

Chapter 19

Digging into the Accountability Gap: Operator's Civil Liability in Healthcare AI-systems

Kostina Prifti, Evert Stamhuis, Klaus Heine

Abstract

The increasing autonomy of artificial intelligence systems (AI-systems) has put the debate about a possible 'accountability gap' in liability law centerstage. The debate is about a possible failure of incumbent liability regimes to pinpoint the accountable agent, if in the wrongdoing an AI-system is involved. A recent attempt to address this 'accountability gap' is a proposal of the European Parliament, which advances laws on civil liability for the entities that control AI-systems. These newly created entities, which have no blueprint yet in liability law, are called 'Operators'. By branching out on the healthcare applications of AI-systems, this chapter analyzes the concept of operator's civil liability. It starts with a description of the liability concepts presented in the proposed legislation, and how they fit doctrinally with the laws and regulations of the current medical liability regime. Complementing the doctrinal analysis, this chapter employs a law and economics analysis, which showcases that the accountability gap is a serious challenge also from a consequentialist point of view. Lastly, this chapter proposes a few legal alternatives that depart from the incumbent concept of strict liability.

Keywords: accountability gap; liability; AI-systems; European Parliament; operator; law and economics.

19.1 Introduction

In October 2020, the European Parliament (EP) accepted a resolution containing a proposal to the European Commission (EC), for the civil liability of operators of AI-systems ('the proposal'),¹ with operators being defined as the actors who are in control of, and benefit from AI-systems.² The proposal introduces this concept in the regulatory system of liability in the EU, along with specific rules of strict or negligence liability for this new actor. These changes are proposed because of a perceived

Kostina Prifti, PhD Researcher, Junior Fellow Jean Monnet Centre of Excellence on Digital Governance, prifti@law.eur.nl; Prof. Dr. Evert Stamhuis, Chair Law & Innovation Erasmus School of Law, Senior fellow Jean Monnet Centre of Excellence on Digital Governance. Other affiliations: Court of Appeal 's Hertogenbosch, University of Aruba, stamhuis@law.eur.nl; Prof. Dr. Klaus Heine, Jean Monnet Chair of Economic Analysis of European Law and Co-Director Jean Monnet Centre of Excellence on Digital Governance, heine@law.eur.nl; Erasmus School of Law, Erasmus University Rotterdam, Burg. Oudlaan 50, Rotterdam.

¹ European Parliament resolution of 20 October 2020 with recommendations for the Commission on a civil liability regime for artificial intelligence (2020/2014(INL)). This proposal can be seen as a continuation and manifestation of other related documents issued by expert bodies of EU institutions, such as the "Liability for artificial intelligence and other emerging digital technologies" by the Expert Group on Liability and New Technologies – New Technologies Formation and the "Ethics Guidelines for Trustworthy Artificial Intelligence" by the High-Level Expert Group on AI. This chapter was basically concluded before publication of the draft AI Act by the European Commission, COM (2021) 206 final, but we could conclude that this latter proposal does not relate to civil liability.

² Article 3 of the proposal.

shift in context factors that AI-systems raise in relation to the incumbent liability regime. Moreover, a timely adaptation of the EU liability regime shall support transactions in the single market and promote the development and deployment of AI in the EU.

There are three ways how AI-systems pose a shift in context factors in relation to the liability regime: The autonomy risk, the association risk, and the network risk.³ Firstly, the increasing autonomy of AI-systems implies a learning process that enables those systems to change their own program and become unpredictable. Secondly, the interaction between humans and AI-systems becomes complex and perplexes the possibility to understand the exact separation of roles and tasks between humans and AI-systems. And thirdly, the network risk considers the complex interactions between AI-systems among each other, creating “an ensemble of concurrently active polymorphic intelligent agents”.⁴

These disruptions relate to the liability regime through key concepts like causation, fault, and deterrence. The inability of operating actors to predict the actions and decisions of AI-systems risks breaking the chain of causation, as foreseeability is a central tenet of causation.⁵ Lack of foreseeability also makes deterrence ineffective, because the actors involved cannot be incentivized to avoid a damage they reasonably cannot foresee. Moreover, the inability to understand what the wrongful actions of each actor are in the complex network makes it difficult to assign responsibility. Such disabilities are also referred to as the ‘responsibility gap’ or ‘liability gap’.⁶ Perhaps a terminology clarification is in order. Firstly, denoting a liability gap expresses a function of location, rather than nature. The term responsibility gap may be more appropriate, however, responsibility has a higher threshold than accountability, due to the cognitive requirements that it implies (e.g., intentionality).⁷ Accountability ensures a proper functioning of the liability regime, even when there is a gap in responsibility, so the nature of the problem is better perceived as an accountability gap, i.e., the inability to pinpoint an accountable agent who needs to compensate the harm, in situations where an AI system is involved.

A full assessment of the proposal comprises both the doctrinal quality as well as its economic efficiency. While there is considerable convergence, as well as some divergence, between the proposal and the various documents mentioned in footnote 1, the proposal adopts and develops the concepts that are present in a more or less systematic framework. As such, it presents itself as a meaningful legislative step on the idea and mechanisms for strict liability of operators, which in turn calls for – *inter alia* – a doctrinal and economic analysis. Consequently, section 19.2 provides a description of the proposal, its main concepts, and relevant mechanisms. In section 19.3 the chapter offers an analysis of the proposal in the dimension of legal criteria as conceptual clarity, systematic consistency and matching the subject matter. Then, in section 19.4, the discussion will be expanded into the economics dimension and it will be asked to what extent the proposal can meet the ambition to allocate liability in a socially optimal way. Thereby the focus will be on a specific sector: The medical application of AI-systems. Medical use of AI-systems will very likely be listed as high risk and fall under the proposal. Moreover, the medical sector provides an excellent context for discussing the intricacies of the proposal as such, because AI-systems become specifically developed and deployed there. Finally, the newly proposed law will - upon acceptance - interact with the incumbent liability laws, that have shown a tendency to diversity, of which medical liability law is one quite specific branch.

³ Teubner 2018, p 107.

⁴ Karnow 1996, p 191.

⁵ Ibid.

⁶ Matthias 2004, p 175.

⁷ Floridi and Sanders 2004, p 364.

19.2 The proposal

The proposal on civil liability for operators of AI-systems takes as its starting point a shift in context factors that are identified in the literature as the ‘accountability gap’. The proposal refers specifically to the opacity of AI-systems (black-box problem), the multitude of actors, and the increasing autonomy of AI-systems.⁸ It then seeks to overcome the liability issues created by AI-systems by introducing four categories of actors: front-end operators, back-end operators, users, and producers. The first two categories are a novelty in EU liability law. The proposal defines the front-end operator broadly as the entity that has control over the AI-system and that benefits from operating the AI-system.⁹ The back-end operator is defined along the same lines; it is the entity that offers continuous back-end services, such as defining the features of the technology and providing data.¹⁰ Both categories of operators are considered to exercise a certain level of control over the AI-systems.

When the same actor is both the front-end and back-end operator, also the legal regime is the same. It is also possible that the operator is the producer of the AI-system, in which case the Directive on Product Liability (PLD)¹¹ becomes relevant. In case that the front-end operator is also the producer of the AI-system, the EP proposes that the liability regime of operators should prevail. If the back-end operator is also the producer, the proposal gives way to the PLD as the prevailing applicable law. When there is only one operator, it is proposed that civil liability of operators should prevail as applicable law. It is important to note that the proposal counts on the inclusion of software as a product under the scope of PLD, along with inclusion of software ‘makers’ as producers.¹²

With this regime the EP expects to guide the liability allocation clearly and predictably to the relevant actors, so that courts will be able to resolve conflicts over it, when they arise. In practice, the exercise to allocate roles can be more complicated than it appears. Two examples taken from health care may showcase this lack of clarity. Consider the Smart Tissue Autonomous Robot (STAR), used for intestinal laparoscopic surgery.¹³ It is able to develop the suturing plan in an automated way, with the surgeon supervising and having the ability to make positional adjustments that in turn impact the suturing plan. The hospital may be considered an operator, as it is in control of and benefits from the use of STAR, but the doctor has also an important supervisory role. It is unclear if the doctor would be an operator, a user, or simply irrelevant, especially considering that the doctor might not be employed by the hospital. In another example, consider Sweet Home, a home automation system typically used to monitor a patient’s wellbeing.¹⁴ The daily operation of Sweet Home may be handled by the producer, which the hospital or another care institution contracts for its patients. It is not clear if the hospital would be considered a user or an operator in this case. And what is the patient? The distinction between users and end-users becomes more relevant through these examples, however this is missing in the proposal.¹⁵ By qualifying the actors involved, the various positions for liability of any harm are supposed to be defined. While the doctor is distinguished from the operator based on

⁸ Recital 3 of the proposal.

⁹ Article 3 of the proposal.

¹⁰ Ibid.

¹¹ Council Directive 85/374/EEC of 25 July 1985 on the approximation of the laws, regulations and administrative provisions of the Member States concerning liability for defective products OJ L 210, 7.8.1985, p. 29–33.

¹² See Point 8 of the Introduction to the proposal.

¹³ Shademan et al. 2016, p 341.

¹⁴ Shishhegar et al. 2018, p 1.

¹⁵ It is however found in the “Ethics Guidelines for Trustworthy Artificial Intelligence” by High-Level Expert Group on AI.

the level of control over the AI-system,¹⁶ it is yet not clear what one may use as a threshold to separate the user from the operator in each context. As soon as the doctor has a level of control, she may fall into the operators' category, with a shift in liability. As said, this distinction is important in order to understand which liability regime is applicable. However, the proposal clarifies in article 2(3) that the victim is not refrained from seeking compensation based on other liability regimes, such as medical malpractice.

Beside the distinction of actors and introducing the new category of operator, the proposal makes a distinction between high-risk and low-risk sectors and uses. This adds another level of sophistication to the liability regime. The outcome of this is a matrix with eight cells of liability regimes (Table 19.1).

Table 19.1 Liability regimes according to the proposal

Actor\risk level	High-risk	Low-risk
User	Negligence liability	Negligence liability
Front-end operator	Strict liability	Negligence liability
Back-end operator	Strict liability	Negligence liability
Producer	Strict liability	Strict liability

It is important to note that under negligence liability the operator must be found in default of a certain action or omission before being held liable, whereas under strict liability the operator is liable merely because the harm or damage has occurred. It is also worth noting that Article 8 of the proposal shifts the burden of proof for negligence liability to the operator, for the latter to prove that one of the circumstances described there applies. Shifting the burden of proof from the victim to the operator facilitates the allocation of liability and eventual compensation for the victim, since the latter may not always be able to prove the negligence of the operator, due to lack of insight, even when there is negligence in the case.

The proposal does not list the high-risk areas and uses of AI-systems; the list is left to be drawn and updated by the EC.¹⁷ However, in a related proposal of the EP on an ethical framework for artificial intelligence,¹⁸ healthcare as an area, and medicine as a use, are listed as high risk. Similar categorizations follow in other legal acts, such as in the GDPR.¹⁹ Therefore, this chapter treats healthcare AI-systems as belonging to the high-risk category, and as a result assumes that operators of healthcare AI-systems would face strict liability whenever a harm or damage arises from these AI-systems.

¹⁶ Recital 11 of the proposal.

¹⁷ Recital 14 of the proposal.

¹⁸ European Parliament resolution of 20 October 2020 with recommendations for the Commission on a framework of ethical aspects of artificial intelligence, robotics and related technologies (2020/2012(INL) and EC White Paper 'On Artificial Intelligence - A European approach to excellence and trust' 19.2.2020 COM(2020) 65 final.

¹⁹ Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC, OJ L 119, 4.5.2016, p. 1–88, Article 9.

19.3 The doctrinal instability

19.3.1 The product/service dichotomy

In this section the doctrinal inconsistencies and deficiencies are highlighted, arising from the proposal. They result from the combination of strict liability with the new category of operator. Those inconsistencies are also the drivers against closing the accountability gap and attaining the behavioral economic objectives of the liability regime, as elaborated in section 19.4.

The actual medical liability framework in the EU is based on two pillars: the PLD and national laws regulating medical malpractice. The former is harmonized, while the latter are not.²⁰

The PLD induces a strict liability regime for damages caused by defective products, which includes medical devices. While the victim is required to prove the defect, the damage, and the causal link between defect and damage, the victim is not required to prove fault or negligence of the producer. Simultaneously, the producer is not shielded from liability by proving lack of fault or negligence. Hence, the producer must compensate whenever a defective product causes a damage, save for some exceptions.²¹

Medical malpractice is not harmonized by EU law. In fact, the EU attempted to harmonize liability of services through a Directive, albeit unsuccessfully.²² As a result, medical malpractice is regulated under the national laws of Member States.²³ Although variations exist, medical malpractice, similarly to liability of other services, is based on negligence liability.²⁴ The potentially wrongful behavior of the medical professional is tested under an objective standard of care.²⁵ Oftentimes, the negligence-based liability is fine-tuned in national law by shifting the burden of proof to the defender, and by variations of vicarious liability or enterprise liability of the hospital.²⁶ The latter is also referred to as “negligence liability with an extra debtor”,²⁷ the result of which, in practice, is that the employer or contractor is strictly liable to compensate the damage caused by its employee or contracted party, but only when negligence is proven on the side of the acting professional.

The co-existence of these two regimes in the medical liability framework is a manifestation of the product/service dichotomy, maintained in case law even in hybrid cases, such as when the service provider supplies both a product and a service. The European Court of Justice (ECJ) has sharply distinguished that the service provider would be strictly liable if the product was defective, while if the service was defective there would be liability based on negligence.²⁸ Such distinctions are much harder to identify in cases where AI-systems are involved, as, for example, microsurgical robots that work under the supervision of a surgeon.²⁹ The hospital would firstly purchase the robot as a product, but may also purchase the service of maintaining and feeding the robot with datasets and learning

²⁰ Koch 2011, p 1-39.

²¹ Article 7 of the PLD.

²² Proposal for a Council Directive on the liability of suppliers of services, COM (90) 482 final — SYN 308 of 20 December 1990.

²³ For a systematic analysis of member state laws on medical liability see Koch 2011.

²⁴ Koch 2011, p 10-21.

²⁵ Koch 2011, p 8-10.

²⁶ Koch 2011, p 10-21.

²⁷ Van Dam 2013, p 300.

²⁸ Ibid.

²⁹ Ficuciello et al. 2019, p 30.

techniques from the same AI company. As one can easily see, the complexity of robots in the medical context makes it difficult to maintain the distinction of the ECJ, leading to very difficult exercises to allocate liability. What failure led to the harm and to what degree? Was it the failing robot or the absence of proper oversight of the doctors and nurses? The tight man-machine cooperation in AI-systems makes it challenging to dissect all components with regard to causal inference.

It boils down to the question: are AI-systems products, services, or both? The proposal is not explicit about the issue, although the terms used to refer to the nature of AI-systems are ‘activities, devices, and processes’,³⁰ which can be interpreted in a way that we have an interplay between a product and a service. This brings forth a major deficiency of the proposal. It presupposes that the distinction between products and services is irrelevant for liability. However, on closer inspection, this stance would lead to infractions on the PLD regime to the detriment of the victim. For example, when the producer of an AI-system also provides updates and maintenance under a comprehensive maintenance (service) contract, the PLD regime would be undermined because negligence liability would be applicable,³¹ with possibly less protection for the victim.³² This is also true when the PLD regime would be extended to software.

19.3.2 The causation turbulence

The element of causation is indispensable for a proper understanding of the concept of liability. It is the ability to distinguish an injury from a misfortune and to isolate the accountable actor(s) from the multitude of events/actions leading to the injury/harm.³³ In the medical context many professionals and devices play an important part in the treatment and care of patients. Causation is therefore an important instrument with which to find the accountable actor(s). If causation cannot be proven, then the accountability gap widens, and the liability framework produces legal as well as economic failure.

The literature explains the likelihood for failure by using two concurring elements: Control and foreseeability.³⁴ For example, the producer of a medical device is assumed to be in control of the manufacturing process, as well as to be able to reasonably foresee all the uses of the device and the expectations of the users (doctor, nurse, laboratory analyst etc.). When the producer can prove that a defect was not reasonably foreseeable,³⁵ or that the defect occurred due to circumstances out of his control,³⁶ the producer may escape liability, because there is no causal link between the harm and the product defect. Therefore, any liability regime must keep a keen eye on control and foreseeability of damages.

The concept of control plays an important role in the proposal. It is one of the elements used to define the operator.³⁷ Recital 10 explains that the rationale of the proposal is based on the fact that the operator exercises a degree of control over the AI-system, comparable to the control that an owner has over its car. However, there appears to be a contradiction between Recital 10 and Article 3, as the

³⁰ Recital 17 of the proposal.

³¹ See Article 8 of the proposal for negligence-based liability rules.

³² Besides defect, harm, and causal link, the negligence-based regime requires a higher threshold for liability, specifically a breach of duty of care, which in turn offers less protection for the victim compared to the PLD regime which would allocate liability regardless of any specified breach in the duty of care.

³³ Karnow 1996, p 191.

³⁴ Matthias 2004 and Karnow 1996. The role of control and foreseeability is also evident in the PLD, respectively Article 7(e).

³⁵ O’Quin 2000, p 287 and Karnow 1996, p 191.

³⁶ Matthias 2004, p 175.

³⁷ Article 3 of the proposal.

latter defines control as *any* action that influences the operation of an AI-system. Perceiving control as *any* action of influence is a lower standard than the control normally required under strict liability. Therefore, it is unclear which version of control the proposal assumes to exist with operators over AI-systems. Is it 'any influence' or the standard level of control?

Both versions of control are problematic. Firstly, if we consider that the standard of control can be lowered to equate influence, then we can support the conclusion that operators are in control over the performance of the AI-systems. They can substantially influence the learning process, the data sets that are fed, and the environment where the AI-system will be deployed.³⁸ However, such an interpretation of control is incompatible with the concept of control that currently justifies the strict liability rationale. Producers are not able to merely "influence" the manufacturing process, but they must control it to such a degree that they get the desired output. The driver of a (non-autonomous) vehicle is not "influencing" but controls every move of the vehicle. To put it briefly, if one accepts that the operator's degree of control over an AI-system is on a level of influencing the system, and not being in control of its performance, the rationale for strict liability of such operators fails its doctrinal requirements. A compromised level of control is insufficient to support causation in all variety of circumstances.

Similarly, if one was to assume that operators have a high degree of control over AI-systems in healthcare, comparable to the driver of a vehicle or the producer of a product, it would fulfil the causation requirements. However, the ability of operators to have such a degree of control is unsubstantiated by research.³⁹ The machine learning rationale, supervised or unsupervised, is based on unpredictability of the learning process as one of its main benefits.⁴⁰ While methods to audit what the algorithm has learned exist, these methods are yet not perfect and often it is not even possible for the programmer to know how the algorithm made the decision, specifically in cases of deep learning.⁴¹ This problem is also known as the 'black-box' problem, which the proposal also recognizes.⁴² Based on this understanding, the assumption that operators have such a degree of control over the AI-systems that they are able to foresee and predict the damage that AI-systems may cause is not mirrored by reality.

To summarize, it is evident that AI-systems cause a turbulence of the elements of causation, leading to a lack of control and foreseeability. The proposal largely overlooks the turbulence in causation by introducing strict liability for operators. An adoption of the proposal would therefore lead to two unwelcomed consequences: Either operators would still be able to escape liability, because the causal link fails, or an overstretch of the liability framework is accepted, according to which causation is no longer conditioned on control and foreseeability of the damage. Both options have the potential to disrupt the intention of the liability regime and the objectives behind it. The former would leave the victim with less recourse vis-à-vis the operator, while the latter makes the actual use of the technology unattractive. Moreover, the economic rationale associated with liability regimes becomes thwarted as will be shown in the next section.

³⁸ Ibid.

³⁹ See Matthias 2004 and the discussion on the accountability gap in the introduction of this chapter.

⁴⁰ Alpaydin 2016, p 1-28.

⁴¹ Annany and Crawford 2018.

⁴² Recital 3 of the proposal.

19.4 Economic analysis

In this section we reflect on the operator's liability more conceptually, supplementing the doctrinal analysis by an economic analysis that sheds more light on the requirements that a modern liability regime must fulfil. To that end it is necessary to distinguish between three analytical steps: 1) the basic function of liability law; 2) the situational context factors for attributing liability; 3) the expected effect of the liability regime and its social desirability.

19.4.1 The two basic functions of liability law

Liability law has two functions: Compensation of victims and deterrence of wrongdoing.⁴³ Providing a fair compensation for victims in case of a mischief is largely with which liability law scholars are concerned. There exists a plethora of legal doctrinal literature that tries to find out what fair compensation is, and under which circumstances a specific rule applies.⁴⁴

Any compensation must come from some funds and becoming sentenced to pay compensation has repercussions on the behavior of agents. Thereby it obviously makes a difference whether one operates under unlimited personal liability or if one is exempted from liability, has limited liability, or is insured. That means that there is, beyond the question of what is seen as a morally fair compensation, another question which asks how the likelihood of a wrongdoing leads to an obligation to pay compensation and how this affects the activity and care level of selling products and services. This question covers the second function of liability law, in which the economic analysis is mostly interested. Which design of liability law drives decision making over risky activities towards the socially optimal amount of damages?⁴⁵

In this chapter it cannot be discussed how liability law affects the activity and care level of (risky) activities in detail. For the purposes of this chapter, it is more important to recollect why, next to individual liability the doctrine of enterprise liability is in place, putting liability not only on individuals but also on companies, for example, hospitals.⁴⁶

The generic logic for enterprise liability follows from a simple syllogism: a) The majority of products and services are delivered by manufacturers organized in enterprises (that is not different in health care); b) holding enterprises liable is setting optimal incentives; c) therefore, holding enterprises liable is socially beneficial.⁴⁷

For the future it can be expected that most AI-systems will be active in enterprises; for example, surgery robots will be around in hospitals or nursery robots in elderly homes. Hence, condition a) of the syllogism is applicable. But one may wonder whether condition b) is still fulfilled. Condition b) assumes that letting enterprises pay compensation sets incentives for the enterprise to control the risk and to do the necessary research and development for improving the product quality. Besides that, it is assumed that considerable profits accrue to enterprises and that, therefore, enterprises have the capacity to pay compensation. But what if this intuition would no longer be true? What if the nexus between liability and incentivizing firms to manage developmental risks is no longer given? What if a surgeon can neither control nor improve a surgery-robot, not even in conjunction with the

⁴³ Posner 1972 and Shavell 2007.

⁴⁴ See, e.g., van Dam 2013.

⁴⁵ Shavell 2007. For an overview, see Cooter and Ulen 2012 and Posner 2011.

⁴⁶ Heine and Grabovets 2015, p 44. For a discussion see also Arlen 1994.

⁴⁷ Engstrom 2013.

technicians of a hospital? Then the enterprise – the hospital - might still be able to pay compensations to victims, thereby fulfilling the first function of liability law, but it would no longer fulfil the second function, which is optimal deterrence and bringing damages down to the optimal social level.

This creates three challenges for the legislator: 1) identifying a target for liability law that is better than the enterprise, while maintaining the two functions of liability and its doctrinal requirements; 2) adapting the doctrines of liability law and thereby embracing AI-systems within a new perimeter of liability law 3) or stepping out of liability law and regulate damages caused by AI by other legal means.

19.4.2 A change in context factors

Whether there is a serious problem for the application of incumbent liability law for cases in which AI-systems are involved depends on whether AI-systems create a change in the context factors in which incumbent liability law usually operates. As already mentioned in the introduction, there exist three risks.⁴⁸ 1) The autonomy risk: if an AI learns, decides, and acts in a non-predictable way, then it comes closer to an autonomous agent that might be treated as an entity with a separate legal personhood. This understanding raises new questions with regard to incentivizing the optimal behavior in AI-systems. 2) The association risk: it is not necessarily clear who the master and who the servant is in each possible situation, for instance a surgeon and a robot learning from each other. 3) The network risk: This risk emerges when, e.g., algorithmic platforms interconnect local AI-systems in hospitals, in order to accelerate machine-learning and to maintain the robots' performance. This risk raises questions on the possibility to design a 'network liability'.⁴⁹

AI-systems change the landscape in which incumbent liability law is used to operate. One may compare that with a situation in which a Formula 1 racing car would be supposed to drive off-road. The importance of the fit between context factors and advanced technology was already highlighted by Nobel laureate Herbert Simon in 1960 when he was asked to sketch out the challenges of automation 25 years ahead in 1985.⁵⁰ Again, this is not an issue of 1 or 0, but a question of overstressing incumbent liability law. This is similarly to a car (liability law) with not enough horsepower (inappropriate doctrinal law) and a weak clutch (not well-prepared courts) cruising through the Alps (a world of digitalization). Then it is only a matter of a few kilometers climbing the hills before the car overheats and the machine collapses.

19.4.3 The expected effect of the liability regime

The foregoing analysis did not highlight all intricacies with which liability law is confronted in a world of AI and Big Data. However, it became clear that the two functions of liability law – compensation and deterrence – are no longer easily fulfilled by incumbent liability law doctrines. They neither set an accurate incentive for doctors in health care, nor for the health care institutions to maintain an optimal level of damages when AI-systems are at play. That has largely to do with the fact that AI-systems cannot fully be controlled by the doctor or the hospital and – that seems even more important – cannot technically be improved by them. That means the deterrence effect of liability law has no effect or may lead to unintended consequences. For example, health care providers might largely abandon advanced AI-systems or marginalizing it to a degree which does not overstretch the incumbent legal doctrines. Or, similarly as in US hospitals, where doctors are not employees of the hospital in order to

⁴⁸ Teubner 2018, p 129-149.

⁴⁹ Teubner 2018, p 129-149.

⁵⁰ Simon 1960, p 17-55.

avoid vicarious liability, strange contractual relations between hospitals, doctors and AI-system operators would be created to somehow fit in the new technologies in the incumbent legal framework.⁵¹ It is apparent that those individual avoidance strategies are not collectively rational and are diminishing potential social welfare gains through the new technology.

What also becomes clear is that the manufacturer or the intermediate providers of AI services must get a more prominent role in the liability regime than it is the case today. It is the manufacturer or provider who has primary control and can monitor the developmental risks of AI-systems.⁵² This becomes prominently apparent in the network risk of AI, but it is also prevalent in the autonomy and association risk. For that background it is intuitively clear to put forward a concept like in the proposal that aims at allocating the liability to the entity that can control and foresee potential harms. But shifting strict liability without any further reflections to a so-called operator or the producers of AI-systems is inherently an overstretch of the incumbent system. This approach makes the same mistake as leaving the current system unchanged, because it is only broadening the incumbent doctrines of liability law without making any substantial changes to it. Operators will find ways to contract around, avoiding a position of being held strictly liable; similar to US hospitals avoiding vicarious liability for the malpractice of their doctors. Or, operators and producers will simply not provide advanced AI-system solutions if they must fear strict liability. A chilling effect on innovation is then the result.

A more comprehensive approach to liability law reform is advisable, which is more likely to bring damages caused by AI-systems on a social optimal level. It comprises changes within liability law as well as regulatory measures outside the classical perimeter of liability law.⁵³ A *first* line of reform would include a debate about introducing legal personhood for AI-systems.⁵⁴ That way a more tailored attribution of responsibility for autonomous decision making would become possible as well as a better alignment with contract and property law. More technically, AI-systems would be integrated into the “nexus of contracts” that a firm basically is.⁵⁵ Thereby it is not meant to create legal personality in analogy to a human, but according to the functions and purposes AI-systems fulfill in a specific context. That is like the invention of company law that has evolved over the last 200 years into a highly differentiated legal tool to facilitate innovation and productivity in modern societies.⁵⁶

A *second* line of reform would have to focus on so-called *design regulations*. Those regulations would prescribe *ex ante* for AI-systems a specific technical design, as for example a robot arm must not be longer than 1 meter for a specific application, or the software must not include a specific sort of algorithm. But a design regulation may also prescribe that a surgeon co-working with a robot needs a

⁵¹ See, e.g., Abraham and Weiler 1994, and Cebul et al. 2008.

⁵² Galasso and Luo 2019, p 493-504.

⁵³ A collection of alternative regulations added to liability law can be found in Rachum-Twaig 2020. Similarly, Galasso and Luo 2019, p 493-504.

⁵⁴ In 2017 the European Parliament was open to this attribute of electronic entities; Res. P8_TA(2017)0051. The sometimes-furious responses (e.g. Robotics, *Open letter to the European Commission Artificial Intelligence and robotics*, <http://www.robotics-openletter.eu/>) appear to have wiped the whole idea off the table. The Commission’s White Paper of February 2020 remains silent; EC White Paper of 19 feb 2020, COM(2020) 65 final.

⁵⁵ The basic idea of conceiving the firm as a “nexus of contracts” has been put forward by Jensen and Meckling 1976. For a legal discussion see also Easterbrook and Fischel 1989.

⁵⁶ For the facilitative effect of corporate law for industrialization in a comprehensive political setting see, for example, Pistor et al 2002.

specific training and license to work with the robot. In that sense design regulations contain a whole spectrum of regulations ranging from hardware requirements⁵⁷ to the law of professions.

A *third* line of reform would have to consider the social relevance of AI-systems vis-à-vis the risk that it may bear for the single individual.⁵⁸ Individuals may realize the enormous benefit of AI for society as such, but may fear at the same time becoming an uncompensated victim of the new technology. In order to avoid that individual rationality inhibits collective rationality, compensation of harm must be guaranteed. There is the possibility of mandating liability insurance for AI-systems (as already indicated in the proposal), whenever it is economically feasible. But that might not be applicable in all cases, and mandating compensation from liability law indifferently may thwart the technology as such.⁵⁹ Thus, as it is already the case with new vaccines, new aircrafts or nuclear technology, a partial and temporary liability exemption for AI-systems might be in order.⁶⁰ This exemption must then be accompanied by a compensation scheme paid out by the government as the most potent insurer of a socially desired innovation that otherwise would not be provided.⁶¹

In summary, the debate about operator's liability in healthcare is part of a larger debate about the applicability of incumbent liability law to cases in which AI-systems are involved. Whether granting legal personhood to AI-systems is the key to avoid a "doctrinal overstretch" of incumbent liability law on the one hand and to avoid an accountability gap on the other hand, will only become clear in future decennia, similarly to the introduction of the limited liability company, in a reaction to the socio-economic and technological disruptions of the colonial trade boost and consecutively, the industrial revolution.⁶² For the time being it is advisable to have a mix of legal reforms that consider very specific design regulations to avoid non-acceptable harm, as well as insurance schemes for compensation – including the government as guarantor of last resort – to facilitate and disseminate the application of AI-systems in society.

19.5 Conclusion and further research

This chapter undertakes a critical analysis of the idea to put strict liability on operators of AI-systems in healthcare, as has been recently proposed by the EP. An important finding is that the legislative proposal does not address the product/service dichotomy in relation to AI-systems. This leads to uncertainties and disruptions in the product liability framework of the EU. Moreover, the analysis highlights a doctrinal inconsistency in the level of control needed for operators of AI-systems to fall under strict liability. Either operators would be able to escape liability because the causal link for wrongdoing is missing, or an overstretch of the liability framework is accepted, according to which causation is no longer conditioned on control and foreseeability of the damage.

⁵⁷ Design regulations are also discussed for 3D-printing, which is another disruptive digital technology. See Heine and Li 2019.

⁵⁸ See Fosch-Villaronga 2019 for a risk-based approach to regulation of healthcare robots.

⁵⁹ The economic effects of granting limited liability are yet not fully understood and are dependent on specific contexts. A recent study by Koudijs and Salisbury 2020 on the effect of limited liability in marital property law sheds more light on the "social innovation" of limited liability.

⁶⁰ For a similar argument see Galasso and Luo 2019, p 97-98

⁶¹ Here it might be interesting to learn from the literature on catastrophe insurance by smart public-private insurance schemes. This literature puts less the deterrence aspect into focus (a catastrophe is to a large degree an unavoidable random shock), but the compensation of victims. See, e.g., Bruggeman et al. 2010.

⁶² For a recent discussion on AI and legal personality see Chesterman 2020.

The economic analysis points out that the deterrence effect, as one of the two rationales behind a liability framework, does not work in cases when the enterprise is not in control of the harm or damage that may occur from the use of AI-systems. That means that as a necessary condition the doctrinal law must be consistent to make the economics work towards a welfare improvement. Moreover, the shift in context factors, namely the autonomy risk, the association risk, and the network risk, implies the need to introduce a new actor in the liability framework, such as the operator of AI-systems, to whom harm and damage can be rightly attributed. However, affording strict liability to the operator of the AI-system as foreseen in the proposal would lead either to attempts of bypassing liability as such by complex contractual arrangements between the operators and users of AI-systems, or a chilling effect for the use of the AI-systems can be expected. Both cases are detrimental to social welfare.

It has not been the aim of this contribution to make normative suggestions how to develop the liability framework in the wake of technological disruption. One line of thought is to further consider legal personality for AI-systems and to create with that a clear target for responsibility. However, this avenue for reform of liability law needs more detailed studies, including an elaborated ethical discourse about how much risk society is ready to bear in order to make advanced AI-systems available. Focus has also to be put on the man-machine interaction in health care. More empirical and practice research will yield valuable insights into the workability of specific medical regulations, but also in the challenges that stem from the product/service dichotomy. Finally, regulatory ex-ante approaches of AI-systems, such as design regulations, also merit further attention.

Reference List

- Abraham K S and Weiler P C (1994) Enterprise Medical Liability and the Evolution of the American Health Care System, *Harvard Law Review* 108:381-436
- Alpaydin E (2016) *Machine Learning*. MIT Press, Cambridge MA
- Annany M and Crawford K (2018) Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability, *New Media & Society* 20: 973-989
- Arlen J (1994) The Potentially Perverse Effects of Corporate Criminal Liability, *Journal of Legal Studies* 23:832-867
- Bruggeman V et al (2010) The Government as Reinsurer of Catastrophic Risks?, *Geneva Papers on Risk and Insurance Theory* 35:369-390
- Cebul R D et al (2008) Organizational Fragmentation and Care Quality in the U.S. Health Care System, *Journal of Economic Perspectives* 22:93-113
- Chesterman S (2020) Artificial intelligence and the limits of legal personality, *International and Comparative Law Quarterly* 69:819–844
- Cooter R and Ulen Th (2012) *Law and Economics*. 6th Ed. Addison-Wesley, Boston
- Easterbrook F H and Fischel D R (1989) The Corporate Contract, *Columbia Law Review* 89:1416-1448
- Engstrom N F (2013) 3-D Printing and Product Liability: Identifying the Obstacles, *University of Pennsylvania Law Review Online* 162:35-42

European Commission (1990) Proposal for a Council Directive on the liability of suppliers of services COM/90/482FINAL - SYN 308. <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:51990PC0482&from=EN> Accessed 16 August 2021

European Commission (2020) White Paper 'On Artificial Intelligence - A European approach to excellence and trust' COM(2020) 65 final. https://ec.europa.eu/info/publications/white-paper-artificial-intelligence-european-approach-excellence-and-trust_en 25 February 2021 Accessed 16 August 2021

European Parliament Resolution (2020) A framework of ethical aspects of artificial intelligence, robotics and related technologies. [https://oeil.secure.europarl.europa.eu/oeil/popups/ficheprocedure.do?lang=en&reference=2020/2012\(INL\)](https://oeil.secure.europarl.europa.eu/oeil/popups/ficheprocedure.do?lang=en&reference=2020/2012(INL)) Accessed 16 August 2021

European Parliament Resolution (2020) Civil liability regime for artificial intelligence. [https://oeil.secure.europarl.europa.eu/oeil/popups/ficheprocedure.do?lang=en&reference=2020/2014\(INL\)](https://oeil.secure.europarl.europa.eu/oeil/popups/ficheprocedure.do?lang=en&reference=2020/2014(INL)) Accessed 16 August 2021

Ficuciello F et al (2019) Autonomy in surgical robots and its meaningful human control. <https://doi.org/10.1515/pjbr-2019-0002>

Floridi L and Sanders J (2004) On the morality of artificial agents, *Minds and machines* 14.3:349-379

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

Galasso A and Luo H (2019) Punishing Robots: Issues in the Economics of Tort Liability and Innovation in Artificial Intelligence. In: Ajay A et al (eds), *The Economics of Artificial Intelligence: An Agenda*. University of Chicago Press, Chicago pp. 493-504

Heine K and Grabovets K (2015) From individuals to organizations: The puzzle of organizational liability in tort law, *Dovenschmidt Quarterly* 2:44-53

Heine K and Li S (2019) What Shall we do with the Drunken Sailor? Product Safety in the Aftermath of 3D Printing, *European Journal of Risk Regulation* 10:23–40

Jensen M C and Meckling W H (1976) Theory of the firm: Managerial behavior, agency costs and ownership structure, *Journal of Financial Economics* 3:305-360

Karnow C EA (1996) Liability for distributed artificial intelligences, *Berkeley Technology Law Journal* 11:147-204

Koch B (2011) *Medical Liability in Europe: A Comparison of Selected Jurisdictions*. De Gruyter, Berlin

Koudijs P and Salisbury L (2020) Limited liability and investment: Evidence from changes in marital property laws in the US South, 1840–1850, *Journal of Financial Economics* 138:1-26

Matthias A (2004) The responsibility gap: Ascribing responsibility for the actions of learning automata, *Ethics and Information Technology* 6:175–183

O'Quinn J C (2000) Not-So-Strict Liability: A Foreseeability Test for *Rylands v. Fletcher* and Other Lessons from *Cambridge Water Co. v. Eastern Counties Leather PLC*, *Harvard Environmental Law Review* 24:287-314

- Pistor K et al (2002) Evolution of Corporate Law: A Cross-Country Comparison, University of Pennsylvania Journal of International Economic Law 23:791-872
- Posner R A (1972) A Theory of Negligence, Journal of Legal Studies 1:29-96
- Posner R A (2011) Economic Analysis of Law. 8th Ed. Wolters-Kluwer, New York
- Rachum-Twaig O (2020) Whose Robot Is It Anyway?: Liability for Artificial-Intelligence-Based Robots, University of Illinois Law Review 4:1141-1175
- Shademan A et al. (2016) Supervised autonomous robotic soft tissue surgery, Science Translational Medicine 8:337–364
- Shavell S (2007) Economic Analysis of Accident Law. Harvard University Press, Cambridge MA
- Shishehgar M et al. (2018) A systematic review of research into how robotic technology can help older people, Smart Health 7:1-18
- Simon H (1960) The Corporation: Will it be Managed by Machines. In: Ansehn M and Bach G L (eds), Management and Corporations. Mc Graw Hill, New York pp.17-55
- Teubner G (2018) Digital Personhood? The Status of Autonomous Software Agents in Private Law. <http://dx.doi.org/10.2139/ssrn.3177096>
- van Dam C (2013) European Tort Law. Oxford University Press, Oxford