Fordham Law Review

Volume 88 | Issue 2 Article 8

2019

Robot, Inc.: Personhood for Autonomous Systems?

Gerhard Wagner
Humboldt University in Berlin, Germany

ROBOT, INC.: PERSONHOOD FOR AUTONOMOUS SYSTEMS?

*Gerhard Wagner**

INTRODUCTION

[R]obots cannot be sued, but they can cause devastating damage.

—Judge A. Leon Higginbotham, Jr.[1]

Since the invention of the steam engine, technological progress has served as a driver of innovation for liability systems.[2] Pertinent examples include the arrival of the railway and the introduction of motor-powered vehicles. Today, the digital revolution challenges established legal axioms more fundamentally than technological innovations from earlier times. The development of robots and other digital agents operating with the help of artificial intelligence will transform many, if not all, product markets. It will also blur the distinction between goods and services and call into question the existing allocation of responsibility between manufacturers and suppliers on one side and owners, operators, and users of such devices on the other.[3]

This Article uses the concepts of "robots" and "autonomous systems" interchangeably. Both concepts share the characteristic of "behavior" determined by a computer code that allows at least some room for "decision-making" by the machine itself. In other words, a robot is not entirely under the control of human actors—particularly in accident situations. This feature distinguishes autonomous systems from more traditional devices that run on software—deterministic computer code that does not allow for machine learning or autonomous decisions by the machine.

---

* Chair for Private Law, Business Law, and Law and Economics at the Humboldt University in Berlin, Germany. This Article was prepared for the Symposium entitled *Rise of the Machines: Artificial Intelligence, Robotics, and the Reprogramming of Law*, hosted by the *Fordham Law Review* and the Neuroscience and Law Center on February 15, 2019, at Fordham University School of Law. For an overview of the Symposium, see Deborah W. Denno & Ryan Surujnath, *Foreword: Rise of the Machines: Artificial Intelligence, Robotics, and the Reprogramming of Law*, 88 FORDHAM L. REV. 381 (2019).

1. United States v. Athlone Indus., Inc., 746 F.2d 977, 979 (3d Cir. 1984).

2. As to the United States, the locus classicus is MORTON J. HORWITZ, THE TRANSFORMATION OF AMERICAN LAW, 1780–1860, at 67–108 (1977). For a more nuanced view, see Gary T. Schwartz, *Tort Law and the Economy in Nineteenth-Century America: A Reinterpretation*, 90 YALE L.J. 1717, 1734–56 (1981).

3. *See* Jeffrey Jones et al., *Mitigating Product Liability for Artificial Intelligence*, JD SUPRA (Mar. 23, 2018), https://www.jdsupra.com/legalnews/mitigating-product-liability-for-77795 [perma.cc/SNP5-NLHR].

Robots and other autonomous systems are not presently recognized as persons or legal entities.[4] This is true for both U.S. and E.U. jurisdictions.[5] To date, no legal system classifies robots or other autonomous systems as legal subjects, instead leaving them to share the classification of other items made by humans: objects. Objects can be owned by legal subjects but cannot themselves be subjects or own objects. This means that the law attributes their actions—and omissions—not to the objects themselves but to the subjects responsible for them. The party responsible for the robot will usually be its owner, but perhaps other attributions are conceivable, particularly in cases where ownership and actual control diverge.

Theories that advocate the classification of robots as subjects rather than objects dispense with notions of ownership and attribution. As legal subjects, robots would not be capable of being owned by someone else. Rather, they themselves could own objects, particularly artifacts.[6] Furthermore, their acts and omissions would be attributed not to another subject but simply and automatically to the robot itself. Under such a theory, attribution of robot behavior to someone else would require reliance on vicarious liability doctrines such as respondeat superior, which would need to apply to humans and robots indiscriminately.[7] For instance, if a robot committed a wrongful act as an employee within the scope of employment, the employer would be liable in tort for the harm caused.[8]

In its resolution Civil Law Rules on Robotics dated February 16, 2017, the European Parliament identified civil liability for damages caused by robots as "a crucial issue" and invited the European Commission to submit proposals for new liability rules addressing harm caused by robot activity.[9] Following another Commission communication entitled "Building a European Data Economy,"[10] the European Parliament considers a choice between different approaches towards the liability of traditional actors who put a robot into the stream of commerce or who control the risks associated with its use.[11] The Parliament also envisioned, at least in the long term, the creation of a special legal status for robots: their recognition as electronic persons, or "ePersons."[12] Such an electronic person would be liable for any

4. *See generally* Shawn J. Bayern, *The Implications of Modern Business-Entity Law for the Regulation of Autonomous Systems*, 19 STAN. TECH. L. REV. 93 (2015).

5. Florian Möslein, *Law and Autonomous Systems Series: Regulating Robotic Conduct—On ESMA's New Guidelines and Beyond*, OXFORD BUS. L. BLOG (Apr. 10, 2018), https://www.law.ox.ac.uk/business-law-blog/blog/2018/04/law-and-autonomous-systems-series-regulating-robotic-conduct-esmas [https://perma.cc/DTJ9-DM29].

6. Matthew U. Scherer, *Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, and Strategies*, 29 HARV. J.L. & TECH. 353, 399 (2016).

7. DAN B. DOBBS ET AL., HORNBOOK ON TORTS §§ 31.1–31.8 (2d ed. 2016).

8. *See id.*

9. European Parliament Resolution of 16 February 2017 with Recommendations to the Commission on Civil Law Rules on Robotics, ¶ 49, EUR. PARL. DOC. P8_TA(2017)0051 (2017) [hereinafter Parliament Resolution].

10. *"Building a European Data Economy,"* COM (2017) 9 final (Jan. 10, 2017) [hereinafter *Commission Communication*].

11. *See* Parliament Resolution, *supra* note 9, ¶ 53.

12. *See id.* ¶ 59(f).

damage it caused.[13] The underlying policy rationale seems to provide legal certainty to manufacturers and users of robots and other autonomous systems.[14] This alludes to the classic idea, cropping up time and again, of using tort law as a subsidy for innovation. To the extent that ePersons could shield manufacturers against liability, they would facilitate the introduction of new digital products.

This Article explores an essential question: whether a reclassification of robots as legal subjects, rather than objects, makes sense. Part I offers a brief account of the general requirements for personhood as defined in philosophy and law. Part II focuses on the narrower issue of whether robots and other autonomous devices should be recognized as "liability subjects." Part III provides an outline of liability regimes for traditional liability subjects, namely manufacturers and users, in order to measure the gap in the liability system that robot liability could fill. On this basis, Part IV directly discusses the pros and cons of robot liability.

The conclusions weigh the costs and benefits of according autonomous systems the status of ePersons against each other. As this Article argues, it seems difficult to identify the real benefits of robot liability, as compared to strict product liability of manufacturers and fault-based liability of users. In essence, the only tangible benefit would be of evidentiary nature: if the robot itself were liable, victims would not bear the burden of figuring out whether the accident was caused by the malfunctioning or the defective design of the robot or through abuse, misuse, or other negligent behavior of the user. As long as autonomous systems are distributed as bundles of hard- and software, inaccessible and unchangeable to the user, these evidentiary problems do not seem very serious. The downside of robot liability is the full or partial externalization of accident costs to the detriment of victims. While the full externalization of accident risk may be avoided through minimum asset requirements or insurance mandates, there are clear limits to these safeguards. For this reason, partial externalization remains. It benefits the manufacturers and users of the robot, who assume the role of "quasi shareholders." Partial risk externalization would also subsidize the manufacture, distribution, and use of such robots. This Article argues that there is no good reason for granting such a subsidy.

I. EPERSONS AND APPROACHES TO PERSONHOOD

At the time this Article was written, no court had ever asked the question of whether robots should be classified as legal entities, and no statute, legal precedent, or doctrine personifying robots existed. Unsurprisingly, therefore, scholars who embrace the personhood of robots rely on philosophical and sociological theories for support.

13. *See id.*
14. *See Commission Communication*, *supra* note 10, at 14.

A. Laundry Lists

The debate over the limits of the concept of personhood, and possible expansions beyond its traditional scope towards inanimate objects, began with a classic article aiming at establishing "rights of the environment"—that is, natural objects.[15] In "Should Trees Have Standing?," published in 1972, Christopher D. Stone indicated that the same analysis that he applied to trees and other environmental entities could also be applied to "humanoids" and "computers."[16] Stone also admitted that the criteria for the legal system to recognize something as a rights holder are anything but clear.[17]

As early as 1992, Lawrence B. Solum examined the case for legal personhood for artificial intelligences.[18] Solum began with the premise that an artifact could qualify for legal personhood if, and only if, it could be shown to possess "intelligence."[19] Drawing on the work of John C. Gray,[20] Solum argued that an artifact could qualify as a legal person if it possessed intelligence and will.[21] He next examined whether a system with artificial intelligence was in command of the competences necessary to serve as trustee, in terms of the capability to understand and fulfill its responsibilities and duties, to make judgments, and to exercise discretion.[22] As a second step, Solum argued that an artificial intelligence system might qualify for personhood on constitutional grounds.[23] It is here that Solum discussed several additional properties necessary for an entity to qualify for legal personhood: namely, the possession of a soul, consciousness, and intentionality and the capacity to feel emotions, have interests, set one's own goals, and possess free will.[24] While all of these qualities admittedly play a role in the constitutional analysis, Solum does not argue that any one factor is either dispositive or indispensable in the sense that, without it, there can be no personhood.[25]

Taken together, the criteria for personhood—intelligence, free will, consciousness, intentionality, and emotions—amount to little more than a laundry list of features that distinguish human beings from other creatures, and from objects.[26] It remains unclear, however, which features are

15. *See generally* Christopher D. Stone, *Should Trees Have Standing?—Toward Legal Rights for Natural Objects*, 45 S. CAL. L. REV. 450 (1972).

16. *See id.* at 456.

17. *See id.* at 458–59.

18. *See generally* Lawrence B. Solum, Essay, *Legal Personhood for Artificial Intelligences*, 70 N.C. L. REV. 1231 (1992).

19. *See id.* at 1235–40.

20. JOHN C. GRAY, THE NATURE AND SOURCES OF THE LAW 27–64 (2d ed. 1921).

21. *See* Solum, *supra* note 18, at 1239–40. For Gray, the capacity to form a will seems to be the decisive element for recognizing an entity as a legal person. *See* GRAY, *supra* note 20, at 28 ("In the case of normal human beings we are not troubled with any question as to the actual presence of a will. The normal man or woman has a will.").

22. *See* Solum, *supra* note 18, at 1243–53.

23. *See id.* at 1255–80.

24. *See id.*

25. *See id.* at 1281–83.

26. *Cf.* SAMIR CHOPRA & LAURENCE F. WHITE, A LEGAL THEORY FOR AUTONOMOUS ARTIFICIAL AGENTS 162–81 (2011).

necessary conditions for the recognition of personhood and which are not. Furthermore, the laundry list approach obscures the fact that human beings' characteristic features are not of a binary nature but a matter of degree. The will of an individual is hardly ever free in the full sense of the term, and many people do not have much of a will to act upon in particular situations. Further, animal rights theorists have pointed out that many animal species possess the central features of humans, if only to a certain degree.[27]

The fact that the properties that define human beings are a matter of degree and do not allow for "either/or" determinations has led scholars Bert Jaap Koops, Mireille Hildebrandt, and David-Olivier Jaquet-Chiffelle to a gradual concept of legal personhood.[28] First they embrace a very broad definition of the criteria for the recognition of artificial intelligence as a person in the legal sense.[29] Personhood is "associated with the legal capacity to act," which in turn may involve the capacity to contract or the capacity to commit a wrong.[30] While autonomous systems may be subjected to civil sanctions, they are not enough of a "person" to also subject them to criminal sanctions.[31] Still, reference to the "legal capacity to act" remains obviously circular, as the capacity to act depends on the recognition as a legal subject.

B. Philosophical Theories

In ethics and the philosophy of morals and justice, approaches to personhood differ between different theoretical frameworks. Thus, one cannot answer the philosophical question of what constitutes a person without examining competing philosophical worldviews. In a deontological theory like Kant's, the crucial element of personhood is the ability to act as an autonomous subject:

> The fact that man can have the "I" in his representations raises him infinitely above all other living beings on earth. Because of this he is a person, and by virtue of the unity of consciousness through all the changes that happen to him, one and the same person—i.e., through rank and dignity an entirely different being from *things*, such as irrational animals, with which one can do as one likes.[32]

In Kant's philosophy, personhood is closely associated with moral agency:

---

27. *See* PETER SINGER, ANIMAL LIBERATION (updated ed. 2009) [hereinafter ANIMAL LIBERATION]; Richard Dawkins, *Gaps in the Mind*, *in* THE GREAT APE PROJECT: EQUALITY BEYOND HUMANITY 80, 81–87 (Paola Cavalieri & Peter Singer eds., 1993); Peter Singer, *Ethics Beyond Species and Beyond Instincts: A Response to Richard Posner*, *in* ANIMAL RIGHTS: CURRENT DEBATES AND NEW DIRECTIONS 78, 79 (Cass R. Sunstein & Martha C. Nussbaum eds., 2004).

28. *See generally* Bert-Jaap Koops et al., *Bridging the Accountability Gap: Rights for New Entities in the Information Society?*, 11 MINN. J.L. SCI. & TECH. 497 (2010).

29. *See id.* at 518–20.

30. *See id.* at 550.

31. *See id.*

32. IMMANUEL KANT, ANTHROPOLOGY FROM A PRAGMATIC POINT OF VIEW 15 (Robert B. Louden trans., Cambridge Univ. Press 2006) (1798). For a thorough exposition of Kant's theory, see Ludwig Siep, *Person and Law in Kant and Hegel*, *in* THE PUBLIC REALM: ESSAYS ON DISCURSIVE TYPES IN POLITICAL PHILOSOPHY 82, 83 (Reiner Schürmann ed., 1989).

> Person is that subject whose actions can be *imputed* to him. *Moral* personhood is then nothing other than the freedom of a rational being under moral laws . . . . The consequence is that a person is subject to no laws other than those which he gives himself (either alone or at least jointly with others).[33]

Thus, personhood requires moral agency in the sense of the ability of a subject to choose moral laws for him- or herself. This choice may be exercised either as a member of a group or society at large or as the subject itself.

Painting with an extremely broad brush, the alternatives to deontological theories of law and personhood are consequentialist approaches that focus not on the principle on which an action was based but on the outcomes that the action achieves in the real world. Utilitarianism is the most prominent consequentialist theory—it uses the balance of aggregate pleasure and pain as the measure of morals.[34] It follows easily that, for a utilitarian, an entity's ability to actually feel pleasure and pain must be the criterion for its recognition in the moral universe.[35] On this basis, it is not easy to explain why entity status should be reserved for human beings while excluding other sentient beings, namely animals.[36] Accordingly, the exclusive focus of classical utilitarianism on humans has come under attack from animal rights advocates who have accused the anthropocentric approach as "speciesism."[37]

From the point of view of utilitarianism, everything depends on whether autonomous systems qualify as sentient beings capable of feeling pleasure and pain.[38] This criterion dominates the utilitarian perspective on animal rights,[39] and nothing changes where the rights of digital autonomous agents are in question. As long as digital autonomous systems lack the capacity to experience pleasure and pain, they do not count in the overall utilitarian calculus. Once they develop this capacity, however, they would have to be taken into account, much like animals should.

In the alternative moral universe of Kant, animals have no rights because they are not human beings capable of rational reasoning or of self-governance (autonomy).[40] Kant developed the moral obligation to treat human beings as

---

33. *Id*. at 84 (quoting IMMANUEL KANT, *Vorbegriffe zur Metaphysik der Sitten*, *in* 6 KANT'S GESAMMELTE SCHRIFTEN: HERAUSGEGEBEN VON DER KÖNIGLICH PREUßISCHEN AKADEMIE DER WISSENSCHAFTEN 221, 223 (1902)).

34. For a classic exposition, see JEREMY BENTHAM, AN INTRODUCTION TO THE PRINCIPLES OF MORALS AND LEGISLATION 14 (Batoche Books 2000) (1789) ("Nature has placed mankind under the governance of two sovereign masters, *pain* and *pleasure*. It is for them alone to point out what we ought to do, as well as to determine what we shall do.").

35. Laura Palazzani, *Person and Human Being in Bioethics and Biolaw*, *in* LEGAL PERSONHOOD: ANIMALS, ARTIFICIAL INTELLIGENCE AND THE UNBORN 105, 106 (Visa A. J. Kurki & Tomasz Pietrzykowski eds., 2017).

36. *Cf*. ANIMAL LIBERATION, *supra* note 27, at 5–10.

37. PETER SINGER, PRACTICAL ETHICS 48–55 (2d ed. 1993).

38. BENTHAM, *supra* note 34, at 1–2.

39. ANIMAL LIBERATION, *supra* note 27, at 10–17.

40. *See* Alexander Broadie & Elizabeth M. Pybus, *Kant's Treatment of Animals*, 49 PHILOSOPHY 375, 379–83 (1974); Lara Denis, *Kant's Conception of Duties Regarding Animals: Reconstruction and Reconsideration*, 17 HIST. PHIL. Q. 405, 406 (2000). *See*

ends, not as means, and as subjects rather than objects. One could claim that autonomous systems should also be treated as ends, not as means, on the ground and under the condition that they acquire a state that could be regarded as autonomous, i.e., rational self-governance.[41] Once technology passes this threshold, robots will have to be recognized as subjects in the moral sense, and as persons in the legal sphere.

Needless to say, neither Bentham, Kant, nor the other founding fathers of utilitarianism or other deontologists ever discussed the classification of digital autonomous systems. Whether normative theories developed at a time when such technology was unimaginable can be applied to new phenomena remains questionable. All the classical writers took it for granted that only humans would qualify for personhood in the moral sense. This is hardly surprising given that intelligent artifacts did not exist at the time of their writings. Therefore, it seems problematic to assume that manufactured digital creatures may have a moral claim against humanity to be treated like their human creators, i.e., as subjects, not objects, and as ends rather than means.

C. Legal Concepts of Personhood: Pragmatism Rather Than Dogmatism

The brief survey of approaches to personhood sketched above has revealed rather a range of theories. What the competing theories have in common is the assumption that an entity applying for recognition as a person must fulfill certain criteria, and that the legal system, in turn, is bound to accord entity status once the pertinent criteria have been satisfied. The differences of opinion begin only when it comes to the definition of the relevant criteria. Utilitarians and Kantians, for instance, disagree about the relevant criteria, but they agree that one criterion or a limited set of criteria controls the world of ethics. To the extent that basic principles of ethics are included in modern legal systems in the form of constitutional guarantees, ethical principles are able to determine legal outcomes as well.

Contrary to the impression that the ethical discussion conveys, the legal concept of personhood is rather complex and nuanced. The legal system makes its own choices that hardly follow any one of the theoretical accounts sketched above. First, all human beings qualify as subjects, or persons, regardless of whether they satisfy any of the criteria offered by the competing philosophical accounts of personhood. A human being is regarded as a person and must be treated as such, even if, for some medical reason, he or she is unable to experience pleasure or pain. Thus, the concept of legal personhood does not subscribe to utilitarianism. On the other hand, it is also immaterial whether a particular human being is capable of rational self-governance in the sense of fully developed autonomy—whether a person is able to define his or her interests, to set his or her own goals, to apply the appropriate means to reach those goals, and to obey moral commands.

generally Christine M. Korsgaard, *A Kantian Case for Animal Rights*, *in* THE ETHICS OF KILLING ANIMALS 154 (Tatjana Višak et al. eds., 2016).

41. *See* Solum, *supra* note 18, at 1270.

Human beings who are incapable of rational self-determination and also of moral self-governance in the Kantian sense nonetheless qualify as persons. This is as true for infants as it is for adults suffering from serious brain injuries, dementia, or other illnesses affecting the central nervous system. In conclusion, the criteria used by the legal system to define personhood are primarily biological. Everything turns on whether the life form under examination belongs to the human species or not.

The decision to accord entity status to all human beings is not without alternative. Throughout human history, legal systems distinguished between groups of humans on the basis of race, sex, heritage, or political attitude. It took more than 2000 years to overcome these discriminations. History provides enough evidence that moral theories of personhood and the legal distinction between subjects and objects must not be short-circuited. The legal system makes its own choices.

The waywardness of the legal system also accounts for the institution of the "legal person." Modern legal systems do not confine entity status exclusively to human beings. Rather, they distinguish between natural persons and corporations—organizations that qualify as so-called legal persons. A corporation is a collective body that is separate from the natural persons associated with it as owners, agents, and employees. Corporate law, as embodied in codes such as the Model Business Corporation Act, endows a corporation with "the same powers as an individual to do all things necessary or convenient to carry out its business and affairs."[42] The forces that led to the recognition of corporations as distinct legal persons were not primarily philosophical in nature.[43] Rather, important economic interests led to the development of institutions designed to allow individuals to make small-scale contributions to large-scale investment projects without entailing the risk of personal bankruptcy.[44] The legal tools to achieve this objective are the separation of ownership and control and the separation of the assets of the corporation from the assets of the individual shareholders. Even though the corporate form is a standard element of modern legal systems, its pragmatic and functionalistic basis remains visible. Up to the present day, the U.S. Supreme Court has been called upon to decide whether corporations enjoy the right to free speech[45] or the right to exercise "their" religion.[46]

If corporate law holds a lesson for broader theories of personhood, it is that entity status cannot be conferred on the basis of philosophical, sociological,[47]

42. MODEL BUS. CORP. ACT § 3.02 (AM. BAR ASS'N 2016); *see also* STEPHEN M. BAINBRIDGE, CORPORATE LAW 2 (2d ed. 2009).

43. For a historical overview, see Henry Hansmann et al., *Law and the Rise of the Firm*, 119 HARV. L. REV. 1335 (2006).

44. STEPHEN M. BAINBRIDGE & M. TODD HENDERSON, LIMITED LIABILITY: A LEGAL AND ECONOMICS ANALYSIS 5–6 (2016); REINIER KRAAKMAN ET AL., THE ANATOMY OF CORPORATE LAW 5–8 (3d ed. 2017).

45. *See generally* Citizens United v. FEC, 558 U.S. 310 (2010).

46. *See generally* Burwell v. Hobby Lobby Stores, Inc., 573 U.S. 682 (2014).

47. A sociological theory of personhood for "software agents" is offered by Gunther Teubner in *Rights of Non-humans?: Electronic Agents and Animals as New Actors in Politics*

or other extralegal theories alone. Rather, the question of legal personhood is of a pragmatic nature.[48] It is not that the legal system "must" recognize digital autonomous systems as legal persons; rather, it is free to decide to accord or withhold entity status. Moreover, a legal system that aspires to general norms of rationality and consistency will grant or withhold entity status not arbitrarily but on the basis of reasons.

Another lesson to be drawn from corporate law is that the recognition of legal personhood does not involve a binary choice.[49] Corporations undoubtedly enjoy entity status in the commercial realm, which leads to mutual asset shielding between the corporation and its owners, managers, and employees.[50] However, whether a corporation also enjoys the full range of constitutional protections is less clear and is certainly a matter of debate. If corporations were denied the protection of fundamental guarantees, such as the right to free speech and religious freedom, this would not be the end of the corporate form. It would not even amount to a logical or normative contradiction to accord entity status but to withhold the application of constitutional guarantees.

II. LIMITED PERSONHOOD AND FUNCTIONS OF LIABILITY

This Part develops the concept of a "liability subject." Liability subjects are not entitled to the full range of rights that human beings and legal persons enjoy. They are recognized as persons for the sole purpose of serving as a defendant in tort suits. In addition, this Part explores the functions of the liability system, as these functions provide the normative framework for thinking about robot liability.

A. Discretion of the Legal System

The legal system is free to confer legal personhood on autonomous systems or to withhold it. It is not the perceived resemblance between humans and robots that forces the legal system to confer entity status on autonomous systems. The real issue is not whether robots must be treated like humans but whether there are good reasons to treat them as legal persons. Where such reasons exist, the legal system is still not compelled to promote robots to persons in every respect. Rather, the legal system recognizes a gradual concept of personhood that allows for the recognition of an autonomous system as a separate legal entity only within certain fields or dimensions.

These propositions clear the way for a rational and pragmatic discussion of the pros and cons of personhood for autonomous systems. The mere fact that they are not human, and even fail to qualify as biological life forms, does

and Law, 33 J.L. & SOC'Y 497 (2006). For an interpretation of Teubner, see Tom Allen & Robin Widdison, *Can Computers Make Contracts?*, 9 HARV. J.L. & TECH. 25, 36–40 (1996).

48. *See* CHOPRA & WHITE, *supra* note 26, at 154.

49. *See id.* at 154–57.

50. *See* BAINBRIDGE, *supra* note 42, at 6; BAINBRIDGE & HENDERSON, *supra* note 44, at 7–15; KRAAKMAN ET AL., *supra* note 44, at 8–9.

not pose an unsurmountable obstacle to their classification as legal persons. The legal universe is free to even classify artifacts such as robots as persons. However, there must be good reasons to accord them this status, and these reasons must be tailored to the specific function that the new candidate for legal personhood is meant to serve.

B. The Concept of a Liability Subject

The question investigated in this Article is not whether robots should be raised to the same level as humans with the full range of rights that humans may hold in modern legal systems. Instead, the subject matter of the inquiry is much more limited: the role of robots as wrongdoers and their eligibility to serve as defendants in tort-based lawsuits. These capabilities may be captured in the term "liability subject." Only if robots qualify as liability subjects can they commit a wrong and be sued for damages in a court of law.

Robots cannot be classified as liability subjects without good reason. This rationale will likely be economic in nature, as is in the case of corporations. Whether or not a robot comes close to, or equals, a human being in philosophical or sociological respects or under a pluralistic theory of legal personhood is irrelevant to whether to recognize autonomous systems as liability subjects. Instead, the objectives of the liability system must control the decision as to which entities qualify as wrongdoers.

C. Functions of the Liability System

This Article takes for granted the traditional goals of the liability system (compensation and deterrence)[51] and accordingly asks whether robots should be classified as legal persons on compensation and deterrence grounds. While the compensation goal gives little guidance on how to structure legal rules, deterrence provides the necessary normative orientation.[52] Accordingly, the following analysis takes a law and economics perspective.

The objective is to maximize the net surplus for society, i.e., the difference between the gain from activities involving robots and the costs of producing and operating them, including the costs of precautions and the costs of accidents that occur in spite of cost-effective precautions.[53] Cost internalization is achieved only if the cost of harm caused by dangerous activities is attributed to the actor engaging in such activities, so that the price of the activity in question reflects its full costs.[54] Where all or part of the risk remains externalized, as it continues to fall on third parties, the cost of the

51. *See* DOBBS ET AL., *supra* note 7, §§ 2.4–2.5.

52. On the relationship between compensation and deterrence, see generally Mark A. Geistfeld, *The Coherence of Compensation-Deterrence Theory in Tort Law*, 61 DEPAUL L. REV. 383 (2012).

53. Guido Calabresi places these two factors together under the rubric of "primary accident costs." GUIDO CALABRESI, THE COSTS OF ACCIDENTS: A LEGAL AND ECONOMIC ANALYSIS 26–27, 68–94 (1970); *see also* STEVEN SHAVELL, FOUNDATIONS OF ECONOMIC ANALYSIS OF LAW 179–81 (2004).

54. *See* SHAVELL, *supra* note 53, at 193–97.

activity is too low and individuals will engage in such activity to an excessive extent.[55]

The question of whether it makes sense to promote robots to liability subjects cannot be answered without a clear view of who else may be held responsible. A brief account of traditional categories of liability for harm caused by technological artifacts will inform the decision on whether robot liability is necessary to fill gaps in the liability system.[56]

III. TRADITIONAL LIABILITY SUBJECTS

This Part explores the actors who would serve as natural defendants in damages suits involving the failure of a robot. As will be seen, the digital innovations will remove much of the control users now have over technical appliances. This decline will be compensated by a proportionate concentration in the hands of the manufacturers. This shift in control will help to push the product liability of manufacturers to center stage and to diminish the liability exposure of product users.

A. The Range of Responsible Parties

The various actors involved in the creation and the operation of autonomous systems can be separated into two distinct groups: manufacturers and users. The manufacturer group includes all those actors, usually businesses, who contribute to the development, design, and production of autonomous systems, including software developers, programmers, and suppliers of component parts. The user group comprises everyone who interacts with robots or other autonomous systems after they are put into circulation, including owners, keepers, and operators of such devices. The distinction between these two groups is not purely descriptive; it also pays tribute to the fact that, within each group, it is fairly easy to allocate the costs of liability to any one member or to share it between several members. The obvious tool for such reallocations is a contractual agreement. Today, standard supply agreements among the members of the manufacturer group, i.e., assemblers and component suppliers of different layers, routinely include clauses that provide for the allocation of the expenses associated with product recalls and other costs caused by defective components.[57]

The mechanism of risk allocation via contract is available for the group of users, as well, i.e., between owners, users, and operators of a robot. Take the example of motor cars. Here, the owner of a car is required to take out liability insurance under state law. If the car is rented out to somebody else, the cost of such insurance is shifted to the lessee-driver, as a component of the price he or she has to pay for the lease of the vehicle. The same happens where a business operates a digital smart machine in its production process:

---

55. *See id.*

56. *See infra* Part III.

57. Omri Ben-Shahar & James J. White, *Boilerplate and Economic Power in Auto Manufacturing Contracts*, 104 MICH. L. REV. 953, 959–60 (2006).

the prices for products manufactured with the help of the device will include a component reflecting the expected costs of any potential harm. Again, costs are shifted within the group of entities that operate or benefit from the use of the robot or other autonomous system. In all cases, as long as responsibility can be attributed to one member of the group, the reallocation of accident costs within the group may be left to the parties.

B. Shifts in Control Induced by Technology

While it is difficult—and not without serious risk of error—to predict the safety characteristics of robots and other autonomous systems, it seems safe to predict that the arrival of such technology will shift control over these machines and appliances away from users and towards manufacturers. Legacy products rely on mechanical technology that is designed, manufactured, and distributed by the parties familiar from product liability suits but that needs to be operated by users.[58] While the manufacturer controls the product's safety features and provides the interfaces between the product and its user (such as buttons, steering wheels, pedals, and the like), it is the user who exercises control in real-world situations and determines the "behavior" of the mechanical device. Cars provide the most obvious example. In contrast to conventional cars, autonomous vehicles will be steered and controlled not by a human driver but by an algorithm developed and installed into the car by its manufacturer. Fully autonomous cars that satisfy "Level 5" of the SAE International classification system for automated vehicles do not require any human intervention when in operation.[59] On the contrary, the intervention of the passenger into the process of driving is prevented through technical safeguards. As a consequence, the "behavior" of the autonomous car is not in the hands of the human driver but in those of the manufacturer. Autonomous cars of the future will transform the user from a driver into a passenger, i.e., into a person who travels inside the car but has no control whatsoever over it. Manufacturers of robots will be able to exercise much more control over the performance and behavior of their creatures than manufacturers of mechanical products are able to exert.

The shift in control is particularly obvious in the case of a closed software system that prevents third parties, including the user, from tampering with the algorithm that runs the device. Here, the manufacturer is the only party who is in a position to determine and improve the safety features of the device; nobody else can. Phrased in economic terms, the manufacturer is

58. DAVID G. OWEN, PRODUCTS LIABILITY LAW 778–79, 938–39 (3d ed. 2015).

59. *See* NAT'L HIGHWAY TRAFFIC SAFETY ADMIN., FEDERAL AUTOMATED VEHICLES POLICY 9 (2016), https://www.transportation.gov/sites/dot.gov/files/docs/AV%20policy%20guidance%20PDF.pdf [https://perma.cc/GP3J-Z23B] (referring to the SAE International definitions for levels of automation and defining Level 5, i.e., full automation, technology as an "automated system [that] can perform all driving tasks, under all conditions that a human driver would perform them").

clearly the cheapest cost avoider.[60] Carmakers will accept the challenge and offer bundles of hardware and software that remain closed to users to ensure their safety. The liability of manufacturers will increase in size and relevance, and the responsibility of users will diminish proportionally.[61]

The example of autonomous cars must not obstruct the view of other autonomous devices. There is a serious possibility that producers of low-risk autonomous systems, such as lawnmowers, will follow an open-system approach that allows users to intermingle with the software that operates the device. Hardware and software would not be marketed in a bundle, allowing users to decide what software product to combine with which kind of hardware. Further, users will be authorized and enabled to modify the software running a robot or other autonomous digital device. In this scenario, it no longer makes sense to place the manufacturer of the original product in the center of the liability stage. The attribution of responsibilities may become rather complex and complicated, as it must be shared or divided between original equipment manufacturers, suppliers of component parts, including software programs, and users, i.e., owners and operators. Whatever principle is adopted in this area, it will almost certainly make it more difficult for the victim to identify the responsible party and to furnish proof that the requirements of liability are in fact satisfied for this party.

It seems that this is the situation that the European Parliament had in mind when it articulated the idea to accord the status of a legal entity, or "ePerson" to the autonomous system or robot itself.[62] Doing so would relieve the victim of the burden of identifying the responsible party and would spare courts the task of allocating liability between a multitude of defendants. This proposal is impossible to evaluate without a clear understanding of the existing liability regimes that determine the responsibilities of the human actors involved in potential accident scenarios for autonomous systems, such as manufacturers[63] and operators.[64]

C. Product Liability as the Default System

Manufacturers, including assemblers and suppliers of raw materials and components, are liable under product liability law. In the United States, a complex cathedral of case law has developed over many decades.[65] In Europe, Directive 85/374/EEC supplies a statutory framework for product liability claims that closely mirrors its U.S. counterpart.[66] One crucial question that deserves analysis concerns the scope of product liability law. Traditionally, it has been limited to "products" in the sense of corporeal

60. *See* CALABRESI, *supra* note 53, at 136–50. *But see* SHAVELL, *supra* note 53, at 189–90.

61. Mark A. Geistfeld, *A Roadmap for Autonomous Vehicles: State Tort Liability, Automobile Insurance, and Federal Safety Regulation*, 105 CALIF. L. REV. 1611, 1691 (2017).

62. *See* Parliament Resolution, *supra* note 9, ¶ 59.

63. *See infra* Part III.C.

64. *See infra* Part III.D.

65. *See generally* OWEN, *supra* note 58.

66. Council Directive 85/374, 1985 O.J. (L 210) (EC).

objects, as defined in Article 2 of Directive 85/374/EEC.[67] Where software has become "embedded" in hardware and both components are marketed together as a bundle of hardware and software, product liability law is applicable even if the defect only affects software.[68] Pertinent examples include accidents involving airplanes, motor cars, and machinery. Today, however, software is no longer distributed with the help of corporeal storage devices such as hard drives, floppy disks, CDs, DVDs, or USB sticks but downloaded from a server located in a virtual space called the cloud. Thus, no corporeal asset is ever placed into the stream of commerce. Not surprisingly, then, the application of product liability law to "mere" software failures is an open question that has arguments on both sides.[69] The better view is to include software in the scope of product liability law. The risks associated with mass-produced software are no different from the ones created by more traditional, physical products. As always, the law must be developed and adapted to new facts and circumstances.

Not surprisingly then, product liability claims involving the use of robotics are not entirely new. For example, the automated surgical system called "da Vinci," often used for prostate cancer surgery (prostatectomy), gave rise to lawsuits that were resolved in the courts.[70] Some of the more prominent published opinions limit themselves to rather peripheral questions, such as the learned intermediary doctrine[71] and issues involving statutes of limitations.[72] Courts were also confronted with the alleged malfunctioning of an autopilot system in an aircraft.[73] Furthermore, there are cases in which workers were killed or injured by robots they were interacting with at work.[74]

The major element in a product liability suit is the establishment of a product defect, and the plaintiff-victim bears the burden of proof. It is common knowledge that so-called manufacturing defects may be established rather easily as, here, the product that caused the harm in question differs from the blueprint that the manufacturer used when producing the item. Examples involving digital products include the incomplete installation of software in an autonomous car or internet-of-things device, as well as accidents caused by software bugs that were inadvertently written into the computer code.

---

67. *Id.*

68. Thomas G. Wolpert, *Product Liability and Software Implicated in Personal Injury*, 60 DEF. COUNS. J. 519, 521 (1993).

69. *See* OWEN, *supra* note 58, at 1082.

70. *See, e.g.*, Pohly v. Intuitive Surgical, Inc., No. 15-cv-04113-MEJ, 2017 WL 900760 (N.D. Cal. Mar. 7, 2017); Reece v. Intuitive Surgical, Inc., 63 F. Supp. 3d 1337 (N.D. Ala. 2014); *In re* Intuitive Surgical, Inc., da Vinci Robotic Surgery Sys. Prods. Liab. Litig., 883 F. Supp. 2d 1339 (J.P.M.L. 2012); Mracek v. Bryn Mawr Hosp., 610 F. Supp. 2d 401 (E.D. Pa. 2009); Taylor v. Intuitive Surgical, Inc., 389 P.3d 517 (Wash. 2017).

71. *See, e.g.*, *Taylor*, 389 P.3d at 524–26.

72. *See, e.g.*, *Reece*, 63 F. Supp. 3d at 1338–39.

73. *See, e.g.*, Ferguson v. Bombardier Servs. Corp., 244 F. App'x 944 (11th Cir. 2007).

74. *See, e.g.*, Payne v. ABB Flexible Automation, Inc., No. 96-2248, 1997 WL 311586 (8th Cir. June 9, 1997).

Design defects that affect a whole series of products are much more important but also more difficult to establish. A product is defectively designed if its layout, chosen by the manufacturer during the research and development process, is found wanting. The consumer expectations test, supported by the language of Article 6(1) of Directive 85/374/EEC[75] and the risk utility test, embraced by the Restatement (Third) of Torts on product liability,[76] offer competing standards for design defects.[77] These two alternative tests may be combined and reconciled in the concept of reasonable consumer expectations—that is, the notion that rational consumers would expect precisely the standard of safety that a cost-benefit analysis generates.[78]

Robots and other autonomous systems will offer new challenges for these two approaches, but the crucial issue will remain the same: given that we are living in an imperfect world, absolute safety is out of reach. The mere fact that a product caused the harm in question does not render it defective. Applying the standard for design defects to autonomous systems requires an inquiry into software programming. The courts, or other decision makers, will need to identify shortcomings of the software that could have been avoided by an alternative program that would have performed as well as the one that was used—but would have avoided the accident in question.

Autonomous systems will pose a new challenge to the concept of design defect. One may be tempted to compare the performance of an autonomous system to the one of a legacy product operated by a human being. In the case of autonomous cars, this solution would amount to a "human driver test": whenever the autonomous system caused an accident, which a reasonable human driver would have been able to avoid, the algorithm would be found defective in design. Intuitive as the human operator test may seem, its application to autonomous systems is misguided.[79] Autonomous systems are expected to decrease the number and severity of accidents dramatically,[80] but accidents will continue to occur. The critical point is that the pool of accidents that an autonomous system still causes will not be the same as the pool of accidents a reasonable driver is unable to avoid. For instance, an autonomous car operating in an orderly manner will never speed, and it cannot be drunk. However, it might fail to observe and account for a freak event that any human would have recognized and adapted his or her behavior to. This might have been the case when a Tesla car operated by its "automatic pilot" hit a truck in glaring sunlight because it failed to recognize the truck

75. *See supra* note 66 and accompanying text.

76. RESTATEMENT (THIRD) OF TORTS: PRODUCTS LIABILITY § 2(b) cmts. f–g (AM. LAW INST. 1998).

77. *See* OWEN, *supra* note 58, at 482–503. For European courts, see SIMON WHITTAKER, LIABILITY FOR PRODUCTS: ENGLISH LAW, FRENCH LAW, AND EUROPEAN HARMONISATION 487–88 (2005).

78. MARK GEISTFELD, PRINCIPLES OF PRODUCTS LIABILITY 37–54 (2d ed. 2011).

79. *See* Geistfeld, *supra* note 61, at 1644–47.

80. *See* NAT'L HIGHWAY TRAFFIC SAFETY ADMIN., *supra* note 59, at 5.

as such.[81] To subject autonomous systems to a human operator test would miss the mark as it would hold the system to a standard up to which it cannot live.

What is required, therefore, is a system-oriented concept of design defect.[82] The crucial question must be whether the system in question, e.g., a fleet of cars operated by the same algorithm, causes an unreasonable number of accidents overall. Whether the individual accident in question would have been avoided by a reasonable human driver should be irrelevant. Sadly, however, the system-oriented concept of design defect may cause adverse effects for competition in the marketplace. If the finding of a design defect requires a comparison between the algorithm that caused the harm and the ones used by other manufacturers, the result will be an "optimal algorithm test" that discriminates against all but the best algorithm in the market. The algorithm that caused the harm will be found defective whenever there is an algorithm in the market that would have avoided the particular accident in question. And even applied to the full class of accidents caused by any one fleet of autonomous products operated by the same algorithm, this method would lead to a finding of all the algorithms in the market as defective—except for the safest of them all. Assuming that one algorithm operates a whole fleet of cars or other products marketed by a particular manufacturer, only the manufacturer with the best algorithm would be spared, while all the other manufacturers would be saddled with the full costs of accidents caused by their products. This outcome would be problematic, as it would exacerbate the first-mover advantage and stifle competition in the respective product market.

D. User Liability as a Supplement

There is no special liability regime for users of products of any kind, let alone robots or other autonomous systems. This does not mean that users go scot-free. Rather, they are subject to general tort law—primarily the tort of negligence, which offers individuals civil redress for personal injury as well as damage to property. The central element of liability is the breach of a duty of care. While the question of whether such a breach occurred occupies center stage in ordinary tort suits, the question of whether a duty is owed by a particular defendant to a particular plaintiff is less prominent and more difficult to answer. Courts treat the imposition of a duty of care on a particular defendant as a policy issue that requires the balancing of a multitude of factors.[83] The Restatement (Third) of Torts on physical and emotional harm stipulates that "[a]n actor ordinarily has a duty to exercise

81. Will Oremus, *The Tesla Autopilot Crash Victim Was Apparently Watching a Movie When He Died*, SLATE (July 1, 2016, 7:43 PM), http://www.slate.com/blogs/moneybox/2016/07/01/tesla_autopilot_crash_victim_joshua_brown_was_watching_a_movie_when_he_died.html [perma.cc/LZ77-BAFN].

82. *See* Geistfeld, *supra* note 61, at 1645–47.

83. *See* Marshall v. Burger King Corp., 856 N.E.2d 1048, 1057 (Ill. 2006); 532 Madison Ave. Gourmet Foods, Inc. v. Finlandia Ctr., Inc., 750 N.E.2d 1097, 1100–01 (N.Y. 2001).

reasonable care when the actor's conduct creates a risk of physical harm," and it continues with a list of factors to be taken into account, including social norms, respect for contractual risk allocations, conflict with the defendant's other rights and obligations, and respect for the courts' institutional role and for the other branches of government.[84]

In spite of these difficulties, it seems clear that operators of vehicles, machines, and other technical appliances are subject to a tort-based duty of care. As users of robots and other autonomous systems will have only very limited control over the devices they operate, the duty of care must be restricted accordingly. The user of such an appliance is answerable in damages only where he or she misused or abused the device, causing harm to others. For example, if the user of an autonomous car overrides the software's protections in order to steer the vehicle off the streets or to go faster than the software allows, this constitutes negligence. Further, in the case of unbundled products or open systems, the user is responsible for any software installed subsequent to the purchase of the original system, and for any modifications made to the original product.

Some legal systems have gone beyond fault-based liability and subjected users to strict liability for harm caused in the operation of an installation, appliance, or machine. This is true for France where strict liability exists for any "keeper" (beneficial owner) of a "thing," regardless of whether the "thing" was defective or not.[85] In the area of motor traffic, France has even moved beyond strict liability by providing a compensation system for traffic accidents that allows parties to settle on mere involvement (implication) rather than fault.[86] While the French solution of strict liability for keepers of any "thing" did not win the approval of groups of legal experts working towards harmonization of European tort law,[87] the advent of autonomous systems may force lawmakers to reconsider it.

If markets develop towards unbundling, and original equipment manufacturers lose control over the safety features of the products they put into circulation, responsibilities will become blurred. It will thus become increasingly difficult for victims to single out the actor responsible for the accident in question. To the extent that the victim fails to pinpoint the responsible party, the damages claim fails and incentives to take care are lost. Such outcomes could be avoided if operators were held strictly liable for any harm caused in the course of the operation of an autonomous system. The question of who bears responsibility for a particular accident would then be

84. RESTATEMENT (THIRD) OF TORTS: PHYSICAL AND EMOTIONAL HARM § 7(a) cmts. c–g (AM. LAW INST. 2010).

85. Gerhard Wagner, *Custodian's Liability*, *in* 1 THE MAX PLANCK ENCYCLOPEDIA OF EUROPEAN PRIVATE LAW 441, 441–43 (Jürgen Basedow et al. eds., 2012).

86. CEES VAN DAM, EUROPEAN TORT LAW 408–11 (2d ed. 2013); Geneviève Viney & Anne Guédan-Lécuyer, *The Development of Traffic Liability in France*, *in* 5 THE DEVELOPMENT OF TRAFFIC LIABILITY 50, 67–69 (Wolfgang Ernst ed., 2010).

87. EUROPEAN GRP. ON TORT LAW, PRINCIPLES OF EUROPEAN TORT LAW 101–04 (2005); 4 PRINCIPLES, DEFINITIONS AND MODEL RULES OF EUROPEAN PRIVATE LAW: DRAFT COMMON FRAME OF REFERENCE (DCFR) 3544, 3558 (Christian von Bar et al. eds., 2008).

shifted towards the user and his insurers who, in turn, would seek recourse against hardware and software manufacturers.

IV. ROBOTS AS LIABILITY SUBJECTS

Moving beyond philosophical analyses of personhood, the following Part explores the essential functions of robot liability in light of the functions of the tort system, and of the proper role of the traditional parties, namely manufacturers and users. As will be seen, there is much to be learned from corporate law when analyzing the potential promotion of robots to ePersons. The danger of risk externalization and internalization strategies, such as minimum asset requirements and insurance mandates, needs to be explored. This leads to the question of whether it will ever be possible to incentivize robots in much the same way that the liability system motivates human beings to take care in order to avoid harm to others.

A. The Function of Robot Liability

It has now become clear that the recognition of robots as legal persons for purposes of civil liability is not a philosophical question that can be answered by examining the characteristics of a digital device and asking whether it is sufficiently similar to a human being. Rather, accepting that robots can be liable calls for a functional explanation that is in tune with the general principles and goals of tort law, namely compensation and deterrence.

It is not easy to find a positive explanation for robot liability, given the range of responsible parties that already exist, namely manufacturers, suppliers, owners, and users of such devices. As the preceding surveys of the liability regimes for manufacturers and users revealed,[88] current tort law generates powerful incentives for the manufacturers and operators of autonomous systems to take due care. So far, the creation of an additional liability subject, the ePerson, is simply superfluous.

There seems to be only one niche where robot liability could serve a useful role: markets for unbundled digital products. In the case of unbundling, people injured by a robot may face serious difficulties in identifying the party who is responsible for the misbehavior of the device.[89] The fact that the robot malfunctioned is no evidence that the hardware put into circulation by one manufacturer or the software downloaded from another manufacturer was defective. Likewise, the responsibility of the user may be difficult to establish. Thus, in a market of unbundled products, the promotion of the robot to a liability subject may serve as a tool for "bundling" responsibility and attributing liability to a single entity to which the victim may turn for compensation. The burden of identifying the party responsible for the malfunction or other defect would then be shifted away from victims and onto the robot's liability insurers. These insurers, in turn, are professional players who may be better able to investigate the facts, evaluate the evidence,

88. *See supra* Part I.B.
89. *See supra* Part III.B.

and pose a credible threat to hold hardware manufacturers, software programmers, or users accountable in exercising their rights of recourse against them. The question remains whether the benefits of promoting robots to liability subjects would outweigh the costs.

B. The Danger of Cost Externalization

As a first approximation, the answer to the question of whether robots should qualify as entities capable of bearing liability must be in the negative. Robots lack the assets necessary to pay off damages claims. If they were nonetheless accepted as legal entities, victims would receive nothing and incentives to take care would be lost. Entity status would result in a complete externalization of accident risk.

In this context, it is important to note that recognizing robots as ePersons would protect all actors "behind" the robot from liability. In the example of company shareholders, the creation of a distinctive legal entity, such as a corporation, works as a shield against liability for the actors who created the entity.[90] This shield acts to stimulate risk-bearing; shareholders cannot lose more than the money they invested into the corporation.[91] Applying the principle of limited liability to ePersons, manufacturers and users of robots would be exempt from liability, as they qualify as quasi shareholders of the robot. The robot's manufacturers, programmers, and users would no longer be liable, as the "behavior" of the robot would no longer be ascribed to them, but instead to the robot itself. This could be tolerated, in the sense of a price worth paying, if the newly created legal entity itself were capable of responding to the threat of liability. This is emphatically not true for robots. Under the proposition of ePerson liability, no one responsive to the financial incentives of the liability system would in fact be exposed to it.

This outcome is intolerable, both under the compensation and the deterrence rationale of tort law. Victims who incurred harm caused by the robot would receive nothing in compensation. The quasi shareholders of the robot would have no financial incentive to manufacture the robot and operate it in a way that reduces the risk of harm. No incentives to take precautions would exist. Furthermore, the price charged for the robot would not reflect the true social cost of its creation and operation, as the harm caused to third parties would remain with the victims. As a consequence, too many robots would be produced and marketed and existing robots would be used excessively—even where the social cost of their operation exceeded the value they generated.

---

90. *See* FRANK H. EASTERBROOK & DANIEL R. FISCHEL, THE ECONOMIC STRUCTURE OF CORPORATE LAW 40–62 (1991); *see also* BAINBRIDGE & HENDERSON, *supra* note 44, at 44–85; KRAAKMAN ET AL., *supra* note 44, at 5–6.

91. *See* BAINBRIDGE & HENDERSON, *supra* note 44, at 2; EASTERBROOK & FISCHEL, *supra* note 90, at 40.

C. Internalization Strategies

The problem of risk externalization, together with the frustration of incentives to take care, may be addressed by the legal system, and the remedies are similar to the ones employed in corporate law. The robot could be required to be endowed with a minimum of assets in order to qualify as a legal entity. Such a minimum asset requirement would force other parties to provide the funds necessary to satisfy potential damages claims. These funds would then be transferred to the robot and held in its own name. Damages claims would be paid off from this pool of assets.

An alternative to minimum asset requirements that serves the same end is mandatory liability insurance. The law could simply stipulate an insurance mandate as a precondition for incorporation of a robot as an ePerson. Again, the burden of providing the mandatory liability insurance would fall on the natural and legal persons who operate the robot or put it into circulation. They would have to supply the insurance contract and pay the premiums, as the robot would have no assets to pay them from.

Within the scope of the insurance cover or asset cushion, the crucial issue is who will be liable to contribute. The robot cannot pay for insurance, so somebody else will need to provide the necessary funds. The usual suspects are already familiar: the manufacturers of the robots and their users. If the manufacturers have to front the costs of insurance, they will pass these costs on to the buyers or operators of the robot. In one form or another, they would end up with the users. The same outcome occurs if users contribute directly to the asset cushion or become liable for insurance premiums. In the end, therefore, the robot's producers and users must pay for the harm the robot causes. The ePerson is only a conduit to channel the costs of coverage to the manufacturers and users.

Whatever tool would be chosen by the legal system, both minimum asset requirements and mandatory insurance are well suited to avoid risk externalization. Victim compensation is assured at least up to the insurance ceiling or the value of the minimum assets. Beyond this amount, however, risk externalization would persist.[92] Again, the essential point about entity status for robots is that this move helps to shield other parties from liability, namely manufacturers and users. Within the corporate context, the protective function of limited liability is acceptable for voluntary creditors who can easily protect themselves against risk externalization, but it is much more problematic for involuntary creditors like tort victims who lack any means to do so.[93]

92. *See* EASTERBROOK & FISCHEL, *supra* note 90, at 49–50.

93. Henry Hansmann & Reinier Kraakman, *Toward Unlimited Shareholder Liability for Corporate Torts*, 100 YALE L.J. 1879, 1920–21 (1991); David W. Leebron, *Limited Liability, Tort Victims, and Creditors*, 91 COLUM. L. REV. 1565, 1601–02 (1991).

D. Incentivizing Robots?

In the case of limited shareholder liability, the corporation is not immune from liability. As an organization that ties together individuals through a nexus of contracts, it may respond to the incentives generated by the liability system.[94] It is essential to understand that matters are different when it comes to ePersons. The reason is that robots, however "intelligent" they may become, will never be able to respond specifically to the incentives generated by the liability system. Sure enough, an autonomous software system can be programmed to "learn" from past experience in the sense that the algorithm improves with every accident it becomes involved in. However, the capacity of the software for improvement is based on its programming—that is, on decisions made by software programmers. Whether or not the autonomous system will be held liable for the consequences of an accident it has caused is irrelevant to the algorithm's learning curve (or lack thereof). Obviously, software can be programmed to improve itself even without the concept of an ePerson and without the threat of robot liability. Thus, autonomous software agents are immune from the financial incentives generated by a credible threat of being held liable for harm caused, at least as long as they do not live to the standards of full autonomy and self-governance in the Kantian sense.[95] The fact that potential ePersons are unreceptive to financial incentives to avoid harm raises serious concerns with a view to deterrence, even if minimum asset requirements or insurance mandates apply.

CONCLUSION

This Article explores the role of robots as wrongdoers and their eligibility to be qualified as liability subjects based on tort law. As scholars have identified, it is conceivable to develop tools that aim to preserve, or restore, the incentives generated by the liability system when addressing the ePersons context. In particular, the law could define minimum asset requirements for ePersons and force the parties that control the safety features of the device and the frequency of its use to contribute to this asset pool. In this way, manufacturers and users would pay indirectly for the costs of accidents caused by the autonomous system. The same outcome would occur if these parties were obligated to pay the premiums for an insurance policy covering third-party losses caused by the robot. Such a system could be supplemented by rights of recourse that the robot, or rather its liability insurer, would have against the manufacturer of the robot, and perhaps also its user. If the funding obligations were accurately calibrated and such rights of recourse generously granted, manufacturers and users would be exposed to the exact same incentives that they would have faced if the robot had not been promoted to a liability subject.

The positive rationale for promoting robots to liability subjects remains weak, however. The basic explanation is that ePersons would help to solve

94. *See* EASTERBROOK & FISCHEL, *supra* note 90, at 40–41.
95. *See supra* Part I.B.

evidentiary problems in markets of unbundled hardware and software products. At present, there is no apparent need to respond to the potential problems of unbundled product markets by allowing the incorporation of robots and autonomous systems as separate legal entities. It is doubtful whether the evidentiary problems expected from markets with unbundled products will ever be worth the cost of creating a new legal entity. The losses associated with the partial externalization of risk need to be balanced against the benefits from improved internalization of risk through the supply of a palpable defendant who could be held fully responsible for the actions and omissions of the robot. Whether the scale will tip in the direction of robot incorporation is far from clear, as the difficulties in terms of initiating claims and their enforcement caused by unbundled products may turn out to be less serious than expected. After all, digital technology offers unique opportunities to record evidentiary data and to provide access to them at zero or trivial cost. It may well be that the information stored in the "black boxes" that will be installed in robots, self-driving cars, and other autonomous systems will allow victims to identify the responsible party easily and accurately. Until it has been proven that these hopes will not materialize, however, legislation to create ePersons as liability subjects is not desirable.

These conclusions admittedly fly in the face of ePerson proponents who aim for meaningful protection of manufacturers and users from excessive liability with a view to stimulate innovation. It is true that robot liability would provide a shield for manufacturers and users as "stakeholders" of the ePerson. Provided that manufacturers and users would be forced to contribute to a pool of assets needed to pay off damages claims or to front the costs of an insurance policy (which is essentially the same), the ePerson supplies a partial shield against their personal liability. Economically, exposing the ePerson to liability would work like a cap on the liabilities of manufacturers and users. They would bear the costs of accidents caused by the robot up to the amount of their contributions to the asset pool or the insurance cover, but no more. Thus, conferring liability onto robots would result in limited liability of the robot's quasi shareholders (its manufacturers and users). Here, as in corporate law, the creation of a legal entity helps to limit the exposure of the individuals who created the entity and thus may stimulate them to take on more risk at lower cost.[96] However, it must be remembered that any "liability subsidy" accorded to certain activities stimulates an excessive amount of such activities. If autonomous systems really generate the great savings in accident costs that they are promised to, then no liability subsidy is needed.

96. BAINBRIDGE & HENDERSON, *supra* note 44, at 2, 47–48, 69.