex technologies.[116] However, this would significantly alter the current distribution of risks to the detriment of the manufacturers. It would also depart sharply from the current principles of the PLD.[117] Given that AI systems may be equipped with event logging or recording systems, victims may moreover be able to get access to better data about the cause of an accident than for non-AI products.[118]

A preferable way to facilitate the burden of proof for the injured party could be introducing a lower standard of proof. This could be accompanied by evidence disclosure duties,[119] cost-shifting rules for expert advice, or – as was as data protection rules permit – requirements to collect data about the functioning of the system, allowing them to retrace possible causes for an error at a later stage.[120]

A review of the PLD should also consider the scope of the defences available to producers, particularly the development-risk defence.[121] AI systems give this defence more relevance: if an AI system with learning capabilities causes harm, the producer may be able to argue that the particular learning and decision-making process was not foreseeable.[122] One could argue that consumers should not bear the risk of gaps in the knowledge about the safety of new technologies. At the same time, the defences are included to maintain incentives to innovate, the requirements for excluding liability are high, and the burden of proof lies with the producer.[123] In light of AI systems, it needs to be assessed what is the 'state of scientific and technical knowledge in relation to machines powered by automated decision making. In a review of the PLD, clarifying the defence could be justified.[124]

5.2. A possible new liability regime for AI products

While AI systems shift the locus of control to producers, producers do not influence the final use of the AI system. It is therefore justified to attribute some liability to the party who owns the AI-powered product (owner) or who uses it (keeper): the "operator".[125] Much of today's AI technology is not fully autonomous and requires at least some level of human supervision. Operator liability encourages them to take precautions in supervising the AI system.[126] Even for highly autonomous AI systems, the operator decides if and how to employ it. Liability provides an incentive for operators to keep an AI device updated and ensure that it is used properly.[127] The operator moreover benefits from employing AI. Holding producers liable for every case of harm, even those they have no control over and are not capturing the benefits from, may harm beneficial innovation.[128]

Finding an appropriate analogy for AI systems in existing rules on strict liability is not straightforward. At the same time, Member States differ considerably in what contexts they subject people to strict liability.[129] This means that an EU approach introducing strict liability for AI would constitute a sharp departure from the standard liability regime in several Member States.

---

[116] Charlotte De Meeus, 'The Product Liability Directive at the Age of the Digital Industrial Revolution: Fit for Innovation?' (2019) 8 Journal of European Consumer and Market Law 149, 151; Wagner (n 71) 747; Seehafer and Kohler (n 65) 216; Roeland de Bruin, 'Autonomous Intelligent Cars on the European Intersection of Liability and Privacy' (2016) 7 European Journal of Risk Regulation, 485, 495, who acknowledges that this could negatively impact innovation.
[117] See further Seehafer and Kohler (n 65) 216.
[118] Koch (n 71) 110; Gerald Spindler, 'Roboter, Automation, künstliche Intelligenz, selbst-steuernde Kfz – Braucht das Recht neue Haftungskategorien?' (2015) 31 Computer und Recht 766, 772.
[119] Schmon (n 80) 6.
[120] Antje von Ungern-Sternberg, 'Artificial Agents and General Principles of Law' in Andreas von Arnauld, Kerstin von der Decken, & Nele Matz-Lück (eds), *German Yearbook of International Law* (Duncker & Humblot 2018) 9.
[121] Benhamou & Ferland (n 30) 13-14. See further Christian Twigg-Flesner, 'Guiding Principles for Updating the Product Liability Directive for the Digital Age' (2021) European Law Institute Innovation Paper Series (Pilot Innovation Paper), 10-11
[122] Seehafer and Kohler (n 65) 215.
[123] Schrader (n 97) 243 and the references therein.
[124] Schmon (n 80). On the defence see further Etzkorn (n 109) 363.
[125] The Expert Group on Liability and New Technologies (n 29) defines as operator as "the person who is in control of the risk connected with the operation of emerging digital technologies and who benefits from their operation". The proposed AI Act defines operator as "provider, the user, the authorised representative, the importer and the distributor" (COM (2021) 206 final, art 3 (8)).
[126] See also Galasso and Luo (n 44) 6.
[127] See also *ibid*.
[128] Benhamou and Ferland (n 30).
[129] See further Ernst Karner Bernhard A. Koch and Mark A. Geistfeld, 'Comparative Law Study on Civil Liability for Artificial Intelligence', Study Prepared for the Directorate-General for Justice and Consumers, 2021.

5.2.1. Scope of a specific AI liability regime

If a stricter standard is to be introduced for certain AI applications, we distinguish two possibilities. The first option consists of aligning the scope of the stricter standard with existing sector-specific rules while the second option amounts to introducing a horizontal liability framework for newly pre-defined high-risk AI applications.

The first option has the advantage of limiting the risks shifting from the victims to the operators as it will limit the strict liability to the sector already defined in the law. Such an option also ensures coherence with safety rules which are also defined at the sector level. Finally, from a legitimacy point of view, this option ensures that the scope of strict liability rule is defined by the legislator when adopting sector regulation and not by the Courts when interpreting criteria to define high-risk applications. Conversely, the second option has the advantage of being more flexible and adaptable for technologies that evolve quickly. This second option is also favoured by the European Parliament in its Resolution of October 2020[130] which recommends that all high-risk AI-systems be exhaustively listed in an Annex to Regulation it proposes.

As AI applications differ in both the benefits and the risks they create for society, it is appropriate to differentiate in the regulatory and liability requirements that apply to different AI applications. However, we identify several potential problems with introducing a horizontal liability framework for high-risk AI applications. Listing "high risk" AI applications may presuppose that AI applications create similar risks regardless of the context in which they are applied. AI encompasses various technologies, which may be used in a wide range of applications, which in turn could be employed in various contexts. Existing sector-specific regulation already reflects the need to differentiate regulation according to the context in which technology is applied. Introducing a horizontal liability regime for high-risk AI would constitute a departure from the existing EU approach to liability. For various reasons, it may be better to streamline liability rules for high-risk AI with existing sector-specific regulation.[131] Including liability rules in sector-specific regulation would allow the Member States to maintain their general rules on strict liability, which would be complemented by enhanced obligations in certain sectors. A horizontal liability regime for a limited number of applications would cut through this system.

The key question is how to define the scope of a horizontal regime for high-risk AI. It needs to be sufficiently clear for users (or other types of operators), and courts to understand what applications are covered by this framework. If a strict liability framework were introduced for high-risk AI systems, cases are likely to revolve around the question of whether a specific device is AI, and whether it is a high risk,[132] as this would determine whether it is covered by the general fault-based or by strict liability. To avoid introducing a new source of legal uncertainty, a European regime would need to clearly define which AI applications are high-risk. The risk of a certain technology may be different depending on the context. Therefore, the risk would have to be determined for a certain device and, in the case of general-purpose devices, for each particular use of that device.

According to the proposed AI Act, an AI system shall be considered high-risk where both of the following conditions are fulfilled: (a) the AI system is intended to be used as a safety component of a product, or is itself a product, covered by the Union harmonisation legislation listed in Annex II; (b) the product whose safety component is the AI system, or the AI system itself as a product, is required to undergo a third-party conformity assessment with a view to the placing on the market or putting into service of that product pursuant to the Union harmonisation legislation listed in Annex II.[133]

The European Parliament defines high risk as '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.[134] The European Parliament Resolution proposes to list all high-risk AI-systems in the EU legislation that should be reviewed at least every six months. In practice,

[130] See EP Resolution, art 4.
[131] Miriam C. Buiten, 'Towards intelligent regulation of Artificial Intelligence' (2019) 10 European Journal of Risk Regulation 41; Chris Reed, 'How should we regulate artificial intelligence?' (2018) Philosophical Transactions of the Royal Society A 376:2128.
[132] See also Sebastian Lohsse, Reiner Schulze, and Dirk Staudenmayer, 'Liability for Artificial Intelligence' in Sebastian Lohsse, Reiner Schulze, and Dirk Staudenmayer (eds), *Liability for Artificial Intelligence and the Internet of Things* (Nomos 2019), 21.
[133] Proposed AI Act, art 6.
[134] EP Resolution, Annex art. 3(c).

maintaining such a list may be burdensome. The risk of each AI application may need to be reviewed regularly: after it is placed on the market, an AI system may evolve and new vulnerabilities may arise. Risk assessment would then need to be repeated once a product is already placed on the market.

Ultimately, only few AI applications might be covered by a uniform horizontal regime. On the one end, some applications are already covered by specific regulation, such as autonomous vehicles or medical devices. These AI applications can continue to be governed by sector-specific rules. For some AI-driven applications, we may find it useful to extend the existing regime for vehicles or other forms of transportations. One could think of drones. For such devices, compulsory liability insurance schemes may also need to be imposed, similarly as for car owners.[135] On the other end, there is a large group of AI systems that is not high-risk. For this group, strict liability would not be justified if we broadly follow the existing principles for strict liability in the Member States. Accordingly, an EU horizontal strict liability regime would cover the group of AI systems that *should* be covered by stricter liability rules but *would not* be covered by any (extended) sector-specific policy. If this group is small, introducing a horizontal EU liability regime may not be justified.

Given the need to differentiate for AI-driven applications based on the context in which they are used, the question is how broadly a horizontal regime can ultimately apply. If most high-risk AI applications will be regulated elsewhere, following the sector-specific rules with a liability regime would be the preferred approach.

5.2.2. AI liability standard

If a new regime applying to (high-risk) AI applications is to be introduced, three broad options are possible: Option 1 with fault-based regime with a higher duty of care, option 2 with fault-based regime with a rebuttable presumption of the fault and/or causality link and option 3 with strict liability.

Option 1 would require the legislator to establish a duty of care on operators that are sufficiently clear to be uniformly interpreted throughout the EU, and that allows victims to establish that the operator was at fault. The difficulty is to establish what duty of care operators owe to others when they employ a (semi-) autonomous AI system. Establishing fault is easily established when operators use an AI system to deliberately cause harm, but it is much more difficult for unintended harm.[136] For decision-assistance technology relying on AI, a different standard may need to be found: the difficulty there is that it is designed to (partly) replace human decision-making. The standard of care would need to clarify in which cases operators can rely on the technology, and when doing so would constitute a fault. Thus such an option does not relieve victims from the problem of proving fault and causality and likely leads to different interpretations from courts throughout the EU.

Operators of AI devices have to comply with a duty of care in choosing to employ an AI system, maintaining the system, and supervising it.[137] The duties of the operator necessarily relate to the level of autonomy of the AI device. If the operator may reasonably expect the AI device to act fully autonomously, the operator has no duty to monitor an AI device.[138] The responsibility of the operator decreases and that of the producer increases, if a product promises to function fully autonomously when used for its intended purpose. The duty of the operator is then limited to using and maintaining the device properly.

To allow operators to uphold their monitoring and maintenance duties, producers should be obliged to instruct operators properly on the use of the product. If producers face a strict liability standard, they should have an interest in providing precise warnings and instructions to buyers to avoid producer liability. If products do not operate autonomously in all circumstances, we can expect producers to issue warnings urging users to monitor the device. Manuals for vacuum robots, for instance, include extensive safety instructions.[139] If the owner of a

[135] See e.g. Georg Borges, 'Product Liability 2.0 – Mere Update or New version?' in Sebastian Lohsse, Reiner Schulze, & Dirk Staudenmayer (eds), *Liability for Artificial Intelligence and the Internet of Things* (Nomos. 2019); Navas (n 70) 166; David Levy, 'Intelligent no-fault insurance for robots' (2020) 1 Journal of Future Robot Life 35.
[136] See also Von Ungern-Sternberg (n 120) 6; Kowert (n 114).
[137] Expert Group on Liability and New Technologies (n 29) 44. Ruth Janal, 'Extra-Contractual Liability for Wrongs Committed by Autonomous Systems' in Martin Ebers and Susana Navas (eds), *Algorithms and Law* (Cambridge University Press 2020), 193 notes that, "the users of an autonomous system may be held liable for the acts of the system if they have breached a duty of care, particularly in operating and supervising the autonomous system."
[138] See also Janal (n 137) 193.
[139] See <https://prod-help-content.care.irobotapi.com/files/s_Series/s9/ownersGuide/ownersGuide_enUS.pdf> accessed on 5 August 2022.

vacuum robot were to employ the device in the presence of small children without supervising them, the owner acted negligently. More generally, the instructions of a producer on supervising an AI device can act as guidance for the duty of care of the operator. A liability rule that shifts liability to producers for fully autonomous AI devices would promote information disclosure by producers.

However, information disclosure can have drawbacks: endless lists of warnings are likely to be ignored by consumers, in the same way that general term and conditions are not read.[140] To ensure that consumers are effectively informed about their devices, setting standards for the information supplied to consumers may still be desirable. A possible solution to this problem would be to regulate information duties or to make information more accessible by introducing "autonomy labels for AI", akin to the European energy labels. The autonomy labels could be aligned with certification processes and other safety regulations and would indicate to consumers what level of supervision is required when using an AI application. Given that the autonomy labels would provide information about the delegation of decisions to the AI system and the humans involved, these labels can inform courts when assigning liability to producers, operators and users. A drawback of this could be that producers may be discouraged from developing AI systems with increased autonomy if this increases their liability, even if this would be a safer alternative to "semi-autonomous" systems that still require human oversight in crucial situations. Nevertheless, autonomy labels could help resolve information problems of courts, by setting clear standards for the division of responsibility for harm involving AI systems.

Option 2 has the advantage of accommodating victims with a rebuttable presumption of the fault and/or the causality link. From the perspective of victims, this may be the preferred solution (even to strict liability, particularly if a presumption of causality would be included). It may help to establish causality where AI systems display autonomy and the operator has little control.[141] The law would still need to specify the duty of care on the operator, to supervise or monitor the AI system.[142] As AI systems reach higher levels of autonomy, such a duty may become more difficult to establish in concrete cases, as discussed above. At some point, AI systems are arguably no longer tools used by humans, but rather machines deployed by humans that act independently of direct human instruction.[143]

Option 3 may better reflection the risk profile of some AI systems as explained in section 4.1 above. Following the considerations laid out above, three arguments would support imposing strict liability on operators of a selected group of AI systems, for instance in sector regulation. First, the advantage of strict liability to establishing a "duty to supervise" under fault-based liability is that it ensures compensation for victims also in cases where, even if operators monitor an AI system, they may not be able to prevent harm if an AI system acts in completely unexpected ways.[144] In cases where the risks associated with an AI system are high, the "abnormal danger" argument mentioned above could justify decoupling liability from fault, rather than raising the standard of care or reversing the burden of proof. Second, strict liability can save the high transaction costs that injured parties would need to expend to litigate liability issues involving autonomous systems where the fault is difficult to establish.[145] Third, a strict liability regime may be more predictable. It would likely lead to fewer interpretation variations across national courts in Member States.

A strict liability rule for (some) AI systems would have analogies in the existing liability regimes of many Member States. Following the typology of the human-AI relationship, AI liability could follow the liability of parents for their children, owners for their animals or principals for their agents. Under a type of parental liability, operators would evade responsibility only if they can prove it was not possible to prevent a machine's action.[146] A type of employer liability could be justified since, by employing AI systems, operators impose risks on others. Particularly where a corporation operates an AI system, we may think of the corporation as operating the robot on its

[140] See the previous CERRE Report by Alexandre De Streel and Anne-Lise Sibony, Towards Smarter Consumer Protection Rules for the Digital Society (*CERRE Policy Report*, 5 October 2017) <https://cerre.eu/wp-content/uploads/2020/06/171005_CERRE_DigitalConsumerProtection_FinalReport.pdf> accessed on 07 August 2022
[141] See also Lior, A. (2020). AI entities as AI agents: Artificial intelligence liability and the AI Respondeat Superior Analogy. *Mitchell Hamline Law Review*, 46, 1044-1102.
[142] For a proposal (in the U.S. context) see Rachum-Twaig (n 81) 1168-70.
[143] David C. Vladeck, Machines without principals: liability rules and artificial intelligence' (2014) 89 Washington Law Review 117, 121.
[144] Janal (n 137) 199.
[145] Cf. Vladeck (n 143) 146-7 with respect to autonomous vehicles. See also Gerald Spindler, 'Zukunft der Digitalisierung – Datenwirtschaft in der Unternehmenspraxis' (2017) Betrieb (DB) 41, 50; Melinda F. Lohmann, 'Roboter als Wundertüten – eine zivilrechtliche Haftungsanalyse' (2017) 26 Aktuelle Juristische Praxis 152, 169.
[146] Pagallo (2012), p. 56.

behalf.[147] By holding operators strictly liable for their AI systems as a principal, they are properly incentivised to take precautions and identify the optimal level of employing the AI system.[148] The third possible analogy is the owner-animal relationship: In their erratic and unpredictable behaviour, AI systems resemble animals.[149] [150]

Liability for animals or children reflects the risk emanating their independent, not fully controllable behaviour. Strict liability is usually justified with the consideration that a particular danger arises from certain useful and therefore permitted facilities or activities. Those persons who are served by the facility or activity should also be assigned the disadvantages caused.[151] While the actions of advanced AI systems may similarly be uncontrollable, they usually promise a significant increase in safety as compared to their non-AI or human counterparts.[152] For this reason, AI systems may be of important value to society and imposing strict liability on their operators may impose an excessive burden.[153]

In the case of the principal-agent relationship, strict liability has a different justification. The key consideration here is that parties shift risk to another party and should bear the consequences of their actions. This argument also holds in AI context, even if defining “control” over the system may be challenging.[154]

6. Conclusion

EU non-contractual liability rules should not be thought of in isolation but as part of a broader set of rules as they jointly shape the incentives of all parties. In particular, liability rules need to be coherent with EU rules on safety and surveillance, the national non-contractual and contractual liability rules and rules on insurance. As always with regulatory design, rules reflect trade-offs which should be well-identified. Liability rules address possible trade-offs between the interests of the producers (and their innovation) and the interests of the users (and their protection). Liability rules should be based on risks of harm, which may differ depending on the application and the context in which AI systems are used. Liability rules should also ensure an efficient disclosure of information in situation of information asymmetries between stakeholders. In terms of timing, liability rules should balance proactive policymaking, anticipating technological changes, with reactive policymaking, adapting the rules only after having gained some experience from deploying the technologies.

Safety is a shared responsibility and non-contractual liability should provide incentives to all stakeholders (producers, operators and users) to take an efficient level of care in designing, testing and employing AI-based solutions, recognising that care by each party may be essential to avoid a failure (complementary efforts). Liability rules should be technologically neutral, providing the same level of protection of users of a product or services powered by AI as users of the same type of product or service which is not powered by AI.[155]

The paper identified three dimensions relevant to reviewing the EU liability framework for AI systems: (1) who should be liable; (2) the scope of new rules; and (3) the level of harmonisation. On the question of who should be liable, the paper considered the liability of producers and of operators. Given the upcoming review of the EU Product Liability Directive, we considered what challenges posed by AI could be addressed within the Product Liability Directive. We recommend clarifying that software is included in the definition of a product; in the age of digitalisation, differentiations between tangible and intangible objects of use are more difficult to justify. Concerning the safety expectations consumers are entitled to have of a product, a reform should consider the

[147] Rachum-Twaig (n 81), p. 1151.
[148] Lior (n 141) 4.
[149] Ibid 18.
[150] The liability standard for animals varies across the Member States, and within some Member States, e.g. depending on the purpose of the animal/AI system. See with respect to Germany e.g. Borges, G. (2018). Rechtliche Rahmenbedingungen für autonome Systeme. *Neue Juristische Wochenschrift, 71*(14), 977-982, 981-982.
[151] Pehm, J. (2018). Systeme der Unfallhaftung beim automatisierten Verkehr. Eine *rechtsvergleichende Analyse der Haftungsrisiken. Zeitschrift für Internationales Wirtschaftsrecht, 3*(6), 259-265, 260, referring to Koziol, H., Apathy, P. & Koch, B. A. (2014). *Österreichisches Haftpflichtrecht Bd. III: Gefährdungs-, Produkt- und Eingriffshaftung*. Jan Sramek Verlag, 7-8.
[152] Wagner, G. (2020). BGB § 823 Schadenersatzpflicht. In F. J. Säcker et al. (eds.), *Münchner Kommentar zum BGB* (pp. 789-791). C.H. Beck, 791; Rachum-Twaig (n 81 1158.
[153] Rachum-Twaig (n 81) 1145.
[154] Rachum-Twaig (n 81) 1151. However, a challenge would be that employer liability usually requires that the principal has control over the agent, which may be more difficult to establish for unpredictable behaviour by an AI system.
[155] Other reasons may lead policymakers to follow different ethics and, thus, depart from technological neutrality.

dynamic nature of AI systems. The notion of the defect should be clarified and may need to be defined in terms of overall failure rate rather than individual error. The standard of proof for proving a defect and the causality between the defect and the harm may need to be lowered to facilitate injured parties obtaining compensation for harm.

Next to manufacturers, the paper identified several reasons to keep operators of AI systems accountable. Firstly, liability for operators encourages them to take precautions in supervising AI systems that are not fully autonomous. Secondly, for highly autonomous AI systems, liability provides an incentive for operators to keep the system updated and ensure that it is used properly. Thirdly, operators benefit from employing AI, making it appropriate for its costs to be internalised.

In terms of liability standards that should apply to operators, introducing strict liability for AI would constitute a sharp departure from the standard liability regime currently in place in several Member States. Given that the risks of employing AI systems depend on the type of device and the context in which it is used, it is recommended to consider strict liability only in certain sector-specific or application-specific contexts. Lowering the standard of proof is an alternative means to facilitate injured parties' access to compensation, which is already happening at Member State level in other contexts.

The Commission's intention to introduce liability rules for (high-risk) AI raises the question to what extent the characteristics of AI justify more EU harmonisation of liability rules, beyond the context of producer liability. The level of harmonisation of liability rules and the scope of such harmonised rules present trade-offs between ensuring legal certainty for injured parties and operators with a uniform framework, while preserving the internal coherence of Member States' national liability rules allowing for learning effects to be delegated to the Member States. We recommend limiting harmonisation to the sectors where there is a set of EU safety rules that liability rules may usefully complement. This would have three main benefits. First, introducing a harmonised operator liability regime for AI systems, or high-risk AI systems, may lead to delimitation difficulties. It may be difficult to find a clear-cut, yet general criterion for distinguishing between "ordinary" and "autonomous" systems. Secondly, listing "high risk" AI applications may presuppose that AI applications create similar risks regardless of the context in which they are applied. Such a regime would therefore need to define not only the high-risk technologies but also the applications or contexts that it covers. Thirdly, existing sector-specific regulation already reflects the need to differentiate regulation according to the context in which technology is applied. Overall, it appears that many AI applications would fall in the non-high-risk category, and those that are high-risk will predominantly be covered by sector-specific regulation. The added value of a horizontal liability regime for high-risk AI, as compared to specifying liability rules in sector-specific regulation, may therefore be limited if sector-specific regulation is adequate.

References

Agrawal, A., Gans, J., & Goldfarb, A. (2018). *Prediction Machines: The Simple Economics of Artificial Intelligence*. Harvard Business Review Press.

Al-shamasneh, A. R. M., & Binti Obaidellah, U. H. (2017). Artificial intelligence techniques for cancer detection and classification: Review study. *European Scientific Journal, 13*(3), 342.

Alawadhi, M., Almazrouie, J., Kamil, M., & Khalil, K. A. (2020). Review and analysis of the importance of autonomous vehicles liability: A systematic literature review. *International Journal of System Assurance Engineering and Management, 11*(6), 1227-1249.

Alheit, K. (2001). The applicability of the EU product liability directive to software. *Comparative and International Law Journal of Southern Africa, 34*(2), 188-209.

Allain, J. S. (2013). From Jeopardy! to jaundice: the medical liability implications of Dr. Watson and other artificial intelligence systems. *Louisiana Law Review, 73(*4), 7.

Allcott, H., & Gentzkow, M. (2017). Social media and fake news in the 2016 election. *Journal of Economic Perspectives, 31*(2), 211-36.

Alter, A. (2017). *Irresistible: The Rise of Addictive Technology and the Business of Keeping Us Hooked*. Penguin Books.

Anderson, T., Torreggiani, W. C., Munk, P. L., & Mallinson, P. I. (2020). The impact of the introduction of artificial intelligence in radiology and its potential legal implications in the UK and Ireland. BJR| Open, 2, 20200030.

Amodei, D., Olah, C., Steinhardt, J., Christiano, P., Schulman, J., & Mané, D. (2016). Concrete problems in AI safety. *arXiv preprint arXiv:1606.06565*.

Anderson, T., & Torreggiani, W. C. (2018). The impact of the introduction of artificial intelligence to Ireland. *Irish medical journal*, *111*(8), 799.

Asaro, P. M. (2008). From mechanisms of adaptation to intelligence amplifiers: The Philosophy of W. Ross Ashby. In M. Wheeler, P. Husbands & O. Holland (eds.), *The Mechanical Mind in History* (pp. 149-184). MIT Press.

Asaro, P. M. (2016). The liability problem for autonomous artificial agents. *AAAI Spring Symposia*, 190-194.

Athey, S. C., Bryan, K. A., & Gans, J. S. (2020). The allocation of decision authority to human and artificial intelligence. *AEA Papers and Proceedings*, 110, 80-84.

Barfield, W. (2018). Liability for autonomous and artificially intelligent robots. *Paladyn, Journal of Behavioral Robotics*, *9*(1), 193-203.

Bartneck, C., Lütge, C., Wagner, A., & Welsh, S. (2021). *An Introduction to Ethics in Robotics and AI*. Springer Nature.

Belfield, H., Hernández-Orallo, J., Ó hÉigeartaigh, S., Maas, M. M., Hagerty, A., and Whittlestone, J. (2020). *Consultation on the White Paper on AI: a European approach*. Report by the Centre for the Study of Existential Risk. Available at <https://www.cser.ac.uk/news/response-european-commissions-consultation-ai/>.

Benhamou, Y., & Ferland, J. (2020). Artificial intelligence & damages: Assessing liability and calculating the damages. In P. D'Agostino, C. Piovesan & A. Gaon (eds.), *Leading Legal Disruption: Artificial Intelligence and a Toolkit for Lawyers and the Law* (pp. 165-197). Thomson Reuters Canada.

Bertolini, A. (2020). *Artificial Intelligence and Civil Liability: Legal Affairs*. Study for the JURI Committee by the Directorate-General for Internal Policies, Policy Department for Citizens' Rights and Constitutional Policies.

BEUC (2019). AI right for consumers. Available at <www.beuc.eu/publications/beuc-x-2019-063_ai_rights_for_consumers.pdf>.

BEUC (2020). The European Consumer Organisation, PRODUCT LIABILITY 2.0 How to make EU rules fit for consumers in the digital age. Available at <www.beuc.eu/publications/beuc-x-2020-024_product_liability_position_paper.pdf>.

Bibal, A., Lognoul, M., de Streel, A., & Frénay, B. (2020). Legal requirements on explainability in machine learning. *Artificial Intelligence and Law*, 1-21.

Boeglin, J. (2015). The costs of self-driving cars: reconciling freedom and privacy with tort liability in autonomous vehicle regulation. *Yale Journal of Law & Technology*, *17*(1), 171-203.

Borges, G. (2018). Rechtliche Rahmenbedingungen für autonome Systeme. *Neue Juristische Wochenschrift, 71*(14), 977-982.

Borges, G. (2019). Product Liability 2.0 – Mere update or new version? In S. Lohsse, R. Schulze, & D. Staudenmayer (eds.), *Liability for Artificial Intelligence and the Internet of Things* (pp. 145-164). Nomos.

Borghetti, J.-S. (2019). How can artificial intelligence be defective? In S. Lohsse, R. Schulze, & D. Staudenmayer (eds.), *Liability for Artificial Intelligence and the Internet of Things* (pp. 63-76). Nomos.

Buiten, M. C. (2019). Towards intelligent regulation of artificial intelligence. *European Journal of Risk Regulation, 10*(1), 41-59.
Buiten, M.C., de Streel, A. & Peitz, M. (2020). Rethinking liability rules for online hosting platforms. 28 *International Journal of Law and Information Technology*, 28, 139-166.
Buiten, M.C., de Streel, A. & Peitz (2021), EU liability rules for the age of artificial intelligence. CERRE Report 2021.
Calo, R. (2015). Robotics and the Lessons of the Cyberlaw. *California Law Review, 103*(3), 513-63.
Carpenter, B. M. & Collins, C. G. (2012). The Shirt Off My Back: Using the Relationship Between a Product and a Service to Your Advantage, International Association of Defense Counsel Newsletter. Available at <www.bakerdonelson.com>.
Cath, C., Wachter, S., Mittelstadt, B., Taddeo, M., & Floridi, L. (2018). Artificial intelligence and the 'good society': the US, EU, and UK approach. *Science and Engineering Ethics*, *24*(2), 505-528.
Cauffman, C. (2018). Robo-liability: The European Union in search of the best way to deal with liability for damage caused by artificial intelligence. *Maastricht Journal of European and Comparative Law, 25*(5), 527-532.
Chinen, M. A. (2016). The co-evolution of autonomous machines and legal responsibility. *Virginia Journal of Law & Technology, 20*(2), 338-93.
Chung, J., & Zink A. (2018). Hey Watson—Can I sue you for malpractice? Examining the liability of artificial intelligence in medicine. *Asia Pacific Journal Health Law & Ethics, 11*(2), 51-80.
Coase, R. (1960). The problem of social cost, *Journal of Law and Economics*, 3(1), 1–44.
Cofone, I. N. (2018). Servers and waiters: What matters in the law of AI. *Stanford Technology Law Review*, *21*(2), 167-197.
Cooter, R. B., & Porat, A. (2007). Total liability for excessive harm. *Journal of Legal Studies, 36*(1), 63-80.
Cooter, R. B., & Ulen, T. (2012). *Law and Economics* (6th edition). Pearson.
Datta, A., Sen, S., & Zick, Y. (2016). Algorithmic transparency via quantitative input influence: Theory and experiments with learning systems. *2016 IEE Symposium on Security and Privacy*, 598-617.
de Bruin, R. (2016). Autonomous intelligent cars on the European intersection of liability and privacy. *European Journal of Risk Regulation*, *7*(3), 485-501.
De Meeus, C. (2019). The Product Liability Directive at the age of the digital industrial revolution: Fit for innovation? *Journal of European Consumer and Market Law, 8*(4), 149-154.
De Streel, A. & Sibony, A.L. (2017). Towards smarter consumer protection rules for the digital society. CERRE Policy Report.
De Streel, A., Bibal, A., Frenay, B., & and Lognoul, M. (2020). Explaining the black box: When law controls AI. CERRE Policy Report.
Diakopoulos, N. (2016). Accountability in algorithmic decision making. *Communications of the ACM, 59*(2), 56–62.
Domingos, P. (2015). *The master algorithm: How the quest for the ultimate learning machine will remake our world*. Basic Books.
Doshi-Velez, F. et al. (2017). Accountability of AI under the law: The role of explanation. *arXiv preprint arXiv:1711.01134*.
Ebers, M. (2017). Autonomes Fahren: Produkt- und Produzentenhaftung. In B.H. Oppermann & H. Stender-Vorwachs (eds.), *Autonomes Fahren* (pp. 93-126). C.H. Beck.
Ebers, M. (2020). Regulating AI and robotics. In M. Ebers & S. Navas (eds.), *Algorithms and Law* (pp. 37-99). Cambridge University Press.
Edwards, L., & Veale, M. (2017). Slave to the algorithm: Why a right to an explanation is probably not the remedy you are looking for. *Duke Law & Technology Review*, *16*, 18-84.
EPRS (2021), Cost of Non-Europe report on artificial intelligence in road transport. Available at < https://www.europarl.europa.eu/thinktank/en/document.html?reference=EPRS_STU(2021)654212>.
Etzkorn, P. (2020). Bedeutung der» Entwicklungslücke «bei selbstlernenden Systemen–Rechtliche Fragen zur fortdauernden Softwareentwicklung durch maschinelles Lernen im Praxiseinsatz. *Multimedia und Recht, 23*(6), 360-365.
European Law Institute, Twigg-Flesner, C. (2021). Guiding principles for updating the Product Liability Directive for the digital age, *ELI Innovation Paper Series*. Available at <https://europeanlawinstitute.eu/fileadmin/user_upload/p_eli/Publications/ELI_Guiding_Principles_for_Updating_the_PLD_for_the_Digital_Age.pdf>.

Fairgrieve, D. (2019). Product liability in the United Kingdom. *Journal of European Consumer and Market Law, 8*(4), 170-72.

Fairgrieve, D., Howells, G., Møgelvang-Hansen, P., Straetmans, G., Verhoeven, D., Machnikowski, P., A. Janssen & Schulze, R. (2016). Product liability directive. In P. Machnikowski (ed.), *European Product Liability: An analysis of the state of the art in the era of new technologies*. Intersentia, 17-108.

Fink, C. (2018). Dangerous speech, anti-muslim violence, and Facebook in Myanmar. *Journal of International Affairs, 71*(1.5), 43-52.

Förster, C. (2017). BGB § 823 Schadensersatzpflicht. In H.G. Bamberger & H. Roth (eds.), *BeckOK BGB*. C.H. Beck.

Fry, H. (2018). *Hello World: How to be Human in the Age of the Machine*. Random House.

Galasso, A., & Luo, H. (2018a). Punishing robots: Issues in the economics of tort liability and innovation in artificial intelligence. In A. Agrawal, J. Gans & A. Goldfarb (eds.), *The Economics of Artificial Intelligence: An Agenda* (pp. 493-504). University of Chicago Press.

Galasso, A., & Luo, H. (2018b). *When does product liability risk chill innovation? evidence from medical implants* (No. w25068). National Bureau of Economic Research.

Gans, J., & Goldfarb, A. (2018). *Prediction Machines: The Simple Economics of Artificial Intelligence*. Harvard Business Review Press.

Gasser, T. (2012). Legal issues of driver assistance systems and autonomous driving. In A. Eskandarian (ed.), *Handbook of Intelligent Vehicles* (pp. 1519-1535). Springer.

Gasser, U., & Almeida, V. A. F. (2017). A layered model for AI governance. *IEEE Internet Computing, 21*(6), 58-62.

Gerstner, M. E. (1993). Liability issues with artificial intelligence software. *Santa Clara Law Review, 33*(1), 239-70.

Giuffrida, I. (2019). Liability for ai decision-making: Some legal and ethical considerations. *Fordham Law Review, 88*(2), 439-456.

Goodman, B., & Flaxman, S. (2017). European Union regulations on algorithmic decision-making and a "right to explanation". *AI Magazine, 38*(3), 50-57.

Gulshan, V. et al. (2016). Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *Jama*, *316*(22), 2402-2410.

Hey, T. (2019). *Die außervertragliche Haftung des Herstellers autonomer Fahrzeuge bei Unfällen im Straßenverkehr*. Springer Gabler.

Howells, G., Twigg-Flesner, C., and Willett, C. (2017). Product liability and digital products. In T-E. Synodinou, P. Jougleux, C. Markou, & T. Prastitou (eds.), *EU Internet Law: Regulation and Enforcement* (pp. 183-195). Springer Nature.

Hubbard, F. P. (2014). Sophisticated robots: Balancing liability, regulation, and innovation. *Florida Law Review*, *66*, 1803-1872.

Janal, R. (2020). Extra-contractual liability for wrongs committed by autonomous systems. In M. Ebers & S. Navas (eds.), *Algorithms and Law* (pp. 174-206). Cambridge University Press.

Karner, E., Koch, B.A. and Geistfeld, M.A., Comparative law study on civil liability for artificial intelligence, Study Prepared for the Directorate-General for Justice and Consumers, 2021.

Karnow, C. E. (2016). The application of traditional tort theory to embodied machine intelligence. In R. Calo, A.M. Froomkin & I. Kerr (eds.), *Robot Law* (pp. 51-77). Edward Elgar Publishing.

Kelley, R., Schaerer, E., Gomez, M., & Nicolescu, M. (2010). Liability in robotics: An international perspective on robots as animals. *Advanced Robotics, 24*(13), 1861-1871.

Kim, H.-E. et al. (2020). Changes in cancer detection and false-positive recall in mammography using artificial intelligence: a retrospective, multireader study. *The Lancet. Digital Health*, *2*(3), e138–e148.

Kitchin, R. (2017). Thinking critically about and researching algorithms. *Information, Communication & Society, 20*(1), 14-29.

Klein, E. (2020). *Why We're Polarised*. Avid Reader Press/Simon & Schuster.

Koch, B. (2019). Product Liability 2.0 – Mere update or new version? In S. Lohsse, R. Schulze, & D. Staudenmayer (eds.), *Liability for Artificial Intelligence and the Internet of Things* (pp. 99-116). Nomos.

Kowert, W. (2017). The foreseeability of human-artificial intelligence interactions. *Texas Law Review*, *96*(1), 181.

Koziol, H., Apathy, P. & Koch, B. A. (2014). *Österreichisches Haftpflichtrecht Bd. III: Gefährdungs-, Produkt- und Eingriffshaftung*. Jan Sramek Verlag.

Kremer, M. (1993). The O-ring theory of economic development. *Quarterly Journal of Economics, 108*(3), 551-575.

Lagasse, J. (2015). Faut-il un droit des robots? Les Notes du CREOGN, Centre de Recherche de l'Ecole des Officiers de la Gendarmerie Nationale, 2015, N° 12. ffhal-03096766, Available at <https://hal.archives-ouvertes.fr/hal-03096766/document>.

Lemley, M. A., & Casey, B. (2019). Remedies for robots. *The University of Chicago Law Review, 86*(5), 1311-1396.

Levy, D. (2020). Intelligent no-fault insurance for robots. *Journal of Future Robot Life, 1*(1) 35-57.

Lior, A. (2020). AI entities as AI agents: Artificial intelligence liability and the AI Respondeat Superior Analogy. *Mitchell Hamline Law Review*, 46, 1044-1102.

Lohmann, M. F. (2017). Roboter als Wundertüten – eine zivilrechtliche Haftungsanalyse. *Aktuelle Juristische Praxis*, *26*(2), 152-162.

Lohmann, M. F. (2016). Liability issues concerning self-driving vehicles. *European Journal of Risk Regulation*, *7*, 335-340.

Lohsse, S., Schulze, R., & Staudenmayer, D. (2019). Liability for artificial intelligence. In S. Lohsse, R. Schulze, & D. Staudenmayer (eds.), *Liability for Artificial Intelligence and the Internet of Things* (pp. 99-116). Nomos.

Lutter, C. (2017). Fragen der Produktehaftung im Hinblick auf den Betrieb unbemanter Schiffe. Recht der Transportwirtschaft, *5*(8), 281-286.

Macgregor, D. (2002). Hand injuries in young children from contact with vacuum cleaners. *Emergency Medicine Journal, 19*(1), 80-81.

Machnikowski, P. (2016). *European Product Liability. An Analysis of the State of the Art in the Era of New Technologies*. Cambridge, Intersentia.

Marchant, G.E. & Lindor, R.A. (2012). The coming collision between autonomous vehicles and the liability system. *Santa Clara Law Rev*iew, 52, 1321.

Marcus, J. S. (2018). Liability: When things go wrong in an increasingly interconnected and autonomous world: A European view. *IEEE Internet of Things Magazine*, *1*(2), 4-5.

McCarthy, J. (2007). *What is artificial intelligence?* Available at <www.formal.stanford.edu/jmc/whatisai.pdf>.

McCarthy, J. J., Minsky M. L., & Rochester, N. (1959). *Artificial intelligence.* Research Laboratory of Electronics at the Massachusetts Institute of Technology.

Mendoza-Caminade, A. (2016). Le droit confronté à l'intelligence artificielle des robots: vers l'émergence de nouveaux concepts juridiques? *Recueil Dalloz, 192*(8), 445-448.

Minsky, M. L. (1959). Some methods of artificial intelligence and heuristic programming. *Proc. Symposium on the Mechanization of Thought Processes, 1*, 5-27.

Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society*, 3(2), 1-20.

Nash, I. (2021). Cybersecurity in a post-data environment: Considerations on the regulation of code and the role of producer and consumer liability in smart devices. *Computer Law & Security Review*, *40*, 105529.

Navas, S., (2020). Robot machines and civil liability. In M. Ebers & S. Navas (eds.), *Algorithms and Law* (pp. 157-173). Cambridge University Press.

Nemeth, K., & Carvalho, J. M. (2019). Time for a change? Product liability in the digital era. *Journal of European Consumer and Market Law, 8*(4), 160-161.

Pagallo, U. (2011). Three roads to complexity, AI and the law of robots: On Crimes, contracts, and torts. In M. Palmirani et al. (eds.) *AI Approaches to the Complexity of Legal Systems* (pp. 48-60). Springer.

Palmer, C. (2020). The Boeing 737 Max Saga: Automating Failure. *Engineering (Beijing, China)*, *6*(1), 2–3.

Palmerini, E. et al. (2014). RoboLaw. *Guidelines on Regulating Robotics. Regulating Emerging Robotic Technologies in Europe: Robots facing Law and Ethics*. Available at <www.robolaw.eu>.

Pasquale, F. (2015). *The Black Box Society: The Secret Algorithms That Control Money and Information*. Harvard University Press.

Pehm, J. (2018). Systeme der Unfallhaftung beim automatisierten Verkehr. Eine *rechtsvergleichende Analyse der Haftungsrisiken. Zeitschrift für Internationales Wirtschaftsrecht, 3*(6), 259-265.

Perrow, C. (2011). *Normal accidents: Living with high risk technologies* (Rev. ed.). Princeton University Press.

Price, W. N. II. (2017). Artificial intelligence in health care: Applications and legal implications*. The SciTech Lawyer, 14*(1), 10-13.

Rachum-Twaig, O. (2020). Whose robot is it anyway?: Liability for artificial-intelligence-based robots. *University of Illinois Law Review*, (4), 1141-1175.

RAND Corporation, Kalra & Groves (2017). The enemy of good: Estimating the cost of waiting for nearly perfect automated vehicles. Available at <www.rand.org/pubs/research_reports/RR2150.html>.

Reece, L. H. (1987). Defective expert systems raise personal injury liability issues. *The National Law Journal*, *9*, 24-29.

Reed, C. (2018). How should we regulate artificial intelligence? *Philosophical Transactions of the Royal Society A, 376*(2128).

Reed, C., Kennedy, E., & Nogueira Silva, S. (2016). Responsibility, autonomy and accountability: Legal liability for machine learning. Paper presented at the *Microsoft Cloud Computing Research Centre 3rd Annual Symposium, September 2016*, Queen Mary University, London.

Scherer, M. U. (2016). Regulating artificial intelligence systems: Risks, challenges, competencies, and strategies. *Harvard Journal of Law & Technology, 29*(2), 353-400.

Schmidt, P., Biessmann, F., & Teubner, T. (2020). Transparency and trust in artificial intelligence systems. *Journal of Decision Systems, 29*(4), 260-78.

Schmon, C. (2018). Product liability of emerging digital technologies. *Zeitschrift für Internationales Wirtschaftsrecht, 3*(6), 254-259.

Schönberger, D. (2019). Artificial intelligence in healthcare: A critical analysis of the legal and ethical implications. *International Journal of Law and Information Technology*, *27*(2), 171-203.

Schrader, P. T. (2016). Haftungsfragen für Schäden beim Einsatz automatisierter Fahrzeuge im Straßenverkehr. *Deutsches Autorecht, 86*(5), 242-246.

Seehafer, A., & Kohler, J. (2020). Künstliche Intelligenz: Updates für das Produkthaftungsrecht? *Europäische Zeitschrift für Wirtschaftsrecht*, 213-218.

Selbst, A. D. (2020). Negligence and AI's human users*. Boston University Law Review, 100*(4), 1315-1376.

Smart, W. D., Grimm, C. M., & Hartzog, W. (2017). An education theory of fault for autonomous systems. *Proceedings of We Robot*.

Smith, H., & Fotheringham, K. (2020). Artificial intelligence in clinical decision-making: Rethinking liability. *Medical Law International, 20*(2), 131-154.

Spindler, G. (2011). Haftung im IT-Bereich. In L. Egon (ed.), *Karlsruher Forum 2010: Haftung und Versicherung im IT-Bereich* (pp. 3-112). Verlag Versicherungswirtschaft.

Spindler, G. (2015). Roboter, Automation, künstliche Intelligenz, selbst-steuernde Kfz–Braucht das Recht neue Haftungskategorien? *Computer und Recht*, *31*(12), 766-776.

Spindler, G. (2018). Zukunft der Digitalisierung–Datenwirtschaft in der Unternehmenspraxis. *Betrieb (DB)*, *71*, 41-50.

Steege, H. (2021). Auswirkungen von künstlicher Intelligenz auf die Produzentenhaftung in Verkehr und Mobilität. *Neue Zeitschrift für Verkehrsrecht*, 6-13.

Stone, P. et al. (2016). *Artificial Intelligence and Life in 2030: One Hundred Year Study on Artificial Intelligence*, Report of the 2015 Study Panel. Available at <https://ai100.stanford.edu/2016-report>.

Stöber, M., Pieronczyk, M.-C., & Möller, A. (2020). Die Schadensersatzhaftung für automatisierte und autonome Fahrzeuge. *Deutsches Autorecht, 90*(11), 609-613.

Sullivan, H. R., & Schweikart, S. J. (2019). Are current tort liability doctrines adequate for addressing injury caused by AI? *AMA journal of ethics*, *21*(2), 160-166.

Surden, H., & Williams, M. A. (2016). Technological opacity, predictability, and self-driving cars. *Cardozo Law Review*, *38*, 121-182.

Svoboda, E. (2020). Deep learning delivers early detection: Artificial-intelligence systems look set to make early detection of lung cancer more accurate and widely Aavailable. *Nature, 587*, 20-22.

Timan, T. et al. (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 SMART 2016/0071: Final Study Report regarding CAD/CCAM and Industrial Robots*, Study for the European Commission. Available at <https://op.europa.eu/en/publication-detail/-/publication/aad6a287-5523-11e9-a8ed-01aa75ed71a1>.

Tutt, T. (2017). An FDA for algorithms. *Administrative Law Review, 69*(1), 83-123.

Ullrich, F., Classen, J., Eger, J., & Hollick, M. (2019, August). Vacuums in the cloud: analyzing security in a hardened IoT ecosystem. In *Proceedings of the 13th USENIX Conference on Offensive Technologies*, 7.

Vladeck, D. C. (2014). Machines without principals: liability rules and artificial intelligence. *Washington Law Review*, *89*(1), 117-150.

Von Ungern-Sternberg, A. (2018). Artificial agents and general principles of law. In A. von Arnauld, K. von der Decken, & N. Matz-Lück (eds.), *German Yearbook of International Law* (pp. 240-267). Duncker & Humblot.

Von Westphalen, F. (2020). Produkthaftungsrechtliche Erwägungen beim Versagen Künstlicher Intelligenz (KI) unter Beachtung der Mitteilung der Kommission COM(2020) 64 final. *Verbraucher und Recht, 35*(7), 248-254.

Wachter S., Mittelstadt, B., & Floridi, L. (2017). Transparent, explainable, and accountable AI for robotics. *Science Robotics, 2*(6).

Wagner, G. (2017). Produkthaftung für autonome Systeme. *Archiv für die civilistische Praxis*, *217*(6), 707-766.

Wagner, G. (2019a). Robot liability. In S. Lohsse, R. Schulze, & D. Staudenmayer (eds.), *Liability for Artificial Intelligence and the Internet of Things* (pp. 25-62). Nomos.

Wagner, G. (2019b). Robot, inc.: Personhood for autonomous systems? *Fordham Law Review, 88(*2), 591-612.

Wagner, G. (2020). BGB § 823 Schadenersatzpflicht. In F. J. Säcker et al. (eds.), *Münchner Kommentar zum BGB* (pp. 789-791). C.H. Beck.

Waltl, B., & Vogl, R. (2018). Increasing transparency in algorithmic-decision making with explainable AI. *Datenschutz und Datensicherheit-DuD*, *42*(10), 613-617.

Watanabe, A. T. et al. (2019). Improved cancer detection using artificial intelligence: A retrospective evaluation of missed cancers on mammography. *Journal of Digital Imaging*, *32*(4), 625–637.

Weber, R. H. (2017). Liability in the Internet of Things. *Journal of European Consumer and Market Law, 3*(5), 207-212.

Wendehorst, C. (2016). *Sale of goods and supply of digital content – two worlds apart?: Why the law on sale of goods needs to respond better to the challenges of the digital age*. Study for the Juri Committee by the Directorate General for Internal Policies, Policy Department C. Available at <www.europarl.europa.eu>.

Wendel, W. (2019). Technological solutions to human error and how they can kill you:Understanding the boeing 737 max products liability litigation. *Journal of Air Lawand Commerce, 84*(3), 379-446.

Williams, J. (2018). *Stand Out of Our Light: Freedom and Resistance in the Attention Economy*. Cambridge University Press.

Wuyts, D. (2014). The Product Liability Directive–more than two decades of defective products in Europe. *Journal of European Tort Law, 5*(1), 1-34.

Yoshikawa, J. (2019). Sharing the costs of artificial intelligence: Universal no-fault social insurance for personal injuries. *Vanderbilt Journal of Entertainment & Technology Law*, *21*(4), 1155-1188.

Zech, H. (2021). Liability for AI: public policy considerations. *ERA Forum, 22*(1), 147-158.

Zweigert. K., & Kötz H. (1996). *Einführung in die Rechtsvergleichung auf dem Gebiete des Privatrechts* (3rd ed.). Mohr Siebeck.