

AM I AN ALGORITHM OR A PRODUCT? WHEN PRODUCTS LIABILITY SHOULD APPLY TO ALGORITHMIC DECISION-MAKERS

Karni A. Chagal-Feferkorn*

INTRODUCTION.....	62
PART I: WHY NOT USE AUTONOMY TO DISTINGUISH PRODUCTS FROM THINKING ALGORITHMS?	70
A. On Automation and Autonomy	70
B. The Many Faces of Autonomy	72
PART II: PRODUCTS LIABILITY RATIONALES.....	77
PART III: IS IT A PRODUCT? IS THERE A DEFECT?	82
A. On “Products”	83
B. On “Defects”	84
PART IV: DISTINGUISHING PRODUCTS FROM THINKING ALGORITHMS	87
A. Examples of Products Versus Thinking Algorithms	88
B. Which Products Liability Rationales Apply to Which Algorithms	90
1. Promoting Safety.....	90
a. The OODA Loop—General	94
b. “Observe” (“Information Acquisition”).....	95
c. “Orient” (“Information Analysis”)	97
d. “Decide” (“Decision Selection”).....	99
e. “Act” (“Action Implementation”).....	101
f. Interim Summary and Measurability of Success Rates	101
2. Avoiding a Chilling Effect	102
3. Victims’ Compensation.....	104
C. Keep It Simple, System.....	107
CONCLUSIONS	113

* Fellow, Haifa Center for Law & Technology, University of Haifa Faculty of Law. I thank Niva Elkin-Koren for her endless support that made this research possible. I would also like to thank Yonatan Arbel, Michal Gal, Yaron David, Itiel Dror, Orr Dunkelman, Rita Osadchy, Amnon Reichman, Catherine Sharkey, and Avinoam Zimmerman. This research was supported by the Center for Cyber Law & Policy (CCLP), established by the University of Haifa in collaboration with the Israeli National Cyber Bureau. Any mistakes or omissions are the author’s.

INTRODUCTION

Culminating in the Industrial Revolution, machines and tools replaced or assisted humans in performing physical tasks.¹ Drills, engravers, weaving machines and the like employed machines' physical advantages to free human beings from repetitive physical labor. In general, damage caused by such tools has been governed, since the middle of the 20th century, by the legal framework of products liability.² Under this system the seller, manufacturer, distributor, or any other party in the distribution chain of a defective product is liable for the physical harm caused to the user or her property.³ Thus, victims of tools that have caught fire or come apart,⁴ of motor vehicles that have crashed,⁵ of food containing external undesired objects,⁶ etc., could all have brought successful products liability claims against the products' manufacturers or sellers.

Over time, and with the advance of technology, the products used by humans became more sophisticated. Machines were no longer used merely to replace humans in performing physical tasks. Instead their superior computational abilities were utilized to assist or replace humans in processing data. The electronic calculator, for example, allowed engineers, merchants,

1. R.M. HARTWELL, *THE INDUSTRIAL REVOLUTION AND ECONOMIC GROWTH* 295-97 (1971); Nick von Tunzelmann, *Time-Saving Technical Change: The Cotton Industry in the English Industrial Revolution*, 32 *EXPLORATIONS ECON. HIST.* 1, 8-13 (1995); see also STEVEN KING & GEOFFREY TIMMINS, *MAKING SENSE OF THE INDUSTRIAL REVOLUTION: ENGLISH ECONOMY AND SOCIETY 1700-1850* 68-97 (2001) (arguing that the adoption of technological measures compared to "hand technologies" was associated with more difficulties and inefficiencies than commonly perceived).

2. Donald G. Gifford, *Technological Triggers to Tort Revolutions: Steam Locomotives, Autonomous Vehicles, and Accident Compensation*, 11 *J. TORT L.* 71, 117-18 (2018). For an overview of products liability history, see generally JANE STAPLETON, *PRODUCT LIABILITY* 9-29 (1994). For further discussion on the development of the products liability framework, see *infra* Part II.

3. RESTATEMENT (SECOND) OF TORTS § 402A (AM. LAW INST. 1965).

4. See, e.g., *Bilenky v. Ryobi Tech.*, 666 F. App'x. 271 (4th Cir. 2007) (affirming that the manufacturer of a lawn-tractor that caught fire and as a result killed its owner was liable under products liability); *Bass v. Phoenix Seadrill*, 562 F. Supp. 790 (E.D. Tex. 1983) (finding the manufacturer of a drilling rig that fell and struck plaintiff liable for damages, referring among other causes of actions to products liability).

5. See, e.g., *Collazo-Santiago v. Toyota Motor Corp.*, 149 F.3d 23 (1st Cir. 1998) (holding Toyota's design of its air bags was defective); *Gray v. Lockheed Aeronautical Sys. Co.*, 125 F.3d 1371 (11th Cir. 1997), *vacated* 524 U.S. 924 (1998), *rev'd per curiam on other grounds* 155 F.3d 1343 (11th Cir. 1998) (holding a military contractor liable for a fatal jet aircraft crash resulting from defective aileron servo); *Four Corners Helicopters, Inc. v. Turbomeca, S.A.*, 979 F.2d 1434 (10th Cir. 1992) (applying strict liability on the engine manufacturer of a helicopter that crashed).

6. See, e.g., *Coca-Cola Bottling Co. v. Parker*, 451 So. 2d 786 (Ala. 1984) (affirming verdict for plaintiff who swallowed particles of glass found in a soft drink bottle); *Selens v. Wakefern Food Corp.*, No. CV044000648, 2005 WL 1331154 (Conn. Super. Ct. May 2, 2005) (denying defendants motions to strike products liability counts associated with the existence of a plastic tab inside a chicken).

accountants and other professionals to provide better and quicker outputs;⁷ autopilots were installed in airplanes to improve flight safety through an automated system capable of processing huge amounts of information in split seconds,⁸ while cruise-control and auto-parking devices were similarly installed in cars.⁹ Despite the increasing level of “sophistication” of these machines and devices, manufacturers and sellers of these products were generally held to the traditional products liability legal framework,¹⁰ akin to the liability applied to “simpler” or “less sophisticated” products.¹¹

Technology is ever advancing, and in addition to relinquishing physical and computational tasks to machines, algorithms’ self-learning abilities now allow them to reach their own conclusions based on databases of previous cases.¹² This in turn enables humans to both entrust machines with making complex decisions that until lately required human discretion and even replace professional human judgment in matters of expertise where there is no clear right or wrong answer.¹³

In the field of law, for example, virtual attorneys—such as ROSS Intelligence’s cognitive computing platform that works with IBM’s Watson—

7. *Electronic Calculators—Handheld*, THE NAT’L MUSEUM OF AM. HISTORY, <http://americanhistory.si.edu/collections/object-groups/handheld-electronic-calculators>; Nick Valentine, *The History of the Calculator*, THE CALCULATOR SITE, <https://www.thecalculatorsite.com/articles/units/history-of-the-calculator.php> (last updated Mar. 24, 2014).

8. M.C. Elish & Tim Hwang, *Praise the Machine! Punish the Human! The Contradicting History of Accountability in Automated Aviation* (Data & Society Research Institute, Comparative Studies in Intelligent Systems, Working Paper 1, May 18, 2015), <http://dx.doi.org/10.2139/ssrn.2720477>; Kyle Colonna, *Autonomous Cars and Tort Liability*, 4 CASE W. RES. J.L. TECH. & INTERNET 81, 93-97 (2012).

9. *Ralph Teetor and the History of Cruise Control*, AM. SAFETY COUNCIL (Dec. 10, 2014); *A Brief History of Cruise Control*, FRONTIER CHRYSLER (Jul. 20, 2016), <https://www.frontierchrysler.ca/history-cruise-control/>; Stephanie Levis, *Look, Ma, No Hands: Self-Parking Cars Take the Wheel*, GEICO INS. (May 5, 2016), <https://www.geico.com/more/driving/auto/car-safety-insurance/look-ma-no-hands-parking-technology-takes-the-wheel>.

10. See, for example, *Richardson v. Bombardier, Inc.*, No. 8:03-cv-544-T-31MSS, 2005 WL 3087864, at *14 (M.D. Fla. Nov. 16, 2005), where plaintiffs (unsuccessfully) raised, among other things, manufacture defect claims pertaining to the autopilot system of an airplane that had crashed. Similar products liability claims were also raised in connection with a plane-crash in *Moe v. Avions Marcel Dassault-Breguet Aviation*, 727 F.2d 917 (10th Cir. 1984). For additional examples, see Colonna, *supra* note 8, at 91-102.

11. Products liability may be imposed on the basis of strict liability, requiring only the showing of a defect of the product, regardless of fault. Other cases require negligence by the defendant for liability to apply. The type of products involved may affect the specific type of products liability applied. See *infra* Part II for a more detailed discussion.

12. For further discussion on how “self-learning” works, see *infra* Part IV.

13. As will be discussed in more detail, such decisions are based on probabilities, such that only in hindsight may it be discovered whether the decision was beneficial or damaging (and even then professional disagreements will probably arise as to the best course of action at the time the decision was made).

are utilized by law firms in conducting legal research,¹⁴ algorithmic ODR mechanisms solve disputes online (often without any human facilitator),¹⁵ and bail algorithms determine whether defendants awaiting trial may post bail to be released.¹⁶ Physicians, too, rely more and more on algorithms in order to diagnose medical conditions and select optimal treatment.¹⁷ Meanwhile, algorithms significantly assist, or sometimes replace, tax advisors, company directors, and even priests.¹⁸

Yet advanced as such machines and algorithms may be, occasionally they are still bound to cause damage.¹⁹ A futuristic transition to a world of robo-doctors, for example, could never achieve perfect rates of patient recovery, and some patients' condition would inevitably deteriorate as a result of decisions made by algorithms. Litigants relying on strategic advice provided by ROSS Intelligence or the like, are also bound to lose cases or negotiations from time to time, while some of the individuals filing tax reports prepared by a virtual tax advisor would certainly still be subject to tax investigations and sanctions due to the algorithm's choices. Indeed, where such decision-makers have already been put to test—in the field of driving—we have witnessed that advanced algorithms are not immune to reaching damaging decisions.²⁰

14. See, e.g., John Mannes, *ROSS Intelligence Lands \$8.7M Series A to Speed Up Legal Research with AI*, TECHCRUNCH (Oct. 11, 2017), <https://techcrunch.com/2017/10/11/ross-intelligence-lands-8-7m-series-a-to-speed-up-legal-research-with-ai/>; Anthony Sills, *Ross and Watson Tackle the Law*, IBM WATSON BLOG (Jan. 14, 2016), <https://www.ibm.com/blogs/watson/2016/01/ross-and-watson-tackle-the-law>.

15. Michael Legg, *The Future of Dispute Resolution: Online ADR and Online Courts*, 27 AUSTRALASIAN DISP. RESOL. J. 227 (2016).

16. A.J. Wang, *Procedural Justice and Risk-Assessment Algorithms* 1 (June 21, 2018), <https://ssrn.com/abstract=3170136>. Cf. Tom Simonite, *How to Upgrade Judges with Machine Learning*, MIT TECH. REV. (Mar. 6, 2017), <https://www.technologyreview.com/s/603763/how-to-upgrade-judges-with-machine-learning/> (discussing how algorithms may assist judges to predict which defendants will fail to show to court).

17. See, e.g., Vinod Khosla, *Technology Will Replace 80% of What Doctors Do*, FORTUNE MAG. (Dec. 4, 2012), <http://fortune.com/2012/12/04/technology-will-replace-80-of-what-doctors-do/>; Alina Shrourou, *Deep Learning in Healthcare: A Move Towards Algorithmic Doctors*, NEWS MEDICAL: LIFE SCIENCES (Mar. 15, 2017), <https://www.news-medical.net/news/20170315/Deep-learning-in-healthcare-a-move-towards-algorithmic-doctors.aspx>.

18. Richard Susskind & Daniel Susskind, *Technology Will Replace Many Doctors, Lawyers, and Other Professionals*, HARV. BUS. REV. (Oct. 11, 2016), <https://hbr.org/2016/10/robots-will-replace-doctors-lawyers-and-other-professionals>; Sasha A. Q. Scott, *Algorithmic Absolution: The Case of Catholic Confessional Apps*, 11 ONLINE - HEIDELBERG J. OF RELIGIONS ON THE INTERNET 254 (2016). For a forecast on the percentage of actions currently performed by human professionals that could be replaced by automation see *Automation Potential and Wages for US Jobs*, MCKINSEY GLOB. INST. (Oct. 1, 2018), <https://public.tableau.com/profile/mckinsey.analytics#!/vizhome/AutomationandUSJobs/Technicalpotentialforautomation>.

19. Be it damage caused as a result of sub-optimal decisions by the algorithm, or because the underlying circumstances made damage inevitable. See *infra* Part III.

20. In March 2018, a self-driving Uber car in autonomous mode hit and killed a woman in Arizona. Sam Levin & Julia Carrie Wong, *Self-Driving Uber Kills Arizona*

Should the traditional products liability framework continue to apply to the new generation of decision-making tools,²¹ those that replace human discretion and enjoy rising levels of “autonomy” and self-learning abilities?²²

Woman in First Fatal Crash Involving Pedestrian, GUARDIAN (Mar. 19, 2018), <https://www.theguardian.com/technology/2018/mar/19/uber-self-driving-car-kills-woman-arizona-tempe>. Less than a week later, a Tesla vehicle running on autopilot crashed and killed its driver. *Tesla Car That Crashed and Killed Driver Was Running on Autopilot, Firm Says*, GUARDIAN (Mar. 31, 2018), <https://www.theguardian.com/technology/2018/mar/31/tesla-car-crash-autopilot-mountain-view>. In May 2016, a Tesla driver died as the vehicle's sensors failed to detect a tractor-trailer crossing the highway. Danny Yadron & Dan Tynan, *Tesla Driver Dies in First Fatal Crash While Using Autopilot Mode*, GUARDIAN (June 30, 2016), <https://www.theguardian.com/technology/2016/jun/30/tesla-autopilot-death-self-driving-car-elon-musk>. While the accident rate of autonomous vehicles is expected to drop significantly, accidents and fatalities cannot be avoided completely. JOHN VILLASENOR, BROOKINGS INST., PRODUCTS LIABILITY AND DRIVERLESS CARS: ISSUES AND GUIDING PRINCIPLES FOR LEGISLATION 6-7 (2014), https://www.brookings.edu/wp-content/uploads/2016/06/Products_Liability_and_Driverless_Cars.pdf (“[V]ehicle automation technologies can perform amazingly well. But they aren’t perfect, and they never will be. Sometimes autonomous vehicles will become involved in accidents due at least in part to a defect in the autonomous vehicle technology.”); see also Bryant Walker Smith, *Automated Driving and Product Liability*, 2017 MICH. ST. L. REV. 1, 18-20 (identifying several reasons why automated systems may not provide absolute safety).

21. This Article does not distinguish between robotic decision makers or algorithmic decision makers that have no physical embodiment. Rather, it uses phrases such as “sophisticated systems,” “self-learning algorithms,” or “autonomous robots” interchangeably. This is because the Article focuses on the decision-making process of the system and is generally indifferent to the existence (or lack thereof) of any physical embodiment. In that context, legal scholars, such as Yale Law Professor Jack M. Balkin, have suggested that both algorithms and robots are similar members of the “Algorithmic Society” and might be treated alike. Jack M. Balkin, *2016 Sidley Austin Distinguished Lecture on Big Data Law and Policy: The Three Laws of Robotics in the Age of Big Data*, 78 OHIO ST. L.J. 1217, 1226 (2017). For further discussion on the decision to disregard the difference between algorithms and robots for the purpose of discussing applicable liability framework, see Karni Chagal-Feferkorn, *The Reasonable Algorithm*, 1 U. ILL. J.L. TECH. & POL’Y 111, 116-17 (2018).

22. This question is of even greater importance due to the shift from “human-afflicted damages” to “algorithmic-afflicted damages” (caused precisely because machines now replace humans for increasingly more tasks), which enlarges the number of instances where said issue is expected to be invoked. In other words, technology now offers to shift decisions and actions from humans to machines, leading to a much larger share of cases that would potentially involve products liability claims. See Smith, *supra* note 20, at 30 (“[I]t is widely accepted that design issues will play a much greater role in automated driving crashes than in today’s conventional driving crashes.”). First, in certain fields technology dispenses with any human involvement. For example, while in the past damages associated with cleaning the house were the fault of a cleaning-person, the shift to cleaning robots will now invoke claims against a “product” or a “machine” once the cleaning has caused damage (such as the case where a cleaning robot in South Korea sucked up its owner’s hair while she was sleeping on the floor, mistaking it for dirt). Justin McCurry, *South Korean Woman’s Hair ‘Eaten’ by Robot Vacuum Cleaner as She Slept*, GUARDIAN (Feb. 8, 2015), <https://www.theguardian.com/world/2015/feb/09/south-korean-womans-hair-eaten-by-robot-vacuum-cleaner-as-she-slept>. See also Colonna, *supra* note 8, at 102-04 (referring in general to the replacement of negligent human drivers with hardware or software that have caused a car accident). Secondly, even when humans remain part of the process, and will merely be assisted by sophisticated algorithms, a larger share of the damaging decisions will be reached by algorithms rather

Several scholars have argued that certain sophisticated or autonomous decision-makers require treatment different from their traditional predecessors.

Professor Jane Bambauer, for example, suggested that certain medical applications should be regulated similarly to how human doctors are regulated.²³ Attorney Jessica S. Allain too compared IBM's Watson units used in the medical field to a "consulting physician,"²⁴ while attorneys Jason Chung and Amanda Zink likened it to a "medical student," explaining that Watson was not a typical medical device and that the products liability regime would not suit it.²⁵ Professor of Law and Health Science Ryan Abbott distinguished a conventional automobile from a driverless car, noting that it might warrant a separate treatment of scrutinizing its actions as compared with those of a reasonable human driver.²⁶ In a previous paper I too argued that, in general, algorithms replacing a human's professional judgment should be subject to the reasonableness standard that currently applies to humans, rather than being treated as a product.²⁷

The European Parliament, to give another example, has issued a report to the E.U. Commission on Civil Law Rules on Robotics explaining that ordinary liability rules are insufficient for autonomous robots, since they can no longer be considered tools in the hands of other actors.²⁸ The report suggested granting autonomous robots an independent legal status of "electronic persons," which might even allow these robots themselves to pay damages for the harm

than the person; thus again, the focus is shifted from human fault to products liability. Smith, *supra* note 20, at 29-30, gives an example from the world of driving, where in the past damaging decisions often stemmed from a combination of both human and machine failures. For instance, car accidents involving some sort of manufacture defect were frequently caused or made worse because, in addition to the car defect, the human driver was speeding or was acting in some other form of faulty driving. Driverless vehicles, on the other hand, will themselves account for most or all real-time driving decisions, and any defect in their manufacturing or design will therefore play a greater—or sole—role in future car accidents.

23. Jane R. Bambauer, *Dr. Robot*, 51 U.C. DAVIS L. REV. 383, 393 (2017).

24. Jessica S. Allain, *From Jeopardy! to Jaundice: The Medical Liability Implications of Dr. Watson and Other Artificial Intelligence Systems*, 73 LA. L. REV. 1049, 1062 (2013).

25. Jason Chung & Amanda Zink, *Hey Watson—Can I Sue You for Malpractice? Examining the Liability of Artificial Intelligence in Medicine*, 11 ASIA-PAC. J. HEALTH L. & ETHICS 51, 68-70 (2018).

26. Ryan Abbott, *The Reasonable Computer: Disrupting the Paradigm of Tort Liability*, 86 GEO. WASH. L. REV. 1, 37-39 (2018).

27. See generally Chagal-Feferkorn, *supra* note 21.

28. "[W]hereas the more autonomous robots are, the less they can be considered to be simple tools in the hands of other actors (such as the manufacturer, the operator, the owner, the user, etc.); whereas this, in turn, questions whether the ordinary rules on liability are sufficient or whether it calls for new principles and rules to provide clarity on the legal liability of various actors concerning responsibility for the acts and omissions of robots where the cause cannot be traced back to a specific human actor and whether the acts or omissions of robots which have caused harm could have been avoided." European Parliament Resolution of 16 February 2017 with Recommendations to the Commission on Civil Law Rules on Robotics (2015/2103(INL)), EUR. PARL. DOC. P8_TA(2017)0051, ¶ AB, <http://www.europarl.europa.eu/sides/getDoc.do?type=TA&reference=P8-TA-2017-0051&language=EN&ring=A8-2017-0005> [hereinafter European Parliament Resolution].

they cause (for instance, through a compulsory insurance scheme developed for specific categories of robots).²⁹ The Committees on Transport and Tourism, on Employment and Social Affairs, and on the Environment, Public Health and Food Safety have all agreed that new liability rules ought to be developed to account for the new character of robotic decision makers.³⁰

Indeed, several regimes of liability rules have been offered in the context of such autonomous robots or systems. Many of them focus on the similarities between these systems and human beings, and propose to apply similar legal treatment to both, either by subjecting systems to the standards of reasonableness as discussed above,³¹ or by treating such systems as agents,³² employees,³³ or children³⁴ for the purposes of applying tort liability on their manufacturers or users (similar to the tort liability that would apply to a

29. “[C]reating a specific legal status for robots in the long run, so that at least the most sophisticated autonomous robots could be established as having the status of electronic persons responsible for making good any damage they may cause, and possibly applying electronic personality to cases where robots make autonomous decisions or otherwise interact with third parties independently.” *Id.* ¶ 59f.

30. Opinion of the Committee on Transport and Tourism for the Committee on Legal Affairs with Recommendations to the Commission on Civil Law Rules on Robotic (2015/2103(INL)), EUR. PARL. DOC. A8-005/2017, ¶ B, <http://www.europarl.europa.eu/sides/getDoc.do?type=REPORT&reference=A8-2017-0005&language=EN#title4> (“[W]hereas, for the purpose of civil liability, a distinction should be drawn between automated vehicles (incorporating devices allowing the automatic execution of some driving operations) and autonomous vehicles (which perform all such operations); . . . whereas in the former case the civil liability regime remains unchanged compared to that with conventional vehicles, while it needs to be adjusted in the latter case.”); Opinion of the Committee on Employment and Social Affairs for the Committee on Legal Affairs with Recommendations to the Commission on Civil Law Rules on Robotic (2015/2103(INL)), EUR. PARL. DOC. A8-005/2017, ¶ 4, <http://www.europarl.europa.eu/sides/getDoc.do?type=REPORT&reference=A8-2017-0005&language=EN#title6> (“[C]onsidering the increasing level of autonomy of robots, this should be accompanied by amending the rules on liability concerning the consequences associated with the actions or inaction of robots; is concerned by the lack of general framework and legal provisions with regard to work automation in this new and ongoing industrial revolution and considers it to be essential for the Union to specify a legal framework that reflects the complexity of robotics and its numerous social implications.”); Opinion of the Committee on the Environment, Public Health and Food Safety for the Committee on Legal Affairs with Recommendations to the Commission on Civil Law Rules on Robotic (2015/2103(INL)), EUR. PARL. DOC. A8-005/2017, ¶ 26, <http://www.europarl.europa.eu/sides/getDoc.do?type=REPORT&reference=A8-2017-0005&language=EN#title7> (“[C]alls on the Commission and on the Member States to promote the development of assistive technologies, also through liability schemes that are different from those currently applicable, in order to facilitate the development and adoption of these technologies.”).

31. *See generally* Abbott, *supra* note 26; Chagal-Feferkorn, *supra* note 21.

32. *See* SAMIR CHOPRA & LAURENCE F. WHITE, A LEGAL THEORY FOR AUTONOMOUS ARTIFICIAL AGENTS 5-28 (2011); Peter M. Asaro, *A Body to Kick, But Still No Soul to Damn: Legal Perspectives on Robotics*, in ROBOT ETHICS: THE ETHICAL AND SOCIAL IMPLICATIONS OF ROBOTICS 169, 178-80 (Patrick Lin et al. eds., 2012); Sam N. Lehman-Wilzig, *Frankenstein Unbound: Towards a Legal Definition of Artificial Intelligence*, 13 FUTURES 442, 451-52 (1981).

33. Lehman-Wilzig, *supra* note 32.

34. *Id.*; CHOPRA & WHITE, *supra* note 32, at 180.

principal, an employer or a parent, respectively). Other propositions focused on developing an insurance scheme adapted to the capabilities and potential danger posed by “sophisticated” or “autonomous” systems.³⁵

In any event, a preliminary question not yet discussed in depth is *when* the system becomes different than a ‘traditional product’ such that products liability is no longer a sufficient framework to treat damages caused by it.

Automated machines of different kinds were described in detail as early as 800 years ago.³⁶ Algorithms themselves date back more than 2,000 years.³⁷ What is it that separates these and other traditional algorithms and machines that have been, and may continue to be, subject to products liability rules from what I will generally refer to as “*thinking algorithms*” that seem to warrant their own custom-made treatment? Why have auto-pilots, for example, been traditionally treated as products³⁸ while autonomous vehicles are suddenly seen as a more human-like system that requires different treatment? Where is the fine line drawn between products and decision-makers?

While several scholars have touched on distinguishing traditional from sophisticated technologies for the purpose of applying products liability, no in-depth discussion specifically on this question has yet been investigated. Rather, the discussions focused on different tort frameworks that ought to be developed for “sophisticated” or “autonomous” technologies, with only anecdotal references to what it is that renders such technologies “sophisticated” or “autonomous.” Moreover, the potential parameters that have been mentioned to classify such technologies indeed related to the system’s level of autonomy (while some referred to autonomy expressly,³⁹ other scholars discussed whether the system is able to wholly replace humans⁴⁰ or whether it outperforms humans⁴¹—all aspects are associated with the system’s level of

35. European Parliament Resolution, *supra* note 28, ¶¶ 57-59.

36. *See generally* IBN AL-RAZZAZ AL-JAZARI, THE BOOK OF KNOWLEDGE OF INGENIOUS MECHANICAL DEVICES (Donald R. Hill trans., Pakistan Hijra Council 1989) (1974).

37. *See generally* *The Euclidian Algorithm*, KHAN ACAD., <https://www.khanacademy.org/computing/computer-science/cryptography/modarithmetic/a/the-euclidean-algorithm> (The Euclidean algorithm, for example, was invented around 300 BC and is still in use in existing technologies). Jeffrey Shallit, *Origins of the Analysis of the Euclidean Algorithm*, 21 HISTORIA MATHEMATICA 401-19 (1994).

38. The case of the 2013 Asiana Airlines crash in San Francisco is an example where the underlying legal actions consisted of products liability claims raised against Boeing, the manufacturer of the auto-throttle that allegedly failed. *See* Denise Johnson, *Fingers Point to Different Defendants in Asiana Airlines Plane Crash*, CLAIMS JOURNAL (Aug 6, 2018), <https://www.claimsjournal.com/news/national/2013/08/06/234344.html>; Brett Snider, *83 Asiana Passengers to Sue Boeing over Crash*, FINDLAW (July 17, 2013), <https://blogs.findlaw.com/injured/2013/07/83-asiana-passengers-to-sue-boeing-over-crash.html>.

39. European Parliament Resolution, *supra* note 28.

40. *See generally* Abbott, *supra* note 26 at 23.

41. *Id.* at 27-30, 39-41; *see also* Jason Millar & Ian Kerr, *Delegation, Relinquishment and Responsibility: The Prospect of Expert Robots*, in ROBOT LAW 102 (Ryan Calo et al. eds., 2016) (“[W]e suggest that a robot can be considered an expert when there is strong evi-

autonomy, as will be discussed below). This Article, however, demonstrates that autonomy is not a desirable differentiator between products and thinking algorithms, given its excessive complexity, the likelihood said differentiator will yield absurd or inconsistent results, and the non-practical nature of the results obtained using autonomy as a differentiator. A Roomba vacuum robot, for example, does replace a person in its action of cleaning; it (arguably) does so in a manner that outperforms the human cleaner,⁴² and it possesses sufficient levels of autonomy that enable it to decide on its own which direction to turn in order to continue cleaning. Yet the Roomba is probably not a thinking algorithm in the sense of requiring a specially-tailored tort legal framework that would apply to it rather than the traditional products liability framework.

This Article therefore proposes a new approach for distinguishing traditional products from “thinking algorithms” for the purpose of determining whether products liability should apply. Instead of examining the system’s characteristics in isolation, I propose a “purposive interpretation” approach: one that analyzes the system’s characteristics vis-à-vis the rationales behind the products liability legal framework, and identifies those associated with promoting said rationales versus ones adversely affecting the ability to accomplish them. The Article will thus offer a novel, practical method for differentiating traditional products from thinking algorithms, based on fulfilling the rationales behind products liability laws and hence provide decision-makers with tools to better decide when products liability should apply and when it should not.

Part I explores the concept of autonomy in algorithmic decision-makers and shows why the different aspects of autonomy are insufficient for determining which systems should and which should not be subject to products liability laws. Part II provides background to the legal framework of products liability and its rationales. Part III then addresses a more preliminary discussion of whether thinking algorithms even fall within the realm of products liability, given that they do not necessarily reflect a product and that damaging decisions reached by them are not necessarily the result of a defect. Part IV analyzes how different characteristics of sophisticated systems affect the ability to fulfil the rationales behind the products liability framework. It therefore points out the characteristics that render the accomplishment of these rationales more difficult to achieve, and demonstrates which specific system might indeed warrant a different legal treatment. Part V mentions alternative suggestions to imposing liability on damages caused by advanced technology and concludes that in either event, a preliminary step is evaluating if indeed the system calls for such an alternative or may continue to be subject to existing tort frameworks.

dence that it is capable of consistently performing a well-defined set of tasks traditionally associated with human expertise, with results that are, on average, better than the average human expert.”).

42. Though not of statistical significance, it certainly outperforms the author of this article.

PART I: WHY NOT USE AUTONOMY TO DISTINGUISH PRODUCTS FROM THINKING ALGORITHMS?

A. *On Automation and Autonomy*

In the context of the recent calls to develop a specialized tort legal framework for “sophisticated” or “autonomous” technologies, it is important to remember that automated machines, characterized by different levels of sophistication, have been utilized by humankind for centuries.⁴³ In addition to open-loop control systems, which executed automated tasks on an injective predetermined and unchangeable trajectory,⁴⁴ closed-loop control systems have for centuries been capable of automatically choosing among predetermined options, based on real-time feedback. Open-loop control windmills, for example, employed the power of wind to grind grain automatically whenever the wind blew. So long as the system was intact it would continue automatically to grind the same predetermined amount of grain on windy days, over and over again.⁴⁵ Closed-loop control windmills, however, were also able to increase the output of the windmill by automatically enlarging the amount of grain poured into the mill as the wind blew stronger (based on a hopper suspended on ropes that received more “knocks” the stronger the wind blew, and automatically caused more grain to be dispensed).⁴⁶ Modern examples of closed-loop automation systems are ubiquitous. They include Roomba vacuum robots (altering their routes on the ground based on feedback such as striking obstacles);⁴⁷ automated parking systems (basing their motion on feedback received by their sensors and cameras)⁴⁸ or even automated snore stopper pillows (a microphone embedded in the pillow identifies sounds of snoring, and based on that feedback automatically inflates the pillow to change the snorer’s

43. IBN AL-RAZZAZ AL JAZARI, *supra* note 36. *See also Automation*, MERRIAM-WEBSTER, <https://www.merriam-webster.com/dictionary/automation> (last visited Aug. 5, 2018) (“1: the technique of making an apparatus, a process, or a system operate automatically; 2: the state of being operated automatically”); *Automatic*, MERRIAM-WEBSTER, <https://www.merriam-webster.com/dictionary/automatic> (last visited Aug. 5, 2018) (“2: having a self-acting or self-regulating mechanism”).

44. JOSEPH DISTEFANO ET. AL., *SCHAUM’S OUTLINE OF THEORY AND PROBLEMS OF FEEDBACK AND CONTROL SYSTEMS* 3-4 (1990).

45. *Id.*; *see also* OTTO MAYR, *THE ORIGINS OF FEEDBACK CONTROL* 6, 90-93, 129-31 (1970). A boiler set by timer to turn on at a certain timing is an example of a modern open loop control mechanism, as the automation of turning the boiler on is independent of any variable other than the timing which was originally set.

46. *Id.* at 90-93. A boiler operating with a thermostat is a closed loop control mechanism, as its continuous operation depends on feedback pertaining to the current water temperature.

47. Julia Layton, *How Robotic Vacuums Work*, HOWSTUFFWORKS (Nov. 3, 2005), <https://electronics.howstuffworks.com/gadgets/home/robotic-vacuum2.htm>.

48. Laura McQuarrie, *The Valet Park4U Is a Smart Self-Parking System from Valeo*, TRENDHUNTER (June 3, 2014), <https://www.trendhunter.com/trends/automatic-parking>.

posture).⁴⁹ Countless other old and new closed-loop control systems allow us to enjoy the benefits of automation daily.⁵⁰ The mere fact that the system adjusts its actions based on external feedback, however, has existed for generations and in itself certainly does not warrant legal treatment different from that of traditional products. What then might assist in making such a distinction?

Legal scholars engaged in products liability and sophisticated systems have suggested various directions to answer said question. The above report by the European Parliament, for example, proposed defining a “smart robot” as one whose *autonomy* is established by its interconnectivity with the environment (potentially through the use of sensors), and its ability to adapt its action to changes in it.⁵¹ Millar and Kerr refer to “expert” robots, which would be classified as such when on average they consistently performed a well-defined set of tasks traditionally associated with human expertise—and do so better than the average human expert.⁵² Abbott focuses on the system’s ability to replace humans, more particularly on its ability to determine for itself how to complete tasks as set by humans.⁵³ Bambauer separates applications based on mere “measurement” from applications whose function is “knowledge-based”;⁵⁴ Chung and Zink focus on the higher level of the system’s duties, which require abilities of interpretation and analysis, while also distinguishing

49. Michael Plishka, *Pillow Adjusts Your Sleep Position*, TRENDHUNTER (Feb. 7, 2014), <https://www.trendhunter.com/trends/pillow-snoring-automatic>.

50. To mention a few amusing examples, automated doors for pets open automatically when receiving feedback indicating pet movement. Rahul Kalvapalle, *The Petwalk Pet Door Is Both Pet-Friendly and Eco-Friendly*, TRENDHUNTER (Mar. 14, 2014), <https://www.trendhunter.com/trends/pet-door>. Special utensils to reduce eating speed are available too; based on motion feedback received by the fork’s sensors, the utensil vibrates whenever the user is eating too fast. Jordan Minor, *HAPIfork is a Smart Fork For Your Dumb Mouth*, GEEK.COM (Aug. 2, 2018), <https://www.geek.com/tech/hapifork-is-a-smart-fork-for-your-dumb-mouth-1730322>.

51. Annex to the Resolution: Recommendations as to the Content of the Proposal Requested, EUR. PARL. DOC. P8_TA(2017)0051, ¶ AB, <http://www.europarl.europa.eu/sides/getDoc.do?type=TA&reference=P8-TA-2017-0051&language=EN&ring=A8-2017-0005#BKMD-13>.

52. Though said definition may also suit a weaving machine replacing a human weaver in the nineteenth century, Millar and Kerr also explained that in general, experts are not measured by their ability to follow known instructions; instead they prove their expertise by “filling in the blanks” so that an identical set of instructions is more likely to succeed when performed by an expert, and more likely to fail when performed by someone else. Millar & Kerr, *supra* note 41, at 110.

53. Abbott, *supra* note 26, at 23 (“What distinguishes an ordinary product from a computer tortfeasor in this system are the concepts of independence and control. Autonomous computers, robots, or machines are given tasks to complete, but they determine for themselves the means of completing those tasks.”). Interestingly, Abbott rejects the test of whether or not the machine’s actions were foreseeable as distinguishing traditional from autonomous products. *Id.* at 23-24 (“But the difference between ordinary products and autonomous computers should not be based on predictability.”).

54. Bambauer, *supra* note 23, at 387 (“Measurement apps will have traditional instruments as their nearest conceptual neighbor, while knowledge apps emulate doctors or, perhaps, the patients’ informal networks of health advisers.”).

different systems according to their ability to make decisions as well as implement them.⁵⁵

As the following review shows, these diverse and seemingly wildly different tests all relate to different aspects of the system's autonomy, whether or not the term autonomy is stated expressly. The problem in classifying a system as a product or a thinking algorithm based on its level of autonomy, however, is the complexity of the term "autonomy," which is much greater than might be intuitively assumed.⁵⁶ Moreover, applying autonomy level or different aspects of it as a differentiator might in many cases lead to absurd or inconsistent outcomes—it merely provides an imprecise test whose results are not necessarily practical.

B. *The Many Faces of Autonomy*

Using autonomy level as means for distinguishing traditional products from thinking algorithms for the purpose of applying products liability laws is undesired. One of the reasons is that such a differentiator would be difficult to implement, as its level of complexity is excessively high.

First, autonomy is a spectrum rather than a binary classification.⁵⁷ As will be discussed below, autonomy consists of various attributes rather than a single one. In addition, many of these attributes—for example, the system's adaptability to changing conditions—are in themselves measured on a scale and cannot be determined on a binary basis.

Secondly, various spectrums of autonomy exist, each focusing on completely different aspects. One common measurement of autonomy is the system's *freedom to act without human involvement* (or the allocation of decision-making power to humans or machine). Sheridan's spectrum, for example, offers ten levels of autonomy, the lowest being a situation where all processes are accomplished by a human being, without any machine assistance; the highest level is where the machine selects the desired courses of action but also executes them, not even informing the human of its choice and ignoring the human altogether.⁵⁸ Adding even more to the complexity of said spectrum

55. Chung & Zink, *supra* note 25, at 77-78.

56. See William C. Marra & Sonia K. McNeil, *Understanding "The Loop": Regulating the Next Generation of War Machines*, 36 HARV. J.L. & PUB. POL'Y 1140, 1151 (2013) ("Autonomy is a complex concept with many components that cannot be captured simply by parsing decision-making independence.").

57. Marra & McNeil, *supra* note 56, at 1158.

58. The full list of Sheridan's spectrum levels is as follows: "Level 1: The computer offers no assistance, human must do it all. Level 2: The computer offers a complete set of action alternatives. Level 3: Narrows the selection down to a few, or Level 4: Suggests one, and Level 5: Executes that suggestion if the human approves, or Level 6: Allows the human a restricted time to veto before automatic execution, or Level 7: Executes automatically, then necessarily informs the human. Level 8: Informs the human after execution only if the human asks, or Level 9: Informs the human after execution if it, the computer, decides to do so. Level 10: The computer decides everything and acts autonomously, ignoring the human

is that, even with identical systems employing identical technology, the division of tasks to those performed by humans and those offloaded to the system varies between particular scenarios.⁵⁹ A GPS system, for example, merely assists the driver in navigation when driving in a familiar neighborhood. But once the driver finds herself in unfamiliar surroundings, the same GPS suddenly does much more than occasional assistance. Rather, the allocation of decision-making power then changes such that the driver tends to accept the system's directions without second-guessing it,⁶⁰ even in cases where it is clear the driver should have overridden the GPS's discretion.⁶¹

A second method of assessing autonomy is based on the system's *ability to replace humans*.⁶² This measurement is in itself branched and complex, as several sub-analyses have been suggested in that context. Most sub-analyses focus on the system's ability to adapt to changing conditions, although the

completely." Marra & McNeil, *supra* note 56. In between levels 1 and 10 are levels where the machine does not execute actions without human approval; executes only after the human has had time to veto the action; executes and then informs the human, based on the human's request or executes and then informs the human on the machine's own discretion of whether to inform or not. Raja Parasuraman et al., *A Model for Types and Levels of Human Interaction with Automation*, 30 IEEE TRANSACTIONS ON SYS. MAN & CYBERNETICS—PART A: SYS. & HUMS. 286, 287 (2000).

59. See generally Itiel E. Dror & Jennifer L. Mnookin, *The Use of Technology in Human Expert Domains: Challenges and Risks Arising from the Use of Automated Fingerprint Identification Systems in Forensic Science*, 9 LAW PROBABILITY & RISK 47, 47-48 (2010) ("These technologies may distribute cognition in a number of ways: [t]hey can work to help and support the human expert; they can work alongside the human expert in a collaborative partnership; or the technology can play a more critical and significant role than the human.").

60. *Id.* at 49 ("If the person is in an unfamiliar city, we may think that . . . the human is 'subservient' to the GPS, following the route uncritically without exercising independent judgment.").

61. For discussion of cases where GPSs sent drivers off a cliff, into a river, or to the wrong side of the street, see *infra* Part IV.

62. See *Automated Vehicles for Safety*, NHTSA, <https://www.nhtsa.gov/technology-innovation/automated-vehicles-safety> (outlining the Society of Automotive Engineers' (SAE) spectrum for autonomous vehicles). Level 0 of the SAE spectrum is no automation, where human drivers perform all tasks required for driving. Level 5 is full automation where the vehicle itself is capable of performing all driving functions under all circumstances. In between, automation (not autonomy) is used as a tool that increasingly assists the human driver. The full list of the SAE's spectrum levels is as follows: Level 0: No Automation. Zero autonomy; the driver performs all driving tasks. Level 1: Driver Assistance. Vehicle is controlled by the driver, but some driving-assist features may be included in the vehicle's design. Level 2: Partial Automation. Vehicle has combined automated functions, like acceleration and steering, but driver must remain engaged with the driving task and monitor the environment at all times. Level 3: Conditional Automation. Driver is a necessity, but it is not required to monitor the environment. Driver must be ready to take control of the vehicle at all times with notice. Level 4: High Automation. Vehicle is capable of performing all driving functions under certain conditions. Driver may have the option to control the vehicle. Level 5: Full Automation. Vehicle is capable of performing all driving functions under all conditions. Driver may have the option to control the vehicle. SAE's highest level of automation therefore still grants the human driver discretion to maintain or take control over the vehicle.

specific tests vary significantly. They include the relatively simple test of whether the machine is limited to choosing among pre-programmed options or is capable of choosing options that are not fully pre-programmed.⁶³ They also include more complex tests, namely if the following three attributes of autonomy are met: *frequency of human operator interaction*, *machine's tolerance for environmental uncertainty* and *level of assertiveness of the machine*.⁶⁴ Another approach focuses on whether the system is able to perform all the following types of activity: *skill-driven*, *rules-driven* and *knowledge-driven*.⁶⁵

A third method for evaluating autonomy refers to a stronger or substantial measure of autonomy that will probably be developed more in the future, focusing on the system's own cognitive-awareness and real freedom of choice.⁶⁶ The spectrum proposed by the Air Force Research Lab (AFRL), for example, refers at the highest levels of autonomy to systems that are cognizant of their environment and not merely possess "knowledge" on it.⁶⁷

63. Kenneth Anderson & Matthew Waxman, *Law and Ethics for Autonomous Weapon Systems: Why a Ban Won't Work and How the Laws of War Can*, HOOVER INSTITUTION: JEAN PERKINS TASK FORCE ON NATIONAL SECURITY AND LAW 5-6 (May 17, 2013), http://media.hoover.org/sites/default/files/documents/Anderson-Waxman_LawAndEthics_r2_FINAL.pdf.

64. According to Marra and McNeil, *supra* note 56, at 1152-55, a machine is the more autonomous the less frequently a human operator must intervene and give it instructions; the more adaptability it shows in the face of scenarios it is not fully programmed to encounter; and the more it is able to change its operating plan in order to achieve its pre-programmed task, for instance, when the machine is "stuck."

65. See Antonio Chialastri, *Automation in Aviation*, in AUTOMATION 79, 83-84 (Florian Kongoli ed., 2012) (illustrating which of the system's abilities are required in order to perform). The last example addresses its ability to make decisions when the anticipated "rules" or circumstances change and the system has to come up with the optimal reaction to a scenario it was not prepared for. *Id.* In more detail, human activity may roughly be divided into three groups: *skill-driven* activities, namely the ability to accomplish physical tasks; *rules-driven* activities, namely the ability to comply with pre-determined rules; and *knowledge-driven*, namely the ability to make decisions when the rules previously mentioned are inadequate. *Id.* An autopilot, for instance, is used for *skill-driven* activities (such as maintaining very precise altitude); for *rules-driven* activities (for instance, landing a plane at the correct angle in certain wind and weather conditions—based on instructions or rules predetermined for such conditions); and theoretically for *knowledge-driven* abilities, such as in a case of unfamiliar malfunction. *Id.* The famous "Miracle on the Hudson," where U.S. Airways pilot Sully Sullenberger safely force-landed an airplane after two of its engines suddenly failed, is an example of human's superiority in "knowledge-driven" decisions, given that pilot Sullenberger made the "right decision" as opposed to flight algorithms which would have seemingly led to a catastrophic crash. Clint Eastwood's movie, *Sully*, 2016 focuses on that exact point. *SULLY* (Flashlight Films 2016). See also Adam Smith, *The Miracle on The Hudson: How It Happened*, TELEGRAPH (Nov. 22, 2016), <https://www.telegraph.co.uk/films/sully/miracle-on-the-hudson-how-it-happened>.

66. See, e.g., Eliav Lieblich & Eyal Benvenisti, *The Obligation to Exercise Discretion in Warfare: Why Autonomous Weapon Systems Are Unlawful*, in AUTONOMOUS WEAPONS SYSTEMS: LAW, ETHICS, POLICY, *supra* note 56, at 250-51.

67. At its lowest level, the AFRL's spectrum refers to a system that is remotely controlled by a human or executes missions entirely pre-planned by humans. At its mid-level

Thirdly, each of the tests discussed above must also consider the specific stage of the machine's decision-making process.

According to military strategist John Boyd, the decision-making process comprises a continuous cycle of decision-making stages, consisting of four steps: observe, orient, decide and act.⁶⁸ This cycle, known as the "OODA Loop," has been used in both military, business and litigation contexts, to explain the process of decision-making and to assist in gaining tactical advantages over opponents.⁶⁹ The OODA Loop is not limited to the decision-making process of humans alone, and could apply to machines. Thomas Sheridan, for example, has developed a model for machine's information-processing, which includes very similar stages of the OODA Loop, including: Information Acquisition, Information Analysis, Decision Selection and Action Implementation.⁷⁰ Naturally, as a machine becomes capable of performing more of the four steps on its own, the better the odds are that it will achieve a higher autonomy score on the spectrum of whether it can replace humans or not.⁷¹ But in addition to this partial overlap with the measures of autonomy discussed above, the OODA Loop may be independently used as a separate dimension on each of the autonomy measures. For instance, the "information acquisition" stage of a certain machine or algorithm may be characterized by high levels of autonomy based on the above spectrums, while the same system's "decision selection" stage might involve very little autonomy.

Also, the information acquisition stage performed by a military drone might include gathering data on potential targets without the need for any human involvement—thus the drone receives a high autonomy score for the first test of autonomy. It might do so even in the face of changing weather conditions or new disguise methods used by the potential targets—thus receiving a high score for the second spectrum of autonomy as well. At the same time, the decision-making stage for exactly the same drone may involve very little autonomy—at least with respect to the first spectrum—as a decision

stages of autonomy, the system itself may respond to real-time events. Marra & McNeil, *supra* note 56, at 1155-58. The full list of the AFRL's spectrum levels is as follows: Level 0: Remotely piloted vehicle. Level 1: Execute pre-planned mission remotely. Level 2: Changeable mission. Level 3: Robust response to real time faults/events. Level 4: Fault/event adaptive vehicle. Level 5: Real-time multi-vehicle coordination. Level 6: Real-time multi-vehicle cooperation. Level 7: Battlespace knowledge. Level 8: Battlespace single cognizance. Level 9: Battlespace swarm cognizance. Level 10: Fully autonomous. *Id.* at 1157.

68. FRANS P.B. OSINGA, *SCIENCE, STRATEGY AND WAR: THE STRATEGIC THEORY OF JOHN BOYD* 1-3 (2007).

69. *See, e.g.,* A.S. DREIER, *STRATEGY, PLANNING & LITIGATING TO WIN: ORCHESTRATING TRIAL OUTCOMES WITH SYSTEMS THEORY, PSYCHOLOGY, MILITARY SCIENCE AND UTILITY THEORY* 74-85 (2012).

70. Parasuraman et al., *supra* note 58; Marra & McNeil *supra* note 56, at 1153-55.

71. Albeit not necessarily in the context of whether or not it is good at adapting to new conditions, which is not necessarily related to being able to execute all four stages of the OODA Loop.

to hit a target will very likely require human authorization and not be executed at the drone's own discretion.⁷²

As demonstrated above, the different measurements of "autonomy" are numerous and complex and at times overlap. Should decision-makers decide to rely on all these measurements when determining whether products liability laws ought to apply, they would have to develop a complex matrix, accounting for all the different aspects of autonomy discussed. Moreover, since the matrix will likely not be of a one size fits all type, decision-makers would have to decide in advance, for each specific case and for each product type, how much weight to give each measurement.⁷³ Such a differentiator has additional disadvantages, including the fact that the measurements themselves are imprecise. For example, when determining the system's tolerance to environmental changes, the decision maker would not be facing a yes or no question, and would have to come up with an out-of-context numerical or qualitative estimation of the system's tolerance of such conditions. Moreover, factoring in all these imprecise estimations into a combined outcome that classifies the system's level of autonomy would lead to a more general sensation of whether the system is autonomous or not. Should the outcome be situated somewhere in the middle of the scale rather than near any of its ends, it would not be a useful tool for determining if we are in the traditional product kingdom or have found ourselves in the realm of thinking algorithms. In fact, a differentiator based on the level of autonomy might lead to absurd results or results incompatible with current classifications. Furthermore, the score or outcome received, even if placed on one of the ends of the autonomy scale, would merely give us an indication of the system's autonomy. It would not, however, give us any indication of the desirability of applying products liability laws to the system.

Choosing but a few autonomy measurements to rely on rather than the whole matrix might render the process less complicated, but would still suffer from imprecision and lack of correlation to the desirability of applying a products liability legal framework. Such an approach would also increase the risk of reaching absurd or inconsistent results.⁷⁴

72. See, e.g., Michael N. Schmitt & Jeffrey S. Thurnher, "Out of the Loop": *Autonomous Weapons Systems and the Law of Armed Conflict*, 4 HARV. NAT'L SEC. J. 231, 255 (2013); Markus Wagner, *Taking Humans Out of the Loop: Implications for International Humanitarian Law*, 21 J.L. INFO. & SCI. 155, 155-57 (2011); Kenneth Anderson & Matthew Waxman, *Law and Ethics for Robot Soldiers*, POL'Y REV. (Dec. 1, 2012), <https://www.hoover.org/research/law-and-ethics-robot-soldiers>.

73. Certain measurements will likely be relaxed and others more scrutinized, based on the specific type of system, its capabilities, and its potential usages.

74. As demonstrated above, analyzing a drone's level of autonomy based on its ability to act without humans' involvement while leaving aside the underlying OODA Loop stage would be meaningless. Similarly, focusing on a system's ability to replace humans or outperform them might render a coffee machine or even a 19th century engraver "autonomous."

Therefore, I introduce a different approach to distinguishing products from thinking algorithms. Rather than focusing on a general out-of-context analysis of autonomy, I analyze how different yes or no features of traditional products and their self-learning counterparts reconcile with the rationales behind products liability laws.

PART II: PRODUCTS LIABILITY RATIONALES

The birth of products liability, among the most popular of all case types in the U.S.,⁷⁵ is attributed to technological advances: The shift from local craftsmen to mass production factories caused a “lack of privity” problem that eliminated victims’ means of redress.⁷⁶ The introduction of products liability resolved that discrepancy by eliminating the requirement of privity of contract between the injured and the tortfeasor.⁷⁷ Under products liability laws, the seller or manufacturer of a defective product in a condition that is unreasonably dangerous is liable for the physical harms caused to the user or her property, even when there is no contractual relationship between them.⁷⁸ One of the main rationales behind products liability laws therefore is compensating the victim, which stems from corrective justice principles, under which the tortfeasor is required to correct the wrong she has committed based on justice and fairness considerations.⁷⁹ Comprising the general rationale of compensating the

75. Tens of thousands of products liability cases are filed annually, more than any other case type. RONALD C. PORTER, *LEX MACHINA, LEX MACHINA—2018 PRODUCT LIABILITY LITIGATION REPORT* 2 fig. 2 (2018). In 2017, for example, 42,789 products liability cases were filed in the US. *Id.* at 1. The total number of commercial, IP, employment, antitrust, securities and bankruptcy cases filed that year combined was lower than said figure. *Id.* at 1-2.

76. *See, e.g., Winterbottom v. Wright*, 152 Eng. Rep. 402, 405 (Exch. 1842) (holding that where the English Court of Exchequer decided a case where a mail coach collapsed and injured the plaintiff, “[t]here is no privity of contract between the parties; []if the plaintiff can sue, every passenger, or even any person passing along the road, who was injured by the upsetting of the coach, might bring a similar action[, and u]nless we confine the operation of such contracts as this to the parties who entered into them, the most absurd and outrageous consequences, to which I can see no limit, will ensue.”).

77. Gifford, *supra* note 2 (recounting that in 1916, the New York Court of Appeals in *MacPherson v. Buick Motor Co.*, 111 N.E. 1050, 1051, held that Buick owed a duty of care to its users despite not having privity of contract with them: “If to the element of danger there is added knowledge that the thing will be used by persons other than the purchaser, and used without new tests, then, irrespective of contract, the manufacturer of this thing of danger is under a duty to make it carefully.”).

78. RESTATEMENT (SECOND) OF TORTS § 402A (AM. LAW INST. 1965).

79. In *Escola*, Justice Traynor of the Supreme Court of California wrote that “[t]he cost of an injury and the loss of time or health may be an overwhelming misfortune to the person injured, and a needless one, for the risk of injury can be insured by the manufacturer.” *Escola v. Coca Cola Bottling Co.*, 150 P.2d 436, 441 (Cal. 1944) (Traynor, J., concurring). *See also* John C.P. Goldberg & Benjamin C. Zipursky, *The Easy Case for Products Liability Law: A Response to Professors Polinsky and Shavell*, 123 HARV. L. REV. 1919, 1944 (2010) (“In fact, we think that the case for allowing persons injured by defective products to obtain redress is very easy. It rests on the idea that a manufacturer bears a

victims—which is one of the rationales behind the general legal framework of tort law⁸⁰—are sub-rationales that are more specific to products liability in particular. First, a manufacturer or seller marketing their product for public consumption assumes a special responsibility toward consumers who might be injured by the product.⁸¹ A related aspect is that consumers who bought products that caused injuries did so relying on manufacturers' advertisements regarding their products, including their safety.⁸² So it is only fair that when the sense of safety and security customers relied on turns out to be unjustified, they will be compensated.⁸³ Another sub-rationale relevant in the context of the fairness of compensating the victim is that manufacturers or sellers are generally in a better position to absorb or spread the costs of damages caused by their products⁸⁴ or to insure against those costs.⁸⁵ Thus, rather than having the unfortunate victim shoulder the overwhelming weight of the entire damages she has suffered, the costs may be insured, allocated among the entire group of consumers (in the form of a price increase), or marked as business expenses in manufacturers' budgets.⁸⁶

A second main rationale behind tort law in general is deterrence.⁸⁷ In the context of products liability, deterrence refers to deterring manufacturers from creating dangerous products, that is, *promoting safety*. Naturally, the threat of liability encourages manufacturers to improve the safety of their products: the safer the product, the less likely it is to cause damage, and the less likely manufacturers are to be sued and to pay damages.⁸⁸ Said rationale is indeed

responsibility to avoid causing injury by sending a dangerously defective product into the stream of commerce and is supported by principles grounded in negligence and warranty”).

80. See, e.g., ERNEST J. WEINRIB, *THE IDEA OF PRIVATE LAW* 3 (1995).

81. RESTATEMENT (SECOND) OF TORTS, § 402A cmt. c (AM. LAW INST. 1965); see also MARSHALL S. SHAPO, *THE LAW OF PRODUCTS LIABILITY*, 7-31 (3d ed. 1994).

82. John E. Montgomery & David G. Owen, *Reflections on the Theory and Administration of Strict Tort Liability for Defective Products*, 27 S.C. L. REV. 803, 809 (1976).

83. This is regardless of victims' means of redress in the form of breach of warranty, a contractual cause of action that in general does require privity between plaintiff and defendant. U.C.C. § 2-312 (2004); see e.g., *All W. Electronics, Inc. v. M-B-W, Inc.*, 75 Cal. Rptr. 2d 509, 514 (1998); *T.W.M. v. Am. Med. Sys., Inc.*, 886 F. Supp. 842, 844 (N.D. Fla. 1995). The privity condition, however, is not required in all states, see, e.g., *Renaissance Leasing, LLC v. Vermeer Mfg. Co.*, 322 S.W.3d 112, 129 (Mo. 2010).

84. See, e.g., Montgomery & Owen, *supra* note 82, at 809-10; *Escola*, 150 P.2d at 436.

85. George L. Priest, *The Current Insurance Crisis and Modern Tort Law*, 96 YALE L.J. 1521, 1559 (1987) (“Courts justified third-party insurance coverage based on how easy it seemed to be for manufacturers or service providers to aggregate risks by adding an insurance premium to the price of the product or service.”).

86. Montgomery & Owen, *supra* note 82.

87. See, e.g., Richard A. Posner, *The Value of Wealth: A Comment on Dworkin and Kronman*, 9 J. LEGAL STUD. 243, 244 (1980); John C.P. Goldberg, *Twentieth-Century Tort Theory*, 91 GEO. L.J. 513 (2003).

88. See RESTATEMENT (THIRD) OF TORTS: PROD. LIAB. § 2 cmt. a (AM. LAW INST. 1998) (“On the premise that tort law serves the instrumental function of creating safety

sensible in the context of products liability, given that manufacturers are best positioned to eliminate or reduce the risks associated with their products (unlike consumers, manufacturers possess information regarding the product and can ensure inspections and quality control measures).⁸⁹

As will be elaborated below, the safety promotion rationale could be absolute, in the sense that whenever a defect exists the manufacturer will be liable regardless of the precautions she has taken to minimize the damage threat. Such liability is classified as strict liability, which requires no fault.⁹⁰ Alternatively, the safety promotion rationale may be applied only to a limited extent, namely to the extent that it is feasible and economical for the manufacturer to take the extra precautions such that their benefits outweigh their costs⁹¹ (as will be elaborated on in the following paragraphs, this type of liability is reflected mostly in design defect scenarios).

The above rationales, as well as the effects of technological advances, are also reflected in the specific mechanisms that govern the different kinds of defects in products, namely *manufacture defect*, *design defect*, and *failure to warn*. Manufacture defect occurs when the product was not properly manufactured (for example, due to a departure from the product's assembly specifications or use of non-appropriate materials). Under the laws of most states, manufacture defects expose sellers to strict liability, which does not require proof of any negligence by the seller, as long as the existence of a defect is proved.⁹²

incentives, imposing strict liability on manufacturers for harm caused by manufacturing defects encourages greater investment in product safety. . ."). Strict liability for harm caused by manufacturing defects has been supported on the ground that it promotes investment in product safety. *See, e.g., Hoven v. Kelble*, 256 N.W.2d 379, 391 (Wis. 1977) ("Strict liability is an effective deterrent; it deters the creation of unnecessary risks, or to put it positively, strict liability is an incentive to safety."); *U.S. Airways v. Elliott Equip. Co.*, No. 06-1481, 2008 WL 4425238, at *5 (E.D. Pa. Sept. 29, 2008) ("[I]mposing strict liability here would serve as an incentive to safety because Fluidics appears to be involved in these types of contracts on a regular basis and is in a better position than a consumer to prevent circulation of defective products.").

89. Manufacturers possess information on the design and production process of the product, as well as on potential alternatives. The exception is the potential usages consumers might make of the product, which are not always in the scope of the manufacturers' knowledge or expectations. *See* RESTATEMENT (THIRD) OF TORTS: PROD. LIAB. § 2(c) cmt. m. (AM. LAW INST. 1998).

90. It has mostly been applied under "manufacture defect" scenarios, as will be discussed below. *See, e.g., Richard L. Cupp & Danielle Polage, The Rhetoric of Strict Liability Versus Negligence: An Empirical Analysis*, 77 N.Y.U. L. REV. 874, 889 (2002).

91. Naturally, which benefits and which costs should be considered have been the subject of much debate and interpretation. *See, e.g., Cipollone v. Liggett Grp., Inc.*, 644 F. Supp. 283, 289-90 (D.N.J. 1986) (concluding that the benefits to be weighed do not include the benefits to the industry where the defendant operates or to its employees).

92. *See* RESTATEMENT (THIRD) OF TORTS: PROD. LIAB. § 2(a) (AM. LAW INST. 1998) ("[A] product . . . 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"); JOHN C.P. GOLDBERG & BENJAMIN C. ZIPURSKY, *THE OXFORD INTRODUCTIONS TO U.S. LAW: TORTS* 284-88 (2010).

Among other things, the introduction of strict products liability is attributed to technological progress. Increasingly complicated production processes render it difficult for victims to investigate the damaging product and to prove negligence in the manufacturing process. Consequently, courts adopted the strict liability no-fault standard that allows victims' redress even as technology advances.⁹³ In addition to ensuring compensation for victims, strict liability—as explained above—strives to encourage manufacturers to improve the safety of their products. The rationale of safety promotion is also reflected in the second scenario of products liability—design defect—albeit in a more complex manner. As suggested by its title, a design defect is a flaw in the design of the product itself, regardless of how it is manufactured. Importantly, damages caused by algorithmic decision makers, which will be the main focus of our discussion, are most likely to fall under the category of design defect rather than manufacture defect.⁹⁴

In the past, design defect was governed by the consumer expectations test, that is, whether the product was dangerous beyond the expectations of an ordinary consumer.⁹⁵

Since no proof of negligence by the manufacturer or seller was required, the consumer expectations test could have been classified as a form of strict liability, reflecting again the rationales of compensating the victim as well as encouraging safety.⁹⁶ Nowadays however—again among other things for technology-related reasons that will be discussed below—the majority of states have moved away from the consumer expectations test to the risk utility test.⁹⁷ By that approach, a design defect occurs when foreseeable risks associated with the product could have been minimized by using a feasible safer alternative.⁹⁸

93. Gifford, *supra* note 2, at 118 (“As handicrafts have been replaced by mass production with its great markets and transportation facilities, the close relationship between the producer and consumer of a product has been altered. Manufacturing processes . . . are ordinarily either inaccessible to or beyond the ken of the general public. The consumer no longer has means or skill enough to investigate for himself the soundness of a product. . . .” (quoting *Escola*, 150 P.2d at 443 (Traynor, J., concurring))).

94. This is because problems associated with coding do not stem from occasional defects in manufacture but instead are implemented all along the product line, which is compatible with design defects. See F. Patrick Hubbard, “*Sophisticated Robots*”: *Balancing Liability, Regulation and Innovation*, 66 FLA. L. REV. 1803, 1854 (2014).

95. See RESTATEMENT (SECOND) OF TORTS § 402A cmt. c (AM. LAW INST. 1965); DAVID G. OWEN, *PRODUCTS LIABILITY LAW* 292-99 (3d ed. 2014).

96. Gifford, *supra* note 2, at 117-18.

97. See Owen, *supra* note 95, at 292-99; Jeffrey K. Gurney, *Sue My Car Not Me: Products Liability and Accidents Involving Autonomous Vehicles*, U. ILL. J.L. TECH. & POL'Y 247, 262 (2013). Other reasons for not adhering to the consumer expectations test are that it connotes contract-based rather than tort-based liability; that its application to bystanders—who presumably had no expectations of the product—is problematic; and that obvious or patent defects might block victims means of redress as they could have expected the danger. Mary J. Davis, *Design Defect Liability: In Search of a Standard of Liability*, 39 WAYNE L. REV. 1217, 1234, 1236, 1231 (1993).

98. RESTATEMENT (THIRD) OF TORTS: PRODS. LIAB. §2(b) (AM. LAW INST. 1998).

To go back to our rationale of safety promotion, the risk utility analysis strives to improve safety, but does so only to the extent such improvements are feasible and economical. This is because the risk utility test does not require the use of the safest design possible, but of the safest design whose costs will not exceed the safety benefits it contributes as compared with the alternatives.⁹⁹ The calculations required to determine whether a safer alternative is or is not mandated are complex;¹⁰⁰ still, manufacturers are deemed best positioned to perform them.¹⁰¹

Asymmetrical information and manufacturers' better knowledge of the product and its potential risks are also key factors in the third scenario invoking products liability—failure to warn. Under this doctrine, manufacturers must adequately warn consumers of the existence of hidden dangers, as well as instruct them on the safe usage of the product, if they wish to avoid liability.¹⁰² Under the Restatement, the requirement is for reasonable instructions or warnings, which largely subjects this type of defect to the negligence standard as well.¹⁰³

The general trend in products liability law has indeed shifted back toward a more negligence-based approach.¹⁰⁴ Although strict liability fulfills the rationales of a victim's compensation and promotion of safety to a fuller extent,¹⁰⁵ other competing rationales and interests have swung the pendulum

99. Gurney, *supra* note 97, at 263; *see also* Turner v. Gen. Motors Corp., 514 S.W.2d 497, 504 (Tex. Civ. App. 1974) (“If a change in design would add little to safety, render the vehicle ugly or inappropriate for its particular purpose, and add a small fortune to the purchase price, then a court should rule as a matter of law that the manufacturer has not created an unreasonable risk of harm.”); *cf.* David G. Owen, *Toward a Proper Test for Design Defectiveness: “Micro-Balancing” Costs and Benefits*, 75 TEX. L. REV. 1661, 1673 (1997) (noting that products should possess benefits that outweigh their costs).

100. Taking into consideration factors such as the likelihood of damage, severity of damage, the costs of adopting a safer measure, the probability that the safer measure would indeed minimize the risk for damage or the magnitude of said damage and, in such cases, the degree to which the damage was indeed lower, etc. For additional factors rendering said calculations more complex “in the real world” (such as administrative costs and uncertainties as to the application of the law) *see* Keith Hylton, *The Law and Economics of Products Liability*, 88 NOTRE DAME L. REV. 2457, 2495-2497 (2013).

101. *See* Owen, *supra* note 99, at 1675-76.

102. *See* RESTATEMENT (THIRD) OF TORTS: PROD. LIAB. § 17, cmt. i (AM. LAW INST. 1998); Owen, *supra* note 99, at 1666.

103. RESTATEMENT (THIRD) OF TORTS: PROD. LIAB. § 2, reporter's note to cmt. m (AM. LAW INST. 1998) (“[A]n overwhelming majority of jurisdictions supports the proposition that a manufacturer has a duty to warn only of risks that were known or should have been known to a reasonable person.”).

104. The majority of products litigation involves design and warning defects which are subject to the negligence standard. Further, design defects are now subject to the risk-utility test rather than the consumer expectations test in most states, which, as explained *supra*, reflects a shift from a strict liability approach to the negligence approach. *See* Gifford, *supra* note 2, at 119-22.

105. Under strict liability, manufacturers pay for damages caused by defective products regardless of the level of care they have demonstrated. Victims are thus compensated in a larger proportion of cases. At the same time, the cost of damage is internalized by

back to the fault-based negligence approach. First, increased likelihood of liability is generally expected to impede development and innovation, which will create a “chilling effect” on technological advancement.¹⁰⁶ Second, increased levels of safety (potentially stemming from likelihood of liability), are expected to adversely affect different features of the product, including its pricing, ease of operation, appearance, and additional factors related to consumers’ preferences other than safety.¹⁰⁷ The legal framework of products liability therefore purports to strike an optimal balance between contradictory rationales and interests. Naturally, such a balance also depends on the specific product and industry.¹⁰⁸ Thinking algorithms’ unique features may well affect the optimal balance between rationales and interests. But before turning to said analysis, let us check whether such algorithms are at all relevant in the products liability context, or whether their features render the products liability legal framework completely inapplicable from the outset.

PART III: IS IT A PRODUCT? IS THERE A DEFECT?

One could argue that regardless of products liability rationales, algorithms that replace human discretion simply cannot be classified as “products.” Alternatively, it could be argued that damages caused by such algorithms may not be attributed to “defects.” If so, the argument goes, applying products liability rationales to distinguish which algorithms should be governed by products liability would simply be irrelevant, just like discussing free speech rationales in order to determine whether the howling of a wolf should be protected under the First Amendment would be irrelevant—the underlying legal framework simply does not apply to begin with. The next section addresses these preliminary queries and explains why products liability rationales should nevertheless be the basis for analysis.

manufacturers, such that the products’ prices reflect their true risks. A. Mitchell Polinsky & Steven Shavell, *The Uneasy Case for Product Liability*, 123 HARV. L. REV. 1437, 1440 (2010).

106. *See, e.g.*, Gifford, *supra* note 2, at 125; Colonna, *supra* note 8, at 109-11; *see also* COMMUNITY RESEARCH AND DEVELOPMENT INFORMATION SERVICE, FINAL REPORT SUMMARY - ROBOLAW (REGULATING EMERGING ROBOTIC TECHNOLOGIES IN EUROPE: ROBOTICS FACING LAW AND ETHICS), http://cordis.europa.eu/result/rcn/161246_en.html (concluding that products liability law might indeed create a chilling effect on the introduction of automated cars to the market).

107. *Cf.* DAVID G. OWEN ET AL., PRODUCTS LIABILITY AND SAFETY: CASES AND MATERIALS 202-03 (3d ed. 1996) (discussing factors typically taken into account under the risk utility analysis employed by courts).

108. *See, e.g.*, Goldberg & Zipursky, *supra* note 79 (rebutting Polinsky and Shavell’s, *supra* note 105, argument against the efficiency and desirability of products liability laws and emphasizing how different industries are differently affected by such laws). *See also* A. Mitchell Polinsky & Steven Shavell, *A Skeptical Attitude About Product Liability Is Justified: A Reply to Professors Goldberg and Zipursky*, 123 HARV. L. REV. 1949, 1949-50 (2010) (stressing that their contention as to “whether product liability is undesirable depends on the particular product”).

A. *On “Products”*

While several states have clearly defined the term “product” for the purpose of applying products liability,¹⁰⁹ in general it is up to the courts to determine in any given case whether an underlying damaging object is indeed a “product.”¹¹⁰ For many types of systems or machines that have caused damage, the classification as a product does not raise any question marks; but systems based on information—which naturally are the subject matter of this Article—do fall within the gray zone of the “products kingdom” and occasionally have been excluded from it.

First, several courts held that information in itself did not constitute a product for the purpose of applying products liability, because it lacked tangible form.¹¹¹ Secondly, and focusing on the analogy of information-based systems to professional services¹¹² (an analogy even better suited to thinking algorithms which, as discussed above, replace human professionals), courts in the past were sometimes inclined to rule that such services were not to be viewed as “products.”¹¹³

Yet in many other instances courts did treat information as a product and applied products liability laws when errors in the information caused damage,

109. See Joseph L. Reutiman, *Be a “Product” Subject to Products Liability Claims*, 22 CORNELL J.L. & PUB. POL’Y 181, 187 (2012) (noting that, as an example, IDAHO CODE ANN. § 6-1402(3) (2008) defines a product as “any object possessing intrinsic value, capable of delivery either as an assembled whole or as a component part or parts, and produced for introduction into trade or commerce” while excluding “[h]uman tissue and organs, including human blood and its components”).

110. See *id.* Cf. Michael D. Scott, *Tort Liability for Vendors of Insecure Software: Has the Time Finally Come?*, 67 MD. L. REV. 425, 425 (2008) (noting that the RESTATEMENT (THIRD) OF TORTS is drafted broadly, leaving much flexibility as to what things fall within its definition and what things do not); RESTATEMENT (THIRD) OF TORTS: PROD. LIAB. § 19(a) (AM. LAW INST. 1998) (“A product is tangible personal property distributed commercially for use or consumption.”).

111. *AmericaWinter v. G.P. Putnam’s Sons*, 938 F.2d 1033, 1036 (9th Cir. 1991) (acknowledging that the information at hand was embedded in a book, which is a tangible object, but nevertheless ruling that an encyclopaedia of mushrooms which misled a reader to consume poisonous mushrooms was not subject to products liability laws); *Torres v. City of Madera*, No. 09–16573, 2005 WL 1683736, 46–49 (E.D. Cal. 2005) (holding that training materials consisting of a CD-ROM, slide presentation, and paper handouts were not products and therefore not subject to products liability); *Am. Online, Inc. v. St. Paul Mercury Ins.*, 207 F. Supp. 2d 459, 467 (E.D. Va. 2002) (“[T]he plain and ordinary meaning of the word tangible is something that is capable of being touched or perceptible to the senses. Computer data, software and systems do not have or possess physical form and are therefore not tangible property as understood by the Policy.”). See Reutiman, *supra* note 109, for further analysis.

112. Scott, *supra* note 110, at 434–36.

113. *Id.* at 461–62 (“Professional services do not ordinarily lend themselves to the doctrine of tort liability without fault because they lack the elements which gave rise to the doctrine.” (quoting *La Rossa v. Sci. Design Co.*, 402 F.2d 937, 942 (3d Cir. 1968)). See also *Torres*, 2005 WL 1683736 at *13; *Lemley v. J & B Tire Co.*, 426 F. Supp. 1378, 1379 (W.D. Pa. 1997); *Snyder v. ISC Alloys*, 772 F. Supp. 244, 250 (W.D. Pa. 1991).

especially when the information was integrated with a physical object.¹¹⁴ Other considerations for applying products liability included whether the object was mass-produced¹¹⁵ or whether it had dangerous potential.¹¹⁶ Naturally, these considerations tend to exist in the “thinking algorithms” this Article focuses on (given that thinking algorithms are often embedded in physical objects such as cellular phones or computers, that they are often mass-marketed, and that errors in them might cause deadly results).¹¹⁷

It is therefore very likely, at least prima facie, that thinking algorithms too might find themselves classified as products, even if their entire essence is information, and even if their function replaces human services.

B. On “Defects”

It was pointed out in Part II that, as indicated by their titles, manufacture defect and design defect both require the existence of a “defect.” Thinking algorithms, however, are inherently expected to cause damage regardless of any defects. This is because sophisticated systems, in particular self-learning algorithms, rely on probability-based predictions,¹¹⁸ and probabilities by nature inevitably get it wrong some of the time. To take a concrete example: a medical algorithm is designed to diagnose patients and prescribe optimal treatment. Assume that the system is 100% certain that the patient has a type of disease

114. *Retail Sys., Inc. v. CNA Ins. Companies*, 469 N.W.2d 735, 737 (Minn. Ct. App. 1991) (“The data on the tape was of permanent value and was integrated completely with the physical property of the tape. Like a motion picture, where the information and the celluloid medium are integrated, so too were the tape and data integrated at the moment the tape was lost.”).

115. *Saloomy v. Jeppesen & Co.*, 707 F.2d 671, 676 (2d Cir. 1983); *Halstead v. United States*, 535 F. Supp. 782, 791 (D. Conn. 1982).

116. *Fluor Corp. v. Jeppesen & Co.*, 170 Cal. App. 3d 468, 474-75 (1985) (referring to a previous holding that only innately dangerous items might be subject to products and holding that errors in aeronautical charts could fall under said definition).

117. To take Waze as an example (which will be further analyzed in Part IV), Waze is embedded in the user’s cellphone, has 100 million active users, and, as elaborated in *infra* note 145, is of clear potential danger. See Greg Sterling, *Waze Launches ‘Local’ Ads Primarily Aimed at SMBs and Franchises*, SEARCH ENGINE LAND (Mar. 28, 2018, 9:20 AM), <https://searchengineland.com/waze-launches-local-ads-primarily-aimed-at-smb-and-franchises-295285>.

118. A Netflix algorithm recommending movies does so based on a numeric prediction that we would like said choice, relying on our previous taste and an analysis of enormous databases of other consumers’ preferences. Anthony Schneck, *The Subliminal Trick Netflix Uses to Get You to Watch Its Movies & Shows*, THRILLIST, <https://www.thrillist.com/entertainment/nation/how-new-netflix-recommendation-algorithm-works> (last updated Oct. 26, 2018, 10:59 AM). A bail algorithm recommending whom to release and whom to deny bail does so based on the probability that the suspect would break the law or escape if allowed to post bail. An application for choosing an optimal treatment for a patient too is based on the probability that the patient indeed has the medical condition diagnosed, and that she would react to the optimal treatment as most other patients would. Vigmesh Ramachandran, *Are Algorithms a Fair Way to Predict Who’ll Skip Bail?*, FUTURITY (June 5, 2017), <https://www.futurity.org/bail-bias-algorithm-1450462-2>.

that is cured without any intervention in 99% of the cases. But in the remaining 1% of the cases, the patient will die if not given treatment.¹¹⁹ Given the probabilities of success, in 99 of 100 cases the algorithm would be right to choose not to intervene. The algorithm, however, naturally cannot “tell” in advance whether it is dealing with the 99 “ordinary” cases, or whether the patient before it is the 1 in a 100 exception. Choosing the optimal course of action will therefore be based on damage expectancy: which decision, if taken a large number of times, will lead to the best results? Assuming the magnitude of damage caused by no intervention is “0” for 99% of the cases and is “100” for the remaining 1%, and that the magnitude of damage due to unnecessary intervention is “10” (assuming it leads to significant side effects), the optimal decision will be not to intervene (as damage expectancy of intervening is $99 \times 10 = 990$, whereas damage expectancy of not intervening is $1 \times 100 = 100$).¹²⁰ Yet every once in a while our medical algorithm will undoubtedly encounter some “exceptional” patients as well, for whom its optimal choice of no intervention will be catastrophic.

Focusing on the damage caused due to a user being on the bad side of the statistics certainly does not mean that the system was defectively manufactured or defectively designed. On the contrary: the system has reached the decision we would want it to reach. It just so happens that whenever thinking algorithms reach decisions based on probabilities—which is exactly what they are designed to help humans with—inevitable damage will occur when the general rule is applied in cases that in hindsight turned out to be the exceptions.¹²¹ Does this mean that thinking algorithms should never be governed by products

119. In real life, of course, the probabilities that are factored in are not one dimensional as in this example but are reflected in various stages of the decision-making process. Diagnosing the patient’s medical condition is in itself often probability-based. Even if all signs indicate cancer, for example, there is still a chance that the patient suffers from a different disease. See, e.g., Lucy McNally, *Woman Endures Months of Unnecessary Chemo Treatment After Being Wrongly Diagnosed with Cancer*, ABC NEWS (Nov. 17, 2016, 4:08 PM), <http://www.abc.net.au/news/2016-11-18/woman-given-unnecessary-chemo-treatment-bad-cancer-diagnosis-nsw/8036438>. Moreover, even if he does suffer from cancer, diagnosis of the specific type of cancer rather than another one is too probability-based. Even within the same types of cancer, different types of tumors exist, and determining between them might be a matter of probabilities. See, e.g., William B. Schwartz et al., *Pathology and Probabilities: A New Approach to Interpreting and Reporting Biopsies*, 305 N. ENG. J. MED. 917, 917-23 (1981).

120. Ignoring the additional damage of “20” sustained by the 1% of patients who did need the intervention, since for them the other alternative is worse.

121. Note that because choices are made based on damage probabilities, it is also possible that the decisions chosen would in fact be damaging in the vast majority of cases, and that this would still be deemed the “right” one to make. If, for example, there is a 1% probability of catastrophic damage of “100” (such as death), and a 99% probability of an average damage of “1” (such as a minor scratch), a thinking algorithm would be correct in preferring to cause scratches 99% of the times (with damage expectancy of “99”) rather than to cause death 1% of the times (with damage expectancy of “100”). In such cases, the algorithm’s choices would almost always be damaging, but in no way defective.

liability and that our analysis should have nothing to do with said legal framework? Not necessarily.

First, in addition to damages caused through no defect, thinking algorithms may certainly be responsible as well for defect-based damages, which do not stem from the user being on the bad side of statistics.¹²² Secondly, thinking algorithms are not unique in their ability to cause damage in the absence of a defect. Traditional products too may be defect-free yet nevertheless cause damage.¹²³ Indeed, to win a products liability lawsuit a plaintiff must prove—even when subject to strict liability theory—the existence of a defect,¹²⁴ implying that certain damages are not caused by a defect.

Granted, one could argue that in thinking algorithms damage caused without the existence of a defect is *inherent*. This would be unlike traditional products which do not inherently cause damage when no defect is involved. If so, products liability may indeed not be the most efficient framework applicable, especially given that products liability procedures are considered expensive and slow.¹²⁵ While a different legal framework—whose underlying assumption is that no defect exists in the first place—might be more efficient, this does not render our current products liability regime irrelevant altogether.

Thinking algorithms, despite their nature as information-based and although they may frequently cause damage regardless of a defect, may thus nevertheless be governed by products liability. I will therefore now turn to analyzing *when* a system is a thinking algorithm such that products liability rationales are less achievable when applied to damages caused by such systems.

122. If, for instance, the medical algorithm mentioned above had chosen to *intervene*, despite the lower chances of success and the higher damage expectancy, we would probably consider it a defect. By the same token, if the algorithm had reached the probabilities described above because it ignored critical information that was available, this would probably also be classified as a defect.

123. Car tires, for example, may explode after a certain time of usage and cause lethal damage. As long as the manufacturer provides appropriate warning as to their maintenance and frequency of replacement, any damage caused by such worn-out tires will not be attributed to a defect. *See, e.g.*, *Carmichael v. Samyang Tires, Inc.*, 923 F. Supp. 1514, 1518-22 (S.D. Ala. 1996) (pertaining to an accident resulting from tire failure and explaining that in order “[t]o maintain a claim under the AEMLD, a plaintiff cannot simply prove that an accident occurred and that he was injured; rather, ‘a defect in the product must be affirmatively shown’” (quoting *Townsend v. General Motors Corp.*, 642 So. 2d 411, 415 (Ala. 1994))).

124. *Carmichael*, 923 F. Supp. at 1518.

125. *See, e.g.*, Robert W. McGee, *Who Really Benefits from Liability Litigation?* 1-3 (Dumont Inst. for Pub. Policy Research, Policy Analysis No. 24, 1996), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=82596. This is even more so in the context of algorithmic decision makers. *See, e.g.*, Gurney, *supra* note 97, at 262-64; David C. Vladeck, *Machines Without Principals: Liability Rules and Artificial Intelligence*, 89 WASH. L. REV. 117, 137-41 (2014).

PART IV: DISTINGUISHING PRODUCTS FROM THINKING ALGORITHMS

As discussed above, this Article proposes an alternative approach for distinguishing traditional products from thinking algorithms which is not based, per se, on the system's level of autonomy. Rather, this part analyzes how different features or characteristics of different decision-making systems—manifested in the four different OODA Loop stages—affect the achievement of the different rationales behind products liability. Practically, the more the system's features reconcile with achieving products liability rationales, the more inclined we would be to classify them as traditional products. Systems whose features impede products liability rationales, however, should be classified as thinking algorithms that warrant different treatment.

Before delving into recognizing and discussing these features, let us provide some background on algorithms' self-learning abilities and their consequences. Given the enormous amounts of data that algorithms are exposed to and are capable of processing, algorithms' learning abilities allow them to learn from existing information and implement the conclusions in future sets of data. "Supervised learning" refers to a process where algorithmic training is more structured, in the sense that algorithms are fed with right and wrong answers pertaining to existing databases, so that they can develop a model for predicting the right answer for similar data sets that were not included in the training. A greater degree of freedom to come up with their own conclusions or recognize their own patterns is given to algorithms in the unsupervised learning process. Here algorithms are not fed any answers but are free to decipher patterns in the data that may indicate the right answer.¹²⁶ Systems' self-learning abilities do not depend only on the type of learning, supervised or unsupervised. The specific field they operate in also affects their capabilities, resulting from, among other things, the type and volume of available information. In the field of radiation oncology, for example, concerns have been raised regarding the possible pace of developing machine-learning abilities, given the difficulty of collecting standardized data sets that the algorithms could train on.¹²⁷ Other projections, however, view radiation oncology as a very good candidate for personalized treatment based on self-learning algorithms.¹²⁸ Based on said projections, radiation therapy will be one of two concrete sets of examples I will review in the next part that uses specific

126. Harry Surden, *Machine Learning and Law*, 89 WASH. L. REV. 87, 93-95 (2014); Avigdor Gal, Organisation for Economic Co-operation and Development [OECD], *It's a Feature, Not a Bug: On Learning Algorithms and What They Teach Us*, at 3, OECD Doc. DAF/COMP/WD(2017)50 (June 7, 2017), [https://one.oecd.org/document/DAF/COMP/WD\(2017\)50/en/pdf](https://one.oecd.org/document/DAF/COMP/WD(2017)50/en/pdf).

127. Babette Ten Haken, *Radiation Oncology, AI and Machine Learning in Research*, BABETTE TEN HAKEN (July 26, 2017), <https://babettetenhaken.com/2017/07/26/radiation-oncology-ai-machine-learning>.

128. Huan-Hsin Tseng et al., *The Role of Machine Learning in Knowledge-Based Response-Adapted Radiotherapy*, FRONTIERS ONCOLOGY 1, 2 (July 27, 2018), <https://doi.org/10.3389/fonc.2018.00266>.

systems in order to identify features whose existence renders it more difficult or more plausible to achieve the rationales behind products liability laws.

A. *Examples of Products Versus Thinking Algorithms*

One of the most notorious tragedies in the history of medical devices is that of Therac-25, a radiation therapy machine used to destroy cancerous tissues. Between 1985 and 1987, six patients in the United States and Canada were inadvertently given an overdose of radiation, resulting in three fatalities.¹²⁹ Investigation revealed that the system had several “bugs” causing it to accidentally release much higher dosages of radiation than prescribed by the machine’s technician.¹³⁰ Although the lawsuits filed in connection with the Therac-25 accidents were all settled before trial,¹³¹ Therac-25 was considered one of the first cases to give rise to products liability claims in connection with medical devices.¹³²

Therac-25 was useful for administering radiation in a precise and automatic manner; the example we shall use as its sophisticated counterpart is a machine also capable of taking and implementing professional decisions. Nowadays, new generation radiation machines mainly focus on improved precision of the radiation’s distribution. Equipped with infrared cameras and robotic beds, radiation machines now make automatic minor adjustments in the positioning of the patient throughout the radiation process to achieve more precise administering of the treatment.¹³³ An additional feature that could be embedded in radiation machines, however, would also include dose calculation algorithms that would enable the machine to administer radiation beams but also to decide (or recommend) the optimal treatment plan for each patient based on his unique characteristics.¹³⁴ Existing algorithms for calculating radiation dosage are used

129. Jonathan Jacky, *Programmed for Disaster: Software Errors That Imperil Lives*, SCIENCES, Sept.-Oct. 1989, at 22, 22; Nancy G. Leveson & Clark S. Turner, *An Investigation of the Therac-25 Accidents*, COMPUTER, July 1993, at 18, 23.

130. For example, one of the software “bugs” that caused several accidents occurred when the human operator inserting the treatment dosage made a change within the eight-second time window during which Therac-25 set its magnets for operation. In these instances the change was not registered. Leveson & Turner, *supra* note 129, at 21, 27-28.

131. Nancy G. Leveson, *Medical Devices: The Therac-25*, at 11 (Dec. 31st, 1995), <https://www.cs.indiana.edu/classes/p415-sjoh/readings/Therac25/therac-update.pdf>; Sara E. Dyson, *Medical Device Software & Product Liability: An Overview (Part I)*, MEDTECH INTELLIGENCE (Sept. 15, 2017), https://www.medtechintelligence.com/feature_article/medical-device-software-products-liability-overview-part/2.

132. Dyson, *supra* note 131.

133. *Next Generation State-of-the-Art Radiation Therapy System at Netcare N1 City Hospital*, NETCARE, <https://www.netcare.co.za/News-Hub/Articles/articleid/639/next-generation-state-of-the-art-radiation-therapy-system-at-netcare-n1-city-hospital> (last visited Nov. 7, 2018).

134. Which consists of “the prescribed dose level for the tumor, the number of therapeutic beams, their angles of incidence, and a set of intensity amplitudes.” Uwe Oelfke

today,¹³⁵ but the future system we take as an example makes use of learning algorithms that produce personalized dosage calculations based on the type of tumor involved and on other parameters that the system itself deems relevant, after “learning” from large databases of previous cases and deciphering correlations between different parameters and improved outcomes.

Our second example of a traditional product versus a thinking algorithm takes us to the world of driving, more specifically the world of navigation. Global Positioning Systems (GPSs) were gradually put to civilian use at the end of the last century,¹³⁶ and are now commonly used in different modes of transportation, including cars. By automatically determining a user’s current location, the GPS uses its database of maps to calculate a route from the specified location to the end point, as entered by the user.¹³⁷ In recent years, numerous accidents have been reported to be the result of following GPS instructions; in addition to unfortunate encounters of vehicles with speeding trains,¹³⁸ GPS systems have led drivers into a creek or to a cliff edge,¹³⁹ to enter a road in the wrong direction, to drive under a bridge too low for their vehicles to pass under safely,¹⁴⁰ and even guided them straight into a war zone.¹⁴¹ Though the classification of GPS systems as products is not necessarily clear-cut, numerous courts have acknowledged the somewhat analogous object of aeronautical charts as products for the purpose of applying products liability.¹⁴² A GPS is therefore the system we will use as an example of a traditional product for our analysis.

& Christian Scholz, *Dose Calculation Algorithms*, in *NEW TECHNOLOGIES IN RADIATION ONCOLOGY* 187, 187 (Wolfgang Schlegel et al. eds., 2006).

135. *Id.* at 188.

136. John E. Woodard, Comment, *Oops, My GPS Made Me Do It!: GPS Manufacturer Liability Under a Strict Products Liability Paradigm When GPS Fails to Give Accurate Directions to GPS End-Users*, 34 U. DAYTON L. REV. 429, 437 (2009).

137. *Id.* at 438-40.

138. *Id.*; see also Yaron Steinbuch & Hoa Nguyen, *Driver Blames GPS for Crash with Metro-North Train in Bedford Hills*, BHD. OF LOCOMOTIVE ENG’RS & TRAINMEN (Sept. 30, 2008), <https://www.ble-t.org/pr/news/headline.asp?id=23809>; *What if a GPS Causes an Atlanta Accident and Injuries?*, SAMMONS & CARPENTER P.C. (Jan. 19, 2017), <https://www.atlantainjuryattorneys-blog.com/2017/01/gps-causes-atlanta-accident-injuries.html>.

139. Mark Havnes, *Too Much Faith in GPS Leaves Monument Visitors Stranded*, SALT LAKE TRIB., Aug. 5, 2008; Edward Joel Lake, *Defective GPS Systems and Product Liability for Accidents*, AVVO (Nov. 21, 2012), <https://www.avvo.com/legal-guides/ugc/defective-gps-systems-and-product-liability-for-accidents>.

140. Lake, *supra* note 139; *\$15M Products Liability Lawsuit Blames GPS Maker Garmin for 2013 Massachusetts Bus Crash*, ALTMAN & ALTMAN LLP (Jan. 27, 2015), https://www.bostoncaraccidentlawyerblog.com/2015/01/15m_products_liability_lawsuit.html.

141. Woodard, *supra* note 136; Yuval Azoulay, *American Tourist Stoned by Mob After Accidentally Entering Qalandiya*, HAARETZ (June 25, 2008, 12:00 AM), <https://www.haaretz.com/1.4995814>.

142. For further discussion, see Part III above. See also *Cruz v. Talmadge*, 244 F. Supp. 3d 231, 232-33 (D. Mass. 2017) (remanding to the superior court department a case

Our equivalent example of a more sophisticated system is Waze, a community-based navigation application designed for navigation per se, but also to outsmart traffic.¹⁴³ Founded in 2006, the Israeli development makes use of real-time updates from its community of drivers—automatic ones sent by the system itself reporting its location, and traffic submissions actively sent by users—in order to create an ever-updating mapping of the roads system.¹⁴⁴ Thus, Waze’s algorithms are able to calculate not only the shortest way from point A to B, but also the quickest way at any given moment. It does so by learning the landscape in any parts of the globe Waze operates, as well as learning the current traffic condition of each and every route, based on real-time feedback from Waze’s users.

As successful as it is, Waze too has been blamed for leading drivers into dangerous situations that might result in harm.¹⁴⁵ As we will see in greater detail below, various features of Waze render it different from traditional GPSs, when the system’s compatibility with products liability laws is analyzed.

B. *Which Products Liability Rationales Apply to Which Algorithms*

For an orderly analysis of the compatibility of the various rationales of products liability laws with both sets of examples, the Article addresses each rationale separately and examines how the thinking algorithm’s features affect it as compared with the traditional product.

1. *Promoting Safety*

As discussed above, a central rationale behind the framework of products liability is to encourage manufacturers to better their products’ safety. Whether or not products liability indeed achieves said end has been the subject of heated debates, but for the sake of discussion we shall assume that in general it does.¹⁴⁶ But how is promotion of safety affected when the products at hand are sophisticated self-learning systems?

involving products liability claims due to GPS navigation orders that caused a bus to crash into a bridge whose height was below the height restriction for buses); *Turner v. Isecuretrac Corp.*, No. 03CA70, 2004 WL 944386, at ¶ 23 (Ohio Ct. App. Apr. 28, 2011) (accepting that a GPS ankle device was a “product” under products liability laws, without questioning it).

143. WAZE, <https://www.waze.com> (last visited Nov. 3, 2018).

144. *FAQ*, WAZE, (July 8, 2013), <https://web.archive.org/web/20130708113343/http://world.waze.com/faq>.

145. *Israeli Family Accidentally Enters Ramallah*, ARUTZ SHEVA (Feb. 24, 2018, 10:57 PM), <https://www.israelnationalnews.com/News/News.aspx/242361>; Chana Roberts, *Three Israelis Accidentally Enter Ramallah*, ARUTZ SHEVA (Dec. 10, 2016, 4:50 PM), <https://www.israelnationalnews.com/News/News.aspx/221510>.

146. According to opponents of the products liability regime, empirical evidence indicates that products liability has not been shown to significantly increase the safety level of products. This is so, the argument goes, because in most cases market powers as well as regulatory measures are themselves responsible for increased safety, such that the

Generally speaking, manufacturers of such systems might find themselves in a place where enhancing safety is very difficult or expensive, or will render the system inefficient. This point may be explained through the example of a futuristic system of a “robo-doctor” that would potentially be able fully to replace a human physician at all stages of decision-making.¹⁴⁷

First, manufacturers of traditional products have a finite number of parameters to consider when preparing for different scenarios and minimizing risks associated with the actions of their machines.¹⁴⁸ By contrast, robo-doctor manufacturers will have an enormous number of scenarios against which they must try to take precautionary measures, based on innumerable parameters: the patient’s general medical condition (blood type, vital signs, height, weight, etc.), past medical conditions (previous lab results, previous diagnoses, previous success or failure of past treatments, etc.), as well as current medical condition (e.g., for cancer diagnosis: type of cancer, its size, its location, its stage, which receptors it has, etc.); various external parameters (Are any epidemics indicated in that region? Do current weather conditions affect the likelihood of a certain diagnosis or the chances of success of a certain treatment? Should a certain blossom in the air be taken into account when assessing the reason for a symptom?); practical parameters (Are qualified staff available immediately to execute a certain medical choice? What is best

incremental addition in safety attributed to products liability litigation is non-significant. Polinsky & Shavell, *supra* note 105, at 1462-69. Either way, both proponents and opponents of the products liability regime agree that when the public and regulators’ level of awareness of the danger or damages caused by the product is low (for instance, because the product is not widely sold, because the media coverage of the damage is not significant, because the damages are difficult to trace back to the product or are only discovered after a long time, etc.), regulation and market forces will be less effective in promoting optimal safety. Goldberg & Zipursky, *supra* note 79, at 1941; Polinsky & Shavell, *supra* note 105, at 1449. General technological and social trends may result in damage being more easily and quickly discovered, and gaining more public exposure. To take the damage caused by the Therac-25 as an example, time improved safety mechanisms would likely have allowed the machine to check the dosage it administered in practice, and thus immediately discover discrepancies. In addition, nowadays reports on over-dosage caused by a radiation machine would probably rapidly ‘go viral’ due to globalization and the social media effects (as opposed to the Therac-25 accidents that occurred in the 1980s of which the public was not aware for many months). Indeed, systems causing damage nowadays are likely to attract greater attention from both the public and regulators than did systems that caused harm decades ago. In this sense, products liability in general perhaps is less critical today for increasing safety than it was in the past. But said difference does not stem from the difference in function or abilities of traditional products versus sophisticated ones but from general trends. As such, they are beyond the scope of our current discussion.

147. For discussions on such futuristic systems and their possible characteristics see generally Alex Woodie, *The Robo-Doctor Is [In]*, DATANAMI (Aug. 30, 2017), <https://www.datanami.com/2017/08/30/the-robo-doctor-is-in/>; Michael MacRae, *The Robo-Doctor Will See You Now*, AM. SOC’Y OF MECH. ENG’RS (May 2012), <https://www.asme.org/engineering-topics/articles/robotics/robo-doctor-will-see-you-now>.

148. To take a simple example of a coffee machine: verifying that the temperature of the liquid produced by the machine is not too hot, verifying that even if operated by a child the machine could not cause electrocution, etc.

practical choice in cases of understaff, or of shortage of specific medications in the hospital's stock?); ethical parameters (What does the patient truly want? How does it reconcile with ethical standards as well as the relevant legislation?).

Secondly, in addition to all the foregoing parameters and scenarios, manufacturers have to deal with information fed into the system by external systems, especially considering the Internet of Things revolution that is expected to connect "real world objects"¹⁴⁹ and allow machines to communicate with themselves directly, without human involvement.¹⁵⁰ The prospect, therefore, is that in the future, algorithms will have to rely on parameters fed to them by other machines¹⁵¹—such as a robo-doctor factoring blood pressure and heartbeat rates broadcast to it by a separate medical device, and perhaps having to assess the probability that said device has provided erroneous readings.

Thirdly, medicine, like other professional fields, changes constantly. Do we expect manufacturers to have their products updated daily (or even hourly) with every new study, while immediately incorporating that study's findings into the robo-doctor's decision-making process? Who will decide which studies should be updated and which are not convincing enough, or which are less relevant to the specific population treated by that robo-doctor? Having to account for such dynamic developments, the manufacturers' task in minimizing risk of error is undoubtedly far more difficult than in the case of static fields where products need not be subject to frequent updates.¹⁵²

Fourthly, this also leads to another unique challenge for manufacturers of thinking algorithms: medicine, like law and many other complex fields where judgment and discretion are significant, is not black and white. Different experts have different opinions and recommend different solutions under identical circumstances. How can manufacturers be expected to minimize the

149. Objects such as buildings, vehicles, home appliances, wearable electronics, etc. See, e.g., Melanie Swan, *Sensor Mania! The Internet of Things, Wearable Computing, Objective Metrics, and the Quantified Self 2.0*, 1 J. SENSOR & ACTUATOR NETWORKS 166 (2012).

150. "[M]ost of the communication will be automated between intelligent devices. Humans will intervene only in a tiny fraction of that flow of communication. Most of it will go on unsensed and really unknown by humans." Nicholas Gane et al., *Ubiquitous Surveillance: Interview with Katherine Hayles*, 24 THEORY, CULTURE & SOC'Y 349, 349 (2007); Katherine Hayles, *Unfinished Work: From Cyborg to the Cognisphere*, 23 THEORY, CULTURE & SOC'Y 159, 159 (2006) (In "highly developed and networked societies . . . human awareness comprises the tip of a huge pyramid of data flows, most of which occur between machines.").

151. See, e.g., Adam Thierer, *The Internet of Things and Wearable Technology: Addressing Privacy and Security Concerns Without Derailing Innovation*, 21 RICH. J.L. & TECH. 1, 4-17 (Feb. 13, 2015).

152. Complex as an autopilot system might be to design, for example, its manufacturers are not likely to encounter daily updates in the field of aviation (or meteorology, atmospheric science, etc.) that would require them to decide whether the new information is relevant to the design of the system and, if so, how the system should take the updates into account.

risk of certain scenarios by choosing a specific course of action, when the choice is not necessarily obvious? (In any event, if the matter is ever litigated the manufacturer could always be blamed for not choosing a different alternative that some experts preferred.)

Lastly, even if raising the safety level despite the countless scenarios would be possible (and not prohibitively costly), manufacturers of thinking algorithms will still have a very difficult time improving the rates of correct decisions of their systems from high to very high if they wish to keep the system efficient and user-friendly. For example, if the robo-doctor has to include in its decision-making process each and every possible medical condition, including extremely rare ones or implausible ones (in order to reduce the chances of misdiagnosis, even for the most unusual cases), it might not be able to be put to practical use. The extra time required for the system's information gathering and analysis process might make the process excessively long, make the system fail to respond in real-time, make patients refuse to tell their medical history if the process is so slow, or disincentivize hospitals from purchasing robo-doctors due to their inefficiency compared to humans.¹⁵³

In sum, the factors affecting how much a manufacturer of a sophisticated system could indeed increase its safety are these: the size of the matrix of parameters the algorithm must consider before making a decision,¹⁵⁴ the dynamic nature of the relevant professional knowledge, the lack of clear right choices, and the extent of trade-off between safety and efficiency. But these parameters are all very general, and deduced from an extreme example of a robo-doctor—a machine that wholly replaces one of the most complicated human professions. To try to concretize the parameters that affect manufacturers' ability to increase safety, hence to meet the first rationale of products liability, let us now analyze our set of two more delineated examples of radiation and navigation systems. The analysis centers on the level of *foreseeability* of the product's actions, as well as how far the manufacturer is able to *control* said actions, and do so *efficiently* (assuming that the less a

153. On the inherent trade-off between safety and efficiency in the field of robotics, see, for example, Christiana Braz et al., *Designing a Trade-Off Between Usability and Security: A Metrics Based-Model*, in HUMAN-COMPUTER INTERACTION—INTERACT 2007: 11TH IFIP TC 13 INTERNATIONAL CONFERENCE; RIO DE JANEIRO, BRAZIL, SEPTEMBER 2007; PROCEEDINGS, PART II, at 114 (Cécilia Baranauskas et al. eds., 2007). For a discussion on said trade-off in the pharmaceutical field, see, for example, Tomas J. Philipson & Eric Sun, *Is The Food and Drug Administration Safe and Effective?*, 22 J. ECON. PERSP. 85 (Winter 2008).

154. An added complication that increases the magnitude of scenarios a manufacturer must prepare for is when parameters are not binary but can be anywhere on a spectrum; this is even more so when the parameters are not deterministic but in themselves reflect probabilities. The parameter of a patient's past medical history of a stage 3 carcinoma, for example, does not reflect one hundred percent certainty that indeed this was the exact medical condition the patient suffered from. Rather, it reflects a certain probability that indeed the diagnosis was correct.

manufacturer is able to foresee and control the choices of its system, the less will it be able to raise the system's safety levels). To examine this, we shall review separately the system's operation in each of the four OODA Loop stages

a. *The OODA Loop—General*

We might not have thought about it this way, but even primitive automatic systems may be responsible for all four stages of the OODA Loop (provided they yield some sort of physical change). Even a simple drill, for example, acquires information about the fact that it has been turned on, and that the human operator has pressed the right button for the drill to start. It also analyzes the information and makes a decision based on it, in the primitive sense of considering that the start button was pushed, along with the fact that no safety lock mechanism was initiated, and then it decides to go ahead and drill, the drilling being the final OODA Loop stage of execution. But in such examples all three first stages of the process are dictated to the system by its human operator in an injective manner, leaving full foreseeability and full control in the manufacturer's hand (unless a bug occurs).

Thinking algorithms replace humans precisely because not only can they outperform them in the final physical execution stage, they can also automatically access and collect vast amounts of information from various sources, of a magnitude that the human brain could not read in decades. These algorithms can then analyze these enormous amounts of information that are beyond a human's grasp, and can make complex decisions based on probabilities that a human cannot even weigh.¹⁵⁵

The Therac-25 machine, for instance, only acquired information that was 100% dictated to it by the operator (having turned the machine on, chosen a specific treatment mode, inserted the desired parameters for treatment, etc.). It analyzed whether it was in a position to start operating, based on pre-programming that in an injective fashion ordered it to do so when all conditions for beginning an operation mode were met, and then blindly decided to go ahead and administer treatment—again because it was programmed to do exactly that once all the required conditions were met. A new generation radiation machine, however, would have more freedom or independence to conduct the first three stages of the OODA Loop, such that the manufacturer would no longer have full foreseeability or control over them. With respect to information acquisition, for example, it is possible that the system itself will decide which sources of data to harvest to improve its success rates (be they medical publications the manufacturer might not even be aware of, or random statistics on such matters as global warming that the system might suddenly find are correlated to radiation success rates). With regard to the analysis stage

155. See, e.g., Ron Bekkerman et al., *Scaling Up Machine Learning: Introduction*, in SCALING UP MACHINE LEARNING PARALLEL AND DISTRIBUTIVE APPROACHES, 1-6 (2012).

too, the radiation machine itself might determine how much weight to attach to each piece of information it has collected, again based on correlations it itself has discovered through its self-learning process. Facing several courses of action ranked with different probabilities of success and different expected damage, the machine itself might decide which alternative to choose, or otherwise decide that its confidence in the preferred option was not high enough. Then it will rather call a human for further instructions than decide to execute.¹⁵⁶ Naturally, the more OODA Loop stages a system is capable of performing in a manner not fully dictated by the manufacturer, the less foreseeability the manufacturer has as to the final outcome of the process, and the less control it obtains over it.

Therefore, and before delving into each of these stages separately, our first parameter indicating that the system is a thinking algorithm that is less compatible with the first rationale of products liability (given that reduced levels of control and foreseeability would render it more difficult to improve safety, as explained above) is the number of OODA Loop stages the systems performs in a manner not fully dictated by humans.

b. “*Observe*” (“*Information Acquisition*”)

As explained above, the information acquisition stage of Therac-25 was fully dictated by humans in the loop. A new generation radiation machine, depending on its specifications, might raise several separate aspects of lack of foreseeability and control by the manufacturer at the information acquisition stage. Naturally, when the system trains on closed sets of databases that the manufacturer has “fed” it, and later continues to collect information from data provided to it exclusively by the manufacturer, the manufacturer maintains foreseeability and control over the information acquisition stage. However, better results might materialize if the system is free to decide on its own to add additional sources of information (for instance, additional medical journals, readers blogs on radiation, etc.) as well as more types of information the system is interested in (for instance, when reviewing a patient’s medical record, collecting information on less trivial types of information such as the day of the week the patient was released from the hospital, etc.). In such cases, lack of foreseeability by the manufacturer is threefold: first, it cannot anticipate the *type of parameters* the system will choose to collect information on. Secondly, it cannot anticipate the *sources of information* the machine will harvest. And thirdly, it cannot anticipate the *specific content* of the data collected, be it content related to external databases such as medical journals (in which the more dynamic and routinely updated the database is, the less likely a

156. See *supra* note 58 and accompanying text for level 5 of Sheridan’s spectrum.

manufacturer is to have foreseeability over its content) or the values of the specific medical parameters measured for each patient.¹⁵⁷

Foreseeability and control at the information acquisition stage, therefore, are reduced if the system can decide on its own what type of information to seek or which sources of information to cover, and the more dynamic its information sources are.

To move to our second set of examples, from the navigation world, here both the traditional GPS and its sophisticated Waze counterpart are engaged in the information acquisition stage. In the case of the traditional GPS, the manufacturer has full foreseeability and control over the data acquired, since the system is based only on maps uploaded by the manufacturer itself (as a one-off occurrence when the system is released to the market and through updates that the manufacturer sends to the system or the user downloads from the manufacturer's website).

The manufacturer, however, does lack control over information that has *changed* from the time the maps were updated until the actual usage of the system (roads closed since the last update, regions that have become hostile, bridges built with a maximal height cap, etc.). Likewise, manufacturers will probably have limited foreseeability with respect to such changes (major changes may well be planned and advertised in advance and thus are foreseeable, but small or temporary changes might be executed without the manufacturer learning about them in advance).

How do the foreseeability and control elements change when the system is not based on maps uploaded by the manufacturer, but on information sent by the users themselves in real time?

Naturally, the system's being based on real-time reports sent by millions of strangers limits the amount of foreseeability of said reports. (The manufacturer can assume, for instance, that reports are unlikely to come from vehicles driving in mid-ocean, but in general which new routes users will report driving on, etc. cannot be anticipated.) Likewise, the manufacturer has no control over the content of information received (later, at the information analysis stage, it may decide to disregard certain types of information that it suspects are unreliable. But the content of information sent and acquired is within the control of the system's users, rather than its manufacturer's).

In fact, not only does the manufacturer have limited foreseeability as to the content of information the system relies on, and not much control over it, but the communal nature of the information acquisition stage makes the system

157. The manufacturer, for example, cannot foresee that a certain patient will be O+ blood type rather than A+ (particularly if said blood type is not much more common than the others). It can anticipate, however, that the patient will be one of the A+/A-/B+/B-/AB+/AB-/O+/O- blood types, and not, say, a whole new blood type M+. In addition to limited foreseeability, control over said content is also problematic, in the sense that the manufacturer does not dictate the specific values of the patient but also in the sense that the system might be fed erroneous information beyond the manufacturer's control (potentially by other systems connected to it via the internet of things).

vulnerable to intentional false reports. (For instance, users wish to game the system to divert other drivers from their own neighborhoods and reduce traffic;¹⁵⁸ or even, hypothetically, psychopaths may attempt to swamp the system with false reports of a new road built over a lake, only to have the system plunge innocent drivers into it.) The manufacturer can take precautions to try to discard false information (accepting information only from users who previously have provided information that proved true, accepting information only when a set minimal number of users support it, accepting certain types of potentially dangerous pieces of information only after a Waze representative has physically visited the location to verify it, etc.). Precautions such as these are relevant in the context of the next OODA Loop stage of information analysis, rather than information gathering where the system depends on input from users.

On the other hand, compared with traditional GPS, Waze minimizes the lack of foreseeability aspect associated with recent changes that occurred after the system was released, or at least shortens its duration. This is because the system learns of such changes in real time and adds the missing information as soon as a sufficient number of reports is received.

So in general, and as may be intuitively expected, a system based on information received from its users is characterized by less foreseeability and control than in the first OODA Loop stage. On the other hand, a system based on information received in real time (rather than on information collected once, or with long spells between collections) is likely to suffer less from lack of foreseeability with respect to updates and new information.

c. “Orient” (“Information Analysis”)

As with the stage of information acquisition, Therac-25 was not involved in information analysis: the analysis was done externally, by humans. The new generation radiation system, on the other hand, has a significant role in the analysis stage, and as a result lowers the manufacturer’s level of foreseeability and control over said stage. In more detail, the system’s mere computational abilities, which enable it to weigh up the various complex parameters collected in the previous stage, do not impair the manufacturer’s foreseeability or control. Even if the manufacturer cannot perform the computation tasks itself, the information analysis stage is perfectly foreseeable and controllable; the manufacturer is the factor that decides how much weight the system should give each parameter, and which collected pieces of data to disregard altogether. The manufacturer can decide, for instance, that a patient’s age should only be

158. Alissa Walker, *Is It Really Possible To Trick Waze To Keep Traffic Off Your Street?*, GIZMODO (Nov. 18, 2014, 4:42 PM), <https://gizmodo.com/is-it-really-possible-to-trick-waze-to-keep-traffic-off-1660273215>. See also Nicholas Tufell, *Students Hack Waze, Send in Army of Traffic Bots*, WIRED (Mar. 25, 2014), <https://www.wired.co.uk/article/waze-hacked-fake-traffic-jam>.

considered when the patient is very young or old, or decide that much weight should be given to the fact that a patient has a history of prior tumors and the system will conduct its computational analysis accordingly. New generation machines, however, are likely to be more than sophisticated calculators. Rather, the whole uniqueness of advanced systems is that they can learn for themselves, at much better success rates than humans, how much weight to attach to each parameter based on prior experience. In such cases it is the machine that will decide how far to consider a patient's medical history, her susceptibility to allergies, or a new controversial study published in a medical journal. The manufacturer's ability to increase control and foreseeability by setting boundaries is limited in the information analysis stage because it can do so mainly sporadically, and with respect to the weight to be given to certain parameters that the manufacturer knows in advance will be part of the system's analysis. Here too, the more dynamic the database the system draws information from, the less able a manufacturer is to pre-instruct in a broad manner that certain weights be given to certain parameters.

On the other hand, the navigation sets of examples seem less revealing in the context of the second OODA Loop stage, as the differences between the traditional GPS and Waze are less evident.

In general, the manufacturer of a traditional GPS has ample foreseeability and control over the analysis stage of the system. The manufacturer dictates in advance how much weight be given to the different parameters (for instance, if a map shows that a road is closed, the manufacturer will likely instruct the system to attach 100% weight to said piece of information and disregard the existence of this road when calculating a route). Users themselves may also influence the analysis stage of the GPS (for instance, by instructing it to avoid toll roads, or to present the shortest route as opposed to the fastest route, etc.). But it is the manufacturer that designed these choices that users may make, hence these choices are foreseeable and controllable. As mentioned, traditional GPSs' lack the information that has changed since the last update of the system, but precisely because such information is not part of the data that the system analyzes there is no sense in foreseeing or controlling *how* the system would analyze it because it simply will not.

With Waze, the real-time character of the updates might cause the system to attach different weights to similar reports, based on the volume of the reports (for example, the system is likely to give much weight to 300 reports of a traffic jam on a country road usually not taken by many cars, and only little weight to two reports of a traffic jam at a very busy intersection). The decision on how the weight ought to change based on the number of reports and other parameters, however, is predetermined by the manufacturer and in that sense is foreseeable and controllable. In addition, the manufacturer is free to decide to attach zero weight to (in other words fully disregard) reports it deems "suspicious."

Moreover, real-time updates that will add new unknown information and will force the system to decide how much weight to attach to such pieces of

data can be predetermined. For instance, if the new information indicates any sort of danger, the system will attach maximal weight to it and verify that no driver is directed to the danger zone until it is proven safe, if the new information indicates an unfamiliar new road the system will not direct users there until a certain volume of reports has been received, etc.

The crowd-sourced and the real-time nature of Waze, therefore, do not seem highly significant in regards to the foreseeability or controllability of the information analysis stage.

d. “Decide” (“Decision Selection”)

Having attached different weights to the myriad of parameters collected through the information acquisition process, and having analyzed it, a system also responsible for the third OODA Loop stage now has to make a decision based on said analysis. This stage involves more than might meet the eye, at least for thinking algorithms. Unlike a coffee machine, for instance, whose decision-making process is generally based on two deterministic options, pour coffee or do not pour coffee, a thinking algorithm might face numerous alternatives, each based on probabilities and each accompanied by a certain level of confidence that indeed said alternative entails said probabilities. After analyzing the information it has acquired, a new generation radiation machine, for example, might come up with dozens of potential treatment dosages, each entailing different probabilities of success and expected damage. The algorithm’s ranking may, for instance, include an option with success rates of 90%, entailing damages of 3,000 units¹⁵⁹ for the 10% of failure; an option successful 80% of the times, entailing damages of 1,000 units in the 20% cases of failure; and a long list of alternative dosages of different success rates and expected damage.

Moreover, like a human physician, the algorithm cannot be 100% certain that said alternatives indeed reflect the probabilities and expected damages the algorithm assumes they do. The algorithm can, for example, determine that it is 95% confident that the first option indeed has a 90% probability of success and a potential of causing damages of 3,000 units, and is only 70% confident that the second option indeed reflects the rates indicated. Our decision selection stage, therefore, involves some tough questions for an algorithm. First, does it give more weight to potential success rates or to potential damages? (Under the first alternative success rates are 90% and expected damages are 10% X 3,000=300 units. Under the second alternative success rates are only 80%, but on the other hand expected damages are lower: 20% X 1,000=200 units. Which is preferable? When the units do not reflect dollar amounts but, for instance,

159. While in many cases “units” would be translated into dollar amounts (for instance, in litigation fees), in other cases the units of damages themselves would be less concrete, for instance, numerical estimations of the physical damage or suffering caused to a patient whose treatment failed.

pain and suffering to a patient, the decision is no longer a matter of mere economic computation) Secondly, how does confidence level affect the choice among the various options? Should the algorithm defer to a human whenever the rate of confidence of its preferred alternative is lower than a certain threshold? (Will the algorithm itself decide that it has to consult with a human, or will the threshold be predetermined by the manufacturer?)

A manufacturer's level of foreseeability and control naturally depends on how these questions are answered, or, more precisely, who, gets to decide these questions. With respect to the radiation machine the manufacturer may decide, for example, to adopt a more careful approach, where a machine is not free to decide on an alternative whose expected damages are more than a negligible percentage, and that whenever its level of confidence is below a very high threshold it must step back and let a human decide. Such an approach, however, will naturally be at the expense of efficiency (as in many cases the machine will not be able to automatically complete the process but will have to wait for a human to arrive and make her decision). Naturally, if the situation calls for instant treatment (for instance in a system used in a trauma unit) having to wait for a person might cost lives. Also, if from the outset the system's success rates are higher than the human counterpart's, from a utilitarian point of view we would prefer the machine to make such decisions, not the humans involved¹⁶⁰ (leading, again, to greatly reduced extents of foreseeability and control by the manufacturer).

The more the system's response time is critical, and the wider the gap between a human's and a machine's success rates (in favor of the machine), the more likely will the manufacturer be forced to forego foreseeability and control at the decision-making stage, and free the system to make its own choices.

Unlike a traditional GPS, which decides based on deterministic parameters (such as distance and existence of available roads), Waze makes its decisions based on probabilistic parameters that change constantly.¹⁶¹ Being a "real-time service," Waze cannot wait for a human in the loop to assist it whenever its level of confidence is not sufficiently high, which again makes the system less foreseeable and controllable. But unlike a radiation machine, Waze is not intended to save lives, and therefore enjoys the luxury of being able to hedge risks. If the system has indications that a certain road is not safe (is under construction, leads to a danger zone, etc.) its manufacturer may theoretically decide in advance to eliminate such an option entirely until it is proven safe (at the expense of directing users to less efficient routes). Life-saving or medical algorithms, however, do not generally have this option, otherwise any choice they made would entail certain risk. Foreseeability and control (hence compatibility with the rationale of promoting safety) are thus reduced in such systems.

160. Millar & Kerr, *supra* note 41, at 103; Lieblich & Benvenisti, *supra* note 66, at 246.

161. Such as the degree of traffic expected at a certain location at a certain point in time when the user is expected to cross said location.

e. “Act” (“Action Implementation”)

Neither of our radiation and navigation systems, whether the traditional or the “thinking” type, has ever replaced humans in the final OODA Loop stage of action implementation (GPS/Waze because they do not have an element of execution to begin with, Therac-25 and the new radiation system because the projection of the beams was never a process that a human performed). To touch on the stage of action implementation, let us therefore think of a robo-surgeon (for instance, the da Vinci system¹⁶²), and focus solely on its execution rather than the decision-making process of how to execute. With respect to said specific stage, there is no learning element. Just like a traditional product, therefore, the system is foreseeable and controllable as long as it does not encounter bugs.

f. *Interim Summary and Measurability of Success Rates*

An analysis of the four stages of the OODA Loop and the levels of foreseeability and control associated with each in respect of our two examples of learning algorithms, revealed that the following parameters tend to reduce the system’s compatibility with the rationale of encouraging manufacturers to promote safety: the greater number of the OODA Loop stages the system is responsible for, a system’s freedom to decide which sources to draw data from, a system’s freedom to decide which parameters to consider, the dynamic nature of the information sources relevant in the field of the system, use of crowd-sourcing to obtain information, the urgency of the process the system is involved in (or the requirement of real-time actions), whether the system is life-saving (such that reduction of efficiency to increase foreseeability is very problematic), whether the system’s success rates are already higher than those of a human equivalent (which again would render sacrificing efficiency to promote foreseeability more problematic), and whether the system considers real-time updates by its users (interestingly, this characteristic may work both ways: on the one hand its content may surprise the manufacturer, but on the other hand it may decrease unforeseeability associated with data that have changed without the manufacturer’s knowledge).

Naturally, this is only an initial list of parameters, and additional examples might yield additional parameters. But it does give us a notion of the types of systems that should be classified as “thinking algorithms,” in the sense that improvement of their safety by the manufacturer is more difficult to accomplish.

In addition to these parameters that impede improvement, another general parameter to take into account is how easily improvement is measured. The

162. The da Vinci is a minimally invasive robotic surgery system that translates a human surgeon’s hand movements into smaller more precise movements. DA VINCI SURGERY, <http://www.davincisurgery.com> (last visited Dec. 27, 2018).

success rates of certain systems are fully measurable (annual number of car accidents autonomous vehicles are involved in per miles traveled, number of fatalities per miles traveled, etc.), but other systems' success rates may well measure the rates of false negative failures but not the false positive ones, or vice versa (failure rates of a bail algorithm which approved bail for defendants who later broke the bail rules are easily measurable while the rate of defendants who were denied bail, even though they would have not broken its rules if given the opportunity, is not). It may be difficult or almost impossible to measure success rates of other systems (for example, success rates of Waze, of algorithmic company directors, or of algorithmic priests). Measurability of success or failure does not directly contribute to manufacturers' ability to improve their systems' safety, yet it does give them a clear picture of the system's performance, and may encourage them to try alternative designs until better results are measured (even if manufacturers cannot explain why such designs work better than the previous ones, in light of lack of foreseeability and control difficulties).¹⁶³ The extent to which the system's failure rates are measurable, then, may assist manufacturers improve their systems' safety level, therefore indicates compatibility with the first rationale of products liability law.¹⁶⁴

2. *Avoiding a Chilling Effect*

It has been argued that in certain industries, the application of products liability not only failed to achieve an increase in safety (or in victim's compensation, which will be discussed next), but led to higher production prices resulting in a suboptimal level of manufacturing or use of beneficial technologies.¹⁶⁵ Although the avoidance of a chilling effect and the promotion of an efficient level of use are not the main rationales behind products liability and are rather competing interests, the shaping of the products liability

163. When Facebook decided to change its news feed algorithm, for example, the team admittedly did not know how the change would affect the algorithm and its choices. Rather, to implement a change, many trial-and-error experiments are required. Will Oremus, *Who Controls Your Facebook Feed*, SLATE (Jan. 3, 2016), http://www.slate.com/articles/technology/cover_story/2016/01/how_facebook_s_news_feed_algorithm_works.html.

164. For some systems, exact measurement of failure or success rates might be possible, but only after the passage of time. Mortality and morbidity rates associated with a new generation radiation machine, for instance, will only be revealed years after its release. During the time when statistics are unavailable, the "measurability" parameter will incline to the "non-compatible with products liability framework," but as soon as it becomes available the system can be treated as more inclined to products liability compatibility.

165. According to Polinsky and Shavell, for instance, products liability led to an astronomic increase in the price of the DPT vaccine, leading to significant under-vaccination. Similarly, litigation costs associated with products liability claims in aviation led to a dramatic increase in sales prices and to suspension of production by leading manufacturers. *Supra* note 105, at 1474.

framework was significantly affected by said interests.¹⁶⁶ Our analysis will therefore include these interests and examine how different parameters of sophisticated systems affect concern over a chilling effect.

In general, the literature addresses concerns that products liability will have a chilling effect on sophisticated systems.¹⁶⁷ According to said concerns, the volume of products liability claims associated with such systems is expected to be high, given the increasing number of cases where damages are not caused by a human tortfeasor but by a system that has replaced her, and since in general the introduction of new technologies has traditionally resulted in an increase of tort claims associated with it.¹⁶⁸ In addition, as a result of globalization and social media, manufacturers will probably receive more post-sale reports regarding the actual use of their products than in the past, advising them of additional potential risks not anticipated in advance. Such additional reports will impose warning requirements on the manufacturers that, when not followed, will be used against them in failure to warn products liability claims.¹⁶⁹

Naturally the less foreseeable and controllable a system is, the greater the fear of a detrimental effect on technology. First, and as discussed above, lack of foreseeability and controllability (divided into its sub-characteristics) render it more difficult for a manufacturer to improve safety, at least for a subset of the innumerable scenarios possible.¹⁷⁰ In such cases, products liability will not necessarily contribute much to safety but will likely result in higher production costs that will translate into reduced manufacturing of certain systems due to less demand. Moreover, lack of foreseeability is likely to render liability costs less predictable, and in turn again delay development or result in high costs.¹⁷¹

166. Promoting an efficient level of usage, however, was discussed as a sub-rationale in the context of deciding between applying strict products liability versus a negligence-based standard. In general, concern over a chilling effect swung the pendulum of the products liability framework in the latter direction. *See supra* Part II.

167. *See, e.g.*, Colonna, *supra* note 8, at 84; *see also* Kevin Funkhouser, *Paving the Road Ahead: Autonomous Vehicles, Products Liability, and the Need for a New Approach*, 2013 UTAH L. REV. 437, 452-58.

168. Colonna, *supra* note 8, at 110 (focusing on the biotechnological industry as an example of a “severe barrier for innovation” caused by products liability).

169. *See* RESTATEMENT (THIRD) OF TORTS: PROD. LIAB. § 10(b) (AM. LAW INST. 1998) (imposing post-sale warning requirements when the seller knew or should have known of a substantial risk of harm). Alternatively, it could be argued that the volume of products liability claims will not increase with the introduction of sophisticated decision-making systems but in fact might decrease, given the improved safety rates expected of these products, which will reduce the general occurrences of damage. Villasenor, *supra* note 20, at 4-5 (reviewing how in-car automation devices of several sorts have led to a reduction in the frequencies of liability claims).

170. *See supra* Part IV, B.1.

171. Smith, *supra* note 20, at 6. Granted, for systems with a potential to cause damage that may be used by the user repeatedly, such as autonomous cars, we would want their prices to reflect a certain amount of risk for damage, in order for the user to internalize expected damages costs and use the system efficiently (for instance, not send its vehicle for rides around the neighborhood just as a joyride). Polinsky & Shavell, *supra* note 105, at

Secondly, also as discussed above, naturally the better result a system has compared with a human equivalent, and the speedier its response, the more decisions will we be likely to entrust it with, leaving humans outside the loop.¹⁷² So for systems with results that show superiority over humans, fear of a chilling effect (in the form of decreased development, decreased demand due to high price, or reduced efficiency resulting from clumsy safety measures that render the system too slow or non-user-friendly) is of greater concern. To give a concrete example: to decrease products liability litigation risks, the manufacturer of the new generation radiation system may design the system to be more foreseeable. For instance, it will feed it with the information it learns and not allow it to search independently for relevant sources of information, or it will heavily intervene and dictate the weight the system will attach to different parameters, not leaving the system discretion to decide based on its past experience. By doing so, manufacturers might increase foreseeability and control, but will not necessarily improve the system's outcomes compared to both older versions of the system or humans. Assuming society wishes to promote the development of such a better technology (which is perhaps conceptually easier once the system already outperforms humans and is thus expected to save lives or reduce injuries), products liability might be overburdensome on certain technologies.

3. *Victims' Compensation*

Generally, products liability has been accused of not being an optimal regime with respect to compensating victims for their damages, given the high costs associated with products liability litigation that render many damage cases not actionable,¹⁷³ and lead to significantly reduced compensation for the victim in cases that are filed due to legal fees.¹⁷⁴ How is victims' compensation affected by the characteristics of the damaging product? Unlike the case with the first products liability rationale of promoting safety, the particular abilities of each algorithm will likely not play a crucial part in affecting victims' ability for redress, although some more general characteristics will.

First, in many cases insurance of various types may render products liability redundant, as damages claims of insured victims might be covered in

1459-61. Similarly, and as discussed in Part II above, one of the goals of the products liability framework is to have the manufacturer internalize the costs associated with the damages it might cause. But the more unpredicted these costs are, the less likely internalization of them is to reflect the actual value of such costs.

172. Abbott, *supra* note 26, at 18; Millar & Kerr, *supra* note 41, at 103.

173. Given that the expected litigation costs exceed the expected reward for the damage suffered by the product, for instance, in car accidents where no significant physical injury was caused. Hubbard, *supra* note 94, at 1827-29; Gurney, *supra* note 97, at 265-66.

174. Polinsky & Shavell, *supra* note 105, at 1469-70 (reviewing empirical studies that found plaintiffs receive only 37–57% of the amount paid by defendants).

full by insurance, which eliminates the need for litigation.¹⁷⁵ Indeed, in the US certain types of insurance, including life insurance, health insurance, disability insurance, property insurance and car insurance, are prevalent.¹⁷⁶ A significant amount of damage cases caused by products are nevertheless uninsured, whether because not all victims have insurance or because the insurance does not cover the full amount of the damages suffered.¹⁷⁷ From the standpoint of assuring victims' compensation, the general conclusion is that products liability is needed more when the type of activity of the system, or the type of potential damage it may cause, are covered less by insurance than by other systems.¹⁷⁸ But this is true regardless of thinking algorithms or of the debate over which tort legal framework ought to apply to autonomous systems. What may be said more specifically about thinking algorithms? One distinction is that the more foreseeable the system (based on the various parameters discussed above), the more willing insurance companies are likely to be to offer insurance at reasonable premiums. This contrasts with algorithms that are less foreseeable, hence are associated with more uncertainty as to the expected amount of damage claims brought against the manufacturers and their expected outcomes, which might render insurance companies reluctant to insure the product at premiums that are not exceedingly high.¹⁷⁹ The division between more foreseeable and less foreseeable algorithms (again, based on the distinctions made above) in the context of availability of insurance is more relevant when the system is new to the market, and statistics of its performance are not yet available. Once experience is gained as to the system's rates of causing damage, the fact that such a system's choices are not foreseeable in the context of each specific user is insignificant. Rather, insurance companies will look at the "big picture" of damage claims brought and won over time. The effect of lack of foreseeability resulting in less available insurance, resulting, in turn, in more compatibility with products liability rationales, is therefore temporary after the product was launched. Despite this, for certain systems, especially medical-related ones, the duration might be very long considering long-run damages that might appear only years after the user's encounter with the system.

Second, in the absence of full coverage insurance, another major consideration affecting victims' ability to receive compensation is whether an attorney will agree to take their case. Under the contingency fee structure typical of products liability, that likelihood is affected not only by the

175. *Id.* at 1441.

176. *Id.* at 1462.

177. *Id.* at 1463.

178. Damages caused by car crashes, for example, are rarely brought to courts, because such damages are usually covered by automotive insurance. Smith, *supra* note 20, at 33.

179. Hubbard, *supra* note 94, at 1816 ("[T]here may well be no such data available to insurers where a seller seeks liability insurance for an innovative sophisticated robot. As a result, products liability insurers may be very concerned about the potential for high claims. Therefore, insurance may be hard to get, very expensive, or both.").

attorney's estimation of the probability and magnitude of success, but also of the expected costs.¹⁸⁰ The longer the case is expected to last, for instance, due to its technical complexity and the need to hear more expert witnesses, the less likely attorneys will be to take a risk and accept a products liability case.¹⁸¹

Under the prevalent risk utility test for design defects, a victim must show that efficient safer designs of the system were within reach.¹⁸² This might require the involvement of more than one expert witness, thus rendering the procedure cost-prohibitive.¹⁸³ The number of experts needed, hence the complexity, duration and cost of procedures, will probably depend on the system's field of operation, and whether it replaces human professional judgment or not. For example, in a products liability suit against a Roomba vacuum cleaner for injuring a user by catching and pulling her hair,¹⁸⁴ a coding expert would be required to show that the manufacturer could have efficiently programmed the system to prevent such accidents. In fields that involve complex professional discretion, such as medicine, engineering or law, it is likely that in addition to a code expert testifying on the availability of technical programming measures, a professional expert in the underlying field (physician, engineer, lawyer, etc.) would also be required to discuss whether such measures were available and accessible in the field of expertise (naturally, the defendant too will likely arm itself with all types of relevant experts, in trying to show that no such measures existed). It could be argued, therefore, that systems that replace human professional judgment are more likely to be associated with longer and more expensive procedures (due to the additional experts required but, quite intuitively, also to the complexity of the matter in general) and thus less likely to be litigated.

A characteristic of sophisticated systems that might affect their classification as thinking algorithms for the purpose of applying products liability rules is the unforeseeable nature of the system (in each of the OODA Loop stages, as discussed above) while statistics of the damages claims associated with said system are not yet available: the more unforeseeability exists, the more relevant the consideration of compensating the victim becomes because insurance will be less available. A second relevant characteristic is whether the system replaces professional discretion. If it does, the rationale of compensation is less applicable, given that more of these cases will not be picked up by attorneys to begin with. Naturally, both these characteristics may exist at the same time (for instance, systems replacing professional discretion may very likely rely on dynamic sources of information, therefore, as discussed above, be less foreseeable).¹⁸⁵ In such cases they will simply pull in opposite

180. Smith, *supra* note 20, at 38.

181. *Id.* at 38-40; Gurney, *supra* note 97, at 265-66; Hubbard, *supra* note 94, at 1826-28.

182. *See supra* Part II.

183. Gurney, *supra* note 97, at 265-66.

184. McCurry, *supra* note 22.

185. *See supra* Part IV, B.1.

directions with respect to the system's compatibility with the rationale of compensating the victim.

C. Keep It Simple, System

Having analyzed different features affecting systems' compatibility with the rationales behind products liability laws, this Article discovered numerous such features that ought to be considered when deciding whether the system is a traditional product or whether a different legal framework should apply to it. Not only is the list of said features relatively long (and open to additional input), its application in practice is not entirely simple, given for instance that certain features might entail more than one effect,¹⁸⁶ or that the effect depends on additional parameters.¹⁸⁷ As can be seen from the two figures below, however, using said method as a differentiator between traditional products and thinking algorithms is nevertheless much easier and simpler than using a system's autonomy level as a distinguishing method between the two.

First, while using autonomy as a differentiator requires the user to answer different types of questions, many of which are open-ended, and independently decide how to factor in all of her general estimations for these questions, a purposive analysis differentiator simply and clearly requires adding a "+" sign whenever a pre-defined feature exists. Second, while the outcome received as a result of the former classification process is a general impression of the system's level of autonomy—without practical guidelines as to how to apply it (especially in cases where the result is not definite)—the latter classification process yields a clear outcome of the number of "+" signs the system has accumulated. Lastly, said outcome has a practical implication, as it doesn't merely reflect the system's theoretical characteristics, but rather indicates how well the system reconciles with the purposes of the products liability framework.

186. For instance, when looking at the feature of "lack of foreseeability and control" as a whole, said feature is not compatible with the rationale of promoting safety (as it might render improvement of safety excessively difficult, costly or inefficient) but might be compatible with the rationale of compensating victims (given that insurance companies would be less willing to insure, thus rendering products liability more needed as an alternative).

187. Looking at the example above where lack of foreseeability and control lead to greater need for products liability as an alternative to insurance, said effect is mostly relevant during an initial period of time after the system has been introduced to the market, where statistics of its performance and potential damages are not yet available. Assuming sufficient time has elapsed, the lack of foreseeability and control might be less dominant in the context of insurers' willingness to insure. Given that lack of foreseeability and control are likely to lead to longer and more expensive legal proceedings, at a certain point in time the dominant effect of this feature on victims' compensation would be limiting accessibility to litigation and thus rendering it again less compatible with products liability.

FIGURE 1: Autonomy Level as a Classifier for Thinking Algorithms

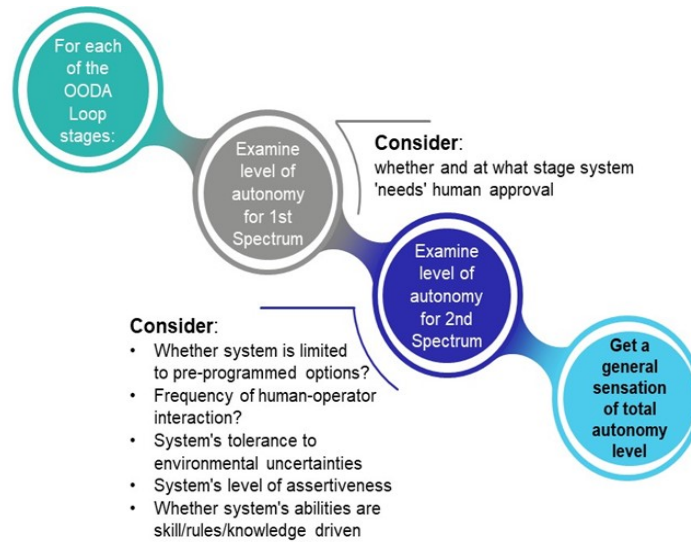
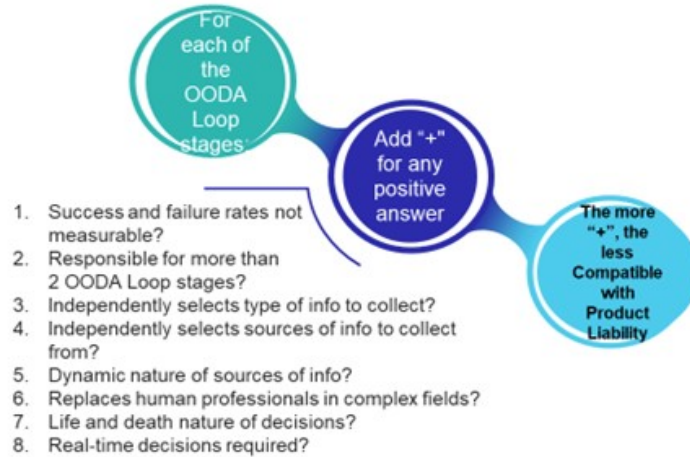


FIGURE 2: Purposive Interpretation as a Classifier for Thinking Algorithms



To demonstrate the relative simplicity of the proposed differentiator, as well as its de facto value, let us look at the table below comparing different types of systems based on this Article’s purposive analysis: Roomba vacuum robots, autopilots, autonomous vehicles, and a futuristic “robo-doctor” which replaces all functions of a human physician. The analysis used to decide which systems deserve which sign in each category is certainly not exhaustive and not conclusive—for instance, different types of autopilots may possess very different sets of features and thus render the analysis different for each type. It

is also based on many assumptions regarding the capabilities and modes of operations of the underlying systems. It does, however, demonstrate how the features listed in the first column could indeed be used to differentiate different types of systems in the context of whether they are compatible with achieving products liability rationales.

TABLE 1: Comparing Systems' Purposive Analyses

	Roomba Robot	Autopilot	Autonomous Vehicle	Robo- doctor
Success rates not measurable?				
Responsible for more than 2 OODA Loop stages?			+	+
Independently selects type of info to collect?			?	+
Independently selects sources of info to collect from?				+
Dynamic nature of sources of info?				+
Replaces professionals in complex fields?		?	?	+
Life and death nature of decisions?		+	+	+
Real time decisions required?		+	+	?

Success rates not measurable?

Generally speaking, the success rates (or, alternatively, damage-causing rates) of all the four systems are measurable. The measurement of Roombas, autopilots, and autonomous vehicles may be conducted based on accident rates associated with the use of the system.¹⁸⁸ The success rates of a robo-doctor may be measured based on mortality and morbidity statistics. None of these systems, therefore, render it especially difficult for a manufacturer to measure improvement, and in that sense to not impede the safety promotion rationale.

188. Naturally, in the latter two cases the statistics would be less unequivocal, given that certain accidents would be inevitable regardless of how the autonomous system functioned. Nevertheless, clear improvement (or dis-improvement) trends may be learned and taken into account by looking into the total number of accidents caused while using the system (an alternative or additional measurement may naturally be the number of fatalities, or the gravity of injuries suffered as a result of using the systems).

Responsible for more than two OODA Loop stages?

As explained above, when analyzing which OODA Loop stages a system is “responsible” for, the focus is on stages in which the system’s choices are not dictated in an injective manner.¹⁸⁹ A Roomba cleaner’s sensors collect predetermined information on its environment (the existence of dirt on the floor, the existence of obstacles in its path, etc.) and analyze the information collected in a predetermined manner (if the robot touches an object, this means an obstacle was encountered, etc.). While there is a feature of randomness to the Roomba cleaner’s decisions, which leads to execution that is not entirely predetermined (the system is designed to randomly try new angles and vectors of movement, to increase the chances of overcoming obstacles and to achieve even coverage),¹⁹⁰ the system is not “responsible” for more than two OODA Loop stages. A similar analysis applies to autopilots as well, which collect predetermined types of information (altitude, wind velocity, etc.) and conduct an analysis of the information gathered in a pre-determined manner (calculation of the optimal parameters such as angles and speed needed for the air vessel to stay on course or land). Unlike airplanes travelling through almost empty skies in terms of potential obstacles, autonomous vehicles must respond to numerous (and often) unexpected obstacles as part of their routine tasks. The information that the system gathers is therefore not entirely predetermined (as the cameras and sensors the vehicle is equipped with might encounter new types of terrain, of road conditions or of obstacles that the manufacturers did not foresee). The analysis of the information gathered is also not predetermined such that it is based on mere calculations. Rather, it is conducted through more human-like trial and error self-learning processes that allow the system itself to decipher the optimal way to interpret the data received.¹⁹¹ When deciding how to drive next, an autonomous vehicle does not have a bottom line calculation indicating precisely the right angle for landing or the right speed for preserving altitude. Rather, faced with scenarios of potential obstacles interfering with its path, the system must decide the *least* dangerous next move. An autonomous vehicle, therefore, seems to be responsible for more than two OODA Loop stages and thus gains a “+” sign under that category. A robo-doctor system would also likely receive a “+” sign, based on its likely ability to collect information from various sources based on its own consideration, to analyze the information gathered based on its self-learning (and not in an injective manner based on predetermined abilities), and its decision about how to act upon the analyses which will reflect a choice between stochastic alternatives based on its level of confidence in the results.

189. See *supra* Part IV, B.1.i.

190. Joshua A. Kroll et al., *Accountable Algorithms*, 165 U. PA. L. REV. 633, 655 (2017).

191. Loz Blain, *AI Algorithm Teaches a Car to Drive from Scratch in 20 Minutes*, NEW ATLAS (July 5, 2018), <https://newatlas.com/wayve-autonomous-car-machine-learning-learn-drive/55340>.

Independently selects type of information to collect?

As discussed above, the type of information gathered by Roomba cleaners and autopilots is limited to the type of information they were programmed to collect: pre-determined parameters concerning their environment. Autonomous vehicles, however, respond to scenarios where they encounter new types of information it was not instructed to collect (new types of terrain that were not foreseeable, extreme weather conditions not foreseeable, a new type of obstacle on the road, etc.). While the sources of information the system collects data from are indeed predetermined (the vehicle's environment), the type of information is not necessarily so. With respect to a robo-doctor, such a system will likely independently choose the type of information it collects (such as a patient's medical history, her current complaints, environmental parameters that might affect the diagnosis or prognosis, etc.) as explained above.¹⁹²

Independently selects sources of information from which to collect?

As discussed in more detail above, a robo-doctor might indeed be very independent in choosing which sources to collect information from, while a Roomba cleaner as well as an autopilot would only collect environmental information on its surroundings, which their manufacturers programmed them to collect. Unless given access to journals and articles focusing on futuristic features of automotive or physics that might somehow alter the decision-making of autonomous vehicles, they too seem to be limited to sources of information dictated to them by their programmer (unlike the *type* of information it may encounter).

Dynamic nature of sources of information?

Roomba cleaners, autopilots, and autonomous vehicles all acquire information from their environment, which is characterized by a certain degree of dynamism, especially with respect to the latter, where road-related parameters change constantly. The need for an ever-updating system, however, which adversely affects manufacturers ability to minimize risk as discussed above, is more prevalent in the robo-doctor system, as a result of the need to take into account new studies and findings that might affect the entire conclusion-reaching process of the system.

Replaces professionals in complex fields?

As discussed above, litigation over damage caused by systems that replace professional discretion is expected to be longer and more expensive, because of the likely need for additional experts in the underlying professional field. This, in turn, would decrease the chances of such cases being brought to court, and obstruct the victim compensation rationale. Though it cannot be anticipated in advance which experts attorneys would think are necessary for the case, it is

192. See *supra* Part IV, B.1.

safe to assume that the need for a professional expert (in addition to a programming expert) would be more likely in the case of a robo-doctor than in the case of a robo-cleaner. As to autopilots and driverless vehicles, while the functions fulfilled by the latter are familiar to all drivers, the functions fulfilled by autopilots require technical knowledge.¹⁹³ As a result, it is plausible to imagine that accidents involving autopilot would require a longer, more costly, and technical presentation to the court.

Life and death nature of decisions?

Other than the Roomba cleaner, the other three systems entail a significant risk to human lives (as a result of an accident or as a result of administering non-effective or dangerous treatment). The desire to produce the maximal benefit from these three types of systems, therefore, might come at the expense of allowing manufacturers to better foresee and control their results, as explained above.¹⁹⁴ The rationale of safety promotion, would therefore be more difficult to achieve for the three latter types of systems.

Real time decisions required?

As discussed, when real-time decisions are required, manufacturers' flexibility to dictate human intervention (above a certain threshold of risk, below a certain threshold of certainty, etc.) is undermined, and, as a result, so is the ability to promote safety. To operate smoothly, a Roomba cleaner does need to decide and act continuously (for instance, on which direction to turn to next), but given the nature of its decisions and these decisions' lack of urgency, a manufacturer is fully free to require human intervention in any scenario that has the potential to cause damage. Autopilots and autonomous vehicles, however, do not enjoy such flexibility. With urgent decisions that must be made within split seconds, both these systems receive a "+" sign. As to robo-doctors, the need for urgent real-time decisions is case sensitive, as an ER treatment process is very different than an annual physical (the latter allowing the system to call for a human decision-maker whenever a damaging scenario is possible). Robo-doctors therefore receive a question mark under that category.

Having demonstrated how the different features of Roomba cleaners, autopilots, autonomous vehicles, and robo-doctors render them more or less compatible with the rationales of products liability, we see how such a differentiator is relatively simple and quick to use. Granted, the details of the above analyses may vary, thus leading to somewhat different results. In addition, the analysis above does not conclusively tell us which products ought to be subject to products liability and which ought not to be. It does, however,

193. For the number of licensed drivers in America, see U.S. DEP'T OF TRANSP. & FED. HIGHWAY ADMIN., *Total Number of Licensed Drivers in the U.S. in 2016, by State*, STATISTA, <https://www.statista.com/statistics/198029/total-number-of-us-licensed-drivers-by-state/> (last visited Nov. 27, 2018).

194. See *supra* Part IV, B.1.

give us a simple and clear indicator: the fewer “+” symbols there are, the more the system can continue to be subject to the traditional products liability framework.

CONCLUSIONS

Focusing on the legal framework of products liability, this Article offers a novel method for determining when damages caused by sophisticated algorithms ought to be subject to traditional products liability laws, and when an alternative treatment is warranted.

Naturally, preferring a certain alternative legal framework over products liability depends on the characteristics, advantages, and disadvantages of such a framework (including, of course, the question of whether manufacturers would still be liable for the damage caused or not). Proposals for such substitute options have been discussed for decades,¹⁹⁵ focusing on aspects of determining liability as well as how and by whom damages will be paid. For instance, several analogies have been drawn between sophisticated systems and their creators and the legal relationship between other entities characterized by diminished foreseeability and control over the actions of tortfeasors that are subject to them. Examples were the relationship between parents and minor children who caused damage,¹⁹⁶ between principals and agents¹⁹⁷ (specifically employers and employees),¹⁹⁸ between owners of dangerous animals and their pets,¹⁹⁹ and even between masters and their slaves.²⁰⁰ Other directions were subjecting the systems themselves to a reasonable analysis in order to determine whether liability by the manufacturer exists,²⁰¹ while a different approach focused on no fault insurance schemes to cover the damages caused by such systems.²⁰²

A comparison of these different approaches, as well as an assessment comparing them to products liability, is beyond the scope of this Article. Given the increasing calls to stop treating sophisticated or autonomous systems as mere products and subjecting them instead to a legal framework other than products liability, the focus of this Article was on the more preliminary stage of examining which systems warrant the development of such legal frameworks

195. *See, e.g.*, Lehman-Wilzig, *supra* note 32.

196. *Id.* at 450-51; CHOPRA & WHITE, *supra* note 32, at 180.

197. Lehman-Wilzig, *supra* note 32, at 451-52; Asaro, *supra* note 32, at 178-79; *see, e.g.*, CHOPRA & WHITE, *supra* note 32, at 5.

198. Lehman-Wilzig, *supra* note 32, at 451-52.

199. *Id.* at 448-49; Asaro, *supra* note 32, at 176-78; Sophia H. Duffy & Jamie Patrick Hopkins, *Sit, Stay, Drive: The Future of Autonomous Car Liability*, 16 SMU SCI. & TECH. L. REV. 453, 467-71 (2013); *see generally*, Richard Kelley et al., *Liability in Robotics: An International Perspective on Robots as Animals*, 24 ADVANCED ROBOTICS 1861 (2010).

200. Lehman-Wilzig, *supra* note 32, at 449-50.

201. *See, e.g.*, Chagal-Feferkorn, *supra* note 21; *see also* Abbott, *supra* note 26.

202. EUR. PARL. DOC., *supra* note 28, at 5, 11.

and which may continue being subject to existing ones. In doing so, this Article found that the current differentiator between traditional products and systems warranting a new legal framework—the system’s level of autonomy—is insufficient. It found that said differentiator is unduly complex, provides an imprecise and impractical tool for differentiation, and might yield inconsistent results. This Article then proposed a different approach for distinguishing products from non-products for the purpose of applying products liability, focusing on the rationales behind the products liability framework and whether different features of sophisticated systems are compatible with them or hinder them.

Certainly, the parameters discussed in the Article warrant additional individual analysis, as they represent general directions of compatibility with products liability rationales, and in specific scenarios might yield different results. The examination of additional types of systems from various sectors employing various technology features will also likely add to the list of relevant parameters. This Article, however, assembled a list of central parameters that are relatively easy to identify and classify as existent or not for each system, and analyzed whether or not they point to compatibility with the rationales of products liability. Decision-makers and scholars encountering a system whose parameters clearly indicate one direction or another can therefore use the proposed analysis to determine whether to apply products liability or to adopt or develop alternative frameworks.