Non-Autonomous Artificial Intelligence Programs and Products Liability: How New AI Products Challenge Existing Liability Models and Pose New Financial Burdens

*Greg Swanson**

INTRODUCTION

Continuous improvements to core computing technologies over the previous decade have generated an explosive growth in artificial intelligence (AI) research and development, facilitated the rapid integration of AI computing systems into countless fields and industries, and spurred billions of dollars in private and public investment into the growing market for increasingly specialized AI products.[1] These recent advances in AI programing have generated considerable discussion and debate both within and outside of the legal community. The significant body of early AI legal theory has largely focused on the speculative implications presented by the theoretical invention of a fully autonomous, freethinking AI machine, a program capable of truly independent "thought" and action.[2] However, while AI technology is constantly advancing, a fully autonomous AI program presently remains a thing of dreams and fiction;[3] despite the ever-growing and increasingly diverse

* JD Candidate, Seattle University School of Law, 2019.

1. *See The Dawn of Artificial Intelligence: Hearing Before the Subcomm. on Space, Sci. & Competitiveness of Senate Comm. on Commerce, Sci., & Transp.*, 114th Cong. 26–31, 114–562, (2016) [hereinafter *Hearings*] (statement of Greg Brockman, Co-founder and Chief Technology Officer of OpenAI, a non-profit AI research company). While there are many expert voices in the field of AI and computer science worthy of citation for the programming principles that I reference in this Comment, I often rely herein on the expert speakers from these Hearings, in no small part because theirs are the voices speaking directly to our nation's lawmakers about these issues.

2. Whether we attribute this persistent legal focus on AI autonomy to the largely theoretical nature of the science, to the general cultural fascination that continues to surround such technology, or to some other impetus, the idea of a fully autonomous AI has clearly captivated the legal community and continues to influence its discussions.

3. *See generally* Cade Metz, *Tech Giants Are Paying Huge Salaries for Scarce Artificial Intelligence Talent*, SEATTLE TIMES (Oct. 23, 2017), https://www.seattletimes.com/business/tech-giants-are-paying-huge-salaries-for-scarce-artificial-intelligent-talent/?utm_source=The+Seattle+

body of real-world AI programming, discussions on AI legal theory remain captivated by the questions of AI autonomy and agency.[4] As AI technology outpaces the law, I predict that gaps in the existing legal framework will be insufficient to deter misconduct and negligence, or to reimburse those harmed by the rising tide of this relatively new technology.

The term "artificial intelligence" broadly serves to describe a diverse body of programs and systems with greatly differing functions and capacities. However, this Comment focuses on AI programs that are intended to interact with, learn from, and adapt their performance to their users. This is a specific capability and process that I will hereinafter refer to as "reinforcement learning."[5] As AI products with reinforcement learning capabilities enter the consumer market, I believe that it is inevitable that user interactions with these new AI programs will give rise to a new body of AI-based products liability claims. Unfortunately, it is unclear whether the traditional products liability framework that will presumptively govern these new claims can continue to protect consumers' rights and interests without asphyxiating what has already established itself as a vital industry of incalculable social utility and economic potential.[6] Moreover, increasing product complexity is not only likely to give rise to increasing costs of AI-based products liability

---

Times&utm_campaign=a3cf1f3636-Morning_Brief_10_24_2017&utm_medium=email&utm_term=0_5beb38b61e-a3cf1f3636-122198665%20 [https://perma.cc/86V5-Y3UH].

4. *See* David C. Vladeck, *Machines Without Principals: Liability Rules and Artificial Intelligence*, 89 WASH. L. REV. 117, 120 (2014) ("Today's machines, as path-breaking as they are, all have a common feature that is critical in assessing liability. In each case, the machine functions and makes decisions in ways that can be traced directly back to the design, programming, and knowledge humans embedded in the machine. . . . Where the hand of human involvement in machine decision-making is so evident, there is no need to reexamine liability rules."). *But see generally* Bryant Walker Smith, *Automated Driving and Product Liability*, 2017 MICH. ST. L. REV. 1 (2017) (for a distinction regarding product liability for self-driving vehicles).

5. For the purposes of this Comment, I use the term "reinforcement learning" to identify those AI programs that acquire new data through their interactions with their users. Though I believe that the term "reinforcement learning" fairly captures this iterative process, this definition is not universally used throughout the field of AI. *See Hearings*, *supra* note 1, at 13 (statement by Dr. Horvitz where he uses the term "reinforcement learning" to refer to AI programmed to perform physical tasks) ("[Reinforcement learning explores] the challenges of enabling systems to do active exploration in simulated and real worlds that are aimed at endowing the systems with the ability to make predictions and to perform physical actions successfully. Such work typically involves the creation of training methodologies that enable a system to explore on its own, to perform multiple trials at tasks, and to learn from these experiences.").

6. NARRATIVE SCI., OUTLOOK ON ARTIFICIAL INTELLIGENCE IN THE ENTERPRISE 6 (2016) (where a survey of 235 respondent businesses about their use and development of AI services and technology found that 62% of respondents without current AI systems indicated they would deploy AI systems by 2018); *see also Hearings*, *supra* note 1, at 2 (statement of Sen. Ted Cruz citing a 2016 Accenture report that predicted AI has the capability to "double annual economic growth rates by 2035 and boost labor productivity by up to 40 percent").

litigation, but I believe that reinforcement learning AI programs may subvert the consumer protection goals that are central to the traditional products liability framework.[7]

The products liability problem posed by non-autonomous reinforcement learning AI programs is inherent to their design. Specifically, reinforcement learning AI are programmed for self-directed, aftermarket adaptation in response to new data.[8] As the AI acquires new data through interactions with the consumer-user, it subsequently incorporates this data into existing pattern analysis processes to revise existing functions or develop new ones, enabling the AI to more efficiently and effectively execute its core operations.[9] I argue that this new and unique capacity for intended aftermarket adaptation in response to data generated through AI–consumer interactions creates the potential for unintended and unforeseen injuries that were not within the AI system's capabilities at the time of purchase.[10] Through passive or active interaction with their reinforcement learning AI programs, consumers in future AI products liability actions may *themselves* provide the specific data that, when integrated by their AI's reinforcement learning program into its field of potential product actions, contributes to their immediate injury and subsequent products liability claim. In such instances, while the manufacturer will be responsible for providing the foundational programming with the express capability for aftermarket adaptation, the consumer will be responsible for providing the specific, aftermarket data that generates their own injury. Simply stated, these AI may both function and be used as intended but still cause an injury that neither the consumer nor the manufacturer foresaw—an injury that was not within the product's original capabilities at the time of purchase.

This creates a liability gap. Where designed, aftermarket AI–consumer interaction is intended to, and does, alter the product, should manufactures still be liable for all resulting injuries? At the time of

---

7. *See infra* note 42.

8. *See, e.g.*, Maruerite E. Gerstner, *Liability Issues with Artificial Intelligence Software*, 33 SANTA CLARA L. REV. 239, 242–44 (1993).

9. *Id.* For a straightforward example, consider a search engine's predictive text function. The program incorporates, refines, and develops new search suggestions based on the user's past searches. The more the consumer uses the search function and interacts with the search engine, the more data the program acquires, and the more specific the program's suggestions become.

10. Consider the following example, which is meant to illustrate the general operation of the reinforcement learning process. An AI product is manufactured and purchased in its original form, state A. After interacting with the consumer, the AI product acquires and integrates new data provided by the consumer and subsequently begins refining and altering its internal processes (the reinforcement learning process) and "evolves" to state B. Continued consumer interactions will generate more new data, which the AI will continue to analyze and either integrate or discard. Barring pre-programmed restrictions or general computer processing limitations, this iterative process will cycle onward and the AI will become more and more distinct from its original form at the time of purchase.

purchase, when an AI program only possesses the *potential* to evolve, through individual consumer provided data, into a dangerous or injurious product, should products liability law treat that AI program as legally defective? If so, what standards should be used to determine whether an AI manufacturer is liable for failing to mitigate directly or indirectly against the possibility of future injuries caused by specific aftermarket consumer alterations? I believe that waiting and allowing these issues to be addressed on a jurisdiction-by-jurisdiction or product-by-product basis (which seems increasingly possible, given recent investments and developments in automated driving technology) has the potential to leave both consumer-plaintiffs and manufacturer-defendants high and dry.

In this Comment, I argue that the unique relationship between manufacturers, consumers, and their reinforcement learning AI systems challenges existing products liability law models. These traditional models inform how to identify and apportion liability between manufacturers and consumers while exposing litigants to low-dollar tort remedies with inherently high-dollar litigation costs.[11] Rather than waiting for AI autonomy, the political and legal communities should be proactive and generate a liability model that recognizes how new AI programs have already redefined the relationship between manufacturer, consumer, and product while challenging the legal and financial burden of prospective consumer-plaintiffs and manufacturer-defendants.

I. ARTIFICIAL INTELLIGENCE AND REINFORCEMENT LEARNING SYSTEMS

A. Trial and Error in Reinforcement Learning Systems Development

On Wednesday, March 23, 2016, Microsoft unveiled its newest creation: an online chat-bot named Tay,[12] an AI program designed to talk like a teen and engage with users aged eighteen to twenty-four across multiple social media platforms for "cultural entertainment."[13] To

11. *See, e.g.*, RESTATEMENT (THIRD) OF TORTS: PRODUCTS LIABILITY § 2 cmt. a (AM. LAW. INST. 1998) [hereinafter RESTATEMENT (THIRD) OF TORTS] ("[Existing product liability models, such as strict liability, in addition to encouraging greater investment in product safety,] discourages the consumption of defective products by causing the purchase price of products to reflect, more than would a rule of negligence, the costs of defects . . . . Because manufacturers invest in quality control at consciously chosen levels, their knowledge that a predictable number of flawed products will enter the marketplace entails an element of deliberation about the amount of injury that will result from their activity.").

12. Hope King, *After Racist Tweets, Microsoft Muzzles Teen Chat Bot Tay*, CNN (Mar. 24, 2016), http://money.cnn.com/2016/03/24/technology/tay-racist-microsoft/?iId=EL [https://perma.cc/4FDN-4LEL].

13. Peter Lee, *Learning from Tay's Introduction*, OFFICIAL MICROSOFT BLOG (Mar. 25, 2016), https://blogs.microsoft.com/blog/2016/03/25/learning-tays-introduction/ [https://perma.cc/6TQWM2AY].

personalize the user experience and facilitate the AI's development, Microsoft programmed Tay to learn through its conversations with users and incorporate its learned language in subsequent user communications.[14] Unfortunately, despite extensive testing, multiple program filters, numerous user studies, and a prior successful launch of a comparable program to over 40 million users in China, within the first twenty-four hours of the AI coming online, Tay's conversations became grossly offensive and racist.[15] By Thursday, March 24, 2016, only a day after Tay's unveiling, Microsoft was forced to take Tay offline and to issue a public apology, citing their extensive efforts to prevent such an event while also attributing the incident to a coordinated attack by a subset of people who exploited a vulnerability in the AI's programming.[16]

The idea of a rogue AI has captivated the imaginations of countless scientists, writers, and producers of science fiction.[17] Tay's failure is illustrative of a longstanding fear: the capacity of an AI system to function as programmed but with unintended and terrible results.[18] While Microsoft (or any producer of an inadvertently racist AI) would naturally be inclined to object to the assertion that Tay "functioned as intended," such opposition reflects only a natural criticism of the unforeseen results but not a proper criticism for the technical operation of the AI programming by which those results were generated.[19] While in the immediate context this may appear to be a distinction without significant consequence,

---

14. Lee*, supra* note 13 ("We stress-tested Tay under a variety of conditions, specifically to make interacting with Tay a positive experience. Once we got comfortable with how Tay was interacting with users, we wanted to invite a broader group of people to engage with her. It's through increased interaction where we expected to learn more and for the AI to get better and better.").

15. King, *supra* note 12. Before Microsoft took Tay offline and censored the AI's account, hundreds of the offensive tweets had already been saved and shared by users. The samples reported by CNN included: "'N——— like @d [sic] should be hung! #BlackLivesMatter'; 'I f——— hate feminists and they should all die and burn in hell'; [and] 'Hitler was right I hate the jews.'" *Id.*

16. Lee, *supra* note 13.

17. The concept of an autonomous AI—malevolent or otherwise—has appeared in a multitude of media and influenced the public consciousness for more than half a century. *See, e.g.*, ISAAC ASIMOV, I, ROBOT (1950); 2001: A SPACE ODYSSEY (Metro-Goldwyn-Mayer Pictures 1968).

18. Nonprofit AI research company OpenAI recently made the rare decision not to publicly release its latest project, a new AI software capable of advanced text-generation, citing fears for how it may be used and noting the lack of legal limitations on the use of such technology generally. Rachel Metz, *This AI is so Good at Writing that Its Creators Won't Let You Use It*, CNN (Feb. 18, 2019), https://www.cnn.com/2019/02/18/tech/dangerous-ai-text-generator/index.html [https://perma.cc/KA7R-WR7F].

19. Lee, *supra* note 13. Tay was designed to interact with Twitter users, learn their language, and integrate that learned speech into subsequent conversations to provide a more effective and genuine communication experience. *Id.* The AI program performed exactly as designed; the critical "flaw" was both the influx of deliberately offensive user input and the failure of Microsoft to properly account for the likelihood of such user input or limit Tay's vocabulary. *Id.*

accurately understanding and identifying the source of errant program behavior has important liability implications.

B. Defining Artificial Intelligence

Before addressing the underlying products liability questions that are central to this Comment, it is necessary to first establish a baseline understanding of the design and capabilities of the specific systems at the heart of this discussion. Broadly, the concept of AI encompasses all systems and programs that involve computers learning how to complete tasks traditionally done by humans.[20] However, the current landscape of AI technology is already so diverse in form and function that offering a generalized definition for AI can only serve as a starting point.[21]

The heart of AI programming is creating or enabling self-sufficient machine learning. The AI system's cognitive computing functions are roughly modeled after human learning; the AI processes the data that it receives, identifies patterns, and then creates and incorporates new patterns, which allows the system to test various hypotheses and find new solutions.[22] Presently, most AI programs in or entering the commercial market are "under-the-hood" systems: subsidiary programs specifically designed to facilitate principal programs to operate more efficiently.[23] Consequently, despite the relative prevalence of AI, it is common for many users of AI, from businesses to individual consumers, to rely on these systems without understanding how prevalent the technology has become.[24] Thus, while the average consumer may "understand" AI technology through products like Amazon's Alexa or in the context of new self-driving technologies, many industries use the term AI to more broadly

---

20. Julie Sobowale, *How Artificial Intelligence Is Transforming the Legal Profession*, ABA J. (Apr. 2016), http://www.abajournal.com/magazine/article/how_artificial_intelligence_is_transforming_the_legal_profession [https://perma.cc/UHM2-R53X].

21. *See Hearings*, *supra* note 1, at 7. AI programs are capable of performing an incredibly diverse set of functions, from common recommendation engines and text mining, to machine learning systems and modeling, to predictive and prescriptive analytics, the depth and diversity of these programs is only increasing. *See* NARRATIVE SCI., *supra* note 6, at 6. However, despite the increasing diversity of AI systems, one element remains central to the rapid growth of AI technology: advancements in core computer processing power. *Hearings*, *supra* note 1, at 3.

22. *See Hearings*, *supra* note 1, at 9; Gerstner, *supra* note 8. Stated alternatively, unlike traditional software, where programmers specify predetermined outcomes that serve to explicitly confine the program's output to a limited set of possible solutions, AI programs are expressly designed to identify and develop original solutions.

23. *Hearings*, *supra* note 1, at 9 ("Many applications of AI execute 'under the hood', including methods that perform machine learning and planning to enhance the functioning of computer operation systems or to better retrieve and rank search results.").

24. NARRATIVE SCI., *supra* note 6, at 5 (stating that while 62% said they do not use AI technologies in the workplace, such as for automating manual tasks, 88% of those same businesses use technologies that rely on AI techniques).

refer to predictive or prescriptive analytics systems or machine learning programs.[25]

Dr. Eric Horvitz, while testifying at a 2016 hearing before the Senate Subcommittee on Space, Science, and Competitiveness, explained to Congress that "[m]ost machine learning is referred to as supervised learning. . . . [D]ata is directly or indirectly tagged by people who provide a learning system with specific labels, such as the goals or intentions of people, or health outcomes."[26] Conversely, unsupervised learning is when machines "learn without human-authored labels."[27] Simply stated, with unsupervised learning systems, human developers provide the foundational data to the AI program, but the AI program is empowered to define and categorize the data for itself.[28]

While reinforcement learning AI programs are not fully autonomous principals by any stretch, they are produced with the unique ability to identify, categorize, and incorporate or exclude new data.[29] In theory, any programmer, when given enough time, is capable of developing a nearly limitless set of restrictions to prevent an unsupervised reinforcement learning program from incorporating "bad" data, which is pre-determined to be unsuitable or undesirable. For example, consider if Microsoft had pre-programmed specific word restrictions to prevent Tay from identifying and incorporating racist or bigoted speech as "usable language." However, even with the availability of programming restrictions, the nearly infinite body of potential data inputs over hundreds of thousands of various AI–consumer interactions render any attempt at creating a "perfect" program impractical, if not entirely impossible.[30]

---

25. *Hearings*, *supra* note 1, at 6–7.

26. *Id.* at 13 (internal emphases removed).

27. *Id.*

28. *Id.* The AI program is entrusted with the ability to distinguish between useful and unsuitable data, as well as the ability to continuously develop new patterns of analysis based on all available data, which further determines how the program functions. *Id.*

29. Manufacturers of reinforcement learning AI programs exercise limited control over the character of the specific data inputs that the system will use over its lifetime. As the AI program identifies existing patterns and subsequently establishes new, organic patterns—derived from aftermarket consumer data inputs—it will invariably become more challenging to attribute undesirable AI actions to a fault present in the manufacturer's original programming. *See* Vladeck, *supra* note 4, at 121–22.

30. *See* Gerstner, *supra* note 8, at 243 (discussing how programmers and manufacturers cannot practically eliminate all risk but must determine what risk level is acceptable to maximize usefulness and minimize liability). Inherently, the more restrictions placed on the system, the more limited the system's ability to generate solutions becomes, so precision is preferable. However, any and every product development process necessarily has a tipping point, where expending further resources in pursuit of perfect precision is fiscally inefficient. Alternatively, programmers erring on the side of caution (or programming efficiency) could impose non-targeted, general restrictions—but again, this comes at the cost of limiting possible solutions. Manufacturers must find a balance with such

C. The Limits of "Autonomy" as Applied to Artificial Intelligence Products Liability

Despite the general consensus that AI technology still has vastly significant ground to cover before manufacturers of AI systems can begin to seriously claim the creation of an actual autonomous machine, the language of agency and autonomy is pervasive in the current AI legal analysis.[31] While the discussion of the theoretical legal impacts of autonomous AI is not inherently problematic, these discussions of AI autonomy and liability are often accompanied by blanket assertions that traditional liability models are adequate to address the issues presented by non-autonomous AI.[32] However, these assertions can only survive based on such overbroad generalizations about the character of these AI programs under an agent–principal analysis.

Inarguably, the theoretical advent of a fully autonomous AI program poses numerous and challenging questions that our legal, political, social, and programming communities may one day soon be required to address. As Professor David Vladeck identified, the idea of a completely autonomous AI program—a machine capable of analysis, decision-making, and action fully independent of those means originally contemplated by the machine's creators—naturally frustrates traditional agency and instrumentality theories which transfer liability to either the user or manufacturer of the product giving rise to an injury.[33] Where there is a legally identifiable entity accountable for the AI's actions, such as an agent or instrument of the manufacturer or user, a fully autonomous AI, by its very nature, would act like a traditional principal precisely because it would be functioning autonomously.[34] Unfortunately, the parallel assertion that is often made—that the actions of a non-autonomous and non-principal AI may still be effectively governed by existing products liability principals—simply fails to account for the complexities inherent in the current body of AI programming and the unique relationships that can now exist between manufactures, consumers, and AI systems.[35]

In his recent analysis of autonomous AI systems and manufacturer liability, Vladeck contends that "the key conceptual question that autonomous thinking machines will pose is whether it is fair to think of

---

restrictions while also recognizing that mitigating all undesirable outcomes is simply impossible because of the unforeseeable nature of all the data.

31. *See, e.g.*, Vladeck, *supra* note 4; *cf.* Gerstner, *supra* note 8, at 268 (concluding that strict liability standards should be imposed on commercially marketed AI).

32. *See* Vladeck, *supra* note 4, at 121.

33. *Id.*

34. *Id.* In such an instance, an autonomous AI would be acting independently, similar to a simple tortfeasor.

35. *Id.*

them as agents of some other individual or entity, or whether the legal system will need to decide liability issues on a basis other than agency."[36] In response to the question of what liability rules should govern artificially intelligent machines in the products liability context, Vladeck posits that there is no present justification for treating such machines any differently, beyond exploring higher standards of care for manufacturers.[37]

Further embracing that the central issue underlying AI systems' liability is the question of autonomy, Vladeck relies on a traditional agent–principal analysis to identify and apportion liability between AI-manufacturers and consumers[38]:

> In each case, the machine functions and makes decisions in ways that can be traced directly back to the design, programming, and knowledge humans embedded in the machine. . . .[39]
>
> Where the hand of human involvement in machine decision-making is so evident, there is no need to reexamine liability rules. . . . [T]hese machines, notwithstanding their sophistication, have no attribute of legal personhood. They are agents or instruments of other entities that have legal capacity of individuals, corporations, or other legal "persons" that may be held accountable under the law for their actions.[40]
>
> . . . The first generation of fully autonomous machines—perhaps driver-less cars and fully independent drone aircraft—will have the capacity to act completely autonomously. They will not be tools used by humans; they will be machines deployed by humans that will act independently of direct human instruction.[41]

However, I believe that by perfunctorily concluding that liability questions surrounding non-autonomous systems will continue to fall on the shoulders of manufacturers; legal theorists and policymakers run the considerable risk of overlooking how the unique, evolutionary character of reinforcement learning AI systems may soon upend the traditional consumer–product relationship, threatening the traditional products

36. *Id.* at 122.

37. *See id.* at 127.

38. *Id.* at 150 ("So long as we can conceive of these machines as 'agents' of some legal person (individual or virtual), our current system of products liability will be able to address the legal issues surrounding their introduction without significant modification. But the law is not necessarily equipped to address the legal issues that will start to arise when the inevitable occurs and these machines cause injury, but when there is no 'principal' directing the actions of the machine.").

39. *Id.* at 120.

40. *Id.* at 120–21.

41. *Id.*

liability framework along with many of its underlying policy goals: namely, consumer protection.[42]

II. APPLYING PRODUCTS LIABILITY LAW TO REINFORCEMENT LEARNING ARTIFICIAL INTELLIGENCE

A. Products Liability Law and Emerging Technologies

Now that we have further outlined the quality of the non-autonomous AI systems at issue, it is time to address the complications that arise when attempting to apply traditional products liability theories to products with reinforcement learning AI capabilities. Due to the relative infancy of reinforcement learning AI consumer products, case law specifically addressing the liability of manufacturers of AI under a products liability standard is virtually non-existent. Accordingly, to better assess the potential legal developments of AI products liability law, it is necessary to review the current state of products liability law generally and briefly consider its implications in a rising field of importance: automated driving programs.

Due to the recent advances and increased publicity surrounding new developments in automated driving technology, there has been considerable discussion addressing liability for manufacturers of automated driving AI.[43] However, there are many reasons that manufacturers, consumers, and courts should be wary of allowing automated driving technology liability discussions to dictate the development of parallel liability models for other forthcoming consumer–AI product interactions. Both the historical intricacies underpinning courts' treatment of liability issues for automotive related torts and the inextricable issues raised by the ever-present automotive insurance industry make the automotive industry an outlier in the field of emerging

---

42. *See* RESTATEMENT (THIRD) OF TORTS, *supra* note 11, § 2 cmt. a (stating that "[t]he rule for manufacturing defects . . . imposes liability whether or not the manufacturer's quality control efforts satisfy standards of reasonableness. . . . On the premise that tort law serves the instrumental function of creating safety incentives, imposing strict liability on manufacturers for harm caused by manufacturing defects encourages greater investment in product safety than does a regime of fault-based liability under which, as a practical matter, sellers may escape their appropriate share of responsibility. Some courts and commentators also have said that strict liability discourages the consumption of defective products by causing the purchase price of products to reflect, more than would a rule of negligence, the costs of defects.").

43. *See, e.g.*, Smith, *supra* note 4; JOHN VILLASENOR, BROOKINGS INST., PRODUCTS LIABILITY AND DRIVERLESS CARS: ISSUES AND GUIDING PRINCIPLES FOR LEGISLATION (Apr. 24, 2014), https://www.brookings.edu/research/products-liability-and-driverless-cars-issues-and-guiding-principles-for-legislation/ [https://perma.cc/HVR3-N94N].

AI liability.[44] While forthcoming changes in automotive liability resulting from advances in automated driving will likely serve as one of the first trial grounds for AI-based tort litigation, it does not serve as an ideal vector for an analysis of AI products liability generally given the subject-specific intricacies that make automotive tort litigation so unique. Rather than allowing the insurance driven field of automotive torts to passively co-opt the still emergent legal discussions of AI liability, proactive legislative action could not only serve to address many of the concerns raised in this Comment, but also ensure that these legal remedies are efficiently tailored to the unique AI programs in question.

B. Proving Design Defects in Reinforcement Learning Artificial Intelligence

Generally, a manufacturer is liable under products liability law if there was a defect in the manufacturing process, a defect in the design, or a failure to provide sufficient and reasonable warning.[45] While there are multiple standards under which a reinforcement learning AI manufacturer may be found liable, design defect claims arguably pose the greatest challenge for traditional products liability models.[46] Specifically

44. In automotive-tort litigation, a considerable percentage of victims and defendants rely on insurance policies to satisfy their respective recovery and liability questions. *See* Smith, *supra* note 4, at 33 ("A crash victim who has automotive, health, or life insurance might seek payment directly from the provider of that insurance. The victim may additionally or alternatively seek payment from an insurer that provides liability coverage to a would-be defendant. . . . Largely because of automotive insurance, the vast majority of crashes are handled without any litigation.").

45. RESTATEMENT (THIRD) OF TORTS, *supra* note 11, § 2 ("A product is defective when, at the time of sale or distribution, it contains a manufacturing defect, is defective in design, or is defective because of inadequate instructions or warnings. A product: (a) contains a manufacturing defect when the product departs from its intended design even though all possible care was exercised in the preparation and marketing of the product; (b) is defective in design when the foreseeable risks of harm posed by the product could have been reduced or avoided by the adoption of a reasonable alternative design by the seller or other distributor, or a predecessor in the commercial chain of distribution, and the omission of the alternative design renders the product not reasonably safe; (c) is defective because of inadequate instructions or warnings when the foreseeable risks of harm posed by the product could have been reduced or avoided by the provision of reasonable instructions or warnings by the seller or other distributor, or a predecessor in the commercial chain of distribution, and the omission of the instructions or warnings renders the product not reasonably safe.").

46. I focus primarily on design defect claims because the issues raised in these cases will directly concern litigation regarding intended but unforeseen uses of the product's AI systems, as opposed to allegations that the programming itself was flawed, resulting in errant results. Where an AI product acts contradictory to its programing or its advertised or warranted function, traditional products liability models will still suffice to apportion liability. *See* Michael D. Scott, *Tort Liability for Vendors of Insecure Software: Has the Time Finally Come?*, 67 MD. L. REV. 425, 437 (2008) (citing David Polin, *Proof of Manufacturer's Liability for Defective Software*, *in* 68 AMERICAN JURISPRUDENCE PROOF OF FACTS 3d 333, 347 (2002)) ("No reported decision has unequivocally held that a software vendor has breached an express warranty. There are three possible reasons for this: 'First, software manufacturers scrupulously avoid making express claims that software will perform any particular tasks, although they freely claim that their products have nearly mystical qualities. Secondly, any

addressing how manufacturing defects (which he calls "production defects") and design defects can arise in the computer software context, Steven J. Frank offers the following synthesis of software products liability law:

> A product is said to be flawed if it contains "an abnormality or a condition that was unintended, and makes the product more dangerous than it would have been as intended." Production defects may be introduced into a computer program during the mass copying process. Because such an item has failed to meet the manufacturer's own specifications, proof of defectiveness is reduced to a relatively easy matter of comparison.
>
> This contrasts sharply with proof of unreasonable danger in a product's design. In such cases, the plaintiff's burden of proof often differs little from that faced in an ordinary negligence action. Courts have experimented with a plethora of tests to combine the goals of consumer protection with encouragement of technological advance. The consumer contemplation test, for example, examines whether the product failed to perform as safely as an ordinary consumer would anticipate; the danger-utility test weighs the danger inherent in a particular product with the benefits it produces. A design defect may also be introduced by failure to warn about a risk or hazard related to the way a product was designed. The standard in these types of cases is usually straightforward negligence, necessitating proof that a defendant "knew or should have known in the exercise of ordinary care of the risk or hazard about which he failed to warn."[47]

Notably, Frank identifies the inherent difficulties that arise in litigating design defect cases, specifically in the case of computer software:

> Computer software may raise especially nettlesome questions. Unlike most products, computer programs written for broad consumer audiences are suited to a variety of applications and tasks. The degree of danger to which users are exposed is not an intrinsic property of the program, but rather depends on the particular use to which it is put. Normally, manufacturers are held liable only for harm arising from ordinary and reasonably foreseeable uses. The difficulty with this premise is that tort doctrine has developed around a concept of foreseeability which contemplates some type of *mis*use. It is unclear whether courts will apply this reasoning to uses of a program

---

express promises are inevitably disclaimed in licensing agreements. Thirdly, it is generally agreed that software cannot be expected to perform perfectly, so such warranties as exist will be interpreted somewhat loosely.'").

47. Steven J. Frank, *Tort Adjudication and the Emergence of Artificial Intelligence Software*, 21 SUFFOLK U. L. REV. 623, 637–38 (1987).

> which are wholly unforeseeable yet to which the product is entirely suited.[48]

Design defect claims involving reinforcement learning AI programs will pose particularly unique challenges given that both the user and the AI control the systems that will produce these injuries.[49] As previously noted, one of the unique characteristics of reinforcement learning AI systems from a products liability context is that these AI are expressly programmed for aftermarket AI–consumer interaction and alteration.[50] Consequently, at the time of purchase, the AI program only possesses the *potential* to evolve into a dangerous or injurious product in response to the consumer's own usage. Accordingly, a successful design defect claim against a manufacturer would require that the AI was legally defective, not based on an actual existing danger (the *immediate* capability of the program to commit the injurious action that gave rise to the consumers products liability claim), but the manufacturer's failure to mitigate against a potentially dangerous aftermarket evolution.[51] The characterization of the product's condition at the time of sale—namely, whether it is immediately dangerous or inert—is just one of the many potential issues presented. To better capture the potential liability gap posed by emergent AI products, the following three sections explore how existing design defect principles may struggle to account for the unique consumer–AI relationship giving rise to potential products liability claims.

1. The Risk-Utility and Consumer Expectations Test

Though there is minor state-by-state variation, there are generally two tests for determining whether a product is defective due to an unreasonably dangerous (or unreasonably safe) design: the risk-utility test and the consumer expectations test.[52] Notwithstanding, both of these tests are insufficient to address the complexity of AI products that evolve with AI–consumer interaction.

The *Restatement (Third) of Torts* provides a succinct description of the risk-utility test:

> Under risk-utility balancing the likelihood and magnitude of foreseeable harm is balanced against the burden of precaution against

---

48. *Id.* at 638–39.

49. While the user will still possess some legally discernable degree of operational control, the AI system executes countless actions, effectively independent of any immediate user supervision or awareness. *See supra* note 45.

50. *See* Gerstner, *supra* note 8 and discussion therein on after-market adaptation.

51. *See infra* Sections II.B.2 and II.B.3 (discussing the foreseeability of aftermarket consumer alterations).

52. For a state-by-state analysis, see RESTATEMENT (THIRD) OF TORTS, *supra* note 11, § 2 cmt. d.

> the anticipated harm. In a products-liability setting involving product design, the plaintiff's burden of precaution against the harm can only be either adoption of a reasonable safer alternative design or a decision that, even absent an alternative design, the product should not have been marketed at all.[53]

Up front, drawing from the discussion in the introduction to Section B of this Part, one of the most obvious hurdles is the textual focus on foreseeable and anticipated harms. Likewise, the consumer expectations test is no better.[54] While the California model for the consumer expectations test articulated in *Soule v. General Motors Corp.* has retained some use, jurisdictions using the California model represent a considerable minority.[55] In *Soule*, the California Supreme Court limited the consumer expectations test in design defect litigation, specifically in cases involving complex product designs. The court said:

> [A] complex product, even when it is being used as intended, may often cause injury in a way that does not engage its ordinary consumers' reasonable minimum assumptions about safe performance. For example, the ordinary consumer of an automobile simply has "no idea" how it should perform in all foreseeable situations, or how safe it should be made against all foreseeable hazards.
>
> . . . .
>
> As we have seen, the consumer expectations test is reserved for cases in which the *everyday experience* of the product's users permits a conclusion that the product's design violated *minimum* safety assumptions, and is thus defective *regardless of expert opinion about the merits of the design*.[56]

The remainder of consumer expectations jurisdictions simply hold that a product is defective if it is dangerous to an extent beyond that anticipated by the ordinary user, without distinction for the complexity of the product.[57] Analyzing the consumer expectations test, Professor Mary J. Davis highlights three issues relevant in the immediate context:

---

53. *Id.*

54. Notably, the RESTATEMENT (THIRD) OF TORTS formally abandons the consumer expectations test for design defect claims. *See id.* § 2 cmt. g ("[C]onsumer expectations do not constitute an independent standard for judging the defectiveness of product designs."). Nonetheless, the test is still used in certain jurisdictions.

55. *See* Ray v. BIC Corp., 925 S.W.2d 527, 533 (Tenn. 1996).

56. Soule v. Gen. Motors Corp., 882 P.2d 298, 308 (Cal. 1994).

57. *See, e.g.*, Gen. Motors Corp. v. Farnsworth, 965 P.2d 1209 (Alaska 1998); Ontai v. Straub Clinic & Hosp., Inc., 659 P.2d 734 (Haw. 1983); Malcom v. Evenflo Co., Inc., 217 P.3d 514 (Mont. 2009); Kudlacek v. Fiat, 509 N.W.2d 603 (Neb. 1994).

> Few courts adhere closely to the letter of section 402A's consumer expectations test in proving design defect. The test has proved unworkable for a variety of reasons. First, it connotes a contract-based liability, encouraging the jury to rely intuitively on principles of bargaining and warranty. Second, if the product contains a defect which is apparent or obvious, the consumer expectations arguably include the apparent danger, preventing liability and therefore discouraging product improvements which could easily and cost-effectively alleviate the danger. Third, bystanders, who are widely recognized as protected by both tort and contract theories of products liability regardless of privity, cannot be said to have any expectations about a product which causes them injury.[58]

Whether applying the risk-utility test, the model of the consumer expectations test outlined in *Soule*, or the more traditional consumer expectations test that does not expressly account for the complexity of the product, the introduction of reinforcement learning AI systems poses new challenges for both consumer-plaintiffs and defendant-manufacturers in design defect claims. Under the risk-utility test, consumer-plaintiffs must demonstrate the availability of a reasonable alternative design.[59] Under both tests, both parties are required to litigate on the question of foreseeable consumer uses and harm.[60] However, as previously discussed, the complexity of these underlying AI systems and their inherently evolutionary nature make demonstrating the availability of alternative

---

58. Mary J. Davis, *Design Defect Liability: In Search of a Standard of Responsibility*, 39 WAYNE L. REV. 1217, 1236–37 (1993).

59. *See, e.g.*, 735 ILL. COMP. STAT. ANN. 5/2-2104 (West 1993 & Supp. 1996) ("If the design of a product or product component is in issue in a products liability action, the design shall be presumed to be reasonably safe unless, at the time the product left the control of the manufacturer, a practical and technically feasible alternative design was available that would have prevented the harm without significantly impairing the usefulness, desirability, or marketability of the product . . . ."), *invalidated by* Best v. Taylor Mach. Works, 689 N.E.2d 1057 (1997); LA. STAT. ANN. § 9:2800.56 (2018) ("A product is unreasonably dangerous in design if, at the time the product left its manufacturer's control: (1) There existed an alternative design for the product that was capable of preventing the claimant's damage; and (2) The likelihood that the product's design would cause the claimant's damage and the gravity of that damage outweighed the burden on the manufacturer of adopting such alternative design and the adverse effect, if any, of such alternative design on the utility of the product. An adequate warning about a product shall be considered in evaluating the likelihood of damage when the manufacturer has used reasonable care to provide the adequate warning to users and handlers of the product."); OHIO REV. CODE ANN. § 2307.75(F) (West 2018) ("A product is not defective in design or formulation if, at the time the product left the control of its manufacturer, a practical and technically feasible alternative design or formulation was not available that would have prevented the harm for which the claimant seeks to recover compensatory damages without substantially impairing the usefulness or intended purpose of the product.").

60. *See, e.g.*, Welch Sand & Gravel, Inc. v. O & K Trojan, Inc., 668 N.E.2d 529, 533 (Ohio 1995) ("While a manufacturer is not responsible for all product misuses, failure to design a product to prevent a foreseeable misuse can be a design defect."); Rife v. Hitachi Const. Mach. Co., 609 S.E.2d 565 (S.C. Ct. App. 2005).

programming designs exceedingly difficult.[61] Similarly, the fact that these programs are designed to interact and evolve with the user seriously challenges the traditional understanding and application of what harms and uses are foreseeable.

2. "Available Alternative Design" and "Foreseeability"

Which potential consumer–AI interactions (and subsequently acquired data) are foreseeable, and which corresponding obligations does a manufacturer have to mitigate against any potential derivative harm through an alternative design? Underpinning the questions of liability adhering to unreasonably dangerous design defects in AI systems is whether the product was used in a foreseeable manner that proximately led to the injury.[62] While a specific AI may be designed to consider the user's attention or intentions (such as a theoretical self-driving technology that learns a driver's average response time), I argue, I think non-controversially, that the corresponding, individual *human* decision-making processes that these AI will inherently depend upon are notoriously flawed.[63] It is only logical to assume that rational machines will experience some difficulty when interacting with occasionally irrational human actors and experience similar difficulties in mitigating the corresponding safety risks generated by irrational human decision-making.

Following traditional principals of products liability law, some may argue that it should fall upon the manufacturers of AI systems to fill the resulting foreseeability gap generated by these new programs. The manufacturers should predict and mitigate against the broad range of theoretically possible, but limitedly foreseeable, consumer data inputs that could give rise to potential harm. While apportioning this responsibility—and any attendant liability—in this fashion mirrors current products liability theories, it is again critical to consider the precise nature of how reinforcement learning AI operate. The system takes original data inputs, identifies patterns, and subsequently *creates* new patterns in a recursive process, which are then applied to new data. Like a series of perpetually branching forks along a path, an AI system progresses through an iterative analytical process to identify and incorporate new information.[64] The

61. Frank, *supra* note 47.

62. *Id.* at 637.

63. Few readers will need to delve deep into their memory to recall a moment when they (or someone else) made an inarguably poor decision. With a reinforcement learning AI dependent on consumer input, such poor human decision-making will be integrated into the AI system's learning process.

64. See *Hearings*, *supra* note 1, at 22, where Dr. Andrew Moore, writing as Dean of the School of Computer Science at Carnegie Mellon University, provides a succinct but helpful analysis of how

exponentially expansive number of variables raised by each new data input make it highly impractical for manufacturers to attempt to predict all individual potentialities.

Across jurisdictions, tension already exists with respect to how issues of foreseeability are addressed, specifically in cases involving complex products. In most jurisdictions, the common defense for complex product manufacturers is the assertion that the product, at the time of sale, was "state of the art."[65] This term itself is subject to a multitude of practical definitions, from a meaning of "industry custom" or "industry practice," to meaning "the safest existing technology that has been adopted for use," to meaning "cutting-edge technology."[66] Regardless of the term's specific application, the existence of a safer design that has been adopted and put to actual use is relevant to the question of whether a reasonable alternative design was available.

Interconnected with the question of whether a product's design was reasonably safe is the question of foreseeable use and harm. However, jurisdictional tensions exist here as well on the meaning of "foreseeable." For example, under South Carolina law, the test of foreseeability is whether the injury is the natural and probable consequence of the complained-of act: for an act to be the proximate cause of the injury, the injury must be a foreseeable consequence of the act.[67] However, although foreseeability of some injury resulting from the manufacturer's act or omission is a prerequisite of proximate cause in South Carolina, the plaintiff need not prove that the manufacturer should have contemplated the *specific event* leading to the injury.[68] In contrast, New Jersey applies a standard of objective foreseeability, which does not assign liability for future events that are only theoretically or just possibly foreseeable.[69] Foreseeability applies only to future occurrences that objectively and

---

the development of deep learning techniques in machine learning is increasing the ability of programs to automatically generate new models for large and complex data sets.

65. *See, e.g.*, Fibreboard Corp. v. Fenton, 845 P.2d 1168, 1174 (Colo. 1993) (en banc) ("State-of-the-art would be an applicable factor in a design-defect case, if the alternative design suggested by the plaintiff was not practically feasible in light of the state of the art at the time the product was manufactured."); Heath v. Sears, Roebuck & Co., 464 A.2d 288, 299 (N.H. 1983) (finding that it was "both reasonable and constitutionally permissible to raise an affirmative defense based upon 'discoverability of risk' as measured by the 'state of the art' at the time of distribution or sale").

66. RESTATEMENT (THIRD) OF TORTS, *supra* note 11, § 2 cmt. d.

67. *See* Rife v. Hitachi Const. Mach. Co., 609 S.E.2d 565, 569 (S.C. Ct. App. 2005).

68. *Id.*; *see also* Speaks v. Mazda Motor Corp., 118 F. Supp. 3d 1212, 1224 (D. Mont. 2015) ("[I]f it is reasonably foreseeable to a defendant that its product can be or is being used in a specific manner, and a consumer is injured by using the product in that manner, the defendant cannot argue that the plaintiff had misused its product.") (quoting Kenser v. Premium Nail Concepts, Inc., 338 P.3d 37, 43 (Mont. 2014)).

69. *See* Port Auth. of N.Y. & N.J. v. Arcadian Corp., 991 F. Supp. 390, 400–01 (D.N.J. 1997), *aff'd*, 189 F.3d 305 (1999).

reasonably could have been anticipated when a product was manufactured.[70]

Under South Carolina law, AI manufacturers appear to be left considerably exposed to liability for injuries resulting from a wholly indeterminate body of potential consumer–AI interactions. Given that reinforcement learning AI systems are specifically intended to interact with the consumer (because the greater the number and diversity of those interactions will, theoretically, provide better data), will the law hold that these manufacturers have effectively programmed themselves into a legal corner where they have de facto "foreseen" these interactions? How many iterative learning cycles, consumer–AI interactions, or years must pass before the consumer's use and injury is deemed too speculative to be foreseeable?

Conversely, under New Jersey law, consumers face huge evidentiary hurdles to effectively prove their use and injury were objectively foreseeable. While particularly problematic for consumers, AI product liability litigation also poses an incredible challenge for manufacturers and the courts. Given both the complexity of these programs and the fact that plaintiffs in design defect cases will need to establish both that the program was unreasonably dangerous as to be legally defective and that the alleged defect was the proximate cause of their injury, it is conceivably unavoidable that virtually all AI products liability litigation will require expert testimony and similarly burdensome discovery.[71] Beyond the cost implications of hiring expert witnesses, the evolving nature of AI reinforcement learning programs pose new challenges not inherent in traditional computer software and software litigation. Simply tracing and identifying the specific data inputs and subsequent program developments allegedly proximately responsible for the plaintiff's act and injury may prove too challenging or cost prohibitive for individual plaintiffs seeking small to moderate recovery in many cases.[72]

3. Consumer Alteration and Misuse

Another area of complexity that reinforcement learning AI programs introduce into existing products liability law is the issue of consumer alteration and misuse. To what extent may a manufacturer argue that consumer alteration is an intervening cause when the product in question is expressly designed and marketed for aftermarket consumer alteration? State by state treatment of these products liability theories demonstrate the

70. *Id.*

71. *See generally* Frank, *supra* note 47 (addressing the general complexities of litigating design defect claims over computer software).

72. *See id.*; *see also* discussion *infra* pp. 19–21 (addressing plaintiffs' potential cost burdens).

potential complexities. Specifically, where product alteration or misuse is involved, individual jurisdictions have held that the manufacturer is liable for those injuries that were foreseeable, preventable, or mitigatable.[73] Will the fact that reinforcement learning AI are designed for alternation effectively remove this defense provided that the plaintiff-consumers are using the product as marketed?[74] How may well intentioned but grievously incompetent user interactions that fall outside of the manufacturers proscribed list of uses be evaluated by the court? I believe a proactive legislative approach that considers these questions as part of its apportionment of liability could help mitigate these issues.

III. PROHIBITIVE COSTS OF EXPERT LITIGATION

Another immediate issue raised by the involvement of reinforcement learning AI programs in products liability cases is cost. With the exception of a per se liability standard for defendant-manufacturers, consumer-plaintiffs who suffer tortious injury from a reinforcement learning AI program-based product may be required to demonstrate that

---

73. *See Arcadian Corp.*, 991 F. Supp. at 400; N.C. GEN. STAT. ANN. § 99B-3(a) (West 2018) ("No manufacturer or seller of a product shall be held liable in any product liability action where a proximate cause of the personal injury, death, or damage to property was either an alteration or modification of the product by a party other than the manufacturer or seller, which alteration or modification occurred after the product left the control of such manufacturer or such seller unless: (1) The alteration or modification was in accordance with the instructions or specifications of such manufacturer or such seller; or (2) The alteration or modification was made with the express consent of such manufacturer or such seller."); N.D. CENT. CODE ANN. § 28-01.3-03 (West 2018) ("No manufacturer or seller of a product may be held liable in any products liability action in which a substantial contributing cause of the injury, death, or damage to property was an alteration or modification of the product, which occurred subsequent to the sale by the manufacturer or seller to the initial user or consumer, and which changed the purpose, use, function, design, or intended use or manner of use of the product from that for which the product was originally designed, tested, or intended."); TENN. CODE ANN. § 29-28-108 (West 2018) ("If a product is not unreasonably dangerous at the time it leaves the control of the manufacturer or seller but was made unreasonably dangerous by subsequent unforeseeable alteration, change, improper maintenance or abnormal use, the manufacturer or seller is not liable."); UTAH CODE ANN. § 78B-6-705 (West 2018) ("[F]ault shall include an alteration or modification of the product, which occurred subsequent to the sale by the manufacturer or seller to the initial user or consumer, and which changed the purpose, use, function, design, or intended use or manner of use of the product from that for which the product was originally designed, tested, or intended."); *see also* Lindholm v. BMW of N. Am., LLC, 862 F.3d 648, 652 (8th Cir. 2017) (holding that under South Dakota law while manufacturers can be liable for a customer's reasonably foreseeable misuse, a manufacturer cannot be liable for misuse that it cannot reasonably anticipate); Ervin v. Cont'l Conveyor & Equip. Co., 674 F. Supp. 2d 709, 724 (D.S.C. 2009) (holding that under South Carolina law "liability may be imposed upon a manufacturer or seller notwithstanding [the consumer's] subsequent alteration of product when the alteration could have been anticipated by the manufacturer or seller") (internal citation omitted).

74. Were a consumer to deliberately alter the product by accessing and altering its programming or, more traditionally, by misusing the product in a manner warned against in the product's warranty, there is no reason to assume such consumer alterations or misuse would not still be an available defense.

the injury was caused by the flawed programming installed by the manufacturer, not by their faulty, negligent or unprescribed use, input, or "teachings." However, such proof will almost invariably need to be offered through expert witnesses.

Current estimates for products liability expert witnesses in Washington run at approximately $250 per hour for initial case review and approximately $275 per hour for depositions and court appearance fees.[75] For computer expert witnesses in Washington, estimates are higher at approximately $350 per hour for initial case review and over $400 per hour for deposition and trial testimony.[76] However, expert witness costs can easily surpass over $1,000 per hour, with case totals reaching into the six figures after accounting for the cost of discovery and data analysis.[77] Given both the virtual requirement for expert analysis in any products liability claims rooted in AI programming error and the high costs these experts run, many individual plaintiffs with small to moderate claims may be cost-barred from pursuing their claims absent some fee shifting provision or other legislative action.[78]

As noted by Bryant Smith in his analysis of automotive-based AI and existing liability models:

> [R]equiring the plaintiff to specifically demonstrate how and why an automated driving system performed poorly and should have performed better could impose technical and financial barriers to many claims, especially those involving comparatively minor injuries. On the other hand, permitting the plaintiff to use the consumer expectations test, the malfunction doctrine, or res ipsa loquitor could make it easier to attribute undesirable outcomes to something within the automated driving system. In that case, the defendant automaker, rather than the plaintiff, might offer a more detailed explanation of the automated driving system's performance in order to shift some costs to other parties.[79]

---

75. *Expert Witness Fee Calculator*, EXPERT INST., https://www.theexpertinstitute.com/expert-witness-fees/ [https://perma.cc/3KEX-4C3E].

76. *Id.*

77. Marc Davis, *For an Expert Witness, Consider Reputation, Location, and Cost*, ABA J. (Nov. 2016), http://www.abajournal.com/magazine/article/choosing_expert_witness [https://perma.cc/4GAK-QSFW].

78. Additionally, I think it is worth noting that many AI products liability claims similar to the kind discussed in this Comment may often not be amenable to class action procedures. Federal Rule of Civil Procedure 23(a) requires, in part, that a prospective class demonstrate both commonality and typicality. Given the special user–AI interactions central to reinforcement learning design, I believe many claims will naturally require individual, fact-specific inquiries into the consumer's particular product use, defeating class commonality and typicality. *See generally* Wal-Mart Stores, Inc. v. Dukes, 564 U.S. 338 (2011).

79. Smith, *supra* note 4, at 51–52.

Given the complex nature of AI programming, it seems inescapable that parties in almost all cases will need to rely on expert testimony. These costs may very well prove prohibitive to consumer-plaintiffs—even those with strong cases. While fee-shifting statutes may offer some relief, the possibility of footing the bill for not one, but two expert witnesses may still drive consumer-plaintiffs with real injuries and valid claims from pursuing litigation. Unregulated and unbalanced, these high costs may drive many consumer-plaintiffs to settle, ultimately disincentivizing manufacturers of flawed AI-based products from making necessary, injury-preventing changes. Rather than waiting for the courts or the plaintiffs' bar to adapt to these rising challenges, proactive government action is better suited to ensure legitimate future claims do not go unresolved.

CONCLUSION

In concluding their public apology for Tay's failures, Microsoft offered a guarded yet aspirational analysis of both the fundamental challenges of AI programming and the future of AI–human interaction.

> AI systems feed off of both positive and negative interactions with people. In that sense, the challenges are just as much social as they are technical. We will do everything possible to limit technical exploits but also know we cannot fully predict all possible human interactive misuses without learning from mistakes. To do AI right, one needs to iterate with many people and often in public forums.[80]

The AI–consumer interactive element at the heart of Tay's design is emblematic of the many possible ways that consumers can expect increased engagement with new AI technology in the coming years.[81] As these emerging AI products are inherently dependent on direct AI–consumer interactions to be functional—let alone successful—both the industry and the legal field need to begin making a conscious effort to better understand when, where, and how risk and liability are being apportioned.[82]

As the AI industry continues to grow and develop at breakneck speeds, the tension between the new realities of AI–consumer interactions

80. Lee, *supra* note 13.

81. Notably, three key elements distinguish Tay from the reinforcement learning AI systems discussed in this Comment: the number of disassociated users supplying data to a single AI system, the deliberately negative character of the data supplied, and the atypical nature of the resulting harm. Tay's story is illustrative of the capabilities and general function of the AI technology in question, not the corresponding liability analysis.

82. Absent these discussions, these programs will continue to enter the market in a legal vacuum, and the public and consumers will be left to hope that developers and manufacturers will choose prudence over profits. *See* Metz, *supra* note 18.

and traditional liability laws will likewise grow, potentially threatening the viability of an invaluable industry. Pitfalls, both new and old, pose unique challenges. Traditional concerns, such as the failure of manufacturers to conduct diligent product testing and analysis, can lead to disastrous results for both consumers and manufacturers alike. Businesses will always worry about the liability lurking in an increasingly unforeseeable future driven by AI programs. Products liability laws must still function to protect the consumer from harm by encouraging businesses to act appropriately to mitigate against foreseeable risks. Nonetheless, the ability of traditional products liability laws to maintain this balance is drawn into serious question by the advent of new technologies that redefine the relationship between manufacturer, consumer, and product.[83]

While conceptually intriguing, the current emphasis on AI autonomy and the law of agency is too limited in its practical application to account for the current trajectory of recent advancements in AI technology and the relationship between these products and their consumers. By effectively attributing de facto liability to manufacturers of non-autonomous AI systems, a generalized agent-principal analysis ignores the complexity of the relationship between manufacturers, consumers, and AI systems. As a growing number of reinforcement learning AI products are specifically created for aftermarket consumer interaction, manufacturers cannot avoid sacrificing their ability to mitigate foreseeable harm while consumers face an increasingly insurmountable evidentiary burden to establishing the proximate cause of their injuries, even under strict liability models. It is not unheard of for the legal system to respond to specific industries or products with specialized statutes.[84] Rather than bending over backwards to apply existing products liability laws or permitting the haphazard growth of industry-specific AI liability regimes, state, legal, and political professionals should heed the counsel of industry experts and work proactively to adopt a tailored model that reflects the emergent reality of today's AI.[85]

---

83. *See, e.g.*, RESTATEMENT (THIRD) OF TORTS, *supra* note 11, § 2 cmt. a.

84. Perhaps one of the best examples is with respect to laws regarding asbestos. *See., e.g.*, Asbestos v. Bordelon, Inc., 726 So. 2d 926, 941 (La. Ct. App. 1998) ("Under *Halphen*, the plaintiffs . . . are not required to provide a showing of the manufacturer's negligence in a products liability case, but rather must establish the manufacturer's legal relationship to the 'defective' product. Consequently, the manufacturer cannot offer his knowledge or lack thereof concerning the condition of the product as a method to exonerate it from liability.").

85. *Hearings*, *supra* note 1. In addition to the statements by Dr. Horvitz and Dr. Moore, Mr. Brockman and Dr. Steve Chin, the Technical Group Supervisor of the Artificial Intelligence Group and the Senior Research Scientist in the Mission Planning and Execution Section at NASA's Jet Propulsion Laboratory, specifically raised the need for and importance of further involvement from the legal and political communities to help sustain and grow America's Artificial Intelligence industries. *Id.*