Varsha Raghavan
Professor Joan Feigenbaum
Topics in Computer Science and Law
16 December 2019

Developing Proper Liability Schemes to Address Artificial Intelligence in the United States

I. Introduction

Artificial intelligence (AI), technology that allows machines to emulate human decision-making processes through techniques like machine learning, is transforming the way society operates and has already impacted many industries, including medicine, transportation, manufacturing, farming, and finance. With AI's ability to out-perform humans in tasks based on a variety of criteria, including speed, objectivity, and accuracy, its positive applications are endless. For example, AI can reduce the risk of vehicle collision by paving the way for autonomous vehicles[1] or help surgeons perform more precise operations.[2] However, similar to how humans are prone to error, so too are AI agents. If a driverless car hits a pedestrian or an algorithm in charge of loans discriminates based on race, how should one go about seeking justice? Such situations raise questions of regulation and remedy for AI-driven harms.

II. The Nature of AI

The difficulty with deciding on a regulatory framework for AI arises from the fact that the algorithms that are used are black boxes. Generally AI systems swallow large

[1]Maddox, Teena. "How Autonomous Vehicles Could Save over 350K Lives in the US and Millions Worldwide." *ZDNet*, ZDNet, 16 Jan. 2019, https://www.zdnet.com/article/how-autonomous-vehicles-could-save-over-350k-lives-in-the-us-and-millions-worldwide/.

[2] Britt, Phil. "How AI-Assisted Surgery Is Improving Patient Outcomes." *Robotics Business Review*, Robotics Business Review, 15 June 2018, https://www.roboticsbusinessreview.com/health-medical/ai-assisted-surgery-improves-patient-outcomes/.

amounts of data, process the data to draw patterns and develop inferences, and then spit out coordinated outputs for specific inputs. However the way in which the data is processed can differ based on the underlying architecture of the system, like the type of neural network. For example, deep learning networks are neural networks with many layers, and because the data points are warped from layer to layer based on different activation functions, it is extremely complicated to understand the mappings between inputs and outputs. Additionally other machine learning algorithms achieve goals that humans simply are unable to do, like analyzing data in higher-dimensional spaces.[3]

The obscurity that accompanies AI systems complicates the matters of intentionality, foreseeability, and knowledge that are important in law. Indeed, complete predictability of AI behavior is counter-intuitive to what it means for something to be artificially-intelligent. Therefore, for the purposes of this paper, it will be assumed that complete and confident foreseeability into the functions of AI systems is out of the question.

III. What is Liability?

It is important to emphasize that this paper is solely dealing with the issue of liability and not accountability. Although the two are often used synonymously[4], there are crucial differences. For the sake of this discussion, the following definitions will be used, as sourced from *thelawdictionary.org*, and developed to clarify the distinction:

[3] Gao, et al. "Learning in High-Dimensional Multimedia Data: The State of the Art." *ArXiv.org*, 10 July 2017, https://arxiv.org/abs/1707.02683.
[4] *thesaurus.com* lists them as synonyms, as do various other online dictionaries.

1. Liability
    a. "The state of being bound or obliged in law or justice to do, pay, or make good something; legal responsibility."[5]
2. Accountability
    a. "When one party must report its activities and take responsibility for them. It is done to keep them honest and responsible."[6]

Liability is a legal term that ties together notions of responsibility and reparation, whereas accountability more generally is a mechanism for attribution. Helen Nissenbaum, a professor at Cornell Tech in New York, offers a closer analysis of these differences. She states that while liability is concerned with the victim's plight and appropriate sanctions, accountability is centered around establishing a relationship between an outcome and an agent, regardless of the existence of a harm, and determining who is answerable for certain undertaken actions.[7] She also addresses the element of moral responsibility and how it features in both of these structures. Legally, entities which have no direct hand in events can be held liable due to their status and ability to offer compensation to the victim, meaning that liability does not necessarily imply blame or fault.[8] Even for accountability, moral responsibility lingers dimly on the horizon[9], because although accountability is about tracing back an action to its origin, the origin-point could be one

[5]"What Is LIABILITY? Definition of LIABILITY (Black's Law Dictionary)." *The Law Dictionary*, 13 Feb. 2014, https://thelawdictionary.org/liability/.
[6]"What Is ACCOUNTABILITY? Definition of ACCOUNTABILITY (Black's Law Dictionary)." *The Law Dictionary*, 6 Oct. 2012, https://thelawdictionary.org/accountability/.
[7]Tavani, Herman T. *Ethics and Technology: Controversies, Questions, and Strategies for Ethical*. John Wiley & Sons, 2016, pg. 120.
[8] Ibid**.**
[9] Ibid.

that cannot be held answerable. For example, legal guardians and parents are answerable for the activities of minors in their care. Therefore it is imperative that one entirely divorces liability from moral culpability and understands that liability chiefly concerns itself with fulfilling an outstanding obligation to a victim rather than fairly attributing blame or establishing causation.

In the realm of AI, questions of accountability and moral responsibility are complicated, especially because the notion of agency becomes obscure and the ability to testify to the results of a system becomes compromised due to the black-box nature of the technology. Although determining liability does not put to rest these concerns, it is a temporary pacifier as it "addresses the needs of the victims"[10], which is important in order to ensure that technological innovation does not proceed ignorant of the woes of those it affects. Presently there exist many theories of liability through which the existence of liability can be proved by a claimant in order to seek restitution. Analyzing how each fares in the face of AI is important to determine which scheme should ultimately be applied to redress harms caused by AI or whether a new one must be crafted. Although each harm should be evaluated on an individual basis to determine liability, this paper is concerned with addressing the likely-case scenario for AI.

IV. Legal Liability: Criminal

Being criminally liable means that one is found responsible for breaking the law and can be prosecuted by the government to pay fines or undertake other penalties, including imprisonment. In order to prove criminal liability, two elements need to be

[10] Ibid, pg. 123

satisfied: mens rea (intention to commit crime) and actus rea (causation of the grievance).[11] Actus rea can include both an undertaken action or a failure to act. Mens rea can be based on acting purposely (conscious intent to cause harm), acting knowingly (aware that a certain action would very likely be followed by harm), acting recklessly (aware of possible risks), and acting negligently (unaware of risks but should have been aware).[12] These elements need to be proved "beyond a reasonable doubt" in order to establish criminal guilt.

The following schemes, as established by Gabriel Hallevy[13], clarify the ways in which criminal liability can be established specifically in regards to artificial intelligence.

1) Perpetrator-via-another: This model is used for cases in which the actus rea was committed by a child or mentally-handicapped individual; in other words, an innocent agent. Because the law deems the actor as incapable of forming mens rea, if the actor was guided by an instructor, that instructor is held liable. Translating this to the issue at hand, if an AI program causes a harm and the law decides to treat AI programs as innocent agents, the user or the programmer shall be held liable.
2) Natural-probable-consequence: This model is often used to prosecute accomplices of crimes. Essentially, even if one did not commit the actus rea, if the action could be constituted as a probable occurrence given the nature of the activities the accomplice was abetting, criminal liability can

[11]Shein, Marcia. "What Is Criminal Liability?" *Federal Criminal Law Center*, 13 Nov. 2019, https://federalcriminallawcenter.com/2015/01/criminal-liability/.
[12]"Mens Rea." *Legal Information Institute*, Legal Information Institute, https://www.law.cornell.edu/wex/mens_rea.
[13]Gabriel. "The Basic Models of Criminal Liability of AI Systems and Outer Circles." *SSRN*, 19 June 2019, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3402527.

be established. In regards to AI, this offers another avenue in which to implicate the user or programmer of the software. Essentially if the user or programmer could foresee the program participating in a criminal scheme and such a thing actually happened, she can be held responsible.

3) Direct liability: This model assigns both elements of criminal liability to the AI system.

Through each of these schemes, it can be seen how the existence of criminal liability can be legally proven for harms caused by AI systems, until one considers the standard of proof in criminal cases.

"Beyond a reasonable doubt" essentially means that jurors are one degree short of completely certainty that each element of the crime occurred in the way presented by the prosecution. As stated above, a key component of criminal liability is mens rea or the intent of the accused to commit the crime. Generally, because of the black-box nature of AI and lack of foreseeability inherent in the design of the machines, proving mens rea "beyond a reasonable doubt" is not possible. For example to prove negligence, one must demonstrate that the accused should have known the risks[14] and therefore failed to act as a reasonable person would have done so. In the case of recklessness, one must demonstrate that the accused was able to appreciate the risks and still act in disregard of them. Ignoring the minor distinctions between the two, assessing risk requires a level of foreseeability that is at odds with AI as it is today. In the case of acting purposely and knowingly, these mental states require being able to understand the consequences of

[14]Schwartzbach, Micah. "What Is Criminal Negligence?" *Www.nolo.com*, Nolo, 10 Sept. 2014, https://www.nolo.com/legal-encyclopedia/what-criminal-negligence.html.

actions, and this is difficult to do because of the inability to understand the relationships between inputs and outputs in AI systems. Therefore even if a person acted in complete malice, proving this mens rea "beyond a reasonable doubt" implicitly confers upon the accused a level of total understanding of the AI system, which is currently impossible. Therefore the accused has a valid defense.

Some crimes require no mental intent, and such acts fall under strict liability. For example, statutory rape, possession of drugs, and speeding[15] require no mens rea proof. Although AI systems that perform crimes held to this standard could then be found criminally liable, the way in which to apply the punishment is unclear. Establishing vicarious liability (like illustrated in schemes (1) and (2) above) once again raises the difficulty of establishing mens rea for those individuals, and punishing the AI system raises the complicated question of to what degree should AI systems be assigned personhood.

Because of the lack of mens rea and the question of how to reconcile personhood and AI systems, establishing criminal liability in cases involving AI is currently not appropriate. However, civil cases are held to a lower standard than "beyond a reasonable doubt," making them the better avenue to handle such issues.

V. Legal Liability: Civil

A civil case can be proved by a "preponderance of evidence," which is often interpreted as being able to prove the defendant is more likely than not responsible and

[15]The Editors of Encyclopaedia Britannica. "Mens Rea." *Encyclopædia Britannica*, Encyclopædia Britannica, Inc., 10 Apr. 2018, https://www.britannica.com/topic/mens-rea.

results in monetary damages being awarded to the winner of the case.[16] In civil suits, the debate centers around torts. Because AI systems are presently not treated as persons before the law, the type of framework used to evaluate civil liability in these cases is dependent upon whether AI is characterized as a product or a service. For this paper, AI shall be categorized as a product because the law in this area is much more developed to address many of the harms that can result from this technology (Lawyer Marguerite E. Gerstner develops this idea further in the Santa Clara Law Review journal article "Liability Issues with Artificial Intelligence Software").[17] Therefore, the following main frameworks exist:[18]

1) Negligence: Similar to criminal negligence, it must be demonstrated that the defendant failed to act as a reasonable person would be expected to do so. Specifically, the following elements must be proved for negligence[19]:
    a) Existence of a duty of care
    b) Breach of that duty
    c) Injury resulting from breach

    In regards to AI, demonstrating negligence is still complicated because of element (a) but doable due to the lower threshold for proof. As Gerstner states, it is obvious that vendors have a duty of care owed to their consumers. However the incomprehensible nature of AI makes it so that

[16]England, Deborah C. "Civil Liability." *Www.criminaldefenselawyer.com*, Nolo, 22 Mar. 2017, https://www.criminaldefenselawyer.com/resources/civil-liability.htm.
[17] Gerstner M.E.: Comment, Liability Issues with Artificial Intelligence Software, 33 Santa Clara L. Rev. 239. http://digitalcommons.law.scu.edu/lawreview/vol33/iss1/7 (1993).
[18] England, Deborah C. "Civil Liability…"
[19]Kingston, J. K. C. "Artificial Intelligence and Legal Liability." *Research and Development in Intelligent Systems XXXIII*, 2016, pp. 269–279., doi:10.1007/978-3-319-47175-4_20.

establishing the appropriate standard of care becomes a Herculean task. Because AI technology is constantly being disrupted by new developments, the baseline on which to evaluate a duty of care is constantly shifting.[20] Therefore it seems most reliable to derive the standard of care owed to the consumer by the vendor based on the terms that were stipulated/ the marketing surrounding the product. Gerstner bolsters this suggestion by remarking that if the system is considered an "expert system," it should be held to the standard of care expected from a professional. This method can also help vindicate users of AI systems who could possibly be implicated in the crimes. For example, for physicians who use AI medical software to help with diagnoses, if the software was marketed as an authority in disease detection (a far-fetched and dangerous claim for any such technology in the present day) and then misdiagnoses a patient, the standard of care of an expert exists between the AI software and the physician (the user) rather than the physician and the patient, and thus the physician is not necessarily liable for the resulting harm. However if the software recommends secondary analysis of any diagnoses and the physician fails to do so, the physician is liable for the harm because the standard of care owed by the AI software to the physician was that of a non-expert.

[20]Rachum-Twaig, Omar. "Liability for Artificial-Intelligence-Based Robots." *The Federmann Cyber Security Center – Cyber Law Program*, 14 Nov. 2019, https://csrcl.huji.ac.il/blog/Liability-for-Artificial-Intelligence-Based-Robots.

This method of determining the standard of care addresses the reality of what a society dependent on the use of AI looks like. Because of the nature of this technology, unpredictability is a necessary tradeoff for overall better results. However holding users of autonomous systems responsible when that unpredictability results in harm is unjust (depending on the use of that AI, of course). For example in the aviation industry, the American government estimates that autonomous systems are used 90% of the time during flights in order to negate poor pilot judgment.[21] The consequence of this, however, is that the manual skills of pilots are eroding, and they are unprepared to handle emergency situations when the systems turn off. Bo Corby, Director of Standards and Training for Future & Active Pilot Advisors, states that pilot training now consists more of teaching pilots to use the automated systems rather than hone their manual flying skills.[22] As a result, pilots are found to develop an overdependence on these systems.[23] This is what many suspect caused the two Boeing 737 MAX crashes, but rather than blaming the pilots, the public has rightfully held Boeing accountable for its negligence in providing pilots with appropriate warning of the autonomous system and a briefing of its

[21]Woodyard, Chris. "On Autopilot: Pilots Are Losing Their Basic Flying Skills, Some Fear after Boeing 737 Max Crashes." *USA Today*, Gannett Satellite Information Network, 27 May 2019, https://www.usatoday.com/story/news/2019/05/25/boeing-737-max-8-autopilot-automation-pilots-skills-flying-hours-safety/1219147001/.
[22]Ibid.
[23]Hawkins, Andrew J. "Deadly Boeing Crashes Raise Questions about Airplane Automation." *The Verge*, The Verge, 15 Mar. 2019, https://www.theverge.com/2019/3/15/18267365/boeing-737-max-8-crash-autopilot-automation.

functionalities. The loss of certain human skills will eventually take place in society due to an active use of AI, and during this transition, it is necessary to ensure that the proper precautions are taken so that liability does not get misplaced.

2) Intentional Torts: There is overlap between intentional torts and crimes, as both require a proof of intent which is usually disregarded in regular tort suits. The main distinction is that intentional torts are judged under civil law standards, meaning that even if the same act was judged as a crime but the case was unsuccessful, it can still be judged as a civil issue and succeed due to the lower standard of proof.[24] This means that damages can be won for harms resulting from intentionally malicious uses of AI and bad actors can still be held accountable on some terms.
3) Breach of Warranty: Because all products are sold with warranties (implicit ones included), plaintiffs can seek damages if they prove injury resulted from a breach of terms.
4) Strict Liability for Products: Under this claim, plaintiffs need to prove the injury happened but do not need to prove negligence or fault of the defendant. This standard is applied to defective products or products that turned dangerous when used in normal, foreseeable circumstances.[25] Manufacturers of the products are normally held liable for strict liability

[24]"Intentional Torts: What Are They? (Definitions & Examples)." *Denver Workers' Compensation Lawyer: Mack Babcock*, https://www.injurylawcolorado.com/legal-library/what-are-intentional-torts.html.
[25]Kingston, J. K. C. "Artificial Intelligence and Legal…"

> offenses because it is recognized that, as a Brookings Institution research paper put it, "consumers have a right to expect safe products."[26]

VI. Proposed Solution

Due to the more relaxed standard of proof necessitated by civil proceedings, lawsuits involving AI systems can best be answered by the schemes detailed in **V.** Indeed, establishing civil liability seems plausible through these contexts, but another question remains unanswered. Because many parties are involved in the production and deployment of AI systems and software, including the materials suppliers, the manufacturer, the vendor, the individual employees in the company, the consumer, etc.[27], who should be held responsible to pay the damages?

The rationale for manufacturers being held as liable defendants in product suits subject to strict liability is best put by litigator Debra England:

> "Courts have reasoned that it is fair to hold manufacturers strictly liable for defective products because public policy is best served by assigning responsibility where it will be most effective in reducing the potential for harm. Manufacturers are in the best position to both address defects in their products and absorb the cost to society of such defects (by spreading it out among all purchasers)."

Applying this logic to AI cases, the manufacturing company of the AI software should be held liable to pay damages, because it has the greatest means to prevent such harms from

---

[26]Villasenor, John. "Products Liability Law as a Way to Address AI Harms." *Brookings*, Brookings, 30 Oct. 2019, https://www.brookings.edu/research/products-liability-law-as-a-way-to-address-ai-harms/.

[27]"Artificial Intelligence - Who Is Liable When AI Fails to Perform?" *Artificial Intelligence – Who Is Liable When AI Fails to Perform? Insight | Technology, Media & Telecommunications | United Kingdom | International Law Firm CMS*, https://cms.law/en/gbr/publication/artificial-intelligence-who-is-liable-when-ai-fails-to-perform.

happening in the first place and developing accountability measures that could serve as legal defense measures when faced with such suits. For example, companies can invest capital in developing comprehensive warranties, devote more attention to vetting and testing software in a variety of environments, and purchase more insurance. They also benefit the most from the sales of the AI systems, so they should bear the brunt of its failures. Individual programmers and users of AI systems do not have the agency or privilege to do such things. Obviously, damages do not only have to be paid by one company and can be divided appropriately along the supply chain with further analysis. This is a temporary solution, however, and more attention should be focused on developing new regulations that better account for the nuances of AI that fall through the cracks of this approach.

VII. Conclusion

As society introduces automation in more of its operations and adopts AI to solve pressing problems, there will be a new class of injuries that begin to infiltrate the courts. While this process is still in its nascent stages, it is important to figure out how to reapply existing legal doctrine to remedy the harm caused by AI. Because AI has not yet reached the sophisticated level needed to demand a claim of personhood, criminal liability frameworks are ill-suited to address these issues, but civil liability schemes have the potential to do so. Companies gain the most profit from developing and deploying AI systems and have more resources to enforce better practices that prevent possible harm and misuse of their products. Thus, they should be held liable in non-intentional tort suits involving AI. Although this position may penalize the industry such that smaller

companies or start-ups may be reluctant to explore AI due to the burden of shouldering costs, it is a necessary protection needed so that society can feel at ease with this transition into a more artificially-intelligent world.

Works Cited

"Artificial Intelligence - Who Is Liable When AI Fails to Perform?" *Artificial Intelligence – Who Is Liable When AI Fails to Perform? Insight | Technology, Media & Telecommunications | United Kingdom | International Law Firm CMS*, https://cms.law/en/gbr/publication/artificial-intelligence-who-is-liable-when-ai-fails-to-perform.

Britt, Phil. "How AI-Assisted Surgery Is Improving Patient Outcomes." *Robotics Business Review*, Robotics Business Review, 15 June 2018, https://www.roboticsbusinessreview.com/health-medical/ai-assisted-surgery-improves-patient-outcomes/.

England, Deborah C. "Civil Liability." *Www.criminaldefenselawyer.com*, Nolo, 22 Mar. 2017, https://www.criminaldefenselawyer.com/resources/civil-liability.htm.

Gabriel. "The Basic Models of Criminal Liability of AI Systems and Outer Circles." *SSRN*, 19 June 2019, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3402527.

Gao, et al. "Learning in High-Dimensional Multimedia Data: The State of the Art." *ArXiv.org*, 10 July 2017, https://arxiv.org/abs/1707.02683.

Hawkins, Andrew J. "Deadly Boeing Crashes Raise Questions about Airplane Automation." *The Verge*, The Verge, 15 Mar. 2019, https://www.theverge.com/2019/3/15/18267365/boeing-737-max-8-crash-autopilot-automation.

"Intentional Torts: What Are They? (Definitions & Examples)." *Denver Workers' Compensation Lawyer: Mack Babcock*, https://www.injurylawcolorado.com/legal-library/what-are-intentional-torts.html.

Kingston, J. K. C. "Artificial Intelligence and Legal Liability." *Research and Development in Intelligent Systems XXXIII*, 2016, pp. 269–279., doi:10.1007/978-3-319-47175-4_20.

Maddox, Teena. "How Autonomous Vehicles Could Save over 350K Lives in the US and Millions Worldwide." *ZDNet*, ZDNet, 16 Jan. 2019, https://www.zdnet.com/article/how-autonomous-vehicles-could-save-over-350k-lives-in-the-us-and-millions-worldwide/.

"Mens Rea." *Legal Information Institute*, Legal Information Institute, https://www.law.cornell.edu/wex/mens_rea.

Rachum-Twaig, Omar. "Liability for Artificial-Intelligence-Based Robots." *The Federmann Cyber Security Center – Cyber Law Program*, 14 Nov. 2019, https://csrcl.huji.ac.il/blog/Liability-for-Artificial-Intelligence-Based-Robots.

Schwartzbach, Micah. "What Is Criminal Negligence?" *Www.nolo.com*, Nolo, 10 Sept. 2014, https://www.nolo.com/legal-encyclopedia/what-criminal-negligence.html.

Shein, Marcia. "What Is Criminal Liability?" *Federal Criminal Law Center*, 13 Nov. 2019, https://federalcriminallawcenter.com/2015/01/criminal-liability/.

Tavani, Herman T. *Ethics and Technology: Controversies, Questions, and Strategies for Ethical*. John Wiley & Sons, 2016.

The Editors of Encyclopaedia Britannica. "Mens Rea." *Encyclopædia Britannica*, Encyclopædia Britannica, Inc., 10 Apr. 2018, https://www.britannica.com/topic/mens-rea.

Villasenor, John. “Products Liability Law as a Way to Address AI Harms.” *Brookings*, Brookings, 30 Oct. 2019, https://www.brookings.edu/research/products-liability-law-as-a-way-to-address-ai-harms/

“What Is ACCOUNTABILITY? Definition of ACCOUNTABILITY (Black's Law Dictionary).” *The Law Dictionary*, 6 Oct. 2012, https://thelawdictionary.org/accountability/.

“What Is LIABILITY? Definition of LIABILITY (Black's Law Dictionary).” *The Law Dictionary*, 13 Feb. 2014, https://thelawdictionary.org/liability/.

Woodyard, Chris. “On Autopilot: Pilots Are Losing Their Basic Flying Skills, Some Fear after Boeing 737 Max Crashes.” *USA Today*, Gannett Satellite Information Network, 27 May 2019, https://www.usatoday.com/story/news/2019/05/25/boeing-737-max-8-autopilot-automation-pilots-skills-flying-hours-safety/1219147001/.