Centre for International
Governance Innovation

Policy Brief No. 177 — July 2023

Addressing the Liability Gap in AI Accidents

Amrita Vasudevan

Key Points

- Traditional tort liability frameworks are unsuitable in addressing accidents involving artificial intelligence (AI) systems due to the unpredictability of AI outcomes and the opacity of AI systems, resulting in a liability gap.
- The European Commission has proposed changes to address the challenges, including measures to alleviate the evidentiary burden on claimants by decreasing the burden of proof and issuing disclosure orders.
- In cases involving high-risk AI systems, strict liability and no-fault liability insurance should be utilized. The presence of insurance covers, however, may encourage actors to increase their exposure to risk.
- The "problem of many hands" has resulted in the diffusion of responsibility among various actors and automated processes in the deployment of AI.

Introduction

In 2016, in a mall in Silicon Valley, a security robot called Knightscope K5 accidentally knocked over a toddler during its patrol. Thankfully, the toddler did not suffer any major injuries (Hrala 2016). Unfortunately, this has not always been the outcome of interactions between AI systems and humans.

Just a year earlier in India, a worker in an auto parts factory was fatally injured by the robotic arm of a machine when he mistakenly walked into the radius of one of the machine's sensors. The workers' union asserted that this was a case of negligence by the factory (Sahni and Jha 2015). Over the past few years, there have been multiple reports of cars equipped with automated driver-assist features being involved in accidents (Graziosi 2022). In 2018, one of Uber's self-driving cars struck and killed a pedestrian, Elaine Herzberg, in Tempe, Arizona. Although the car was able to detect Herzberg five seconds before the crash, the braking system failed to engage because it misclassified her as an "other object" rather than a person (Hawkins 2019).

From driving, banking and doing household chores, to medical diagnoses and job recruitment, AI systems are already part of numerous and varied social interactions and activities. As these complex computational systems become embedded in society, accidents that involve AI systems will become increasingly common. However, the

About the Author

Amrita Vasudevan is a CIGI fellow and an independent researcher based in Vancouver, British Columbia. Her research focuses on the political economy of regulating digital technologies and investigating the impact of these technologies through a feminist lens.

Her work falls at the intersection of law, technology and society with a focus on the Global South. She has previously worked for IT for Change and the Digital Futures Lab, both research and advocacy organizations in India. She has written on issues of gender-based cyber violence, platform governance, regulation of artificial intelligence and data governance and trade. She has also worked as a legislative assistant to a member of Parliament in India.

Amrita holds a master of law from Yale Law School, which she attended on a Fulbright-Nehru Master's Fellowship.

unique characteristics of AI systems, particularly their unpredictability and inscrutability, can pose fundamental challenges in meeting the liability standards laid down in the law (Selbst 2020).

There is also the question of who is at fault when AI accidents occur: Is it the user/operator, the manufacturer or the coders who designed the AI algorithms? To determine liability and award compensation to victims, untangling the different intersecting agents involved in the development and deployment of AI systems may become necessary. Alternatively, if unravelling the network of actors and processes becomes unreasonably cumbersome or impossible, accountability may need to be distributed across multiple actors or all actors in the accident network.

This policy brief unpacks the difficulties and inconsistencies that arise from applying traditional liability frameworks to AI accidents. The brief critically analyzes measures that are being suggested by regulators and academics to plug the liability gap that is emerging. The European Union, for example, has been at the forefront of legal innovation aimed at addressing the inconsistencies in the law that digital technologies, particularly AI, have provoked (Edwards 2022). The European Commission recently released two proposals on redressing the harms caused by AI, including a provision on compensating the victims of AI accidents (Dheu, De Bruyne and Ducuing 2022). The proposals deal with non-contractual civil liability, known as "tort law."

The focus of this brief will also be on tort law as it forms the backbone of general accident claims. Tort is a civil wrong that causes injury or harm, whether physical, financial, reputational or emotional, and gives rise to legal liability. Broadly, tort law allows for financial compensation for harm caused by intentional conduct or the failure to meet a duty of care. The duty of care may relate to actions or omissions and is benchmarked against "someone of ordinary prudence."[1]

If an AI system is deliberately used to harm someone (for example, deepfakes that defame an individual), then intentional torts will apply (Thomasen 2021). However, in most cases of AI accidents, such as those mentioned above, the tort of negligence would be invoked. Broadly, the

1 See www.law.cornell.edu/wex/tort.

main elements of the tort of negligence include the existence of a duty owed by the defendant to the plaintiff, breach of that duty by the defendant, injury suffered by the plaintiff, and causation or proof that the defendant's breach caused the injury. In determining whether the duty to act exists, the court will take into consideration if the defendant created a risk that resulted in harm to the plaintiff, knew that their actions would cause harm, volunteered to protect the plaintiff, or entered into a business or other voluntary relationship with the plaintiff.[2]

Critically, in proving causation, the plaintiff needs to establish that the harm caused was a foreseeable consequence of the defendants' conduct. As discussed below, AI systems have raised fundamental doubts as to the viability of this test.

Establishing Liability under Tort Law for AI Injuries

Andrew D. Selbst, a professor of law at UCLA, observes that the most common current use of AI is in "decision assistance" to humans, rather than in fully autonomous robots. When used in medical diagnosis, for instance, AI analyzes large data sets to identify patterns too complex for humans to perceive. Based on its analysis, AI can make predictions and recommendations. Here, the AI system is akin to a tool used by a doctor. To avoid a claim of negligence in case of an injury, medical professionals are expected to understand the capabilities and limitations of these tools and to exercise a standard of care (Selbst 2020). The duty of care, however, is not infinite and is limited by the foreseeability of harm and the feasibility of taking reasonable precautionary measures (Hubbard 2014).

Herein lies the contradiction. The use of AI tools can constrain human intuition and experiential learning by replacing or augmenting "human decision processes with inscrutable, unintuitive, statistically derived, and often secret code" (Selbst 2020, 1321). While this feature can increase certainty and therefore enhance safety, it also obscures the foreseeability of harm. For instance, a doctor may not be able to detect when a misdiagnosis occurs due to an error in an AI tool they are using.

Further, AI systems are often described as black boxes due to the complexity of their internal workings and the inscrutability of their decision-making processes (Lior 2021b). It is sometimes difficult to decipher how input variables are related to each other to reach a final prediction (Rudin and Radin 2019). Corporations that own and deploy AI often double down on their impenetrability by closely guarding the software and the data that are used to train these systems (Burrell 2016). The intractability of AI systems limits foreseeability, which makes liability difficult to establish (Selbst 2020).

In light of these challenges, EU member states are considering revisions to liability standards.

The European Union's Proposed AI Liability Directive

The European Commission's new proposal for non-contractual civil liability (the AI Liability Directive) introduces a rebuttable presumption of causality in the case of injuries caused by AI systems.[3] The directive draws a limited presumption "between the breach of a duty of care by the defendant and the AI system's output" (Wagner 2023, 40).[4] The claimants or court will still have to establish, first, that the defendant did not meet a duty of care as defined by national or EU law and, second, that it is "reasonably likely" that the failure to meet the prescribed duty of care caused damage. Third, a claimant will need to establish that the AI system's act or omission caused injury (ibid.).

2 See www.law.cornell.edu/wex/negligence.

3 *Proposal for a Directive of the European Parliament and of the Council on adapting non-contractual civil liability rules to artificial intelligence (AI Liability Directive)*, 2022/0303 (COD), COM(2022) 496 final, online: <https://commission.europa.eu/system/files/2022-09/1_1_197605_prop_dir_ai_en.pdf>.

4 *Ibid*, art 4(1): "Rebuttable presumption of a causal link in the case of fault: Subject to the requirements laid down in this Article, national courts shall presume, for the purposes of applying liability rules to a claim for damages, the causal link between the fault of the defendant and the output produced by the AI system or the failure of the AI system to produce an output."

But there are some gaps in these proposed measures. For example, many of the terms used in the directive, such as "reasonably likely" and the "duty of care," are unclear or vaguely defined and have raised questions as to how they will intersect with traditional notions of fault. It will be left to national courts to define these terms, but doing so could adversely affect legal certainty and lead to fragmentation across different nations (Dheu, De Bruyne and Ducuing 2022).

The directive also aims to give national courts the power to order the disclosure of evidence pertaining to high-risk AI systems by the defendant. This will help claimants access information from the defendant to build an initial case and therefore make a more informed decision about whether to initiate legal proceedings. If the defendant fails to disclose court-ordered evidence, it will trigger a rebuttable presumption of the relevant duty of care (Hacker 2023).

Prior to a disclosure, the claimant will need to produce evidence demonstrating the plausibility of a damages claim. This, it is argued, is a fairly high threshold for claimants to meet and may effectively prevent disclosures by the defendant (ibid.). Interpretation and analysis of such evidence by the plaintiff is a complex and costly task and thus could be a prohibitive endeavour (Dheu, De Bruyne and Ducuing 2022).

Application of Product Liability Framework to AI Accidents

In addition to the tort of negligence, many jurisdictions have product liability torts to protect consumers from injuries caused by defective products. The defect may relate to the design, the lack of warning or instruction, or manufacturing, and may fall under negligence or strict liability. In the latter case, liability exists regardless of intent or knowledge of harm.[5]

In Canada, a negligence framework is applied to any kind of product liability, which means that the claimant will have to prove that the defendant owed a duty of care that was not met (Thomasen 2021). As in the case of negligence litigation for AI accidents, similar problems with establishing reasonable foreseeability will arise here.

In the United States, manufacturing defects will invoke strict liability for the manufacturer. Usually, on one hand, a comparison is made with the manufacturer's own blueprint for the product to check if there are any inconsistencies with the final product. For design and warning defects, on the other hand, liability will be established through a cost-benefit analysis. In case of design defects, either the "risk-utility test" or the "consumer expectation test" is applied. Product liability torts in the European Union have also adopted these tests (Riordan 2003). The risk-utility test holds defendants liable for a defect if there exists a reasonable alternate design, and the consumer expectation test investigates whether the danger posed by the defect is beyond the contemplation of the user. In essence, these tests seek to balance the benefits of safety against the cost of finding every imperfection (Selbst 2020).

As noted earlier, AI decision-making processes are often obscure and occluded, which makes it difficult to classify whether a defect was due to the manufacturing or design. Further, Selbst notes that in the case of design defects, the claimant must show that the injury was caused by an AI decision that should have been tested and avoided. Breaking into the black box of AI systems to tease out these threads will be near impossible for plaintiffs. It may also be deemed unreasonable under tort law to expect manufacturers and programmers to test for every fact pattern and eventuality that an AI system could reach (ibid.).

The European Union's Proposed Product Liability Directive

The European Union has proposed amendments to its product liability regime (the 1985 Product Liability Directive) to specifically include intangible software and digital manufacturing files, which

5 See www.law.cornell.edu/wex/product_liability.

include AI products, and address some of these challenges. Article 9(4) of the directive notes that if a national court feels that a "claimant faces excessive difficulties, due to technical or scientific complexity" of AI systems to establish defectiveness or the causal link between defectiveness and injury, then both defectiveness and causality may be presumed.[6] This is a rebuttable presumption.[7]

Like the AI Liability Directive, the revised Product Liability Directive also eases the evidentiary burden on the plaintiff and reduces the information asymmetry between the plaintiff and the defendant by empowering national courts to issue disclosure orders. If the defendant does not comply with the order, it will trigger a presumption of defectiveness (Selbst 2020).

Product liability torts usually evaluate the defectiveness of a product based on the standards that exist at the time when the product enters the market. With certain kinds of AI, such as those that are capable of learning and evolving after their introduction to the market, such a static frame is not useful (Wendehorst 2020). Reinforcement learning, for example, is a kind of machine learning that continues to learn through a trial-and-error basis even in its final operating environment (Matthias 2004).

The revised Product Liability Directive has responded to the dynamic nature of AI products by acknowledging that sometimes producers continue to exert control over the product even after the product has entered the market. It is common practice to issue software updates for a product, for example, to fix a safety feature. Here, defectiveness will be judged by standards present at the time when the update was made. Because the emphasis is on "control," the directive will likely exclude software updates provided by a third party, which are not authorized by the producer (European Law Institute 2022). Further, the manufacturer would also be responsible to provide updates for the safety of the product (Buiten, de Streel and Peitz 2023).

The Product Liability Directive will also hold developers responsible for the harm caused by an AI system, which continues to learn after being placed in the market. For example, a self-driving car may learn bad driving habits from other human drivers while on the road. Thus, developers will need to ensure that the adaptive driving features of the car do not result in it breaching the risk-tolerance standard that has been set (Hacker 2023).

Overall, the European Union has taken important steps toward addressing the gap in liability that AI systems have made conspicuous. By making disclosure easier, it is possible to peek under the hood of AI systems and offer explanations for the system's decisions.

Explainable AI could possibly be used to establish a chain of events, to evaluate the AI's propensity for accidents through empirical evidence and to establish reasonable alternatives in design and development through counterfactual evidence (Fraser, Simcock and Snoswell 2022). When an AI system's process and/or output is explainable, then foreseeability and defectiveness can be contested when litigating negligence or product liability.

However, this is still an emerging area of research, and academics have cautioned that explainable AI has its limitations (Lakkaraju 2021). Moreover, the mere availability of data is insufficient as defendants may overload the victim with information, leaving the latter to sift through the noise and make sense of the evidence (Dheu, De Bruyne and Ducuing 2022; Poursabzi-Sangdeh et al. 2021). Therefore, it is vital that the data provided should be presented in a manner that is easy to understand and interpretable by the plaintiff (Ribeiro, Singh and Guestrin 2016).

Application of Strict Liability and No-Fault Liability to AI Accidents

The challenges posed by AI systems to tort liability may lead to situations in which the victims of AI accidents are left without any compensation. The costs of litigation, particularly obtaining the expertise needed to understand AI systems and establish fault, may discourage victims from initiating court cases. Moreover, AI accidents can be unpredictable, occurring despite adhering to reasonable safety standards

6 *Proposal for a Directive of the European Parliament and of the Council on liability for defective products*, 2022/0302 (COD), COM(2022) 495 final, online: <https://single-market-economy.ec.europa.eu/system/files/2022-09/COM_2022_495_1_EN_ACT_part1_v6.pdf>.

7 *Ibid.*

and no fault being apparent (Hubbard 2014). This has resulted in a call for the application of a strict liability regime based on causation rather than establishing fault, defect, malperformance or non-compliance (Wendehorst 2020).

Some commentators have also suggested the adoption of a "no-fault" social insurance scheme to cover all personal injuries arising from AI accidents, which would side-step the tort regime altogether. The insurance scheme, as proposed, could be financed from general tax revenue (Yoshikawa 2020). The main attraction of such an insurance scheme is that it removes the uncertainty of legal proceedings. This, in turn, can incentivize unencumbered innovation while also ensuring compensation without meeting the high threshold of proof required by tort law (Marchisio 2021).

A critical weakness of an all-encompassing, no-fault regime is that it reduces the motivation for designers and manufacturers to internalize the costs of innovation through the adoption of safety measures. To encourage AI developers and users to adopt safety processes and standards, a cap could be placed on how much compensation can be paid. Additionally, premiums could be adjusted depending on the risk AI systems pose (Lior 2020).

Another concern with lowering the threshold of liability is that it could result in multiple lawsuits against producers and users of AI systems. Some academics have noted that small players in the market may not be able to shoulder such high litigation costs, and this could lead to the consolidation of the market around a handful of bigger players (Thomasen 2021). To avoid such an outcome, it is argued that there should not be a blanket application of strict liability and no-fault liability across all AI use cases. Instead, their application should be limited to high-risk AI use cases only (Wendehorst 2020).

However, even without low liability thresholds, the AI market is already moving toward monopolization. Meredith Whittaker, co-founder of the AI Now Institute and president of the Signal Foundation, notes that big tech controls "the tooling, development environments, languages, and software that define the AI research process — they make the water in which AI swims" (Whittaker 2021, 53). Therefore, concerns over the monopolization of AI markets should be separately addressed by policy makers. A good example of this is the European Union's Digital Markets Act, which has sought to keep big tech companies in check by preventing them from abusing their market power (Vallance 2022).

Belling the Cat: Who Should Be Held Responsible for AI Accidents?

Deciding what kind of liability should apply is only a partial inquiry into addressing accountability. It is also important to resolve who should be held liable when an AI accident occurs. The "problem of many hands" can make the allocation of responsibility for the outputs produced by AI systems challenging. The problem refers to the diffusion of responsibility among multiple actors (hardware manufacturers, software developers, data trainers, users and so forth) and automated processes that are involved in the development and deployment of AI systems (Buiten, de Streel and Peitz 2023). Additionally, as Filippo Santoni de Sio and Giulio Mecacci (2021, 1063) argue, the opacity and lack of explainability of the outputs produced by these systems can make it "more difficult for individual persons to satisfy the traditional conditions for moral and legal culpability: intention, foreseeability, and control."

Attempting to force fit AI accidents into the traditional mould for liability can lead to unfortunate results. Take the 2018 accident referred to earlier involving one of Uber's self-driving cars: the backup driver was eventually charged in the fatal crash. However, the National Transportation Safety Board also found that Uber lacked adequate safety assessment processes and, more generally, a culture of safety (Lee 2019). Therefore, the attribution of all fault to the last human actor in the chain of command is inappropriate and can result in scapegoating (Vasudevan, Aneja and Chamuah 2021).

Tesla has previously defended itself when one of its self-driving cars was involved in a fatal crash by arguing that the negligence of the driver was a more immediate reason for the crash than the actions of the programmer (Kowert 2017). Frequently, humans in the loop become the "moral crumple zone," meaning they are blamed even though automated

processes may have influenced their final judgment or other actors may have contributed to the damage that was caused (Elish 2019). Selbst, for instance, notes that the hand-off model in autonomous cars, where the car transfers control to the safety driver during emergencies, is dangerous because humans struggle to re-engage when they are not continuously monitoring a situation (Selbst 2020).

In a blog post titled "Who Pays for AI Injury?", Mihailis Diamantis argues that analogous to liability for an employee's actions, corporations that design and run AI systems should be held liable for injuries caused by AI since they control and benefit from these algorithms. But due to the many hands problem, courts find it difficult to pin liability, allowing corporate actors to escape charges (Diamantis 2020). Diamantis draws attention to a 2015 case in which Wanda Holbrook, a maintenance technician, was crushed by a robot used in an auto parts factory after it breached safety protocols. Her widower found it extremely difficult to find the entity that was responsible for the accident. In the end, he sued five corporations, claiming that "each had a hand in installing, integrating, engineering, servicing, controlling, and/or manufacturing the robot and/or its safety devices" (Diamantis 2023). Unfortunately, he was not able to even make an initial case against them, let alone pierce the corporate veil to hold a particular individual liable. In the end, the court dismissed his suit against four of the companies, and the case against the fifth defendant is also likely to be unsuccessful (ibid., 809).

In 2017, the European Parliament considered the idea of establishing a compensation fund for AI accidents that would either apply to all AI systems in general or to specific robotic categories. The fund, it was suggested, would be financed by different entities involved in the production of an AI tool, thus eliminating the difficult task of finding the entity responsible after an AI accident occurs (Antunes 2021). Further, the liability of the manufacturer, programmer, owner or the user could be limited if they contributed to the fund, or if they jointly purchased an insurance policy. Such a trade-off essentially provides a safe harbour from liability and could disincentivize manufacturers from investing in the safety of their products (Lior 2022). Just like the no-fault social insurance scheme, while the fund may be useful, it should not be a stand-alone solution.

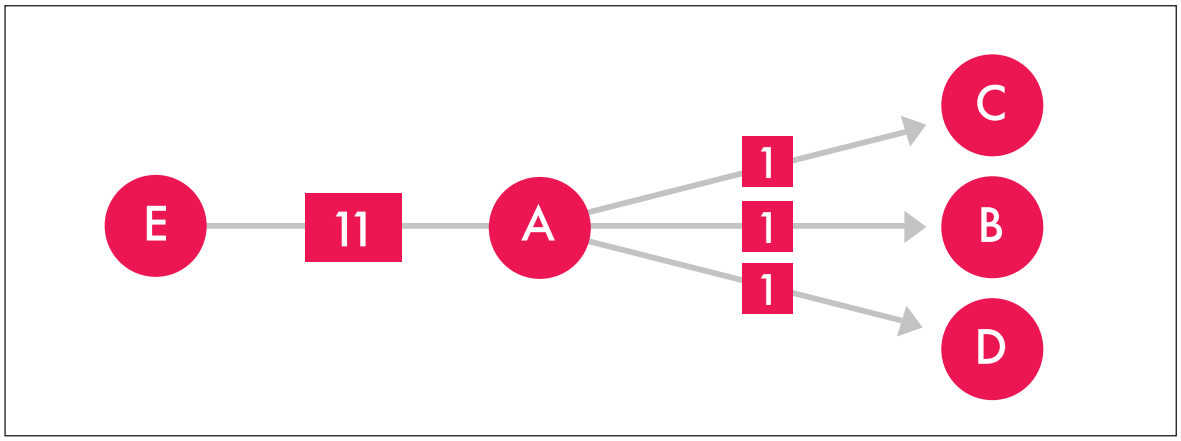
Anat Lior, who has written widely on torts and AI regulation, suggests that network theory can be used to map out and distribute liability among the different actors involved in an AI accident. The actors involved in the accident, whether human (such as producers, designers and manufacturers) or non-human (such as robots and algorithms), are represented as nodes in the network, while their relationships are depicted as edges (see Figure 1). Visualizing the network of nodes and edges can reveal the nature and strength of the relationship between different actors, exposing the humans responsible for the AI actions (Lior 2021a). Lior writes, "The true value of utilizing network theory in the AI liability context is its ability to unmask those neighboring human nodes… and thus make sure they are held liable for their actions" (ibid., 1149). Going through this exercise can reveal which human nodes exert control over the AI system and the extent of their control. This, in turn, can be useful in assigning liability.

Thus, while users have a critical role to play in the operation of AI systems, concentrating on their liability alone does not comport with how these systems function and will ultimately fall short of meaningful accountability. Liability frameworks should reflect this understanding.

Conclusion

AI accidents can no longer be dismissed as isolated incidents that can be overlooked in favour of the benefits AI systems offer. Turning a blind eye to the harms of AI will only mean burdening victims of AI accidents with the costs of innovation. Instead, we need to push for the responsible development and deployment of AI. This will necessarily entail adapting the current legal framework to respond to the uncertainty and opacity of AI systems to deliver accountability and justice to victims of AI accidents. A critical component of the legal framework that will need to be addressed is tort law. As previously discussed, the European Union has put forward substantial measures for adapting tort liability frameworks to AI accidents by reducing the burden of proof on the plaintiff and easing their access to evidence. Because actors will try to minimize liability to avoid paying compensation, tort law can be used to nudge actors to develop safer products (Hubbard 2014).

Figure 1: Accident Network Involving an Autonomous Car



Source: Lior (2021a).

Note: An accident network involving an autonomous car and bystanders with an AI tortfeasor (A), three human victims (B, C and D) and a human principal (E). On one hand, the edge in this figure is characterized by a high weight (11) because E and A share an agency relationship that includes control, guidance and monitoring. On the other hand, the edge between A and B, C and D has a low weight since the accident is a random event.

In devising accountability frameworks, we must work with the understanding that AI consists of networks of actors and automated processes. Otherwise, we risk fixating on individual accountability of the user or operator. As in the case of the crash caused by Uber's self-driving car, this approach could result in inappropriately placing all blame on the driver of the car, thereby making them a scapegoat. Individualized forms of responsibility need to make way for more collective forms of responsibility.

Legal innovation around networked forms of liability is thus crucial. Here, accountability is either shared among several actors within a network, or all actors in the network take on some level of accountability. This approach centres the victim's needs and ensures that they are fairly compensated for harms caused by AI accidents even when it becomes difficult to identify or determine which actor(s) should be held responsible. The primary focus here is not diagnosing causality but redressing harm.

Works Cited

Antunes, Henrique Sousa. 2021. "Civil Liability Applicable to Artificial Intelligence: A Preliminary Critique of the European Parliament Resolution of 2020." *Portuguese Law Review* 5 (1): 1–23. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3743242.

Buiten, Miriam, Alexandre de Streel and Martin Peitz. 2023. "The law and economics of AI liability." *Computer Law & Security Review* 48: 105794. https://doi.org/10.1016/j.clsr.2023.105794.

Burrell, Jenna. 2016. "How the machine 'thinks': Understanding opacity in machine learning algorithms." *Big Data & Society* 3 (1). https://doi.org/10.1177/2053951715622512.

Dheu, Orian, Jan De Bruyne and Charlotte Ducuing. 2022. "The European Commission's Approach To Extra-Contractual Liability and AI — A First Analysis and Evaluation of the Two Proposals." CiTiP Working Paper Series. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4239792.

Diamantis, Mihailis E. 2020. "Who Pays for AI Injury?" *Oxford Business Law Blog*, May 4. https://blogs.law.ox.ac.uk/business-law-blog/blog/2020/05/who-pays-ai-injury.

———. 2023. "Employed Algorithms: A Labor Model of Corporate Liability for AI." *Duke Law Journal* 72 (4): 798–859. https://scholarship.law.duke.edu/dlj/vol72/iss4/2.

Edwards, Lilian. 2022. "The EU AI Act: a summary of its significance and scope." Ada Lovelace Institute. April. www.adalovelaceinstitute.org/wp-content/uploads/2022/04/Expert-explainer-The-EU-AI-Act-11-April-2022.pdf.

Elish, Madeleine Clare. 2019. "Moral Crumple Zones: Cautionary Tales in Human-Robot Interaction." *Engaging Science, Technology, and Society* 5: 40–60. https://estsjournal.org/index.php/ests/article/view/260.

European Law Institute. 2022. *ELI Draft of a Revised Product Liability Directive*. Vienna, Austria: European Law Institute. www.europeanlawinstitute.eu/fileadmin/user_upload/p_eli/Publications/ELI_Draft_of_a_Revised_Product_Liability_Directive.pdf.

Fraser, Henry, Rhyle Simcock and Aaron J. Snoswell. 2022. "AI Opacity and Explainability in Tort Litigation." *2022 ACM Conference on Fairness, Accountability, and Transparency* 185–96. https://doi.org/10.1145/3531146.3533084.

Grazisosi, Graig. "Self-driving car systems were involved in 400 crashes since 2021 — 273 were Teslas." *The Independent*, June 15. www.independent.co.uk/news/world/americas/self-driving-car-crashes-tesla-b2101763.html.

Hacker, Philipp. 2023. "The European AI Liability Directives: Critique of a Half-Hearted Approach and Lessons for the Future." Working Paper. January 20. https://arxiv.org/pdf/2211.13960.pdf.

Hawkins, Andrew J. 2019. "Serious safety lapses led to Uber's fatal self-driving crash, new documents suggest." *The Verge*, November 6. www.theverge.com/2019/11/6/20951385/uber-self-driving-crash-death-reason-ntsb-dcouments.

Hrala, Josh. 2016. "A Mall Security Robot Has Knocked Down And Run Over a Toddler in Silicon Valley." Science Alert, July 15. www.sciencealert.com/a-mall-security-robot-recently-knocked-down-and-ran-over-a-toddler-in-silicon-valley.

Hubbard, F. Patrick. 2014. "'Sophisticated Robots': Balancing Liability, Regulation, and Innovation." *Florida Law Review* 66 (5): 1803–72. https://scholarship.law.ufl.edu/flr/vol66/iss5/1/.

Kowert, Weston. 2017. "The Foreseeability of Human-Artificial Intelligence Interactions." *Texas Law Review* 96 (1): 181–204. https://texaslawreview.org/foreseeability-human-artificial-intelligence-interactions/.

Lakkaraju, Hima. 2021. "Tutorial: Explainable ML in the Wild." March 4. YouTube video, 28:40. www.youtube.com/watch?v=K6-ujR_67eY.

Lee, Dave. 2019. "Uber self-driving crash 'mostly caused by human error.'" BBC News, November 20. www.bbc.com/news/technology-50484172.

Lior, Anat. 2020. "AI Strict Liability Vis-À-Vis AI Monopolization." *Columbia Science and Technology Law Review* 22 (1): 90–126. https://doi.org/10.52214/stlr.v22i1.8055.

———. 2021a. "The AI Accident Network: Artificial Intelligence Liability Meets Network Theory." *Tulane Law Review* 95 (5): 1104–66. www.tulanelawreview.org/pub/volume95/issue5/the-ai-accident-network-artificial-intelligence-liability-meets-network-theory.

———. 2021b. "Artificial Intelligence and Tort Law — Who Should be Held Liable when AI Causes Damages?" Heinrich Böll Stiftung, December 24. https://il.boell.org/en/2021/12/24/artificial-intelligence-ai-tort-law-and-network-theory-who-should-be-held-liable-when-ai.

———. 2022. "Insuring AI: The Role of Insurance in Artificial Intelligence Regulation." *Harvard Journal of Law & Technology* 35 (2): 467–530. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4266259.

Marchisio, Emiliano. 2021. "In support of 'no-fault' civil liability rules for artificial intelligence." *SN Social Sciences* 1 (54). https://doi.10.1007/s43545-020-00043-z.

Matthias, Andreas. 2004. "The responsibility gap: Ascribing responsibility for the actions of learning automata." *Ethics and Information Technology* 6: 175–83. https://link.springer.com/article/10.1007/s10676-004-3422-1.

Poursabzi-Sangdeh, Forough, Daniel G. Goldstein, Jake M. Hofman, Jennifer Wormian Wortman Vaughan and Hanna Wallach. 2021. "Manipulating and Measuring Model Interpretability." *CHI '21: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, Article No. 237, 1–52. https://dl.acm.org/doi/10.1145/3411764.3445315.

Ribeiro, Marco Tulio, Sameer Singh and Carlos Guestrin. 2016. "'Why Should I Trust You?' Explaining the Predictions of Any Classifier." *KDD '16: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* 1135–44. https://doi.org/10.1145/2939672.2939778.

Riordan, Brandon J. 2003. "Unravelling the Mystery — A Comparative Introduction to Product Liability Law in the US and Europe." *South Carolina Journal of International Law and Business* 1 (1): 27–40.

Rudin, Cynthia and Joanna Radin. 2019. "Why Are We Using Black Box Models in AI When We Don't Need To? A Lesson From an Explainable AI Competition." *Harvard Data Science Review*, November 22. https://hdsr.mitpress.mit.edu/pub/f9kuryi8/release/8.

Sahni, Isha and Dhananjay Jha. 2015. "Manesar: Factory worker crushed to death by industrial robot." *Hindustan Times*, August 13. www.hindustantimes.com/gurgaon/manesar-factory-worker-crushed-to-death-by-industrial-robot/story-0Hc7V2uu2L2jlYfo9gEdXK.html.

Santoni de Sio, Filippo and Giulio Mecacci. 2021. "Four Responsibility Gaps with Artificial Intelligence: Why they Matter and How to Address them." *Philosophy & Technology* 34: 1057–84. https://link.springer.com/article/10.1007/s13347-021-00450-x.

Selbst, Andrew D. 2020. "Negligence and AI's Human Users." *Boston University Law Review* 100 (4): 1315–76. www.bu.edu/bulawreview/files/2020/09/SELBST.pdf.

Thomasen, Kristen. 2021. "AI and Tort Law." In *Artificial Intelligence and the Law in Canada*, edited by Florian Martin-Bariteau and Teresa Scassa. Toronto, ON: LexisNexis Canada.

Vallance, Chris. 2022. "New EU law could open up messaging and app buying." BBC News, November 1. www.bbc.com/news/technology-63458377.

Vasudevan, Amrita, Urvashi Aneja and Angelina Chamuah. 2021. "Locating responsibility in socio-technical systems." Paper No. 02. Digital Futures Lab. July. www.responsibletech.in/post/locating-responsibility-in-socio-technical-systems.

Wagner, Gerhard. 2023. "Liability Rules for the Digital Age — Aiming for the Brussels Effect." Forschungsinstituts für Recht und Digitale Transformation Working Paper No. 10. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4320285.

Wendehorst, Christiane. 2020. "Strict Liability for AI and other Emerging Technologies." *Journal of European Tort Law* 11 (2): 150–80. www.degruyter.com/document/doi/10.1515/jetl-2020-0140/html?lang=en.

Whittaker, Meredith. 2021. "The steep cost of capture." *Interactions* 28 (6): 50–55. https://doi.org/10.1145/3488666.

Yoshikawa, Jin. 2020. "Sharing the Costs of Artificial Intelligence: Universal No-Fault Social Insurance for Personal Injuries." *Vanderbilt Journal of Entertainment and Technology Law* 21 (4): 1155–87. https://scholarship.law.vanderbilt.edu/jetlaw/vol21/iss4/8/.

About CIGI

The Centre for International Governance Innovation (CIGI) is an independent, non-partisan think tank whose peer-reviewed research and trusted analysis influence policy makers to innovate. Our global network of multidisciplinary researchers and strategic partnerships provide policy solutions for the digital era with one goal: to improve people's lives everywhere. Headquartered in Waterloo, Canada, CIGI has received support from the Government of Canada, the Government of Ontario and founder Jim Balsillie.

À propos du CIGI

Le Centre pour l'innovation dans la gouvernance internationale (CIGI) est un groupe de réflexion indépendant et non partisan dont les recherches évaluées par des pairs et les analyses fiables incitent les décideurs à innover. Grâce à son réseau mondial de chercheurs pluridisciplinaires et de partenariats stratégiques, le CIGI offre des solutions politiques adaptées à l'ère numérique dans le seul but d'améliorer la vie des gens du monde entier. Le CIGI, dont le siège se trouve à Waterloo, au Canada, bénéficie du soutien du gouvernement du Canada, du gouvernement de l'Ontario et de son fondateur, Jim Balsillie.

Credits

President Paul Samson
Director, Program Management Dianna English
Project Manager Jenny Thiel
Publications Editor Susan Bubak
Graphic Designer Brooklynn Schwartz

For publications enquiries, please contact publications@cigionline.org.

67 Erb Street West
Waterloo, ON, Canada N2L 6C2
www.cigionline.org