The Foreseeability of Human–Artificial Intelligence Interactions[*]

Consider the following hypotheticals:

(1) A hospital uses artificial intelligence software to analyze a patient's medical history and make a determination as to whether he or she needs surgery. One day, the artificial intelligence software incorrectly diagnoses a patient and recommends an unnecessary surgery. In preparation for the surgery, an anesthesiologist applies an incorrect dosage of the surgical anesthetic and kills the patient.

(2) An investment firm uses artificial intelligence software to identify promising stocks for investment. Without any further research, an investment banker negligently recommends stocks off of the software's prepared list. Those stocks go bust, costing their new owners thousands of dollars.

(3) A vehicle with autonomous-driving software is cruising down a two-lane road. The lane to its right is filled with cars driving in the same direction. A human driver is in oncoming traffic and recognizes the autonomous car as being from a notable autonomous car brand. The human driver decides it would be fun to "play chicken" with the car to see how it will react. The human driver proceeds to swerve into the autonomous vehicle's lane and the autonomous vehicle, *thinking* it best to avoid a head-on collision and not realizing the human driver won't hit it, swerves into the right lane, triggering a collision with an innocent third-party car.

(4) A delivery drone, piloted with autonomous-piloting software, is en route to deliver a package. On its way, it passes the home of a paranoid man who is very concerned with his privacy. He proceeds to take a baseball, and with an impressive throw, knocks the drone out of the sky. The drone crashes down and hits a child playing in a nearby park.

(5) A company selling artificial intelligence software sells its product to a racist. The racist proceeds to install the software onto a robot butler, and the robot butler proceeds to learn and develop under the teachings of its owner. One day, a black UPS driver delivers a package to the front door. The now-racist robot answers the door and upon seeing the black UPS driver, *thinks*, "The only reason a

> black person would be on my front porch would be if he were here
> to burgle my owner." The robot proceeds to attack the UPS driver
> under the mistaken assumption that he is a burglar.

In each of the above hypotheticals, the use of artificial intelligence led to the injury of an innocent person. When faced with an injury caused by another, each of these persons may seek a remedy through the tort system. The tort system is designed to provide monetary damages for injured parties when they are harmed by the negligent conduct of another.[1] In this way, the tort system assures that the costs of negligent conduct lie with those responsible for causing the injury.[2] Each injured party in the hypotheticals above can sue the negligent actor who caused the harm—but who (or what) exactly caused the injured party's harm? In the above hypotheticals, there are human actors who cause the injured party's harm through obviously negligent conduct or even intentional conduct. These human actors present themselves as obvious targets, but what about the developers of the artificial intelligence software? When the injured parties sue in court, they are likely to sue whomever has the deepest pockets.[3] This should strike fear into the hearts of many artificial intelligence companies, because in these tort suits, they are likely to be the parties in the best financial position to pay out damages.

If artificial intelligence companies are sued for the negligent development of their software, courts will be faced with a difficult question of foreseeability. When proving a case of negligence, plaintiffs are required to show the harm that occurred was a foreseeable consequence of the defendant's negligent conduct.[4] This is also called satisfying the proximate cause requirement of a negligence case.[5] In each hypothetical, was the interaction between the artificial intelligence software and human actor foreseeable? How does the liability of the software developer fit in? Technology as a whole has grown exponentially over time and artificial intelligence technology will be no exception.[6] New advances in machine learning coupled with other continuing developments in artificial intelligence

---

1. *See* W. PAGE KEETON ET AL., PROSSER AND KEETON ON THE LAW OF TORTS 6 (5th ed. 1984) (describing that the goal of the tort system is to make victims whole again at the expense of tortfeasors).

2. *Id.* at 7.

3. *See* Robert MacCoun, *Is There a "Deep-Pocket" Bias in the Tort System?*, 1993 RAND INST. CIV. J. 1, 2–3 (examining the "deep-pocket bias," which suggests juries award higher damages against corporations).

4. *See* DAVID W. ROBERTSON ET AL., CASES AND MATERIALS ON TORTS 162–63 (4th ed. 2011) (explaining that the trier of fact can "find proximate cause whenever the plaintiff's injury was among the array of risks the creation or exacerbation of which led to the conclusion that the defendant's conduct was negligent").

5. *Id.*

6. Hans Moravec, *When Will Computer Hardware Match the Human Brain?*, J. EVOLUTION & TECH., Mar. 1998, at 1, 1.

software will increase the prevalence of artificial intelligence in our lives, making it important to discuss the question of who will be liable when this new technology causes injury.[7] And in each of these incidents, the presence of artificial intelligence will force us to address the difficult question of whether a human's interaction with artificial intelligence was foreseeable or unexpected.[8]

Many forms of artificial intelligence, including autonomous vehicles, employ machine learning.[9] Machine learning departs from software coding in the conventional sense and begins to look more like coaching than it does programming.[10] As the software interacts with the world, it looks to see which of its actions create the most successful results. It then incorporates its most successful actions into future behavior.[11] In this way, the software *evolves* over time. A new artificial intelligence software is not unlike the brain of a human child—ready to be molded and shaped by its experiences.[12] When the software developer places the artificial intelligence into the real world, the developer cannot predict how the artificial intelligence will solve the tasks and problems it encounters. The machine will teach itself how to solve obstacles in ways that are unpredictable.[13] A side effect of humans coaching machines rather than coding line by line will be an inherent amount of unpredictability and a lack of control over the software by the developer once the software is sold.[14] Due to this unpredictability, some have suggested that the experiences of a learning artificial Intelligence will cause the artificial Intelligence's interactions with humans to be so unpredictable that they "could be viewed as a superseding cause—that is, 'an intervening force . . . sufficient to prevent liability for an actor whose tortious conduct was a factual cause of harm'—of any harm that such systems cause."[15]

---

7. *Id.* at 1, 7–8.

8. *See* Matthew U. Scherer, *Regulating Artificial intelligence Systems: Risks, Challenges, Competencies, and Strategies*, 29 HARV. J.L. & TECH. 353, 365 (2016) (stating that, as artificial intelligence systems develop, it is "all but certain that issues pertaining to unforeseeable AI behavior will crop up with increasing frequency").

9. Jason Tanz, *Soon We Won't Program Computers. We'll Train Them Like Dogs.*, WIRED (May 17, 2016), https://www.wired.com/2016/05/the-end-of-code/ [https://perma.cc/NJ4E-XALV].

10. *Id.*

11. *Id.*

12. *Id.*

13. Scherer, *supra* note 8, at 365–66.

14. *See* Jonathan Tapson, *Google's Go Victory Shows AI Thinking Can Be Unpredictable, and That's a Concern*, CONVERSATION (Mar. 17, 2016), https://theconversation.com/googles-go-victory-shows-ai-thinking-can-be-unpredictable-and-thats-a-concern-56209 [https://perma.cc/3MNK-89L6] (discussing how artificial intelligence technology such as Google's AlphaGo behaves unpredictably).

15. Scherer, *supra* note 8, at 365; *see generally* WENDELL WALLACH & COLIN ALLEN, MORAL MACHINES: TEACHING ROBOTS RIGHT FROM WRONG 197–214 (2009) (discussing liability in relation to artificial intelligence).

Because the conduct of artificial intelligence depends on external influences after the code is out of the hands of the developers, external influences are an actual cause of any bad behavior by artificial intelligence systems. Therefore, the superseding cause doctrine could swoop in, label the situation as unforeseeable, and save the artificial intelligence developer from liability.[16] Even companies selling artificial intelligence products and developing new artificial intelligence software operate as though defects in their own systems are shielded from incurring liability by the superseding cause doctrine.[17] When reliance on the Tesla autopilot system resulted in a fatal crash in June of 2016, Tesla quickly tried to shield itself from liability by pointing out that the negligent interactions of the driver were a more immediate cause of the crash, than the actions of the programmer.[18]

Other commentators suggest that superseding cause will have no place in protecting the developers of artificial intelligence software, and such software developers will be wholly liable for the actions of their systems.[19] Placing artificial intelligence tort cases in the extremes of total liability or no liability at all is unwise. It is likely that an intervening cause won't entirely shield a developer of artificial intelligence software from liability but will reduce liability as a consideration in a comparative fault analysis. The problem of superseding cause is a familiar one, and while new cases may be cloaked in unfamiliar facts with the advent of artificial intelligence, old case law is applicable to give a good idea of how courts will respond to these new problems.[20]

This Note is offered to clarify misconceptions and uncertainty about the interplay between artificial intelligence and the superseding cause doctrine. This Note concludes that the superseding cause doctrine has begun disappearing from tort analysis and therefore will be unavailable to

---

16. Scherer, *supra* note 8, at 365–66.

17. *See* Matt McFarland, *Who's Responsible When an Autonomous Car Crashes?*, CNN TECH (July 7, 2016), http://money.cnn.com/2016/07/07/technology/tesla-liability-risk/index.html [https://perma.cc/B992-4F8P] (discussing Tesla's perspective on liability stemming from self-driving auto accidents); *see also* Bill Vlasic & Neal E. Boudette, *As U.S. Investigates Fatal Tesla Crash, Company Defends Autopilot System*, N.Y. TIMES (July 12, 2016), https://www.nytimes.com/2016/07/13/business/tesla-autopilot-fatal-crash-investigation.html [https://perma.cc/EXM5-NBZ7] (discussing defects in Tesla's self-driving automobiles); Interview with Akshay Sabhikhi, Chief Exec. Officer, Cognitive Scale (Apr. 3, 2017) (audio recording on file with author).

18. The Tesla Team, *A Tragic Loss*, TESLA BLOG (June 30, 2016), https://www.tesla.com/blog/tragic-loss [https://perma.cc/94SX-RJCJ].

19. Omri Ben-Shahar, *Should Carmakers Be Liable When A Self-Driving Car Crashes?*, FORBES (Sept. 22, 2016), https://www.forbes.com/sites/omribenshahar/2016/09/22/should-carmakers-be-liable-when-a-self-driving-car-crashes/#5acffecb48fb [https://perma.cc/U22K-B4H4].

20. *See generally* Nicolas Petit, Law and Regulation of Artificial Intelligence and Robots: Conceptual Framework and Normative Implications (Mar. 9, 2017) (unpublished manuscript), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2931339 [https://perma.cc/9TGQ-TS2F] (proposing a framework to regulate artificial intelligence based on existing legal principles).

completely shield artificial intelligence software developers from liability. Defendants that once would have escaped liability under the shield of the superseding cause doctrine will now likely be subject to a normal proximate cause analysis and will be assigned liability through a comparative fault system. With liability in future tort cases a probable reality for many software companies, those companies will need to take steps to reduce the incidents their software could cause, or figure out how to protect themselves in the legal system. Without taking steps to protect themselves from liability, artificial intelligence companies could be forced to shut their doors. Such a result would be negative not only for the companies but for society as a whole, which benefits from the innovation of artificial intelligence software developers. In order to strike a balance between protecting individuals from the potential harms of artificial intelligence and encouraging companies to develop such technology, companies must carefully evaluate the foreseeable risks of the technology they are entering into the market and take steps to minimize those risks. If companies take these steps, they will not only help to minimize their eventual liability but ensure that their artificial intelligence software is ready for the human world in which we live.

I.   The Law of Superseding Causes

The superseding cause doctrine has a long history in the courts, and over time, substantial case law has developed cataloguing its many changes. While the increasing presence of artificial intelligence brings new factual scenarios where artificial intelligence causes injury, the court system is engineered to resolve new ambiguities in the law.[21] Courts have faced new and disruptive technologies many times before and proved that they are capable of addressing these issues.[22] So despite new factual scenarios that artificial intelligence tort cases will bring, the robust case law on the superseding cause doctrine will likely still be applicable:

> The peculiarities of each innovation have been worked out by the common law on a case-by-case basis until a legal consensus is reached. While legislative bodies and government agencies often end up playing catch-up to technological change, the law is a living thing and is capable of evolving with technology. Amongst legal experts

---

21. Adam Thierer, *When the Trial Lawyers Come for the Robot Cars*, SLATE (June 10, 2016), http://www.slate.com/articles/technology/future_tense/2016/06/if_a_driverless_car_crashes_who_is_liable.html [https://perma.cc/YK7B-DUMG].

22. *See* Dylan LeValley, Note, *Autonomous Vehicle Liability—Application of Common Carrier Liability*, 36 SEATTLE U. L. REV. 5, 9–11 (2013) (discussing the legal reactions to airplane autopilots and automated elevators).

there is already widespread agreement that the current liability system is best-placed to handle innovation.[23]

The superseding cause doctrine impacts the tort negligence analysis in three ways, through (1) proximate cause, (2) breach, and (3) comparative fault.[24] The proximate cause element of the negligence analysis is where the superseding cause doctrine has traditionally been applied.[25] The doctrine can label an intervening cause sufficiently unforeseeable, preventing a finding of proximate cause, and thus preventing liability for the alleged tortfeasor. The breach element is characterized by Learned Hand's B<PL formula.[26] The formula seeks to explain that the tort element of breach is a balance between the burden a defendant would have to take to prevent a harm (B), the likelihood of the harm (P), and the size of the harm (L).[27] A defendant will not be considered negligent if the likelihood and size of harm caused by the conduct are not great enough to justify the burden of reforming the conduct in a way to prevent the harm.[28] The unforeseeability of a superseding cause can reduce the probability of harm to such a low value that breach cannot be found, thus eliminating liability for the alleged tortfeasor.

Comparative fault is the theory that a jury should be able to assign each negligent actor in a case a certain percentage of the fault.[29] Thus an intervening cause of harm will eat a percentage of the fault points that the defendant would otherwise be liable for.[30] This initially sounds like good news for a defendant, but may not serve much of a benefit if the jurisdiction has retained joint and several liability.[31] In jurisdictions with joint and several liability, each defendant with any fault points will be liable for the entirety of the damages.[32] The defendants will have to hold each other responsible for paying their fair share (through judicial means if need be).[33] This Note will analyze each of these areas in detail and examine how the modern outlook on superseding cause will apply to fact patterns involving artificial intelligence.

---

23. AM. ASS'N FOR JUSTICE, DRIVEN TO SAFETY: ROBOT CARS AND THE FUTURE OF LIABILITY 9 (Feb. 2017), https://www.justice.org/sites/default/files/Driven%20to% 20Safety%202017%20Online.pdf [https://perma.cc/HJQ5-BADZ].

24. RESTATEMENT (THIRD) OF TORTS § 34 (AM. LAW INST. 2010).

25. *See* Laurence H. Eldredge, *Culpable Intervention as Superseding Cause*, 86 U. PA. L. REV. 121, 124–25 (1937) (describing the evolution of the last-wrongdoer rule as an aspect of causation and its subsequent dissipation in the early part of the 20th century).

26. United States v. Carroll Towing, 159 F.2d 169, 173 (2d Cir. 1947).

27. *Id.*

28. *Id.*

29. ROBERTSON ET AL., *supra* note 4, at 344.

30. RESTATEMENT (THIRD) OF TORTS § 34 (AM. LAW INST. 2010).

31. *See* ROBERTSON ET AL., *supra* note 4, at 372 (stating that "a pure comparative fault system with full joint and several liability . . . would generally benefit plaintiffs . . .").

32. *See id.* at 372 (explaining that defendants in joint and several liability jurisdictions normally are responsible for all "100 fault points").

33. *See id.* at 371 (explaining that defendants are left to work out apportionment of damages between themselves).

II.   Proximate Cause

The superseding cause doctrine establishes that "[w]hen a force of nature or an independent act is also a factual cause of harm, an actor's liability is limited to those harms that result from the risks that made the actor's conduct tortious."[34] The superseding cause doctrine applies with equal force whether the original act was innocent or tortious.[35] When applied, the superseding cause doctrine will protect the defendant from liability.[36]

The doctrine has the most force when there is "serious misconduct by someone other than the defendant [that] . . . [intervenes] between the defendant's negligent conduct and the injury" suffered by the plaintiff.[37] Consider the following examples as illustrative of the doctrine's intended effect. In *Watson v. Kentucky*,[38] the defendant was a rail carrier who negligently caused a tank car filled with gasoline to derail and spill its contents into a nearby street.[39] The gasoline caught fire and exploded, harming the nearby plaintiff.[40] The court held that if a third party intentionally lit the spilled gasoline with a match, that action would be an unforeseeable, superseding cause that would shield the railroad carrier from liability despite its negligence.[41] In *Kent v. Commonwealth*,[42] a police officer was shot by a convicted murderer who had been paroled from a life sentence by the Massachusetts parole board.[43] The officer sued the state claiming that the parole board was negligent in releasing a dangerous prisoner, but the court deemed the intervening act of the murderer sufficiently unforeseeable, thus shielding the state from liability.[44] In *Braun v. New Hope Township*,[45] a farmer broke a "road closed" sign in the middle of a road with his tractor.[46] The township learned the sign was down and reinstalled it, but reinstalled it on the right side of the road instead of the middle, and installed it in a way that was shorter than the sign had been before.[47] When an accident occurred on the road a month later, the court held the farmer was not liable because the township's negligent repair was a superseding cause.[48]

---

34. RESTATEMENT (THIRD) OF TORTS § 34 (AM. LAW INST. 2010).
35. *Id.* § 34, cmt. b.
36. *Id.*
37. ROBERTSON ET AL., *supra* note 4, at 179.
38. 126 S.W. 146 (Ky. 1910).
39. *Id.* at 147.
40. *Id.*
41. *Id.* at 151.
42. 771 N.E.2d 770 (Mass. 2002).
43. *Id.* at 772.
44. *Id.* at 772, 777–78.
45. 646 N.W.2d 737 (S.D. 2002).
46. *Id.* at 739.
47. *Id.*
48. *Id.* at 743.

However, not all intervening causes are created equal. Courts will not grant all intervening causes the status of a superseding cause.[49] "An intervening act may not serve as a superseding cause, and relieve an actor of responsibility, where the risk of the intervening act occurring is the very same risk which renders the actor negligent."[50] Judge Posner explains the idea:

> [T]he doctrine of [superseding] cause is not applicable when the duty of care claimed to have been violated is precisely a duty to protect against ordinarily unforeseeable conduct. . . . And so a hospital that fails to maintain a careful watch over patients known to be suicidal is not excused by the doctrine of [superseding] cause from liability for a suicide, . . . any more than a zoo can escape liability for allowing a tiger to escape and maul people on the ground that the tiger is the [superseding] cause of the mauling.[51]

Yet even when the intentional act by a third party was arguably foreseeable, the courts have historically struggled with intentional bad acts from third parties.[52] For a while, courts tended to always treat intentional bad acts as a superseding cause.[53] The rail carrier case discussed above is an example. In that case, the court said that had the gasoline been ignited intentionally, the ignition would be a superseding cause.[54] If the fire were started negligently, like by a man attempting to light his cigar, then the intervening act would have been a sufficiently foreseeable cause, and the doctrine would not apply.[55] The idea of always applying the superseding cause doctrine when dealing with intentionally bad actors can be defended on the basis of "steering plaintiffs toward the most obviously and immediately responsible tortfeasors and away from others."[56]

The problem of intentionally bad acts worked itself out as courts began to stop looking at superseding cause as a unique self-contained doctrine and began to look at these cases as a class of foreseeability problems entitled to reasoning no different than that of all foreseeability problems.[57] The majority

---

49. *See* Williams v. United States, 352 F.2d 477, 481 (5th Cir. 1965) (holding that foreseeable intervening acts are not superseding causes).

50. Derdiarian v. Felix Contracting Corp., 414 N.E.2d 666, 671 (N.Y. 1980).

51. Jutzi-Johnson v. United States, 263 F.3d 753, 756 (7th Cir. 2001) (citing to DeMontiney v. Desert Manor Convalescent Ctr., 695 P.2d 255, 259–60 (Ariz. 1985), and City of Mangum v. Brownlee, 75 P.2d 174 (Okla. 1938), respectively).

52. *See* RESTATEMENT (THIRD) OF TORTS § 34 cmt. d, reporters' note (AM. LAW INST. 2010) (stating that courts "give considerably more weight to intervening . . . criminal acts [], even when they might well be within the fuzzy boundaries of foreseeability").

53. *Id.* § 34 cmt. e.

54. Watson v. Ky. & Ind. Bridge & R. Co., 126 S.W. 146, 151 (Ky. 1910).

55. *Id.* at 150–51.

56. *See* ROBERTSON ET AL., *supra* note 4, at 187 (discussing Judge Posner's reasoning behind his decision in *Edwards v. Honeywell, Inc.*, 50 F.3d 484 (7th Cir. 1995)).

57. *See* Coyne v. Taber Partners I, 53 F.3d 454, 460–61 (1st Cir. 1995) (holding that attacks from taxi-union protestors were foreseeable consequences of driving a different transportation service through the protest); Stagl v. Delta Airlines, Inc., 52 F.3d 463, 465–66, 473–74 (2d Cir.

of courts now looks at intervening acts and superseding causes as "simply subsets or particular examples of the basic scope of the risk problem [that] can be resolved under ordinary foreseeability rules."[58] Modern cases do tend to obscure this point as judges have a habit of sticking to specialized legal language, even as core ideas change.[59] Rather than debating the policy of counting intentional third party torts as a superseding cause, courts now just ask whether the intervening cause was foreseeable or not, regardless of the level of culpability within the intervening act.[60]

> Indeed, it would be anomalous and inconsistent to, on the one hand, permit an actor to be negligent because of a failure to take adequate precautions in the face of the foreseeable risk of another's misconduct but to then hold that the intervention of culpable human conduct constitutes a superseding cause that prevents the actor's negligence from being a proximate cause of the harm.[61]

As courts have moved from a refocused-breach analysis of proximate cause to an array-of-risks outlook, more and more events are considered foreseeable.[62] The array-of-risks approach states that in order to call a harm suffered by a plaintiff foreseeable, the harm need only be among the array of potential risks the creation or exacerbation of which would lead to finding a breach of duty.[63] On the other hand, a finding of foreseeability under the refocused-breach approach requires a finding that the harm suffered by the

---

1995) (finding that injury from aggressive luggage retrieval at an airport could be a foreseeable result of a flight delay and inadequate regulation of baggage retrieval); Williams v. United States, 352 F.2d 477, 481 (5th Cir. 1965) (stating that "[t]he negligent act of a third party will not cut off the liability of an original wrong-doer if the intervening act is foreseeable").

58. DAN B. DOBBS, THE LAW OF TORTS 460 (1st ed. 2000).

59. *Id.*

60. *See* Summy v. City of Des Moines, 708 N.W.2d 333, 343 (Iowa 2006) (stating that the intervening act that the defendant had a duty to protect against cannot, as a matter of law, constitute sole proximate cause of plaintiff's harm); City of Cedar Falls v. Cedar Falls Cmty. Sch. Dist., 617 N.W.2d 11, 18 (Iowa 2000) (holding that "an intervening force which falls squarely within the scope of the original risk will not supersede the defendant's responsibility" (quoting Hollingsworth v. Schminkey, 553 N.W.2d 591, 598 (Iowa 1996))); Taylor-Rice v. State, 979 P.2d 1086, 1098–99 (Haw. 1999) (finding that negligence in failing to maintain a safe highway guardrail was not superseded by foreseeable inattentive driving by an intoxicated driver); Cusenbary v. Mortensen, 987 P.2d 351, 355 (Mont. 1999) (stating that foreseeable intervening acts "do not break the chain of causation"); Stewart v. Federated Dep't Stores, Inc., 662 A.2d 753, 759 (Conn. 1995) (holding that whether the murder of a shopper in the parking lot of defendant was a superseding cause was a question of foreseeability for the factfinder); Dura Corp. v. Harned, 703 P.2d 396, 402–03 (Alaska 1985) (concluding that an intervening act that is within the scope of the foreseeable risk is not a superseding cause); Largo Corp. v. Crespin, 727 P.2d 1098, 1101, 1103 (Colo. 1986) (holding that "[a]n intentionally tortious or criminal act of a third party does not break the causal chain if it is reasonably foreseeable"); Moning v. Alfono, 254 N.W.2d 759, 766 (Mich. 1977) (finding the negligent marketing of a slingshot to minors encompasses the foreseeable risk that a child will negligently use the slingshot to cause harm to a bystander).

61. RESTATEMENT (THIRD) OF TORTS § 34, cmt. d (AM. LAW INST. 2010).

62. ROBERTSON ET AL., *supra* note 4, at 163, 166.

63. *Id.* at 163.

plaintiff would alone be sufficient for finding a breach of duty.[64] The modern trend is to use the array-of-risks approach, which results in finding many more harms foreseeable.[65] Intervening causes are no exception; the doctrine of superseding cause has lost much of its strength in tort analysis. "So far as [proximate cause] is concerned, it should make no difference whether the intervening actor is negligent or intentional or criminal. Even criminal conduct by others is often reasonably to be anticipated."[66] A trend in the courts today is to allow a finding of foreseeability even despite the seemingly unexpected act of a culpable third party.[67] For example, acts of rape and sodomy of a student have been held potentially foreseeable consequences of negligent supervision by a teacher during a field trip.[68]

The case of *Derdiarian v. Felix*[69] strongly illustrates the modern view of the superseding cause doctrine.[70] In *Derdiarian*, a third party knowingly chose not to take his epilepsy medication and suffered an epileptic seizure while driving, which caused him to lose consciousness and crash the car into a construction site.[71] The plaintiff was struck by the careening car, which knocked him into a container of boiling-hot liquid used on the construction site, turning the plaintiff into a human ball of fire.[72] The plaintiff sued the contracting corporation for failing to take adequate measures to secure the construction site. The court held that the "precise manner of the event need not be anticipated" and sent the question of the accident's foreseeability to the jury.[73]

While the above outlook describes the current trend of superseding cause, some courts do give considerably more weight to intervening criminal acts.[74] These courts will push borderline cases of foreseeability into the unforeseeable category, preventing liability for less culpable defendants.[75] Courts may also take advantage of foreseeability's inherent flexibility to find particularly egregious intervening causes to be unforeseeable even when an honest answer might point the other way.

---

64. *Id.* at 166–67.

65. *Id.* at 163, 166.

66. FOWLER ET AL., HARPER, JAMES AND GRAY ON TORTS 190–92 (3rd ed. 2007).

67. *See, e.g.*, *id.* at 181 (discussing the willingness of courts to find liability where the act was unforeseeable by the negligent party but was not "highly extraordinary").

68. Bell v. Bd. of Educ., 687 N.E.2d 1325, 1327 (N.Y. 1997).

69. 414 N.E.2d 666 (N.Y. 1980).

70. *Id.* at 671.

71. *Id.* at 668.

72. *Id.*

73. *Id.*

74. *See* Hibma v. Odegaard, 769 F.2d 1147, 1156 (7th Cir. 1985) (finding the intervening acts of prison inmates who raped plaintiff were superseding causes that prevented liability from the sexual assaults for the law enforcement officers who framed plaintiff for crimes he did not commit).

75. *See id.*

The cases described above should give software developers in the artificial intelligence community pause. With artificial intelligence all around us, the software will be involved in incidents that cause harm. This is an inevitable reality of developing technology capable of use in so many contexts.[76] Artificial intelligence systems are able to perform complex tasks, such as building investment portfolios or driving cars, without human supervision.[77] The complexity of artificial intelligence software will continue to increase rapidly, and more and more tasks will be left in the hands of the machines, including most jobs.[78]

In many of the incidents involving artificial intelligence software, the artificial intelligence will have merely set the stage, giving an intervening cause the opportunity to create harm. However, as shown above, when the negligent act creates an *opportunity* for harm caused by a third party, the negligent act can be held liable as long as the third party's conduct was foreseeable. Look again at the examples outlined at the beginning of this Note. The harms created by third-party actors are only possible because of the conduct of the artificial intelligence software. When software developers sell their artificial intelligence, it could be foreseeable that third parties would interact with the software in a way that could cause harm.[79] Because of the foreseeability of these intervening causes, modern courts likely won't use the superseding cause doctrine to protect the artificial intelligence developers. The question of foreseeability of the third-party conduct in response to the artificial intelligence will go to the jury, where a finding of liability is very possible.

Artificial intelligence developers may try to argue that the precise manner in which the software reacts to human influence is unable to be anticipated. Therefore, the interactions could never be foreseeable. After all, with self-learning programs, artificial intelligence is "designed to act in a manner that seems creative, at least in the sense that the actions would be deemed 'creative' or as a manifestation of 'outside-the-box' thinking if

---

76. *But see* Russell Brandom, *Humanity and AI Will Be Inseparable*, VERGE (Nov. 15, 2016), https://www.theverge.com/a/verge-2021/humanity-and-ai-will-be-inseparable [https://perma.cc/YU4E-EAVG] (emphasizing the ways in which artificial intelligence may be able to protect humans from harm).

77. *See* Neil Johnson et al., *Abrupt Rise of New Machine Ecology Beyond Human Response Time*, SCI. REPORTS (Sept. 11, 2013), https://www.nature.com/articles/srep02627 [https://perma.cc/R5X6-M65N] (using a software package to conduct a study of ultrafast extreme events in financial market stock prices).

78. STUART J. RUSSELL & PETER NORVIG, ARTIFICIAL INTELLIGENCE: A MODERN APPROACH 1034 (3d ed. 2010) (explaining the ability of artificial intelligence to quickly replace humans in many jobs).

79. *See* Gary Lea, *Who's to Blame When Artificial Intelligence Goes Wrong?*, CONVERSATION (Aug. 16, 2015), https://theconversation.com/whos-to-blame-when-artificial-intelligence-systems-go-wrong-45771 [https://perma.cc/Z8LU-QWA4] (exploring the foreseeability implications of artificial intelligence programmed by a third party).

performed by a human."[80] Artificial intelligence developers will have to admit that the software carries inherent unpredictability.[81] Once the artificial intelligence is sent off to the buyer, the programmer no longer has control and the artificial intelligence could be shaped by its new owner in uncountable ways.[82] As seen in *Derdiarian*, this argument is likely to fail. The exact manner in which the harm came about, or the exact reaction taken by the artificial intelligence software, will not be the major factor of the foreseeability analysis.[83] Instead the courts will ask if misuse of artificial intelligence by third parties was foreseeable. Not knowing exactly how the artificial intelligence will respond to misuse by third parties is unlikely to serve as any defense. Examining the superseding cause doctrine's use in the proximate cause analysis shows that the doctrine has lost much of its importance. Problems involving an intervening cause are likely to be analyzed under the lens of ordinary foreseeability. The above case law demonstrates that many fact patterns involving artificial intelligence and an intervening cause (like the hypotheticals at the beginning of this Note) will result in liability for the software developers. In many cases, artificial intelligence software will set the stage for what can be considered a foreseeable intervening act.

III. Breach

The background of superseding cause shows it is unlikely to save the developers of artificial intelligence software from potential liability as a matter of proximate cause. But that isn't the end of the road for software developers. While the incident causing harm may have been foreseeable, the artificial intelligence companies can still argue it was not sufficiently foreseeable to justify the burdens required to prevent such harms.[84] As discussed above, third-party misconduct—negligent, reckless, intentional, or criminal—will often be considered sufficiently foreseeable as to render relevant an inquiry into the burden of precautions facing the actor.[85] Yet these burdens can often be extremely high.[86] Under a negligence analysis featuring

---

80. Scherer, *supra* note 8, at 363.

81. *See* Tapson, *supra* note 14 (warning of the potential pervasive application of artificial intelligence).

82. *See* Tanz, *supra* note 9 (describing the limitations of developer control over artificial intelligence and the unpredictability of machine learning).

83. Derdiarian v. Felix Contracting Corp., 414 N.E.2d 666, 671 (N.Y. 1980).

84. Grace & Co. v. City of Los Angeles, 278 F.2d 771, 775 (9th Cir. 1960) (holding that the risk of a burst pipe did not outweigh the burden of digging up the entire pipe system in the area).

85. *See* David C. Vladeck, *Machines Without Principals: Liability Rules and Artificial Intelligence*, 89 WASH. L. REV. 117, 123–28, 135 (2014) (analyzing liability rules as applied to highly intelligent, autonomous machines).

86. *Cf.* Ikene v. Maruo, 511 P.2d 1087, 1088–89 (Haw. 1973) (rejecting negligence claims against public highway agencies for failing to design curves and install guardrails that would protect cars that drive out of control because of the unreasonable burden that would be required of the state).

an intervening cause, the primary factors to be considered for breach are those found in Learned Hand's classic B<PL analysis:[87] L (the magnitude of the foreseeable risk), P (the probability or foreseeability of such a risk), and B (the burden of precautions that might protect against the risk).[88]

It is foreseeable, for example, that some number of motorists, while driving on the state's highways, will speed, drive drunk, or fall asleep, and thereby will fail to navigate curves or otherwise allow their cars to leave the highway. Such an intervening cause of harm is foreseeable enough to prevent the implementation of the superseding cause doctrine to absolve the state of liability. However, if the state is liable for the failure to design curves and erect barriers that would protect against such out-of-control vehicles, the overall burden on the state would be excessive, by way of either bearing and defending against liability or the cost of redesigning highways.[89] At the same time, however, the burden on the state to simply maintain a shoulder would not be so great as to prevent liability.[90]

Imagine driving south down a road. You approach an intersection where only the east–west road is governed by a stop sign, but your north–south road is not. It would be reasonable to expect you to keep an eye on the east–west road in case another driver misses the stop sign. To put it in terms of breach, the burden of monitoring the east–west road is not so great as to prevent you from doing so despite the low probability of a car accident. It would not be reasonable to expect you to slow your car as you approach the intersection and carefully confirm that no other drivers are going to miss their stop sign before you continue driving through the intersection. Such a burden is said to be too great to justify in light of the low probability of an accident. Of course, such an analysis is constantly ongoing. If, while performing your reasonable monitoring of the intersection, you do notice that another driver is going too fast to stop in time, it would be reasonable to expect you to slow your car down in light of the high probability of an accident.[91] The law performs these balancing tests through the breach analysis. We avoid "requiring excessive precautions of actors relating to harms that are

---

87. United States v. Carroll Towing Co., 159 F.2d 169, 173 (2d Cir. 1947).

88. *Id.*

89. *See Ikene*, 511 P.2d at 1089 (finding no obligation of a city to post a speed limit of 35 miles per hour to prevent cars speeding in excess of 40 miles per hour).

90. *See* McKenna v. Volkswagenwerk Aktiengesellschaft, 558 P.2d 1018, 1022 (Haw. 1977) (finding liability when a city failed to maintain a shoulder along the highway).

91. *See* Brockett v. Prater, 675 P.2d 638, 640 (Wyo. 1984) (affirming jury's finding of no negligence in failing to halt to make sure that others will honor the right-of-way); Stirling v. Sapp, 229 So. 2d 850, 852 (Fla. 1969) (referencing Florida law that a driver with the right-of-way can legally assume that an approaching driver on an intersecting road will yield to the right-of-way); *see also* LeJeune v. Union Pac. R.R., 712 So. 2d 491, 495 (La. 1998) (noting that a railroad company can presume approaching vehicles "will obey the law and stop in time to avoid an accident" and is thus not required to slow its trains).

immediately due to the improper conduct of third parties, even when that improper conduct can be regarded as somewhat foreseeable."[92]

Through proximate cause analysis, we know most courts will find incidents involving artificial intelligence to be somewhat predictable, potentially forcing artificial intelligence developers to rely on the argument that the burden of preventing such incidents is too high. The developers are unlikely to succeed when "an accident was caused by a clear defect or malfunction in the [software] design, especially if the defect could have been prevented or fixed by an alternative design."[93] In these situations, the breach analysis does not come down to whether or not the artificial intelligence software is better than what it is replacing, but what the cost would have been to the software developer to tweak the software and make it safer.[94] Errors in software are especially susceptible to hindsight bias. Once a problem is discovered with software, a plaintiff's attorney could easily argue that the burden to the software company would only be typing in new lines of code—a burden that could sound very low to many laypeople. Leaving the complexities and difficulty of software coding up to a jury to appreciate is a dangerous proposition for artificial intelligence developers.

> [H]indsight from the accident that actually occurred will inevitably provide new insights into how the technology could have been made safer, which will then be imputed to the manufacturer. Given the complexity of an autonomous system, a plaintiff's expert will almost always be able to testify (with the benefit of hindsight) that the manufacturer should have known about and adopted the alternative, safer design.[95]

Plaintiffs' experts will be able to point to alternative software codes with the benefit of hindsight, second guessing the coding of software developers. Defendants will have a hard time because the scope of liability (the L factor in the Learned Hand Formula) could potentially "be severe—the loss of one or more lives or other serious injury," compared to a small burden that is only the "cost of the marginal improvement that might have prevented the accident."[96] The complexity of artificial intelligence software will make it very challenging for a developer to win the cost–benefit argument.

---

92. RESTATEMENT (THIRD) OF TORTS § 19 cmt. g (AM. LAW INST. 2005). *Compare* McMillan v. Mich. State Highway Comm'n, 393 N.W.2d 332, 333, 337 (Mich. 1986) (imposing liability when a private utility locates a utility pole close enough to a public highway to create a risk of injury for occupants of cars that lurch off the highway), *with* Gouge v. Cent. Ill. Pub. Serv. Co., 582 N.E.2d 108, 112 (Ill. 1991) (holding no liability in a similar fact pattern).

93. Gary E. Marchant & Rachel A. Lindor, *The Coming Collision Between Autonomous Vehicles and the Liability System*, 52 SANTA CLARA L. REV. 1321, 1333 (2012).

94. *See id.* (discussing the cost–benefit analysis in the context of automobile manufacturing).

95. *Id.*

96. *Id.* at 1334.

The final product of modern artificial intelligence software is often trained and coached rather than coded.[97] Therefore, it is possible that adding new code to artificial intelligence software will not be enough to prevent the software from causing injury.[98] The only way to prevent artificial intelligence from being misused would be to strip the software of its fundamental aspects. One of the primary benefits of artificial intelligence is its ability to learn and mold itself with new experiences, resulting in it taking on almost human characteristics. Without allowing it to continue to do so, artificial intelligence is relatively useless. Artificial intelligence developers could argue that the only way to prevent artificial intelligence from being misused would be to not use the software at all. Abandoning that type of programming altogether would be too great a burden. While the harm caused by artificial intelligence software may be foreseeable, artificial intelligence is still a great substitute for human doctors (who will eventually become monkeys in lab coats by comparison), delivery men, or drivers it replaces.[99] "Robot drivers react faster than humans, have 360-degree perception and do not get distracted, sleepy or intoxicated . . . ."[100] Losing such a valuable asset would be devastating to businesses developing the software, businesses using the software, and society as a whole. In this way, artificial intelligence's benefits to society could be argued to outweigh its costs.

This argument is unlikely to work. Even if artificial intelligence software offers a net safety gain, its developers may still be held accountable if the software malfunctions or third-party interactions with the software result in harm.[101] "There are many examples of products that have a net safety benefit that are still subject to liability when an injury results."[102] The government has already tackled this problem in a similar arena: vaccines.[103] The value of vaccines is enormous and "[t]he public health benefit[s] . . . undeniable, yet they are so frequently the source of lawsuits that federal

---

97. Monaghan, *supra* note 9.

98. *See id.* ("[A]s networks have grown more intertwined and their functions more complex, code has come to seem more like an alien force, the ghosts in the machine ever more elusive and ungovernable.").

99. *See* Matthew Hart, *This Is Why We "Monkey Drivers" Need to Be Replaced by Self-Driving Cars*, NERDIST (Sept. 4, 2016), http://nerdist.com/this-is-why-we-monkey-drivers-need-to-be-replaced-by-self-driving-cars/ [https://perma.cc/9R64-R9Y6] (summarizing a video attributing automobile traffic to poor reactions, lack of coordination, and unpredictable behavior).

100. John Markoff, *Google Cars Drive Themselves, in Traffic*, N.Y. TIMES (Oct. 9, 2010), http://www.nytimes.com/2010/10/10/science/10google.html?mcubz=1 [https://perma.cc/33R6-YXFH].

101. *See* Marchant & Lindor, *supra* note 93, at 1330 (discussing the relative risk of autonomous vehicles compared to conventional vehicles).

102. *Id.* at 1331.

103. *See, e.g.*, National Childhood Vaccine Injury Act, 42 U.S.C. §§ 300aa-1–34 (1986) (establishing "a National Vaccine Program to achieve optimal prevention of human infectious diseases through immunization").

preemption laws had to be passed to protect their manufacturers."[104] Despite the statutory protections, when a consumer contracted polio from an oral polio vaccine, an $8.5 million verdict was held for the injured plaintiff.[105] The arguments that the only way to prevent the harms of vaccines would be to eliminate them did not impress the courts.[106] Luckily for the vaccine producers, their argument worked better with Congress.[107]

The automobile industry has faced a similar problem.[108] General Motors (GM) was sued when a car collision caused a passenger-side airbag to fail to deploy, resulting in injuries to the plaintiff.[109] GM argued that adding the passenger airbag at all was an improvement over the industry standard at the time.[110] That argument fell on deaf ears, as did the argument that the only way for GM to avoid the risk of any malfunction altogether would be to remove the passenger airbags (at the time, that would have met the National Highway Transportation Safety Administration requirements).[111]

GM was the subject of another series of lawsuits concerning its C/K pickup.[112]

> Plaintiffs in these suits alleged that GM's placement of the gas tank on the side of the model, outside the vehicle frame, created an increased risk of fatal fires after side impacts. GM attempted to defend the safety of its vehicle with comparative analyses, contending that the overall crashworthiness of its vehicles was better than most vehicles on the road.[113]

To support its arguments, GM cited to extensive safety reports and argued that the mere existence of the gas tank would increase the likelihood of fires wherever it was located. The only way to prevent the risk of a gas tank fire would be to not have a gas tank in the vehicle at all.[114] GM's

---

104. Marchant & Lindor, *supra* note 93, at 1331.

105. Strong v. Am. Cyanamid Co., 261 S.W.3d 493, 521 (Mo. Ct. App. 2007) (confirming the $8.5 million jury verdict for the plaintiff), *overruled on other grounds*, Badahmon v. Catering St. Louis, 395 S.W.3d 29 (Mo. 2013) (en banc).

106. *See id.* at 506 (focusing on the plaintiff's satisfaction of proving tort elements rather than discussing the value of vaccines).

107. *See* 42 U.S.C. §§ 300aa-1 to -34 (establishing a system of regulations and standards for vaccines rather than banning them).

108. Gen. Motors Corp. v. Burry, 203 S.W.3d 514, 525 (Tex. App.—Fort Worth 2006, pet. dism'd); *see also* Morton Int'l v. Gillespie, 39 S.W.3d 651, 654 (Tex. App.—Texarkana 2001, pet. denied) (detailing a suit brought against a vehicle manufacturer and seller for injuries caused by an airbag after the plaintiff's car accident).

109. *Burry*, 203 S.W.3d at 524–25.

110. *Id.* at 529.

111. *Id.*

112. *See* Terence Moran, *GM Burns Itself*, AM. LAW., Apr. 1993, at 69 (describing the design flaws of GM's popular pickup truck that led to extensive product liability).

113. Marchant & Lindor, *supra* note 93, at 1332.

114. *See* Moran, *supra* note 112, at 78 ("Anywhere a manufacturer puts a tank poses potential hazards . . . .").

arguments did not curry favor with juries.[115] The juries did not care about the overall greater safety and returned damages of $101 million in punitive damages.[116]

While it appears arguments citing an unbearably high burden for artificial intelligence developers are unlikely to succeed, it is important to keep in mind the policy considerations judges may employ when evaluating the burden. Toyota recently accepted liability to the tune of $1.2 billion for a defect in its cars causing sudden acceleration.[117] A $1.2 billion judgment is steep, and artificial intelligence won't be accused of just sudden acceleration, but of any behavior that could be deemed imperfect. If this is the sort of liability carmakers could be facing, frequent suits could prevent companies from entering the market with these products to begin with. The amount at stake in suits against the manufacturers could be tremendous.[118] It is possible that the potential liability is so great that judges would seek to prevent chilling the development of artificial intelligence technologies. A world without artificial intelligence could be the worst result for society as a whole, and so a judge might take the approach of keeping a thumb on the side of the scale requiring a low burden and a finding of high foreseeability.

Judges may be justified in this belief. As Elon Musk said in his Master Plan, Part Deux for Tesla, "[partial driving autonomy] is already significantly safer than a person driving by themselves and it would therefore be morally reprehensible to delay release simply for fear of . . . legal liability."[119] If autonomous driving software can already improve safety on the road, then it would make society worse off to delay the implementation of such software due to liability concerns, and while Musk himself may not be scared away by legal liability, many others might be. The same logic can apply to any form of artificial intelligence—it is able to bring such positive change to society that we should do what we can to encourage its development, including lowering tort costs.

But on the other side, a strong argument exists that large car companies (and for that matter, all large companies dealing with artificial intelligence) are unlikely to be chilled. There is a lot of profit to be made with artificial

---

115. *See* Sam LaManna, *GM Verdict Could Affect Future Cases*, NAT'L L.J., May 3, 1993, at 21 (reporting on a jury award including $101 million in punitive damages against GM).

116. GM v. Moseley, 447 S.E.2d 302, 305 (Ga. Ct. App. 1994) (indicating that the jury assessed $101 million in punitive damages against GM).

117. Carol J. Williams, *Toyota is Just the Latest Automaker to Face Auto Safety Litigation*, L.A. TIMES (Mar. 14, 2010), http://articles.latimes.com/2010/mar/14/business/la-fi-toyota-litigate14-2010mar14 [https://perma.cc/5AEA-3HDD].

118. *See* Thierer, *supra* note 21 (explaining that because America's legal system lacks a loser-pays rule, consumers are incentivized to file potentially frivolous lawsuits at the first sign of trouble).

119. Elon Musk, *Master Plan, Part Deux*, TESLA BLOG (July 20, 2016), https://www.tesla.com/blog/master-plan-part-deux [https://perma.cc/F8A7-NBH4].

intelligence.[120] With profit numbers reaching tens of trillions,[121] it seems silly to imagine the companies getting scared out of the market. However, not every company that is going to want to enter the market will have the safety net titans of industry have. Many smaller competitors, unable to withstand a large judgment, could be scared out of the market for fear of unlimited liability should their technology cause injury.

But even if there were a chilling effect, would that not represent the exact outcome our tort system is designed to create? If an activity is creating more harm than good, then one of the purposes of the tort system is to discourage such activity through civil liability.[122] Until the benefits outweigh the costs, a dangerous activity should be chilled. The market forces will determine value and costs of new artificial intelligence products, and companies won't proceed until the balance of values favors proceeding.[123]

Even in the face of uncertainty and potential liability, car companies don't seem to be hindered. While the fear of liability as an obstacle to innovation remains a common argument,[124] the proponents of robotics and artificial intelligence are moving full steam ahead. "At least 19 companies have announced their goal to develop driverless car technology by 2021."[125] Far from being chilled by potential liability, "Volvo, Google, and Mercedes-Benz have already pledged to accept liability if their vehicles cause an accident."[126] As far as the car industry is concerned, liability does not seem to be a big cause for alarm.

Car companies could stand to benefit from accepting all liability. If autonomous cars are as safe as their creators claim, then the rate of accidents should go down, leading to less liability for manufacturers. Offering full

---

120. Donal Power, *Self-Driving Car Market Worth Trillions by 2030?*, READWRITE (June 8, 2016), http://readwrite.com/2016/06/08/self-driving-cars-speeding-toward-2-6-trillion-market-2030-tl4/ [https://perma.cc/4XKT-VAA2].

121. Jonathan Schaeffer, *Canada Must Focus Its AI Vision If It Wants to Lead the World*, GLOBE & MAIL (Mar. 3, 2017), https://beta.theglobeandmail.com/report-on-business/rob-commentary/canada-must-focus-its-ai-vision-if-it-wants-to-lead-the-world/article34204949 [https://perma.cc/ZU2A-MGT6].

122. *See* Steven Shavell, *On the Social Function and the Regulation of Liability Insurance*, 25 GENEVA PAPERS ON RISK & INS. 166, 166 (2000) (describing how civil liability discourages undesirable behavior).

123. *See* Kyle Colonna, *Autonomous Cars and Tort Liability: Why the Market Will "Drive" Autonomous Cars Out of the Marketplace*, 4 CASE W. RES. J.L. TECH. & INTERNET 81, 97 (2012) (examining how potential civil liability has affected the development and use of autopilot technology on airplanes).

124. *See* F. Patrick Hubbard, *"Sophisticated Robots": Balancing Liability, Regulation, and Innovation*, 66 FLA. L. REV. 1803, 1870 (2014) (examining whether granting immunity to sellers of robots fosters innovation).

125. Danielle Muoio, *19 Companies Racing to Put Self-driving Cars on the Road by 2021*, BUS. INSIDER (Oct. 17, 2016), http://www.businessinsider.com/companies-making-driverless-cars-by-2020-2016-10/#tesla-is-aiming-to-have-its-driverless-technology-ready-by-2018-1 [https://perma.cc/4V4N-8LD9].

126. Ben-Shahar, *supra* note 19.

protection in the case of an accident is also a great marketing tool. Consumers are "irrationally afraid of self-driving cars—55 percent of consumers say that they would not ride in them."[127] A quick way to convince people to give autonomous cars a try would be to give a full warranty.

Another argument favoring finding liability for artificial intelligence developers is that to do otherwise could hurt the incentive to innovate.[128] This effect can be seen in the vaccine industry, where vaccine manufacturers enjoy wide immunity, which has resulted in the failure to update vaccines as new technology arises.[129] If the manufacturers aren't going to be held liable, then they lose much of their incentive to improve their product. Justice Sotomayor explained that insulation from liability can have a negative impact on innovation, stating that expansion of immunity "leaves a regulatory vacuum in which no one ensures that vaccine manufacturers adequately take account of scientific and technological advancements when designing or distributing their products."[130] Just as vaccine manufacturers lost incentive, so too could car manufacturers. "[O]ne disadvantage of these approaches is that by immunizing the internalization of accident costs from vehicle manufacturers, they may reduce the pressure on manufacturers to make incremental improvements in the safety of their autonomous systems."[131]

Everything discussed here considered, developers of artificial intelligence software have several arguments to prevent a finding of liability under the breach element. To rely on these arguments may prove misguided as there are multiple policy reasons and plenty of case history to support a finding of breach.

IV. Comparative Fault

The traditional common law doctrine of contributory negligence (the precursor to modern comparative fault) served as a way for "defendants who were indisputably guilty of seriously negligent conduct [to] escape[] liability. If such a defendant could prove a negligence case against the plaintiff—i.e., if the defendant could prove that the plaintiff, too, was guilty of negligent conduct . . . the defendant would not be liable."[132] In this way, contributory

---

127. *Id.*

128. *See* Marchant & Lindor, *supra* note 93, at 1340 (discussing how reducing liability for vehicle manufacturers decreases the incentive to improve vehicle safety).

129. *See* Meredith Melnick, Bruesewitz v. Wyeth*: What the Supreme Court Decision Means for Vaccines*, Time (Feb. 24, 2011), http:/healthland.time.com/2011/02/024/bruesewitz-v-wyeth-what-the-supreme-court-decision-means-for-vaccines/ [https:perma.cc/CSU7-MQPC] (discussing the Supreme Court's 6–2 ruling shielding vaccine developers from liability after a vaccine that had not been updated since the 1940s caused brain damage and seizures in a teenager).

130. Bruesewitz v. Wyeth, 562 U.S. 223, 250 (2011) (Sotomayor, J., dissenting).

131. Marchant & Lindor, *supra* note 93, at 1340.

132. Robertson et al., *supra* note 4, at 343–44.

negligence worked a lot like a superseding cause if the plaintiff herself was a superseding cause.

This style of no-recovery contributory negligence has its roots in England.[133] The rule made its way into the United States, but did not begin to resemble the modern version of the comparative fault rule until the middle of the 20th century.[134] Because of the slow adoption of comparative fault, the doctrine of superseding cause grew up in its absence and is very much a product of the environment it was raised in.[135] "Thus, the law on intervening acts and superseding causes . . . is a product of rules that did not permit a negligent tortfeasor to obtain contribution from another negligent tortfeasor, nor from even an intentional tortfeasor who was also a cause of the plaintiff's harm."[136] When the only option left to a judge would be to bar recovery, the use of superseding cause to distinguish between a highly culpable actor and a moderately culpable actor would seem very reasonable.[137] The dilemma was summed up nicely by Charles Carpenter in 1932:

> When a damage to the plaintiff occurs through the operation of several factors some of which are more substantial than the one for which the defendant is responsible, it may appeal to most persons as unjust, particularly if the defendant's factor is trivial, to permit the plaintiff to throw the whole loss on the defendant. As there is no human method of properly apportioning the loss between the plaintiff and defendant, either the one or the other having to bear the whole loss, it will in many instances seem more satisfactory to leave the loss where it originally falls.[138]

However, the problem laid out above is mostly a problem of the past as most jurisdictions have moved towards comparative fault.[139] The modern comparative fault system is less draconian and is described as follows:

> Where any person suffers damage as the result partly of his own fault and partly of the fault of any other person or persons, a claim in respect of that damage shall not be defeated by reason of the fault of the person

---

133. *See* Butterfield v. Forrester (1809) 103 Eng. Rep. 926 (K.B.) (introducing the theory of contributory negligence by holding that a plaintiff could not recover for injuries from an accident when he lacked ordinary care in avoiding the accident).

134. W. PAGE KEETON ET AL., PROSSER AND KEETON ON THE LAW OF TORTS 337–38 (5th ed. 1984).

135. Michael D. Green, *The Unanticipated Ripples of Comparative Negligence: Superseding Cause in Products Liability and Beyond*, 53 S.C. L. REV. 1103, 1110–11 (2002).

136. RESTATEMENT (THIRD) OF TORTS § 34 cmt. c (AM. LAW INST. 2010).

137. *See* Green, *supra* note 135, at 1111, 1113 (explaining the use of proximate cause and superseding cause doctrine to limit liability, especially for more culpable or less culpable defendants).

138. Charles E. Carpenter, *Workable Rules for Determining Proximate Cause*, 20 CALIF. L. REV. 229, 233 (1932).

139. *See* Green, *supra* note 135, at 1112, 1114 (explaining the decline of "all-or-nothing" liability in cases with multiple tortfeasors).

suffering the damage, but the damages recoverable in respect thereof shall be reduced to such extent as the court thinks just and equitable having regard to the claimant's share in the responsibility of the damage . . . .[140]

In the modern system then, when multiple tortfeasors exist, like in your standard superseding cause case, both tortfeasors will have a percentage of the fault attributed to them. In many jurisdictions today, the concerns laid out by commentators like Charles Carpenter are much less relevant.[141] With comparative fault, a negligent plaintiff can still partially recover from a significantly more culpable defendant. Under comparative fault, the difficult questions of how much more culpable a defendant may be can now just be answered in percentage terms by the jury.[142] While the superseding cause doctrine is not dead and will not go away in the minds of judges for a long time, the doctrine has been drastically reduced in importance as comparative fault answers similar questions much more cleanly than superseding cause ever did.[143]

The adoption of comparative fault in the tort system is a radical change for the superseding cause doctrine. With the use of comparative fault, there is rarely ever a need to implement superseding cause.[144] "Under a 'proportional fault' system, no justification exists for applying the doctrines of intervening negligence and last clear chance . . . . [C]omplete apportionment between the negligent parties, based on their respective degrees of fault, is the proper method for calculating and awarding damages . . . ."[145] The doctrine of superseding cause has not been completely eliminated in favor of comparative fault—many jurisdictions do still rely on it.[146] But as discussed in the proximate cause section of this Note, the modern view of superseding cause is as a specific category of foreseeability problems, which should be treated no differently than any other foreseeability question.[147]

---

140. *See* ROBERTSON ET AL., *supra* note 4, at 344 (quoting the Law Reform (Contributory Negligence) Act to explain the transition to the modern comparative fault system and its implementation into American jurisprudence).

141. *See, e.g.*, Exxon Co., U.S.A. v. Sofec, Inc., 517 U.S. 830, 836–37 (1996) (noting that comparative fault rules apply to cases of admiralty jurisdiction).

142. *See, e.g.*, *id.* at 840–41 ("The issues of proximate causation and superseding cause involve application of law to fact, which is left to the factfinder, subject to limited review.").

143. *See id.* at 837 (noting that there is no inconsistency between the superseding cause doctrine and "a comparative fault method of allocating damages").

144. Hansen v. Umtech Industrieservice Und Spedition, GBmbH, No. 95-516 MMS., 1996 WL 622557, at *11 (D. Del. July 3, 1996) (rejecting the claim that a plaintiff's conduct was a superseding cause because it was inconsistent with comparative fault).

145. Hercules, Inc. v. Stevens Shipping Co., 765 F.2d 1069, 1075 (11th Cir. 1985).

146. *See Exxon*, 517 U.S. at 838 ("[O]f the 46 States that have adopted a comparative fault system, at least 44 continue to recognize and apply the superseding cause doctrine.").

147. *See* Coyne v. Taber Partners I, 53 F.3d 454, 460–61 (1st Cir. 1995) (holding that attacks from the taxi union protestors might have been a foreseeable consequence of driving a different

The rise of comparative fault and the diminished role of superseding cause likely comes as bad news for artificial intelligence software developers. The software from these developers will inevitably be involved in a considerable amount of incidents moving forward. As was demonstrated earlier in this Note, a finding of negligence is very possible, and thus the jury will have fault points to assign during the comparative fault stage of the negligence analysis.

While on the one hand the modern trend of comparative fault has largely meant the end of the superseding cause doctrine as a complete shield for potential tortfeasors, it also means that the software developer is more likely to avoid being jointly and severally liable for the entire award of damages. The interplay between joint and several liability and comparative fault varies from jurisdiction to jurisdiction. In the jurisdictions where joint and several liability has survived the move to a comparative fault system, the artificial intelligence software developer could be stuck with a lot of liability.[148] If the software developer is given even a single fault point from the jury, then the developer would be liable for the entire damage award and would be responsible for going after the intervening cause to make sure the other culpable party pays its fair share of the damages. In many instances, the more culpable actor may be just an individual who used artificial intelligence software for nefarious purposes. Those individuals won't have very deep pockets, leaving a large chance that the artificial intelligence company will be stuck holding the bill.

However, in jurisdictions where joint and several liability has been abolished, the artificial intelligence companies may see a favorable result after the tort analysis. While the artificial intelligence developer will probably receive some fault points from the jury, the vast majority of fault points will lie with the more culpable intervening cause.[149] Without joint and several liability, those are fault points the software developer won't ever be on the hook for. While not as good of a result as avoiding liability altogether under the old system of superseding cause, at least the developer avoids the risk of being stuck with the entire amount of damages.

Conclusion

Liability for artificial intelligence software developments is a very real possibility. The interactions between a third party and artificial intelligence

---

transportation service through the protest); Stagl v. Delta Airlines, Inc., 52 F.3d 463, 473–74 (2d Cir. 1995) (finding that injury from aggressive luggage retrieval at an airport could be a foreseeable result of a flight delay and an inadequate regulation of baggage retrieval); Williams v. United States, 352 F.3d 477, 481 (5th Cir. 1965) (stating that "the negligent act of a third party will not cut off the liability of an original wrongdoer if the intervening act is foreseeable").

148. ROBERTSON ET AL., *supra* note 4, at 372.

149. *See id.* (describing the benefit to defendants when comparative fault points are "siphoned off" to culpable third parties).

software that have resulted in harm to another will not be a definite shield against liability for the software developer. Many interactions between intervening third parties and the software may be sufficiently unforeseeable if they are risks an artificial intelligence company cannot guard against. However, due to the wide range of potential uses for artificial intelligence, many interactions will be deemed foreseeable. In any event, there is nothing special about an intervening cause that creates a different negligence analysis than any other cause of harm. The negligence analysis will come down to the question of foreseeability, as many cases do. With the rise of comparative fault, juries will be incentivized to assign at least some fault points to the artificial intelligence developer in lieu of focusing on the more culpable intervening force altogether. In jurisdictions with joint and several liability, this could be a disastrous result for artificial intelligence companies as they could be left to foot the bill.

Artificial intelligence developers need to take steps to protect themselves from looming liability. Tesla requires its buyers to sign a contract that mandates they agree to keep their hands on the wheel at all times, even when the autopilot is engaged.[150] Artificial intelligence companies should take a page out of Tesla's book. A contract requiring buyers of artificial intelligence products to use the products responsibly could go a long way. Alternatively, artificial intelligence companies could just exercise tight control over their software post-sale and perform routine patches and updates, which would prevent the software from growing too customized in unforeseen ways. Artificial intelligence companies may also want to lobby their representatives. As discussed previously, vaccine corporations enjoy widespread immunity.[151] Congress has also passed the Protection of Lawful Commerce in Arms Act to give manufacturers of guns and ammunition immunity from tort suits arising from use of their products for criminal purposes.[152] Artificial intelligence companies could find themselves in desperate need of a similar statute to protect them from misuse of their products.

Artificial intelligence companies need to be aware of the very real threat of tort liability. If they do not take steps to protect themselves from liability, these companies could be closing their doors as quickly as they have opened them. Not only would this be bad for the artificial intelligence community, but it would hurt society as a whole to lose innovators of such a promising new technology. The tort system requires a balance between protecting individuals from the potential harms of artificial intelligence and the free

---

150. Ben-Shahar, *supra* note 19.

151. 42 U.S.C. §§ 300aa-1 to -34.

152. Protection of Lawful Commerce in Arms Act, 15 U.S.C. §§ 7901–03 (2012).

development of such technology. Companies must carefully evaluate the foreseeable risks of the technology they are entering into the market and take steps to minimize those risks. If companies take these steps, they will not only help to minimize their eventual liability, but ensure that their artificial intelligence software is ready for the human world in which we live.

Weston Kowert