Bridging the liability gaps: Why AI challenges the existing rules on liability and how to design human-empowering solutions

De Conca, S.

published in
Law and Artificial Intelligence
2022

DOI (link to publisher)
10.1007/978-94-6265-523-2_13

document version
Early version, also known as pre-print

Chapter X - Bridging the liability gaps: why AI challenges the existing rules on liability and how to design human-empowering solutions

Silvia De Conca[*]

Abstract. *This chapter explores the so-called 'liability gaps' that occurs when, in applying existing contractual, extracontractual, or strict liability rules to harms caused by AI, the inherent characteristics of AI may result in unsatisfying outcomes, in particular for the damaged party. The chapter explains the liability gaps, investigating which features of AI challenge the application of traditional legal solutions and why. Subsequently, this chapter explores the challenges connected to the different possible solutions, including contract law, extracontractual law, product liability, mandatory insurance, company law, and the idea of granting legal personhood to AI and robots. The analysis is carried out using hypothetical scenarios, to highlight both the abstract and practical implications of AI, based on the roles and interactions of the various parties involved. As a conclusion, this chapter offers an overview of the fundamental principles and guidelines that should be followed to elaborate a comprehensive and effective strategy to bridge the liability gaps. The argument made is that the guiding principle in designing legal solutions to the liability gaps must be the protection of individuals, particularly their dignity, rights and interests.*

Keywords: Artificial Intelligence; contract law; damages; European law; liability; tort.

X.1 Introduction

In the Autumn of 2019, software developer David Heinemeier Hansson and his wife applied for Apple Card, a brand-new credit card service offered by Apple Inc. and issued by Goldman Sachs. Much to their surprise, when their Apple Cards were issued, they discovered that Mr. Hansson's credit limit was 20 times that of his wife. Some factual circumstances made this difference particularly troubling: the couple had been married for a long time in community-property regime, filed joint tax returns, and Mr. Hansson even had a lower credit score than his wife![1] The couple suspected that the algorithm had learned to discriminate against women. They contacted Apple's customer support, which re-directed them to Goldman Sachs. The latter affirmed they were certain that there was no discrimination of Mrs. Hansson and that some other factors had probably influenced the decision. Which factors, however, they could not tell, because neither they nor Apple knew exactly how the complex algorithm tasked with screening the customers had reached its decision.[2] The couple took it to Twitter to voice their terrible suspicion and discovered they were not alone: several other cases emerged – including the one of Steve Wozniak, Apple's co-founder, whose wife had also been

[*] Dr. Silvia De Conca, Transnational Legal Studies department, Vrije Universiteit Amsterdam, The Netherlands
e-mail: s.deconca@vu.nl

[1] Vincent J (2019) Apple's credit card is being investigated for discriminating against women https://www.theverge.com/2019/11/11/20958953/apple-credit-card-gender-discrimination-algorithms-black-box-investigation Accessed 19 February 2021

[2] ibid.

discriminated against by Apple Card.[3] Eventually, the Apple Card service was put under investigation by the New York State Department of Financial Services.[4]

The case of Apple Card's involuntary algorithmic discrimination against women is paradigmatic: it shows that when an AI system[5] damages a person, uncertainties arise regarding which subject should be liable for the damage and its redress. These uncertainties, as will be explained below, derive from technological and organizational factors that challenge the traditional balances and solutions provided by private law regarding damages and liability. As a result, gaps are created in the existing legal regimes that jeopardize the possibility for damaged parties to obtain redress. It is to these gaps, referred to as the liability gaps, that this chapter is dedicated. The analysis focuses on the issues connected to the liability gaps for damages caused by AI from a European perspective, leveraging the efforts and initiatives of European policymakers, as well as the lively debate happening among experts and scholars.

Section X.2 offers a definition of the liability gaps, with an overview of the main features of AI systems that challenge the traditional allocation of legal liability among the relevant parties. Subsequently, section X.3 explores the main doctrinal solutions suggested so far to bridge the liability gaps, spanning from contract and extracontractual law to corporate law, insurance, and even to the idea of granting a form of legal agenthood to AI. The analysis is developed using hypothetical scenarios describing various AI systems deployed in different settings: financial services, autonomous vehicles, and smart consumers products. Finally, section X.4 concludes the chapter, offering principles and guidelines that should be taken into consideration in elaborating a solution to the liability gaps. In doing so, section X.4 stresses the importance of maintaining a human-centred perspective while designing a combined approach to intervene at different levels, adjusting existing regulations and introducing new ones, where necessary.

X.2 Mind the gap: the disruption to the liability rules caused by AI

The term 'liability gaps' refers to a situation in which, in the presence of damages, the allocation of legal liability to the parties involved is disrupted by certain circumstances. Due to the disruption, the existing legal institutes do not provide satisfactory tools to remedy the harms suffered, and the damaged parties might be unable to obtain redress.[6]
Typically, whenever damages occur, the allocation of liability would depend on some or all of the following elements: the existence of the damage, any relationship existing among the parties involved, possible conducts that the parties were supposed to adhere to according to

[3] ibid.

[4] Vigdor N (2019) Apple Card Investigated After Gender Discrimination Complaints https://www.nytimes.com/2019/11/10/business/Apple-credit-card-investigation.html Accessed 19 February 2021. Please note that as of February 2021 no additional information on the investigation or its outcomes appears to have been made available.

[5] In this chapter, AI system indicates a product, device, service, or machine deploying a form of AI. AI should be intended in this chapter as any Machine Learning or other data analytics techniques presenting the capability to achieve a certain objective with a significant degree of autonomy, following supervised or unsupervised learning or other forms of software learning capability.

[6] The phenomenon is often referred to as responsibility gaps or liability gaps, cfr. among many, Johnson 2014, p 2; Bertolini 2013, p 231; European Commission 2019, p 17

the law, the actual conduct of the parties, the existence of defects if the damage derives from a product, the causal link between the conduct and the damage, the nature of the damage.
As an example, consider the case of a driver that, while backing out of a parking spot, hits and damages a neighbour's letterbox. Temporarily leaving aside the fact that most European countries provide for a special liability regime for car accidents, it is nevertheless possible to identify the main elements based on which the liability is ascribed. As represented in Figure X.1, the two main parties involved are the driver (indicated with Alice in the figure) and the neighbour (indicated with Bob). The dotted line between them represents the relation between the two: they are connected by the event itself, which is accidental, random, and not persistent over time (hopefully). Their personal spheres came into contact because of the damage to the letterbox and are not connected by any form of pre-existing contract or agreement. With regard to Alice's conduct, the law imposes certain duties on drivers, due to the potential dangers connected with cars. By behaving in a more careful way, for instance looking into the car's mirror, Alice could have prevented the accident. Alice's behaviour (driving without checking the mirrors) is the source of the damage (the broken letterbox), which can be easily quantified in the cost of a new letterbox (or of the repairs to the damaged one).[7]

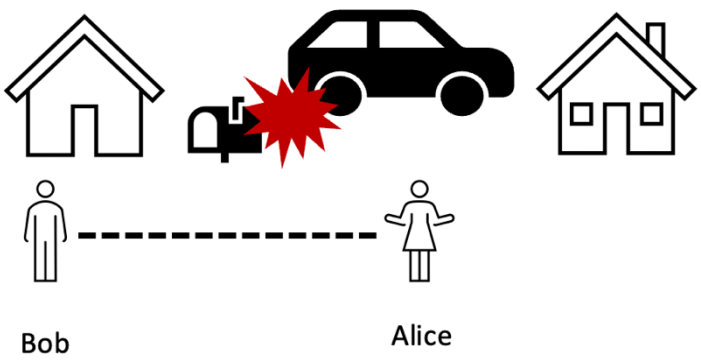


Figure X.1: Alice accidentally crashes into her neighbour Bob's letterbox.

Emerging technologies can disrupt the allocation of liability, interfering with the elements listed above. The arrival of new technologies, as well as the novel application of existing technologies, can often give life to an initial period of regulatory uncertainty,[8] during which liability gaps might become evident, until a proper liability solution is established by the legislator. This was, for instance, the case of industrial machines in the 19th century United States. At the time, the country suffered a staggering increase of accidents involving steam-powered heavy machinery.[9] The numbers and circumstances being unprecedented, it was not clear initially who was liable and how the damaged individuals could obtain redress.

[7] The example is a simplification. In real life, special liability regimes would create a presumption of fault of Alice. Alice would have a mandatory civil responsibility insurance to cover possible damages occurred while driving. Furthermore, additional parties might be involved, such as the manufacturer of the car or of any of its components, that might exempt Alice from any liability or to which she could demand to be indemnified, in case of defects or malfunctioning.

[8] Pagallo et al 2018, p 19

[9] Witt 2001, p 694

Throughout the second half of the 19th and the beginning of the 20th century, attempts were made at deploying existing legal and institutional mechanisms to solve this novel problem, until a branch of tort law prevailed, becoming the main framework for the regulation of damages caused by heavy machinery and other kinds of machines (sometimes referred to as accident law).[10]

In the case of AI, several factors contribute to the emergence of liability gaps and, in particular, the fact that AI systems are composed of several hardware and software parts interacting with each other, sometimes in unpredictable manners; the many actors involved, such as the various manufacturers and providers of the different components, as well as possible intermediaries and developers of connected applications or products; the capability of the algorithm to learn from data sets, changing over time; the changes that updates might induce to the software component sometimes triggering new or unexpected vulnerabilities or problems.[11]

AI systems are particularly complex, because they are composed of several hardware and software parts, and often developed at the crossroads of different industries besides computer science (such as robotics, healthcare, automotive, cloud computing, Internet of Things or IoT, telecommunications, and so on). From this complexity, two disruptive factors emerge: the presence of several actors, potentially responsible for the various components, the programming, the data sets based on which the algorithm elaborates its models; and the additional layer of unpredictability deriving from the ways in which the AI element might interact with the various components, with interconnected technologies (for instance, the IoT, where several 'smart' devices interact with each other), or even with the humans around it.[12]

AI-based products are not as static as their traditional counterparts but change over time. First and foremost, they change because the AI keeps learning and evolving based on the data collected from sensors and other sources. AI is inherently dynamic and designed to learn, respond to external stimuli, and elaborate new solutions to reach a goal. AI systems also change over time because, like many other digital technologies, they require periodical updates to fix vulnerabilities and/or errors in their functioning (the infamous bugs).[13]

Some of these factors are not new (like the presence of several components and parties, or the updates), and merely enhance problems that have already emerged in the past with other technologies. The capability to learn over time, instead, is specific to Artificial Intelligence and represents a new challenge for the law.

Each of the abovementioned factors can have a disruptive effect on the essential elements of the allocation of liability. The presence of several different components, and the possibility of unexpected effects deriving from their interaction, can affect the causal link between the conduct of the parties and the damage.[14] Consider, for instance, a smart thermostat, composed of temperature and humidity sensors distributed around the house and a central thermostat unit: possible errors in the temperature-setting of the thermostat might be the

[10] Witt 2001, p 745

[11] European Commission 2019, p 21

[12] Leenes et al 2017, p 9

[13] European Commission 2019, p 21

[14] Scherer 2016, p 363

consequence of an erroneous data collection in the sensor, of a defect in the physical components of the sensors, or of the software optimizing and analysing all the data collected. If the thermostat is also connected to other IoT devices around the house, errors in the collection and analysis of the data collected by these other devices, or defects in their components, might also affect the smart thermostat. The interaction of several components and devices can give unexpected results, which makes it difficult to identify the origin of the damage. It also multiplies the actors involved. For the smart thermostat, the parties involved might include not only the manufacturer of the thermostat itself, but also the manufacturers of the sensors, other components (like processors or motherboards), or other devices (and of their components). The presence of several actors affects the identification of the conduct that the parties should have adhered to: each party involved is bound by a certain duty of care and professional standards, based on each field and the relative state of the art.

The same disruptive effect on the causal link can derive from the learning capability of AI. The autonomy of AI can challenge the traditional human-centred perspective of liability.[15] If the above-mentioned thermostat has 'learned' autonomously (by elaborating a certain model based on the data collected) how to regulate humidity and temperature in the house, is the manufacturer responsible for possible damages? Is it a defect or a design error if an AI autonomously learns a certain behaviour? If an AI system learns from the data derived from the behaviour of its users, are these users responsible for training the algorithm?[16] These elements appear difficult to reconcile with the traditional approach of the law, which is based on products that do not react to external input in an autonomous (or semi-autonomous) way, but whose behaviour is entirely anticipated and pre-programmed by their manufacturers. The learning capability of AI also increases the possibilities that a machine might cause not only damages that are patrimonial in nature, but also pure economic loss or even psychological and existential (as is the case with the discrimination of women carried out by the algorithm of Apple Card). The nature of the damages can affect the possibility to obtain redress, because not all legal systems regulate psychological, existential, or pure economic loss.[17]

X.3 State of the art: proposed legal solutions to the liability gaps

Several possible solutions to the liability gaps created by AI systems have been explored by experts and European policy makers. Overall, there is consensus on the fact that currently the law offers tools to address damages caused by AI systems, but the existing tools might present shortcomings that lead to undesirable or less optimal results.[18] The application of the main existing legal tools to the liability gaps for damages caused by AI systems is illustrated below, using hypothetical scenarios.

X.3.1 Contractual and extracontractual liability (fault or negligence)

---

[15] In most legal systems only humans that are capable and of age are liable, even if the damage has been caused by a machine, an animal, or an underage or incapable person (with few exceptions). Cfr. Asaro (2011), p 176; Pagallo 2012, p 55

[16] Pagallo 2013b, p 504; Asaro 2011, p 174

[17] There is no universal definition of pure economic loss, and the rules concerning it vary widely among European countries. As a general definition, pure economic loss entails the suffering of an economic loss not connected to a pre-existing harm. Cfr. Bussani and Palmer 2003.

[18] Johnson 2014, p 2; European Commission 2019, p 16

In all European legal systems, civil liability (as opposed to criminal liability) derives from the breach of an obligation. The source of the obligation determines the kind of liability. The two main sources of obligations are contracts and delicts.[19]
By entering into a contract, the parties agree to comply with certain obligations, such as providing a product or service and, in return, paying the price. The breach of contractual obligations gives life, therefore, to the contractual liability of the party that did not comply with the obligation (unless certain excusatory circumstances occur).[20]
Extracontractual liability (tort in common law jurisdictions) occurs when two or more parties are not in a contract or in any other form of sanctioned relationship but, due to the behaviour of one party, the other suffers a damage.[21] In this case, the behaviour of the damaging party breaches a general duty of not causing harm (*neminem laedere*) either intentionally (fault) or because of negligence. In general, extracontractual liability applies when two strangers come into contact and one of the two is damaged by the actions of the other: this is, for instance, the case of a passer-by hit by a vase falling from a window. The person living in the house from which the vase fell is liable to pay the damages, because of negligent behaviour resulting in the breach of the general duty of care (i.e., not harming anyone). Extra-contractual liability can also be applied regardless of fault or negligence, in the form of the so-called strict liability for special categories of activities or products.

Most 'grey area' cases should be traced back to either contractual or extracontractual liability, based on the origin of the obligation: if a contract cannot be identified in the relationship between the parties, then the obligation (and the liability deriving from it) would be put under the extracontractual umbrella. This distinction has concrete consequences, in particular with regard to the burden of proof and prescription. Contractual liability usually puts the burden of proof on the party that has breached the contract (and not on the damaged party), with longer prescription times (up to 10 years in some European countries). Within extracontractual liability the burden of proof is on the subject who brings the action (the damaged party) and prescription is usually of maximum 5 years.[22] The burden of proof is important with regard to AI systems, because their opacity, the fact that they are proprietary knowledge of a party, and the necessity to hire expensive experts to analyse them, represent factual barriers to the access to justice for damaged parties.[23]
Both contractual and extracontractual liability can apply to cases of damages caused by an AI system, depending on the underlying obligations and relationships existing among the parties involved. Currently, however, applying any of these two types of liability to AI systems might give unsatisfactory results, as will be illustrated by the scenarios below and in the next subsection, respectively.

Scenario A: Alice & The Bank

[19] von Bar and Drobnig 2009

[20] Fauvarque-Cosson and Mazeaud 2009

[21] ibid.

[22] Marsh 1994; Farnsworth 2006

[23] Pagallo 2013a, p 135; European Commission 2019, p 51

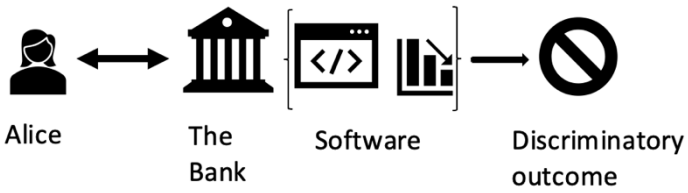
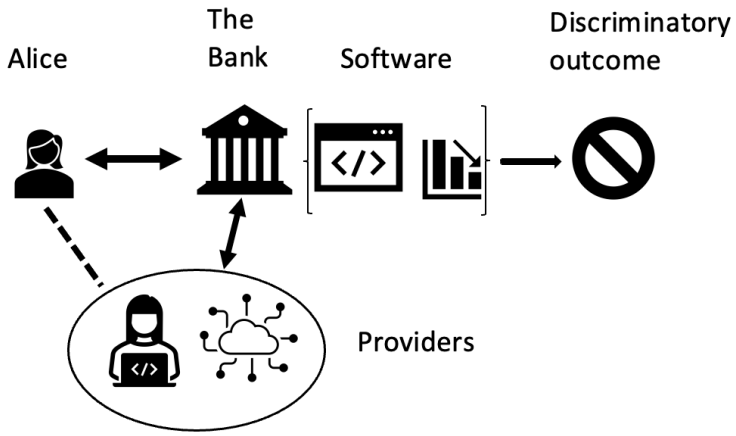


Figure X.2: A simplified representation relationship between Alice and The Bank. Alice has a contractual relationship with The Bank, whose software produces a discriminatory outcome.

The first hypothetical scenario starts with a "simple" two-actor situation: an individual and a company. In a situation similar to the case of Apple Card, Alice applies with a bank (The Bank) for a credit. The Bank uses predictive analytics software to select which applicants can obtain the credit. The software rejects Alice's application. It appears that the reason for the rejection can be traced back to a bias accidentally incorporated by the software in the models created from training data sets. The Bank, however, claims that it does not know the reasons behind the software's decision, denies the existence of biases, and affirms that it cannot be held accountable for the algorithmic outcome. From the perspective of liability, it can be difficult to identify the obligations of The Bank and, consequently, whether The Bank is in breach: The Bank is not obliged to grant Alice the credit.[24] Due to the opacity of the AI technology involved, it is also not easy to identify possible involuntary discrimination or whether The Bank wilfully instructed the algorithm. If the discrimination is identified and traced back to the algorithm learning a bias from the input data, the causal connection between the damage (the discrimination) and the conduct of The Bank is particularly weak.

This scenario is complicated by two additional factors. The first one is that other parties, for instance, the company developing the algorithm or a third-party company providing the dataset to train the AI (called Providers in Figure X.3), might also play a role. The discrimination might be connected to the actions of the Providers, who do not have a contract with Alice (the dotted line in Figure X.3).



[24] Please note that the relationship between Alice and The Bank might be either contractual or pre-contractual, depending on the kind of documents exchanged and the rules existing in a certain jurisdiction. The existence of a contractual or pre-contractual relationship does not automatically imply the existence of contractual liability in this case, because the damage derives from discrimination, which is an action prohibited by the law.

Figure X.3: The relationship between Alice and The Bank is complicated by the presence of the Providers, who provided The Bank with the AI software and the datasets to train it.

The position of Alice vis-à-vis the Providers can be traced back to extracontractual liability. If they provided The Bank with a component or software that results in harm to a third party, they breach a general principle of not harming. The extracontractual liability regime is less favourable to Alice under several perspectives. It affects the burden of proof, the prescription, and in general the kind of compensation Alice might be entitled to. With regard to the burden of proof, for instance, should this case be considered as extracontractual liability, in many jurisdictions Alice would have to prove the damage, its economic value, the causal connection with the behaviour of one of the additional parties, as well as their fault or even intent to damage her.[25] This is a significant burden of proof, especially because Alice would hardly have access to the AI and would most likely need to hire an expert to support her claim, which would be expensive. These circumstances would affect the access to justice and the possibility to obtain redress.

The complexity of this scenario is worsened by the fact that the kind of prejudice suffered by Alice can fall into the pure economic loss (Alice might not be able to purchase a house, or start a company, and so on) and/or into moral damage (discrimination). The possibility to obtain redress for any of these types of prejudice have been and often still are debated in many legal systems and is often excluded in case of extracontractual liability.[26]

It is important to consider that Alice has suffered a prejudice in her rights (economic means, potential lack of chances and the right not to be discriminated against). Furthermore, Alice trusted The Bank, assuming that in assessing her application it would have acted diligently and used professional care, not following biases and discriminatory behaviour. In this scenario, the law offers tools for the damaged party in the form of contractual or extracontractual liability rules. These tools appear, however, insufficient to ensure that the damages or losses suffered by a party are moved back into the sphere of those parties that had control over the cause of the damage and/or that benefited from it.

X.3.2 Another form of extra-contractual liability: strict liability

European legal systems generally provide for special forms of extracontractual/tort liability, with limited application. This is the case for strict liability and its derivatives, namely product liability, liability from driving automobiles (or other forms of registered movable property like airplanes or ships), as well as sector-specific regimes (special provisions exist for doctors, stock market brokers and financial institutions, teachers, and so on).[27]

In the context of these special forms of liability, usually the requisite of negligence or fault is withdrawn, and the burden of proof is often simplified or reversed: it is the damaging party that must prove that its conduct did not cause the damage or that all reasonable efforts have been made to avoid the damage, while the damaged party only needs to prove that the damage occurred, the defect, and the connection between the two.[28] The configuration of these forms of special liability makes them particularly appropriate to deal with dangerous activities (driving) or situations of asymmetry of information or power (consumer products).

[25] Pagallo 2011, p 55; Pagallo 2013a, p 115; Karnow 2016, p 51

[26] European Commission 2019, p 51

[27] Buyuksagis and van Boom 2013, p 609

[28] ibid.

Forms of special or strict liability could be applied to situations in which AI systems cause damages, to correct some of the shortcomings highlighted in the previous section. However, strict liability, product liability, liability from traffic accidents, and the other special regimes are exceptions to the norm. They cannot be used as a general rule, and their applicability is limited to certain circumstances, activities, or industries. Their deployment with damages caused by AI systems could only be possible for certain applications. For instance, the liability regime for automobiles could apply to autonomous vehicles[29] (the so called self-driving cars), while product liability could apply to smart consumer products.[30] The application of special liability regimes to AI is not without uncertainties and grey areas: it might not solve the liability gaps but reiterate them, as shown in the scenarios below.

Alternatively, a solution could be to create a special liability *ad hoc* for AI, to deal specifically with the liability gaps it creates.[31] While the idea appears tempting, it would most likely be unfeasible: AI presents a wide range of applications and uses, in almost every industry and sector. A unique regime applicable to all of this would likely be impossible. For this reason, the idea has been put forward to create special liability regimes for certain applications of AI that are high risk.[32] This would likely offer a solution at least for some liability gaps that emerge in connection with the use of AI. However, in determining what applications pose high-risks, legislators need to consider that, besides the most prominent examples (such as autonomous weapons and autonomous vehicles, due to their potential to injure or kill individuals)[33], trivial and daily AI products might still be prejudicial to the fundamental rights of individuals (as the Apple Card case shows). The more individuals carry out their lives immerged into digital and 'smart' environments, the higher the risks that even minor AI applications present important negative consequences.[34] The proposal for an AI Regulation made public by the European Commission in 2021 seems to acknowledge this, as it lists among the high-risk AI systems not only life-threatening applications, but also systems used to determine the access to education, to welfare and benefits, or AI software for recruitment.[35]

Scenario B: Alice & autonomous vehicles

The beginning of Section X.2 shows the allocation of liability when a driving accident results in damages to property (Figure X.1). What happens if Alice, instead of driving a traditional car, "drives" an autonomous vehicle?

It has been highlighted that the introduction of AI systems can increase the complexity of a situation, because of the opacity of the algorithm and the presence of multiple actors due to interoperable technologies. This is also the case with regard to autonomous vehicles. In a

[29] Royakkers and van Est 2016, p 185

[30] Asaro 2011, p 170; Bertolini 2013

[31] Bertolini 2013, p 219

[32] European Commission 2018 and 2019

[33] UNESCO 2017, p 19

[34] This idea has been brought forward at the Sixth T.M.C. Asser Annual Lecture, given by Prof. Andrew Murrey and entitled 'Almost human: Law and human agency in the Time of Artificial Intelligence', that took place on 26 November 2020, online.

[35] 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, of 21 April 2021 [2021/0106(COD)], Annex III

traditional, 'dumb' car, the driver is assumed to be liable, under a special liability regime. In some cases, if the driver can prove that the damage is the result of a defect or manufacturing of the car, the manufacturer will be liable, either alone or together with the driver.[36]

In this scenario, even though Alice is sitting in the driver's seat, the car is backing out of the parking spot by itself, based on the analysis of the data collected by its multiple sensors. Technically, Alice is not driving the car and does not have control over the decisions taken by the car itself. When the car hits the neighbour's letterbox, different hypotheses can be advanced about who is liable for the damage.
Initially, when the Sci-Fi dream of a car driving itself started to become a reality, the circumstance that the driver was not in control of the driving created some doubts about whether or not the driver could be considered liable.[37] It should, however, be taken into consideration that the driver of the autonomous vehicle, the individual that is being transported by the autonomous vehicle, can be assumed to be liable for the mere fact of owning and/or operating it without driving.[38] This solution is based on the fact that Alice is not in control of the decisions taken by the autonomous vehicle, but benefits from using it. This solution is currently already in use in those jurisdictions that allow for autonomous and semi-autonomous vehicles to circulate on the road. Usually, the driver is also obliged to have an additional insurance, or a traditional civil responsibility insurance that also covers damages deriving from autonomous vehicles.[39]

Other parties might also be involved. The accident might be due to a defect in one of the car's components (e.g., the sensors failed and did not inform the system driving the car of the presence of an obstacle), or to an error of the car's software (e.g., the sensors detected the obstacle, but the AI software did not understand the data correctly). If, as it is currently prospected, autonomous vehicles in the future will circulate in a smart infrastructure, communicating at all times with the road, other autonomous vehicles, and traffic lights, the accidents might also result from an error or defect of the infrastructure. In case of a system failure of the autonomous vehicle or of the infrastructure, the car manufacturer, as well as the entities in charge of the infrastructure or its components might also be liable, either alone or with the driver. It is important to consider that with AI systems it can be difficult to identify defects, malfunctioning, or errors. These systems are complex and opaque, and often even the manufacturers and programmers cannot be certain of the reasons why a particular decision was taken. The following scenario shows how AI systems challenge the way in which errors or defects are defined.

Scenario C: Alice & the smart home
Many AI systems are marketed to consumers, that is, their target buyers are individuals that will use them for personal use and not commercial or work-related purposes.[40] These include an array of IoT devices (smart speakers, fridges, coffee makers, mattresses, light bulbs, children's toys, even toothbrushes and sex toys). If traditional, 'dumb' products are prejudicial

[36] Royakkers and van Est 2016, p 185

[37] Marchant and Lindor 2012, p 1326

[38] European Commission 2019, p 39

[39] Hevelke and Nida-Rümelin 2015; Marchant and Lindor 2012, p 1339

[40] Weatherill 2013

to a consumer or his/her patrimony, the European regime for consumer protection applies. Under this regime, the producers and resellers of a product are liable for the damages. The burden of proof is simplified for the damaged consumer, who must prove the damage, the defect, and the causal link between the two. It is up to the producer then to prove the existence of circumstances that exclude the liability (for instance that the defect came to be after it was put into circulation, or that the state of the art did not allow the producer to identify or foresee the defect)[41], and a favourable prescription term applies. This regime reinforces the protection of the consumer on the basis that he is in an unfavourable position vis-à-vis the manufacturers and resellers, due to the asymmetry of information on the products, their design, features, and manufacturing processes.[42]

What happens if the consumer protection regime is applied to an AI system? Assume that, as shown in Figure X.4, individual consumer Alice purchases an AI-powered coffee maker. The coffee maker uses sensors and data collected to deduce the presence of a mug, analyse Alice's habits and preferences, and have her favourite coffee ready in time for breakfast. One morning, however, the coffee maker's algorithm pours the boiling coffee on the kitchen countertop while Alice is sleeping, severely damaging it. There appear to be no broken parts, no faulty components: the algorithm has made an evaluation based on the data collected and has established that it was the time to brew and pour the coffee.

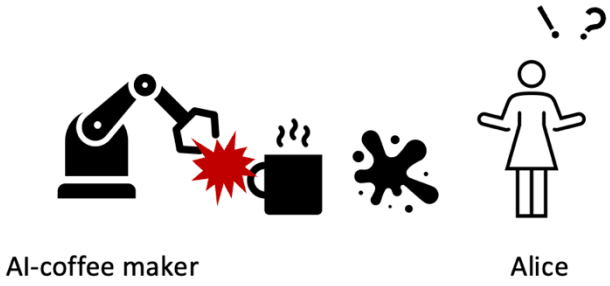


Figure X.4: Alice's coffee maker pours coffee on the countertop based on its own elaboration of Alice's habits.

The application of the consumer protection regime to this scenario is not without doubts.[43] Two elements are particularly challenged by AI: the definition of product, and what constitutes a defect.

It is not clear whether the existing European provisions concerning consumer products include software.[44] For the purposes of product safety, for instance, software is only considered when they represent an integrated and necessary component of hardware, while stand-alone software is not regulated.[45] Furthermore, many digital goods offered via the Internet can be categorized as services, depending on national provisions regulating

[41] 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], article 7

[42] Weatherill 2013

[43] Bertolini 2013; Asaro 2011

[44] 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], article 2

[45] European Commission (2020), p 10

contracts.[46] The applicability of the product liability regime to AI is rendered uncertain by the fact that software is not univocally considered a product.[47]
In addition, product liability applies when a product is defective either by design or due to errors in the assemblage and manufacturing processes, or because of transportation.[48] As the scenario shows, however, AI systems might result in damages even when there is no defect.[49] The algorithm might learn something wrong, or even something in itself not wrong, but something that does not apply to the context in which the AI system operates. Algorithms operate opaquely through correlation, not causation, and therefore might identify patterns in the data that do not match with how humans interpret reality. Whether or not this behaviour of AI systems is a defect is not yet established by the law. As a result, applying product liability to AI systems might still result in liability gaps.

X.3.3 Beyond liability: other legal tools

Contractual or extracontractual liability have not been the only legal tools proposed to mitigate the uncertainty in the allocation of liability or to redistribute the economic damages, with regard to AI. The most notable other solutions proposed are mandatory insurance schemes, using limited liability companies or other company types, and creating an *ad hoc* type of legal personality for AI systems and/or robots.

Regardless of the allocation of liability, insurance schemes are already in use to make sure that individuals carrying out certain activities can keep possible victims indemnified in case of damages, as is the case for cars. It has, therefore, been proposed that owners or operators of AI systems could be obliged to take up insurance, especially for certain sectors or activities considered at high risk.[50] While, however, this could help maintain victims indemnified, insurance schemes do not substitute a liability regime. They can only complement it. Even in the presence of insurance it would still be necessary to identify one or more liable parties.[51]

The idea has also been considered of incorporating a limited liability company (or another form of company) 'around' the AI, having the AI system as its shareholder, so that damaged parties could avail themselves of the patrimony of the company to compensate damages.[52] While it is believed that this solution could be already feasible in some legal systems, it is difficult to see how it can bridge the liability gaps concretely. The company would need to acquire a patrimony: this can only be done if shareholders with legal personality (natural or legal persons) contribute it. The company would still necessitate human intervention, and its incorporation would only add intermediate passages to the possibility of damaged parties to obtain redress.[53]

---

[46] Jaquemin 2017

[47] European Commission (2020), p 4

[48] Weatherill 2013; see also the definition of defect provided by article 6 of the Council Directive 85/374/EEC

[49] European Commission 2020a

[50] Turner 2019, p 112

[51] European Commission 2019, p 61

[52] Bayern 2015

[53] Scherer 2019

Finally, a daring solution has recently gained some momentum: creating a special form of legal personality for AI systems and robots. The idea was already put forward in 1992[54] and has now been revived, becoming the centre of a heated debate. The European Parliament seemed to have embraced it in 2017 when, under the term Electronic Personality, it was proposed as a possible option to be considered by the European Commission to regulate liability issues connected to robotics and AI.[55] The underlying concept is that just like companies have a form of legal personality, different from that recognized to natural persons (that is, human beings), so could artifacts, by virtue of their autonomy and potential societal relevance.[56] The idea is not *per se* impossible or wrong, as the law can identify which entities are entitled to rights and subject to duties.[57] It could provide for a solution to the liability gaps, provided that together with (some form of) legal personhood AI systems are also granted the agenthood necessary to acquire patrimony, enter into contracts, or stand in court. The recognition of legal personhood would open the way to the liability of AI systems, possibly together with other parties, such as their manufacturers or users (a solution necessary to incentivize the latter to ensure high quality and appropriate care in the creation and use of AI systems). To compensate the damaged parties for their losses, AI systems would also need to have a 'wallet', funds and assets. The proposal of the European Parliament has, in any case, been criticized in an open letter signed by over one hundred experts.[58] It has been considered to undermine the dignity of human beings, unethical, and not really based on reality, especially because completely autonomous and intelligent robots do not exist yet, as the ones currently on the market have very limited capabilities. The idea of granting legal personhood to AI systems or robots remains, for the moment, only hypothetical and the Parliament proposal has been disregarded by the Commission.[59]

X.4 Reflections: putting humans back at the centre

Any intervention to tackle liability gaps created by AI systems should take into consideration that the applications of AI are endless, and AI is used in almost every sector and industry. Consequently, one-size-fits-all solutions are unlikely to be satisfactory while the risk of fragmenting the protection of those who suffered damage is high. The transborder nature of many transactions involving AI systems (with multinational corporations offering products and services to individuals and companies located in several different countries) also implies that a harmonized and holistic approach – at least at a European level – could be necessary to prevent damaged parties from suffering further undesired consequences due to, for instance, forum shopping.

This does not mean that liability gaps for damages created by AI systems cannot be solved in a satisfactory fashion, providing effective redress to damaged parties. The phenomenon of

---

[54] Solum 1992

[55] European Parliament 2017, p

[56] Teubner 2006; van den Hoven van Genderen 2019

[57] Hage 2017; Teubner 2018

[58] Open Letter to the European Commission: Artificial Intelligence and Robotics (2017), https://www.politico.eu/wp-content/uploads/2018/04/RoboticsOpenLetter.pdf Accessed 19 February 2021.

[59] The European Commission, in the context of its AI strategy for Europe, has not embraced the view of the European Parliament concerning a special legal personality for artificial agents. Cfr. European Commission 2018 and European Commission's High-Level Expert Group on Artificial Intelligence 2019

liability gaps should be observed from the perspective of European policymaking, that at its core has the protection of the dignity and freedom of individuals, democracy, and equality.[60] Any legal solution to it should put humans at the centre, focusing on the damaged individuals, their rights and interests. To bridge the liability gaps, the starting point should be to acknowledge that the law operates now in a socio-technological ecosystem, in which individuals, companies, and institutions interact both online and offline with a plethora of devices and services that are digital, responsive, (semi)autonomous. A holistic, systemic approach is necessary to predict how the law interacts with this ecosystem, its actors and dynamics. The design of legal solutions to the liability gaps should make sure that "No one is left behind in the digital transformation".[61] To draw order from chaos, attention should be given to human-centred values,[62] to empower the individuals harmed by AI systems, compensating the effects of detrimental factors, such as the asymmetry of information and power, the lack of expertise, or the fragmentation generated by the presence of several parties.

In the light of the above, it is necessary to intervene on multiple fronts, combining existing solutions with special regimes created *ad hoc* for AI systems.[63]
This combined approach implies clarifying the scope and modalities of implementation of existing regimes, such as insurances, product liability or liability for driving accidents, with regard to AI systems. The modalities to apply contractual and extracontractual liability regimes to AI must also be clarified. These two forms of liability represent the baseline of protection, the safety net covering the spaces left empty in-between special regimes or application-specific regulations and can help mitigate the risk of fragmentation of the liability regime. Complementarily, it is also necessary to establish solutions tailored to the reality of AI. *Ad hoc* solutions can focus on specific uses or applications (for instance, regulating the liability for damages deriving from AI software operating in the stock market, or liability and chain of command in the case of autonomous weapons).
Underlying this combined intervention, there needs to be a strong set of values, to guarantee uniformity and maintain internal systemic coherence. Some of these values have been identified at a European level and include: distributing the loss connected to the damages in a fair and efficient way; granting access to justice to the damaged party; addressing in a similar way situations which present comparable risks; the duty of producers to ensure transparency of the technology and its functioning/use; introducing, besides insurances, compensation funds to support victims even in situations in which insurances are absent.[64]

At a practical level, some of these principles can be declined into more detailed guidelines. Allocating the loss in a fair and efficient way, for instance, includes identifying as liable all those parties that benefited from the AI system, were in control of the AI system, and/or operated the AI system.[65] This, for instance, mitigates the risk of a party invoking the

[60] These values are expressly protected by the main tools constituting and regulating the EU, such as the Treaty on European Union, and the Charter of Fundamental Rights.

[61] European Commission 2018, p 2

[62] European Commission 2020b

[63] European Commission 2019, p 32

[64] ibid.

[65] ibid., p 39

autonomy of the AI system to exclude liability, if that party benefited from the use of the AI in the first place. Granting access to justice to a damaged party translates into making sure that the burden of proof is allocated with the party that can actually comply with it (for instance, the company that owns or manufactured the AI system) and not automatically with the damaged party acting as plaintiff in a judicial procedure.[66] Similarly, prescription times are important to ensure the effective access to justice of a damaged party and need to take into consideration elements such as the changes occurring after an update or the self-learning capability of AI. Currently, applying contractual and extra-contractual liability might lead to solutions that do not reflect the above-mentioned principles. Implementing these principles in the context of both liability regimes might imply distancing from a rigid interpretation of the boundaries between contract and delicts. As an example, holding liable the many parties involved (producers, users that benefited from the AI system, operators) implies that less importance is attributed to the existence of contractual obligations among the parties involved. Normally, the lack of a contract would exclude the application of contractual liability. At the same time, moving the burden of proof away from the damaged party would not be always possible based on extra-contractual liability. A more human-centred solution to the liability gaps should contain a mix of contractual and extra-contractual elements, to compensate for the asymmetry of information and power created by the technical and business features of AI systems.[67]

Based on the nature and application of specific AI systems, it is also reasonable to intervene on the opacity of AI systems, establishing a duty of transparency for producers, for example through mandatory logs in which important passages and events are included, that can be revised in case of damages to assess what went wrong, when, where, and, possibly, why.[68]

X.5 Conclusions

In the past, legal solutions – contractual and extracontractual liability, special liability regimes, insurances, company law – have been developed having in mind 'dumb' machines and products. Artifacts that are non-autonomous, not interconnected, and not self-learning. The solutions analysed throughout this chapter revolve around more traditional relationships between two or more parties, that can be traced back to a contract or a delict, at least in most cases.[69]

The abovementioned solutions are challenged by AI's features: autonomy, self-learning capability, opaqueness, complexity, interoperability. In applying the existing legal tools to ascribe liability, such as contractual, extra-contractual, or strict liability, the inherent features of AI systems create uncertainty and unsatisfactory results (liability gaps). Elements such as the absence of a contractual relationship between the damage party and producers, providers, or users of an AI system, or the allocation of the burden of proof might lead to the undesired result of precluding an effective remedy to those damaged by an AI system.

Bridging the liability gaps requires the reintroduction and reaffirmation of those rationales and values that already informed the existing legal tools but were affected by the uncertainty

[66] European Commission 2020b

[67] De Conca 2020

[68] European Commission 2019, p 49

[69] The application of contractual and extracontractual liability in the past has been challenged by complex situations, such as the case of pre-contractual liability, or the duty of care vis-à-vis third parties. Cfr. for instance, Michoński 2015

deriving from the characteristics of AI. Any intervention to ascribe liability for damages caused by AI must be based on one, fundamental value: putting humans at the centre.

Reference List

Asaro P M (2011) A Body to Kick, but Still No Soul to Damn: Legal Perspectives on Robotics. In: Lin P et al, Robot Ethics: The Ethical and Social Implications of Robotics. MIT Press, pp 169-186

Bayern S (2015) The Implications of Modern Business-Entity Law for the Regulation of Autonomous Systems, Stan. Tech. L. Rev. 19:93-112

Bertolini A (2013) Robots as Products: The Case for a Realistic Analysis of Robotic Applications and Liability Rules, LIT 5(2):214-247

Buyuksagis E and van Boom W H (2013) Strict liability in contemporary European codification: Torn between objects, activities, and their risks, Georgetown Journal of International Law 44(2):609-640

Bussani M and Palmer V V (Eds.) (2003) Pure economic loss in Europe. Cambridge University Press

De Conca S (2020) Bridging the liability gap using human-centered legal design: three scenarios to apply the liability from social contact, WeRobot 2020, Ottawa 22-25 September 2020. https://techlaw.uottawa.ca/werobot/papers Accessed on 20 July 2021

European Commission (2018) Artificial Intelligence for Europe: 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 [COM(2018) 237]

European Commission (2019) Liability for Artificial Intelligence and other emerging digital technologies – Report from the Expert Group on Liability and New Technologies

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

European Commission (2020)b White Paper on Artificial Intelligence - A European approach to excellence and trust [COM(2020) 65]

European Commission's High-Level Expert Group on Artificial Intelligence (2019) Ethics Guidelines for Trustworthy AI

European Parliament (2017) Draft Report of the Committee on Legal Affairs with recommendations to the Commission on Civil Law Rules on Robotics [2015/2103(INL)]

Farnsworth E A (2006) Comparative Contract Law. In: The Oxford Handbook of Comparative Law. Oxford University Press

Fauvarque-Cosson B and Mazeaud D (eds) (2009) European Contract Law, Materials for a Common Frame of Reference: Terminology, Guiding Principles, Model Rules. Sellier European Law Publishers

Hage J (2017) Theoretical foundations for the responsibility of autonomous agents, Artif Intell Law 25:255-271

Hevelke A and Nida-Rümelin J (2015) Responsibility for Crashes of Autonomous Vehicles: An Ethical Analysis, Sci Eng Ethics 21:619–630

Jacquemin H (2017) Digital content and sales or service contracts under eu law and belgian/french law, Journal of Intellectual Property, Information Technology and Electronic Commerce Law 8(1):27-38

Johnson D G (2015) Technology with No Human Responsibility?, Journal of Business Ethics 127(4):707-715

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

Leenes R, Palmerini E, Koops B J, Bertolini A, Salvini P, Lucivero F (2017) Regulatory challenges of robotics: some guidelines for addressing legal and ethical issues, Law, Innovation and Technology, 9:1, pp 1-44

Marchant G E and Lindor R A (2012) The coming collision between autonomous vehicles and the liability system, Santa Clara Law Review 52(4):1321-1340.

Marsh P D V (1994) Comparative Contract Law: England, France, Germany. Gower

Michoński D (2015) Contractual or Delictual? On the Character of Pre-contractual Liability in Selected European Legal Systems, Comparative Law Review 20:151-175

Pagallo U (2012) Three Roads to Complexity, AI and the Law of Robots: On Crimes, Contracts, and Torts. In Palmirani M et al, AICOL Workshops 2011, LNAI 7639. Springer, pp 48-60

Pagallo U (2013)a The Laws of Robots: Crimes, Contracts, and Torts. Springer

Pagallo U (2013)b Robots in the cloud with privacy: A new threat to data protection?, CLSR 29(5):501-508

Pagallo U, Corrales M, Fenwick M, Forgó N (2018) The Rise of Robotics & AI: Technological Advances & Normative Dilemmas. In: Corrales M et al (eds), Robotics, AI and the Future of Law. Springer, pp 1-14

Royakkers L and van Est R (2016) Just Ordinary Robots: Automation from Love to War. CRC Press

Scherer M (2016) Regulating Artificial Intelligence Systems: Risks, Challenges, Competence, and Strategies, Harvard Journal of Law & Technology 29(2):354-400

Scherer M (2019) Of Wild Beasts and Digital Analogues: The Legal Status of Autonomous Systems, NEV. L.J. 19:259-291

Solum L B (1992) Legal Personhood for Artificial Intelligences, N.C. L. Rev. 70(4)::1231-1287

Teubner G (2006) Rights of Non-humans? Electronic Agents and Animals as New Actors in Politics and Law, Journal of Law and Society 33(4):497-521

Teubner G (2018) Digital Personhood? The Status of Autonomous Software Agents in Private Law, Ancilla Iuris 106-149

Turner J (2019) Robot Rules: Regulating Artificial Intelligence. Palgrave Macmillan

UNESCO (2017) Report of the World Commission on the Ethics of Scientific Knowledge and Technology (COMEST) on Robotics Ethics [SHS/YES/COMEST-10/17/2 REV]

van den Hoven van Genderen R. (2019) Does Future Society Need Legal Personhood for Robots and AI?. In: Ranschaert E. et al (eds), Artificial Intelligence in Medical Imaging. Springer, pp 257-290

von Bar C and Drobnig U (eds) (2009) The Interaction of Contract Law and Tort and Property Law in Europe, A Comparative Study. Sellier European Law Publishers

Weatherill S (2013) EU Consumer Law and Policy. Edward Elgar Law

Witt J (2001) Toward a New History of American Accident Law: Classical Tort Law and the Cooperative First-Party Insurance Movement, Harvard Law Review 114(3):690-841