NUS National University of Singapore | Centre for Technology, Robotics, Artificial Intelligence & the Law Faculty of Law

NUS Law Working Paper 2021/019

NUS Centre for Technology, Robotics, Artificial Intelligence & the Law Working Paper 21/02

Artificial Intelligence and Agents

Daniel Seng

Tan Cheng Han SC

Associate Professor, Director, Centre for Technology, Robotics, Artificial Intelligence & the Law, Faculty of Law, NUS

Dean and Chair Professor of Commercial Law, School of Law, and Co-Director, Centre for Public Affairs and Law and the Technology Policy and Law Lab, City University of Hong Kong

[October 2021]

Artificial Intelligence and Agents

Daniel Seng[*] and Tan Cheng Han[**]

Abstract

With the increasing sophistication of AI and machine learning as implemented in electronic agents, arguments have been made to ascribe to such agents personality rights so that they may be treated as agents in the law. The recent decision by the Australian Federal Court in Thaler to characterize the artificial neural network system DABUS as an inventor represents a possible shift in judicial thinking that electronic agents are not just automatic but also autonomous. In addition, this legal recognition has been urged on the grounds that it is only by constituting the electronic agents as legal agents that their human principals may be bound by the agent's actions and activities, and that a proper foundation of legal liability may be mounted against the human principal for the agent's misfeasance. This paper argues otherwise. It contends that no matter how sophisticated current electronic agents may be, they are still examples of Weak AI, exhibit no true autonomy, and cannot be constituted as legal personalities. In addition, their characterization as legal agents is unnecessary because their actions (and misapplications) can be legally addressed by the under-appreciated instrumentality principle. By treating the electronic agents as instruments or extensions of the acts of the human principal, issues in contract and tort law can be readily resolved. This essay concludes that until a Strong AI application can be demonstrated, the issue of legal agency of electronic agents ought not to detain the development of technology and of the law in this space.

Introduction

1. It is 5 pm, and your day is done. You check your phone app to verify that the home air conditioner will start the moment you leave the office.[1] You are confident that the smart Internet of

[*] Associate Professor, Faculty of Law, and Director, Centre for Technology, Robotics, AI and the Law (TRAIL), National University of Singapore. I would like to thank Ms Hitomi Yap and Mr Shaun Lim for their help with the research and editing of this paper. All errors and omissions however remain mine.

[**] Dean and Chair Professor of Commercial Law, School of Law, and Co-Director, Centre for Public Affairs and Law and the Technology Policy and Law Lab, City University of Hong Kong.

[1] A smart air conditioner controller mimics the infrared signal from the air conditioner while using the home's WiFi network to connect with an app on the phone. *See* thesmartcave, Here's what you need to know about smart AC controllers, <https://thesmartcave.com/best-smart-air-conditioner-controller/>.

Things (IoT) thermostat had turned the air conditioner off once it detected that no one was home.[2] When you reach the garage, you activate the Smart Summon feature. Your Tesla unparks itself and drives to the pickup spot,[3] whereupon you hop in and drive off. At the garage exit, the gantry sensor communicates with the car's onboard unit and debits the parking fee from your bank account.[4] Coasting out of the garage and onto the expressway, you start the Autopilot feature.[5] As you settle yourself into the driver's seat, a synthetic voice[6] reminds you of your scheduled COVID-19 vaccination registration for tomorrow.[7] You smile, knowing that once vaccinated, you can avoid having your nose swabbed, even if this can be done more comfortably by robots![8]

2. This scenario is no longer the province of science fiction. In fact, modern society would grind to a halt without the use of these systems that are able to "act[] on behalf of [their] user[s] and [try] to meet certain objectives or complete tasks without any direct input or direct supervision from [their] user[s]".[9] These systems, described in technical and marketing literature as "agents", pervade all aspects of our lives in the Fourth Industrial Revolution.[10] Internet searches work through web crawlers, spiders or bots—software programs that systematically browse the World Wide Web to collect and index content.[11] Electronic commerce operates with software to advertise and market products, collect orders, process payments, and initiate distribution and manage logistics.[12] Chatbots and voice assistants answer queries and perform tasks or services for individuals or companies through commands or questions.[13] Sophisticated software programs interpret human speech and respond either by way of a

[2] *See e.g.,* the Nest Learning Thermostat, which is a web-connected, smart thermostat from Nest, a division of Google, that automatically learns about routines and programs itself to set the temperature in your home. *See* Google, *How Nest thermostats learn*, <https://support.google.com/googlenest/answer/9247510?hl=en>.
[3] The Tesla Smart Summon feature allows a Tesla driver to turn her car on remotely and beckon her Tesla to her, to save her the trouble of entering or exiting the vehicle. *See* <https://www.tesla.com/support/autopilot>.
[4] The new Electronic Road Pricing system, which was to start in 2020, includes a feature to pay for parking. *See* <https://www.motorist.sg/article/413/new-erp-system-to-start-in-2020-includes-new-in-vehicle-units>.
[5] *See* Tesla, *Autopilot and Full Self-Driving Capability Features*, <https://www.tesla.com/support/autopilot>.
[6] See Wikipedia, *Speech synthesis*, <https://en.wikipedia.org/wiki/Speech_synthesis>.
[7] *See* Ministry of Health, Singapore, *COVID-19 Vaccination Registration*, <https://www.vaccine.gov.sg/>.
[8] *See* Straits Times, *New Covid-19 swab test robot offers safe, more comfortable procedure for patients*, Sep. 22, 2020, <https://www.straitstimes.com/singapore/robot-that-conducts-swab-tests-for-covid-19-is-safe-faster-and-more-comfortable-for>.
[9] JJ Borking, BMA van Eck, and P Siepel, *Intelligent software agents and privacy* (1999) at 1.
[10] Wikipedia, *Fourth Industrial Revolution*, <https://en.wikipedia.org/wiki/Fourth_Industrial_Revolution> (last accessed 26 June 2021).
[11] Wikipedia, *Web crawler*, <https://en.wikipedia.org/wiki/Web_crawler> (last accessed 28 May 2021).
[12] Wikipedia, *E-commerce*, <https://en.wikipedia.org/wiki/E-commerce> (last accessed 28 May 2021).
[13] Wikipedia, *Virtual assistant*, <https://en.wikipedia.org/wiki/Virtual_assistant> (last accessed 28 May 2021).

dialog, or via synthesized voices or customized messages and responses.[14] Many systems have been implemented either in the form of apps on smartphones, or in standalone devices such as smart speakers and remote controls.[15] They have become part of IoT – an ecosystem where devices with sensors automate many aspects of our homes and businesses.[16]

3. Undoubtedly, these software systems and devices vary greatly in their level of sophistication and complexity. Their ubiquity is also a product of their diverse functionalities and deployments. What they have in common is that they supplant and automate processes that would otherwise require human intervention. However, excluding the human element from economic and business transactions raises questions about the nature and character of these interactions. For instance, what is the legal status of a contract made through an automated system? Does a transacting party have any recourse, in contract or in tort, for any errors, mistakes and omissions that have been made by the automated system? Some academic writers have argued that these issues are best resolved by characterizing these automated systems as legal agents. In this paper, we will review all these issues and arguments.

Types of Artificial Agents

4. A study of automated systems must start by defining what an automated system is. Some have referred to these systems as "automated", "electronic" or "artificial" agents. Others distinguish between "first generation … electronic agents" that are "rules-based"[17] or "exhibit[] a limited intelligence, autonomy, and mobility"[18] and the "second generation of intelligent software agents"[19] or "more autonomous artificial agents"[20] that "use intelligent algorithms"[21] or "exhibit[] a considerable level of autonomy, mobility, and sophistication".[22] There are also references to agents that are "self-modifying and acting according to their own experience",[23] have "an autonomous capacity"[24] or are "capable of

[14] *Id*.
[15] Smartphone users would be familiar with software and devices such as Apple's Siri, Amazon's Echo and Alexa, and Google Assistant.
[16] Wikipedia, *Internet of things*, <https://en.wikipedia.org/wiki/Internet_of_things> (last accessed 26 June 2021).
[17] Suzanne Smed, *Intelligent Software Agents and Agency Law*, 14 Santa Clara High Tech. L.J. 503 (1998).
[18] Emad Abdel Rahim Dahiyat, "Law and software agents: Are they 'Agents' by the way?" 29 Artificial Intelligence and Law 59 (2021), at <https://doi.org/10.1007/s10506-020-09265-1>.
[19] Smed (n 17) 503.
[20] Samir Chopra and Laurence F White, "Artificial Agents and the Contracting Problem", J. of Law, Tech. and Policy 363, 364 (2009).
[21] ibid 365.
[22] Dahiyat (n 18) 62.
[23] ibid 65. *See also* Smed (n 17) at 503.
[24] Dahiyat (n 18) 65.

learning and adapting over time"[25] to respond to changes. These systems can make optimal decisions or "act in some extra-legal manner"[26] and even "take actions that neither the licensor nor licensee anticipated".[27] They "employ more sophisticated decision-making mechanisms using statistical or probabilistic machine learning algorithms [where] there may be no strictly binary rules which determine the outcome [such that it] is the combination of relevant factors, and the relative weights the system accords them, that determines the outcome."[28]

5. These somewhat melodramatic classifications are not helpful, not only because of the arbitrary and fuzzy distinctions made, but also because of the equivalences made using value-loaded terms such as "intelligence" and "discretion".[29] For instance, it used to be thought that a system is "autonomous" or "intelligent" if it is programmed to be capable of "spontaneous learning" and "adjust[ing] its behavior" by "adapting its behavior patterns in response to newly encountered circumstances".[30] Today, we view these systems as implementations of machine learning using neural networks and deep learning, and describe them as implementations of Weak AI or "Artificial Narrow Intelligence" rather than the "intelligence" that is to be equated with human intelligence (also known as Strong AI or "Artificial General Intelligence").[31] This implies that we need a more robust definition as we understand the technology and its limitations better. Thus, we choose to define an "agent" (in a non-legal sense) as "anything that can be viewed as perceiving its environment through sensors and acting upon that environment through actuators."[32] This definition, advanced by Russell and Norvig in the leading textbook on Artificial Intelligence, speaks for the fact that an agent so defined can be a human agent, who has eyes, ears and other organs for sensors, and hands, legs, vocal cords and so on for actuators; a software agent – which receives files, network packets and human input as sensor inputs and acts on

[25] Chopra and White (n 20) 369.
[26] Dahiyat (n 18) 65.
[27] Smed (n 17) 503. *See also* L H Scholz, "Algorithmic Contracts", 20 Stanford Technology LR 128, 150-155 (2017).
[28] Chopra and White (n 20) 369.
[29] Computer scientists themselves have been guilty. For instance, Marvin Minsky defined "Artificial intelligence" as the "science of making machines do things that would require intelligence if done by men." Marvin L. Minsky, Introduction to Semantic Information Processing, at v (1968). Since then, many in the field have resiled from such statements.
[30] *See* Thomas S. Ray, "Evolution and Optimization of Digital Organisms", Scientific Excellence in Supercomputing: The IBM Prize Papers (Keith R. Billingsley, Ed Derohanes, Hilton Brown, III eds., 1991).
[31] *See e.g.,* Eda Kavlakoglu, "AI vs. Machine Learning vs. Deep Learning vs. Neural Networks: What's the Difference?", IBM (27 May 2020), at <https://www.ibm.com/cloud/blog/ai-vs-machine-learning-vs-deep-learning-vs-neural-networks>.
[32] Russell, Stuart and Norvig, Peter, ARTIFICIAL INTELLIGENCE: A MODERN APPROACH (4th ed, 2021 ed), at 36 (*AI: A Modern Approach*).

the environment by writing files, sending network packets and displaying information or generating sounds; or a robotic agent – with cameras and range finders for sensors and various motors for actuators.[33]

6. In trying to set up an artificial agent as one with the same qualities as a human agent, Russell and Norvig do not classify an artificial agent as being advanced or basic, but focus on whether the artificial agent is doing "the right thing". Thus an artificial agent's choice of action at any instance depends on (a) its built-in knowledge and (b) the sequence of content its sensors have perceived (the agent's percept sequence). The choice is effected by mapping every percept sequence to each choice of action, by way of different implementations or combinations of implementations known as "models".[34] While humans, who have desires and preferences of their own, choose actions that produce desirable results from their point of view (or additionally, are morally, ethically and legally correct), machines *do not have desires and preferences of their own.*[35] Instead, a non-human agent (which we will term "an artificial agent") is programmed to maximize its performance based on these models and one that does so successfully is said to exhibit "rationality"[36] or "intelligence".[37]

Artificial Agents as Legal Agents

7. Because artificial agents are perceived to be "intelligent" and are believed to be "exercising discretion" autonomously,[38] proponents have advanced the theory that artificial agents should be treated as legal agents, with their users operating as principals in agency law. They assert:

> Of relevance here is the concept of authority found in agency law, which circumscribes the authority the agent has to make decisions on behalf of the principal. Implicit in the existence of that authority is the discretion to take decisions the principal herself would not have taken. Thus, prima facie, artificial agents, concluding contracts on behalf of the corporations or users that deploy them, function like legal agents.[39]

[33] ibid 36.
[34] ibid.
[35] ibid 39.
[36] ibid 39-40.
[37] ibid 49.
[38] Chopra and White (n 20) 370 (this is so even when agents are said to be "exercising discretion" in a predictable way because "had the principal considered the matter herself, she might have reached a different decision.")
[39] ibid 370.

8. This theory coalesces around the concept of treating artificial agents as legal agents in the law of agency.[40] This is because a legal agent does not need full legal capacity: a legal agent does not need capacity to hold legal rights or be subject to liabilities to bind the principal.[41] Since a child has the capacity to enter into contracts to bind her parent or guardian at common law even though she has no capacity to bind herself to a contract,[42] it has been argued that artificial agents can, too, constitute legal agents to bind their human principals.[43] The theory also draws support from the recognition of ships and companies as artificial legal persons with legal capacities, rights and liabilities.[44] But this premise is flawed because unlike artificial agents, people (and companies) underlie the operation of ships[45] and companies.[46]

9. A variant of this theory recognizes this difficulty and seeks to create a specific legal status for artificial agents – autonomous robots, in this case. This was proposed by the European Parliament's Resolution of 2017.[47] Drawing upon the Resolution, which requires that "autonomous robots" be individually registered and be supported by an obligatory insurance scheme, Dahiyat has proposed yet another variant where all businesses intending to use artificial agents do so through the creation of companies under existing corporations legislation. In this way, artificial agents "would act in the name of the company".[48] Dahiyat argues that this could be considered the first step that might prepare for the

[40] ibid 376-68, 401-402; Smed (n 17) 505; I R Kerr, "Spirits in the Material World: Intelligent Agents as Intermediaries in Electronic Commerce", 22 Dalhousie LJ 190, 239-247 (1999); Scholz (n 27) 165-168; Wein, Leon E., "The responsibility of intelligent artifacts: Toward an automation jurisprudence" (1992) 6 Harvard Journal of Law and Technology 103, 106-107.

[41] *See* Watts and Reynolds, *Bowstead and Reynolds on Agency* (22nd ed), Article 5, paras. 2-012 – 2-013; *Restatement (Third) of Agency*, § 3.05, Comment b, Illustration.

[42] ibid. The rule, which is quite established at common law, has been previously applied to married women (who until recently had no contractual capacity).

[43] Chopra and White (n 20) 401-402; Chopra, Samir and White, Laurence F., *A Legal Theory for Autonomous Artificial Agents* (2011), at 41-42 ("It would be possible to treat an artificial agent as having authority to enter contracts on behalf of its operator, but without itself having legal personality.").

[44] *E.g.*, Chopra and White (n 20) 379; Wein (n 40) 107-108, 118.

[45] Oliver Wendell Holmes, The Common Law, 29 (1923) (reiterating that the reference to ships as legal persons is a tool of legal convenience).

[46] *See* e.g. Dahiyat (n 18) 9; Chen and Burgess, "The boundaries of legal personhood: how spontaneous intelligence can problematise differences between humans, artificial intelligence, companies and animals" (2019) 27 Artificial Intelligence and Law 73, 84-85. It was also pointed out that while companies exist as separate artificial legal persons, the legal doctrine of piercing the corporate veil exists to get at the actual deciding human mind and will that represent a company. ibid 85.

[47] European Parliament's Resolution of 16 Feb 2017, paragraph 59(f) (2015/2103(INL)).

[48] Dahiyat (n 18) 78-81. Such phrases may not be helpful as they presume a state of affairs where artificial agents have some form of personality and therefore can act in the name of or on behalf of another person. It would seem more accurate to think of corporations that deploy artificial agents as potentially bearing responsibility for the software that they use.

introduction into the legal and social framework the idea of sharing responsibility with intelligent computer systems, and could also open doors to the creation of a new type of hybrid personality consisting of a human and software agent operating in tandem.[49] However, this proposal, which requires the agent-proxy companies to be legally subject to additional obligations such as disclosure of the agents' contracting terms and providing information about their functionalities, actually confirms the limitations of artificial agents: they are on their own unable to attest to the scope of their capacities and legal authority, and do so separate from their programmed actions.

10. A frequently-used justification for the "software agent as legal agent" theory is that artificial agents make "autonomous" decisions, which mirror those of human agents, such that "had the principal considered the matter herself, she *might* have reached a *different* decision (emphasis added)."[50] As noted, proponents choose this justification because of the minimized requirement of legal capacity for agents at common law. However, it is worth noting that this does not mean that an agent does *not* need *any* capacity to enter into a binding transaction. For the minor to act as an agent at common law, she must have *sufficient understanding to consent to the agency and to do the act required*.[51] The commentaries to the Third Restatement emphasize that the person/actor/agent's capacity to affect the legal relations of another (the principal and third party) is limited by the person's ability to take action, which is in turn a function of the person's *physical and mental ability as an individual*.[52] Thus, while a minor has the autonomy to enter into a binding contract online against her parent who had allowed the child Internet access (because the child had acted on the access granted to make purchases), the Third Restatement goes on to deny the computer itself a similar status. "The computer itself is … not P's agent, because it is *not a person*" (emphasis added).[53] The relevant section of the Third Restatement reads:

> *To be capable of acting as a principal or an agent, it is necessary to be a person, which in this respect requires capacity to be the holder of legal rights and the object of legal duties. … Accordingly, it is not possible for an inanimate object or a*

[49] ibid 79.
[50] Chopra and White (n 20) 370.
[51] *See Smally v Smally* (1700) 1 Eq.Ca.Abr. 6; *Watkins v Vince* (1818) 2 Starke 368; *Re D'Angibau* (1880) 15 Ch.D. 228 at 246; *Travelers Guarantee Co of Canada v Farajollahi* [2012] B.C.S.C. 1283.
[52] Third Restatement, § 3.05, Comment b.
[53] Third Restatement, § 3.05, Comment b, referring to § 1.04(5).

> *nonhuman animal to be a principal or an agent under the common law definition of agency.*[54]

11. Thus, the need to demonstrate autonomy and personhood as impediments to the theory of agency for artificial agents must be clearly acknowledged.[55] But starting around 2006, the development of deep neural networks modelled after biological neurons enabled their application of systems that approached human-level, and in some cases, exceeded human-level, performance on various tasks such as pattern and image recognition and language translation.[56] Another class of algorithms – evolutionary algorithms inspired by biological evolution – perform well in often complex problems, which do not make any assumptions about the possible solutions.[57] This has led to observations that these machines exhibit, not *automatic*, but *autonomous* or at least semi-autonomous behaviour,[58] which, in the case of Thaler v Commissioner of Patents, appears to have persuaded Beach J to find that the machine in question, DABUS, was the inventor of the patent sought.[59]

12. Beach J's reticence in deciding if the machine in question was autonomous[60] is actually understandable. Few details were given to substantiate the patentee Thaler's claim that his DABUS neural network was a "self-assembling" system that "mimic[s] aspects of human brain function … [in that] DABUS perceives and thinks like a person".[61] It is well accepted that artificial neural networks are nothing more than computational problem solving systems that require instructions defined by a human as to how to solve a problem, and are not truly "autonomous".[62] An examination of the earlier prototype to DABUS[63] and a recent paper published by Thaler[64] confirms this as it suggests that Thaler's scientific claims are more modest: his DABUS system is a new architecture (or topology) of multiple

[54] Third Restatement, § 1.04(5), Comment e.
[55] *See e.g.,* Sommer, Joseph, "Against Cyber-Law" (2000) 15 Berk. Tech. L.J. 1145, 1177-1178 ("[a] programmed machine is not a juridical person and therefore cannot be an agent… it is clearly a machine.").
[56] Wikipedia, Artificial neural network, <https://en.wikipedia.org/wiki/Artificial_neural_network>.
[57] Wikipedia, Evolutionary algorithm, <https://en.wikipedia.org/wiki/Evolutionary_algorithm>.
[58] *Thaler v Commissioner of Patents* [2021] FCA 879, [128]. Cf. Thaler v Comptroller General of Patents, Trade Marks and Designs [2021] EWCA Civ 1374 (holding that DABUS is not an inventor for different reasons).
[59] ibid [131].
[60] Ibid [18], [128], [131], [143].
[61] ibid [42].
[62] Daria Kim, "'AI-Generated Inventions': Time to Get the Record Straight?" (2020) 69(5) *GRUR International* 443. But see Thaler v Commissioner of Patents [2021] FCA 879, that held that a neural network system could be an "inventor", but the decision did not turn on whether the system was "autonomous".
[63] U.S. Patent 5,852,815 (Neural Network Based Prototyping System and Method).
[64] Thaler, Vast Topological Learning and Sentient AGI, (2021) 8 Journal of Artificial Intelligence and Consciousness 81.

deep neural networks with reinforcement learning elements that is capable of representing concepts and ideas.[65] Thus, DABUS is an example of a Weak AI implemented in artificial agents that merely "reflect a conscious and deliberate decision by *a human or humans to achieve a particular end*."[66] However much of a breakthrough DABUS may be, it is still, as Thaler admitted, a paradigm for scaling neural systems to Strong AI.[67] Electronic agents today, however sophisticated they are, are still examples of Weak AI: they are not capable of solving an arbitrarily wide variety of tasks, including novel ones.[68]

13. Hence outside of Strong AI, there is no true autonomous decision-making element within an artificial agent, without drawing on metaphysical and philosophical arguments that equate the models within artificial agents with the minds and wills of individuals.[69] This is because the Weak AI systems, without more, have no concepts of ethics, morality, legality and empathy.[70] If not explicitly programmed, an artificial agent will have no idea, for instance, as to the duty of fidelity expected of agents.[71] It will even undertake an illegal or void act on behalf of its principal[72] because it will reflect any biases that it is programmed to replicate, so long as it is not explicitly programmed to address them – as the COMPAS case illustrates.[73] It is indeed possible to program machines with individual aspects of

[65] ibid 108.
[66] Chen and Burgess (n 46) 88 (emphasis added). The authors describe a system which they referred to as "spontaneous intelligence", distinguished from artificial agents, which mirrors what others have referred to as Strong AI.
[67] Thaler, "Vast Topological Learning" (n 64) 109.
[68] Russell and Norvig (n 32) 981 (citing the distinction first made by John Searle in 1980).
[69] E.g., "Likewise, even if we accept that a software agent enjoys self-consciousness, it is not clear yet that achieving self-consciousness is a sufficient condition of legal personality". Dahiyat (n 18) 10.
[70] Concerns have been expressed by the absence of standards that define ethical design for AI and autonomous systems. *See* Russell and Norvig (n 32) 997. This has prompted the release of numerous AI ethical frameworks by governments and institutions around the world such as the EU Guidelines on Ethics in AI (2019), the Beijing AI Principles (2019), the Singapore AI Governance Framework v1 (2019), v2 (2020), the ACM Code of Ethics and Professional Conduct (2018), Microsoft AI principles (2018) and the Partnership on AI (Amazon, Google, Facebook, Microsoft, IBM) (2018). For a review of these frameworks, *See* Fjeld, Jessica et. al., "Principled Artificial Intelligence: Mapping Consensus in Ethical and Rights-based Approaches to Principles for AI." Berkman Klein Center for Internet & Society, 2020, at <http://nrs.harvard.edu/urn-3:HUL.InstRepos:42160420>.
[71] *UBS AG v Kommunale Wasserwerke Leipzig GmbH* [2017] 2 CLC 584, [91]. It is difficult to see how software code can owe fiduciary duties to its user or how such duties may be meaningfully exercised against software.
[72] *Cohen v Kittell* (1889) 22 QBD 680 (holding an agent not liable to a principal for omitting to carry out his illegal mandate).
[73] *Loomis v. Wisconsin* 881 N.W.2d 749. The defendant in this case contended that that his due process rights were infringed as he was prevented from challenging the scientific validity and accuracy of COMPAS. COMPAS is short for "Correctional Offender Management Profiling for Alternative Sanctions". It was used as a pretrial recidivism risk assessment tool by the courts in Wisconsin to determine if bail should be granted to the accused.

consciousness – awareness, self-awareness and attention – to make them "intelligent".[74] But this does not make them acquire true human consciousness – because they still lack "human subjective experience" – the ability to appreciate the subjective experience of others.[75]

14. The current state of the art is that any artificial agent will behave exclusively within the parameters of its programmed model(s), including those that are poorly programmed.[76] As long as the model exists for the agent – be it a simple reflex agent, a model-based agent, a goal-based agent, a utility-based agent or a hybrid of the above[77] – it is us as humans who set the performance measures for the models that in turn set the bounds within which the artificial agent behaves. This includes the possibility that an agent's response may be uncertain or unpredictable because the wrong performance measure was put into the agent, or where there was initial uncertainty about the true performance measure.[78] For instance, when the very first Ariane 5 rocket was launched, it flipped 90 degrees in the wrong direction 37 seconds after launch and was destroyed at a cost of approximately $370m. The reason for its explosion was initially said to be unknown.[79] But a detailed inquiry subsequently identified a rudimentary software bug that led to data overflow as the cause of the disaster.[80] Another, more recent example, can be found when AlphaGo, a sophisticated computer program that plays the board game Go, was engaged in its first competitive game against a 9-dan professional Go player, Lee Sedol.[81] In its second game of the five game series, it played move 37, which was remarked upon by commentators as "a very strange move … [we] thought it was a mistake".[82] It turned out that the move turned the course of the game and was subsequently described by commentators and researchers as "brilliant" and "beautiful" when they looked at both the code and the board setup in more detail.[83]

[74] Russell and Norvig (n 32) 986.
[75] ibid 986. If and when the machines actually do acquire this ability, they will be characterized as Strong AI, and will qualify as machines that are actually consciously thinking and not engaged in simulated thinking. ibid 985.
[76] The current state of the art is such that artificial agents can only be characterized as weak AI. ibid 981.
[77] Russell and Norvig (n 32) 49-58.
[78] ibid 39-40. This is so especially in environments that are partially observable, nondeterministic, dynamic, continuous and unknown, or involve other agents. ibid 46.
[79] *See e.g.,* CNN, Unmanned European rocket explodes on first flight, June 4, 1996, archived at <https://web.archive.org/web/20000819090542/http://www.cnn.com/WORLD/9606/04/rocket.explode/>.
[80] *See e.g.,* bugsnag, The Worst Computer Bugs in History: The Ariane 5 Disaster, Sep. 7, 2017, <https://www.bugsnag.com/blog/bug-day-ariane-5-disaster>.
[81] AlphaGo ran on a neural network and the original version of the program was the first computer Go program to beat a human professional Go player without handicap on a full-sized board. Wikipedia, AlphaGo, <https://en.wikipedia.org/wiki/AlphaGo> (last accessed 2 June 2021).
[82] *See* Wired, In Two Moves, AlphaGo and Lee Sedol Redefined the Future, Mar. 16, 2016, <https://www.wired.com/2016/03/two-moves-alphago-lee-sedol-redefined-future/>.
[83] ibid.

These are but two examples of how artificial agents are actually behaving according to human-defined parameters, however wrong, unpredictable or intelligent such behaviour may seem.

15. There are additional objections to recognizing artificial agents as legal persons in law. "To be a legal person is to be subject to legal rights and duties."[84] While human agents have a "res" or physical presence – against which legal norms such as holding assets and rights and duties, including rewards, constraints, sanctions, penalties and punishments may apply – there is no such physically referable entity with respect to artificial agents.[85] For instance, where an agent exceeds her authority, she breaches her warranty of authority to the third party.[86] It is hard to see how an artificial agent comprised of software that malfunctions and therefore ostensibly "acts outside its authority" can make a contractual promise to another, or how such a promise, if it exists, can be enforced against the software program. If such claims may be made against the principal on the basis that the principal is the owner of the software (assuming this to be the case) on which the artificial agent operates, it would seemingly make the principal both agent and principal.[87] If, on the other hand, the software is owned by a different party and the claim is brought against the developer, it would have the remarkable effect of making software developers effectively agents. In addition, there would be practical difficulties in separating artificial agents in hardware from those implemented in software[88] and registering them as entities as such.[89] This point is further amplified by the fact that artificial agents are themselves not monolithic entities, since each agent could itself be made up of other software agents and entities.[90] The multiagent environment could also be competitive, co-operative or even a mixture of both,[91] rendering it even more difficult to isolate and identify each discrete artificial agent as an entity.

Artificial Agents as Instruments or Tools

16. It is therefore self-evident that legislative intervention is needed to constitute the artificial agent as a legal agent with or without legal personality. This objection notwithstanding, it has been argued

[84] Smith B, "Legal personality", 37 Yale Law J 283–299 (1928).
[85] Chen and Burgess (n 46) 82-83.
[86] Tan, Cheng-Han, The Law of Agency (2nd ed, 2017), at 281.
[87] If ownership is a sufficient basis on which to bring such claims, perhaps because this gives rise to tortious liability, no recourse to agency would be necessary.
[88] *Artificial Intelligence and Law*, at 9. ***check citation***
[89] Dahiyat (n 18) 69-71. Dahiyat, for instance, talks about using digital signatures to identify artificial agents and confirm their integrity.
[90] Russell and Norvig (n 32) 44-45.
[91] ibid 45.

that an artificial agent should be a legal agent for reasons of expediency. The first is that rendering the artificial agent a legal agent allows for the principal to be bound by the responses of the agent.[92] The second is that the artificial agent as a legal agent enables the principal (user) of the artificial agent to be absolved of unplanned behavior emanating from the agent.[93]

17. We review these reasons in turn.

Enabling the Principal to be Bound by the Agent's Responses

18. As a justification for the "artificial agent as legal agent" theory, the first reason simply recites the raison d'état for agency, which is that the agent enjoys the power to create legal relations between the user as principal and third parties.[94] But if an artificial agent cannot be regarded as a legal agent, this *does not* mean that the user *cannot* be bound by the responses of the artificial agent. Where the artificial agent can be characterized as an instrumentality of the person who uses it, the agent shall be treated as an extension of such person. For instance, an owner who trains a dog to pick up beer from the neighborhood store in exchange for subsequent payment will be bound if his dog proceeds to do so without his prior direction.[95] The dog's actions bind his owner, not because it is a legal agent of the owner, but because the dog or any other animal serves as the principal's instrumentality.[96] Contrary to the theory of legal agency, the relevant act is carried out, not *on behalf of* the principal, but *by* the principal through the instrumentality. As the Commentaries to the Uniform Electronic Transactions Act (UETA) explain, with reference to the definition of "electronic agent":[97]

> This definition establishes that an electronic agent is a machine. As the term "electronic agent" has come to be recognized, it is limited to a tool function…
>
> An electronic agent, such as a computer program or other automated means employed by a person, is a tool of that person. As a general rule, the employer of a

[92] Chopra and White (n 20) 382.
[93] *See e.g.,* Chopra and White (n 43) 120-121.
[94] *See e.g., Scott v Davis* (2000) 204 CLR 333 at [227], [228] per Justice Gummow.
[95] Third Restatement, § 1.04(5), Comment e, Illustration 3. The illustration is based on *Commonwealth v. Tarrant*, 326 N.E.2d 710 (Mass.1975) (holding that a medium-sized German shepherd that accompanied the defendant into the victim's residence during a robbery was the defendant's dangerous weapon for purposes of armed-robbery).
[96] Id.
[97] S 2(6), UETA ("'Electronic agent' means a computer program or an electronic or other automated means used independently to initiate an action or respond to electronic records or performances in whole or in part, without review or action by an individual."). *See also* s 106(3), Federal Electronic Signature in Global and National Commerce Act, 15 U.S.C. § 7001; Singapore's Electronic Transactions Act, s 2(1).

> tool is responsible for the results obtained by the use of that tool since the tool has no independent volition of its own. However, an electronic agent, by definition, is capable within the parameters of its programming, of initiating, responding or interacting with other parties or their electronic agents once it has been activated by a party, without further attention of that party.[98]

19. Thus, section 14 of the UETA[99] and Article 12 of the UNCITRAL UN Convention on the Use of Electronic Communications in International Contracts[100] both proceed on the assumption that the artificial agent is an instrumentality of the principal and binds the principal as such, without reference to the need to subject the artificial agent to rights and obligations. "When machines are involved [as electronic agents for parties to a transaction], the requisite intention [for contract formation] flows from the programming and use of the machine."[101] In an affirmation of the artificial agent as instrumentality approach, the UNCITRAL secretariat made a similar observation as follows:

> Article 12 of the Electronic Communications Convention is an enabling provision and should not be misinterpreted as allowing for an automated message system or a computer to be made the subject of rights and obligations. Electronic communications that are generated automatically by message systems or computers without direct human intervention should be regarded as "originating" from the legal entity on behalf of which the message system or computer is operated.[102]

20. It is on a similar basis that many cases were resolved without controversy and without resorting to complex agency principles. In one of the earliest cases on this issue, the court in the 1933 U.S. case of *McCaughn v American Meter Co.* held that a device which automatically dispensed gas upon receiving a

[98] Uniform Electronic Transactions Act (1999) with Prefatory Note and Comments, Feb. 14, 2000, at 7-8, <https://www.uniformlaws.org/HigherLogic/System/DownloadDocumentFile.ashx?DocumentFileKey=4f718047-e765-b9d8-6875-f7a225d629a8&forceDialog=0>. The Committee did consider the following: "it is conceivable that, within the useful life of this Act, electronic agents may be created with the ability to act autonomously, and not just automatically. That is, through developments in artificial intelligence, a computer may be able to "learn through experience, modify the instructions in their own programs, and even devise new instructions." … If such developments occur, courts may construe the definition of electronic agent accordingly, in order to recognize such new capabilities." At 8 (hereinafter UETA Prefatory Note). However, as noted above, given the current state of technology, there is no change to the treatment of an artificial agent as a tool or instrument of the principal.

[99] UETA, s 14.

[100] UNCITRAL UN Convention on the Use of Electronic Communications in International Contracts, Art. 12.

[101] *UETA Prefatory Note*, at 37 (Commentaries for Section 14).

[102] UNCITRAL, Explanatory note by the UNCITRAL secretariat on the UN Convention on the Use of Electronic Communications in International Contracts (2007), [213] at 70.

coin deposit by a buyer – a vending machine – facilitated a contract to sell gas to the buyer "without any working human agency".[103] In England, in *Thornton v. Shoe Lane Parking*, a parking contract was held to be made with the garage owner when the ticketing machine at the entrance of the garage dispensed the parking ticket to the plaintiff.[104] And in the local case of *Chwee Kin Keong v. Digilandmall.com Pte Ltd* it was held that the plaintiffs had successfully contracted to purchase commercial laser printers (at low prices) when the defendant's website automatically processed the plaintiffs' orders and dispatched confirmation email notes to the plaintiffs' email accounts within minutes of the orders, although the contracts were subsequently vitiated for unilateral mistake.[105] Recently, in *Quoine Pte Ltd v B2C2 Ltd*, the Singapore Court of Appeal held that the respondent's trading software had successfully entered into a purchase of cryptocurrencies on the appellant's trading platform when certain trigger exchange rates were met on the platform, and that the respondent was entitled to retain the purchased cryptocurrencies even though the exchange rates were erroneous.[106] Although one of the judges dissented on the issue of whether there was an operative mistake that vitiated the contracts, the effect of the majority and minority judgements is that so long as users had full knowledge of what they were doing, they could choose to be bound by (a) the automated process under which contracts may arise and/or (b) agreements for which they had incomplete details (presumably as long as such details can be determined reasonably or by law[107]).

21. To support their argument of legal agency, Chopra and White argue that the artificial agent as instrumentality argument has led U.S. courts to hold that innate objects were not "capable of entering into consensual relationships", thus leading to "inconsistent outcomes".[108] They cited Wein who in turn referred to two cases: *Marsh v. American Locker Co.* for the proposition that because a coin-operated locker could not be said to "consent" to an assumption of bailment liability, the court had to discard the possibility that the transaction was a bailment,[109] and *Bernstein v. Northwestern Nat. Bank in Philadelphia* for the proposition that depositing a bag in a night depository did not create a debtor-creditor relationship with the bank as the "inanimate depository cannot provide the requisite act of

[103] 67 F.2d 148 (CA 3rd Cir, 1933).
[104] [1971] 2 Q.B. 163, 169.
[105] [2004] 2 SLR(R) 594, [2004] SGHC 71, [96] ("As most web merchants have automated software responses, they need to ensure that such automated responses correctly reflect their intentions from an objective perspective."); upheld on appeal in [2005] 1 SLR(R) 502, [2005] SGCA 2.
[106] *Quoine Pte Ltd v B2C2 Ltd* [2020] 2 SLR 20; [2020] SGCA(I) 2.
[107] E Peel, *The Law of Contract* (15th ed, 2020) para 2-087.
[108] Chopra and White (n 43) 22.
[109] Wein (n 40) 125-126.

conscious reception, and therefore is incapable of entering into a consensual relationship on behalf of the bank."[110]

22. It is trite law that an instrumentality cannot enter into contractual relationships on its own since it does not have legal personality. Nevertheless, the real issue is whether a person can be bound by the use of an instrumentality. As previously noted, there are clear authorities in the U.S., England and Singapore that support the artificial agent as instrumentality reasoning. Furthermore, a close reading of *Marsh* and *Bernstein* suggests that these cases actually *support* the instrumentality reasoning. In *Marsh*, the court was not satisfied that the coin-operated locker gave the defendant bailee exclusive control over the contents of the missing parcel (as compared with the plaintiff user) and thus denied bailment liability. This means that the court accepted that the coin-operated locker mechanism gave the alleged *bailor* exclusive control over its contents, without any human intervention. In fact, U.S. courts have not denied the user a bailment relationship with the bailee operating storage services because the services were automated (e.g. storage lockers, deposit boxes and garages). Instead, the cases turned on a factual finding of whether the machine providing bailment services afforded exclusive control to the bailor or the bailee.[111] In *Bernstein*, the court found *for* the plaintiff user on the basis that a bailment relationship *was* constituted by the deposit of the bag in the night depository,[112] but denied that a contract was formed, not because of the use of an inanimate depository, but because it characterized the depository deposit as an offer by the user that had to be subject to an additional and unequivocal act by the bank to create the status of debtor and creditor.[113] In any event, after *Bernstein*, the Pennsylvania Supreme Court ruled that the deposit of money in a night depository of the defendant bank *could* constitute a contract of debt with the bank.[114] The "inconsistencies" noted by Wein, Chopra and White turned not on the use of artificial agents as instruments to effect deposits, bailments or contracts, but on the proper characterization of the transactions as bailments, contracts, licences or as a duty of care in

[110] ibid 126.
[111] *See e.g., 1420 Park Road Parking, Inc. v. Consolidated Mut. Ins. Co.*, 168 A.2d 900 (D.C.Mun.App.1961); cf. *Scruggs v. Dennis*, 222 Tenn. 714, 440 S.W.2d 20 (1969).
[112] *Bernstein v. Northwestern Nat. Bank in Philadelphia*, 157 Pa.Super. 73, 77 (Pa. Super. Ct. 1945).
[113] ibid 75-76.
[114] *See Phillips Home Furnishings, Inc. v. Continental Bank*, 467 Pa. 43 (Pa. Super. Ct. 1976) (reversing and remanding on the point of assessment of the exculpatory clause in the bank's night depository agreement). Conversely, in *Employers Ins of Wausau v Chemical Bank,* 117 Misc.2d 601 (1983), the New York court found that the deposit in the night vault constituted a bailment relationship, and that a deposit agreement was only created when the bank opens the deposited contents and credit them to the depositor's account.

negligence. These "inconsistencies" eventually led other courts to decide liability by determining if the deposit service operator was under a duty to take reasonable steps to prevent harm to the contents.[115]

Liability with Automated Agents

23. Another frequently advanced justification for treating automated agents as "legal agents" is that doing so opens up another avenue of redress for malfunctioning agents: the "principal" or operator of such software agents may be held to be vicariously liable in tort for the actions of her agent.[116] Proponents of the "legal agents" theory assert that where decision making is delegated to an artificial agent and it would be inappropriate to assign the artificial agent moral culpability, making the software a "legal agent" enables the imposition of vicarious liability to be focused on the "human as the locus of liability … [where] we are less inclined to attribute the mischief to the machine [without which the human principal will be insulated from liability]".[117]

24. It is true that the law of vicarious liability has been broadened beyond the doctrine of *respondeat superior*. "In principle, liability in tort depends upon proof of a personal breach of duty. To that principle, there is at common law only one true exception – namely, vicarious liability."[118] From a practical standpoint, this means that for liability to reach the principal in vicarious liability, one has to constitute the artificial agent as a legal person and tortfeasor, and demonstrate two elements: a relationship between the principal and the tortfeasor which makes it proper for the law to make one pay for the fault of the other, and the connection between that relationship and the tortfeasor's wrongdoing.[119] In general, this arises in employment situations or where the relationship between the parties is one akin to employment.[120] In some exceptional circumstances, a principal who is not in an employment or employment-type relationship may incur vicarious liability for unauthorized acts of agents.[121] Leaving aside the difficulty of construing an artificial agent as an employee or agent, it is

[115] *See MyGlynn v. Parking Authority of City of Newark*, 86 NJ 551, 560 (S.C. of NJ, 1981); *Garlock v. Multiple Parking Services, Inc.*, 103 Misc.2d 943 (1980) (citing *Basso v Miller*, 40 N.Y.2d 233 (C.A. N.Y., 1976) (abolishing the distinction between licensees, trespassers and invitees, and applying foreseeable reasonable care under the circumstances as a measure of liability").
[116] *See e.g.,* Smed (n 17) 506; Chopra and White (n 43) 120-121. This is in addition to remedies for product liability claims, which are explored further in Chapter X.
[117] Wein (n 40) at 113.
[118] *Woodland v Essex County Council* [2013] UKSC 66, [3] per Lord Sumption, delivering the leading judgment.
[119] See e.g., *Various Claimants v Catholic Child Welfare Society* [2012] UKSC 56 [21] (*Christian Brothers*); *Cox v Ministry of Justice* [2016] UKSC 10, [15]; *Barclays Bank plc v Various Claimants* [2020] UKSC 13, [1].
[120] *Cox* (ibid); *Barclays Bank* (ibid) [27].
[121] *Christian Brothers* (n 119) [47]; *Cox* (ibid) [16]-[17]. *See also* Tan, Cheng-Han, "Vicarious Liability in the Law of Agency" (forthcoming in the Journal of Business Law).

unnecessary to rely on such analysis to hold a person responsible for wrongful acts arising from the use of artificial agents. If the instrumentality theory is accepted, the resolution of the issue of liability turns on negligence: whether the principal had taken reasonable care in the performance of functions entrusted to it, in so far as it performed those functions itself, through its employees, or, as explained above, using its deployed instrumentalities. The principal here is liable not for the failure of the artificial agent (as a separate legal entity) but for its own failure to exercise due care in selecting, testing, operating and monitoring its artificial agent.[122]

25. While this greatly simplifies the cause of action, legal agency proponents may contend that there may be difficulties in proving the *fault* of the principal in this direct cause of action in negligence.[123] And concerns may be made about the absence of well-established models or industry or technical standards to prescribe the requisite standards to be observed. Our response is simply that fault *with* the artificial agent has to be proved to maintain an action in negligence in either instance.[124] And the absence of industry practices does not mean that legal standards of care cannot be prescribed.[125] If the agents are dangerous, principals will owe duties to guard the dangerous machines to prevent injuries,[126] and persons who knowingly use or deal with such instrumentalities must be guarded, covered or protected.[127] Autonomous mobile robots may need to be equipped with passive safety features to prevent them from ever making contact with humans, or to warn humans of their approach, or even to cease their activities when they sense human presence.[128] Likewise, software designers and commercial vendors are negligently responsible for security vulnerabilities in their products and could be held liable for the harm caused by cyber criminals who exploit such

[122] EC, Liability for artificial intelligence and other emerging digital technologies (2019) 44. For examples in U.S. law, *See e.g.,* C Ralph Kinsey Jr., "Nondelegable Duty – Duty and Vicarious Liability for Negligence", 44 N.C. L. Rev. 242, 243 (1965).
[123] See e.g., Pagallo, *The Laws of Robots, Crimes, Contracts, and Torts*, (Springer, 2013) 124 (noting that '…the capacity of such machines to gain knowledge and skills from interaction with human caretakers, suggest that the fault would rarely fall on the designers, manufacturers or suppliers').
[124] Liability for AI (n 122) 46.
[125] ibid 23-24.
[126] *See Holbrook v. Prodomax Automation Ltd* 2019 WL 6840187 (discussing how manufacturer of robots could be held liable for death of technician servicing robots). *See also* the case of Elaine Herzberg, who was killed by a self-driving Uber car in 2018. Bernie Woodall, Uber avoids legal battle with family of autonomous vehicle victim, Reuters, Mar. 29, 2018, <https://www.reuters.com/article/us-autos-selfdriving-uber-settlement-idUSKBN1H5092>.
[127] Chopra and White (n 43) 125.
[128] Autonomous vehicles have been designed to come to a stop when sensing a human in front of their path: *See* <https://www.channelnewsasia.com/news/singapore/driverless-vehicles-safety-concern-testing-extended-12034946>; *See also* https://www.techbriefs.com/component/content/article/tb/stories/blog/37748.

vulnerabilities.[129] A manufacturer of an airplane autopilot system was implicated for negligence in the design of its system in an action by the victims against the airlines and their pilots, who entered the wrong coordinates for their destination airport that led to the eventual death of all on board when the plane crashed.[130] And manufacturers of autonomous vehicles have been the subject of actions in negligence for elements in their autopilot software that have led to the death of drivers or pedestrians.[131] Similarly, if the trading provider Quoine had not reversed the counterparty trades erroneously conducted by its software,[132] the counterparties could have brought a claim in negligence against Quoine for the misconfiguration of its platform which triggered the subsequent abnormal transactions.

26. An extension of this duty in negligence exists – where the duty of care is held to extend to *procuring* the careful performance of work delegated to others, who may be not only agents but also independent contractors. Termed the "non-delegable duty" of care, this is an application of the concept of assumption of responsibility to determine the scope of the duty, including whether the loss is economic or physical.[133] It is triggered when the claimant is vulnerable, where there exists a relationship between the claimant and principal by virtue of which the latter has a degree of protective custody over the former, and the subsequent delegation of that custody to another person.[134] It would apply to the principal's negligence in selecting, supervising, or otherwise controlling or failing to control that other person.[135] this could be pertinent, for instance, where a principal delegates its critical human in-the-loop operations to another, who is at fault in using an automated mechanism to discharge that duty.

[129] Shackleford et al, "*Toward a Global Cybersecurity Standard of Care?*" (2015) 50 Tex. Intl LJ 305; Chopra and White (n 43) 126.
[130] *In Re Air Crash Near Cali, Colombia on December 20, 1995*, 985 F. Supp. 1106 (S.D. Fla. 1997), reversed and remanded on appeal for further proceedings in *Piamba Cortes v. American Airlines, Inc.*, 1777 F.3d 1272 (11th Cir. 1999).
[131] In addition to the settlement against Uber, see *Umeda v Tesla Inc., Slip Copy* 2020 WL 5653496 for the recent lawsuit filed against Tesla for a pedestrian death in Japan. For an analysis of the incident, *see* Eliot, Tesla Lawsuit Over Autopilot-Engaged Pedestrian Death Could Disrupt Automated Driving Progress, Forbes, May 16, 2020, <https://www.forbes.com/sites/lanceeliot/2020/05/16/lawsuit-against-tesla-for-autopilot-engaged-pedestrian-death-could-disrupt-full-self-driving-progress/?sh=69e9a7c071f4>.
[132] *Quoine Pte Ltd v B2C2 Ltd* (n 106) [30].
[133] *See Woodland v Essex County Council* [2013] UKSC 66, [11]. *See also Commonwealth v Introvigne* (1982) 150 CLR 258; *New South Wales v Lepore* (2003) 212 CLR 511; § 7.03(1) of the Restatement (Third) on Agency. For its criticisms, see Glanville Williams, "Liability for Independent Contractors" (1956) CLJ 180; Anthony Gray, *Vicarious Liability: Critique and Reform* (Hart, 2018) ch 10.
[134] *See Woodland v Essex County Council* [2013] UKSC 66, [12], [23].
[135] § 7.03(1) of the Restatement (Third) on Agency Comment b.

27. It is accepted that negligence actions against artificial agents will involve issues of technical complexity and proof of causation, which may require costly expert analysis,[136] especially because the errors could be in the code, the training data, the design or architecture or even its operational safeguards, the interconnections between hardware and software, and interactions therein. However, these issues are neither unique nor insurmountable,[137] especially if they are supported by a robust application for discovery against the human principal as to the workings of the artificial agents.[138] Presumptions such as the res ipsa loquitur rule to place the burdens of producing evidence on the party in control of the evidence may also be usefully relied upon.[139] In any event, it should also not be assumed that these issues will be easier if agency analysis is used. Any question of whether an artificial agent was acting within authority is also likely to require understanding of the algorithms embedded in the artificial agent. In any case, discovery without any human intercession against an artificial agent is unlikely to be practical or conceivable.

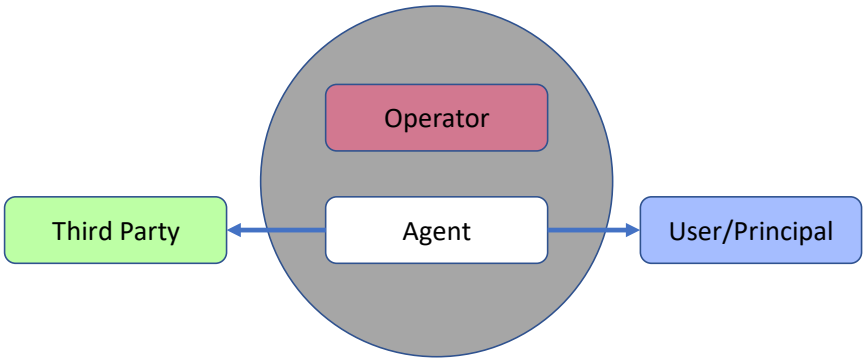

Figure 1: Figure from Chopra and White showing User as Principal (ebay.com example)[140]

28. Finally, Chopra and White cited ebay.com as an example where the risks of erroneous transactions will fall on the correct party where the end user, ostensibly as the principal, uses an

[136] *Liability for AI* (n 122122) 20-21.
[137] ibid 21-22. See e.g., the NTSB's investigation into the Uber accident that involved an autonomous vehicle in Arizona. NTSB, *Collision Between Vehicle Controlled by Developmental Automated Driving System and Pedestrian Tempe, Arizona March 18, 2018 HWY18MH010* (Nov. 19, 2019), <https://www.ntsb.gov/news/events/Documents/2019-HWY18MH010-BMG-abstract.pdf>.
[138] Daniel Seng and Stephen Mason, *AI and Evidence*, 33 Singapore Academy LJ 241 (2021) 274.
[139] Bryan Casey, Robot Ipsa Loquitur, (2019) 108 Georgetown LJ 225. Cf. Report on the Attribution of Civil Liability for Accidents Involving Autonomous Cars (2020) 47-48, Singapore Academy of Law, argued that the rule may be hard to apply to automated vehicles because it cannot be shown the defendant is in control of the situation to trigger the res ipsa loquitur rule. However, this fails to understand with detailed and automated logs, the locus of the component that caused the accident can be dissected and found and the res ipsa loquitur rule applied to the defendant in charge of *that* component. Casey, ibid 274-277.
[140] Chopra and White (n 43) 49.

artificial agent that is characterized as a legal agent that is operated by the company (ebay.com) to "instruct[] the agent to bid up to a specified maximum in an auction being conducted by the third party".[141] The authors characterized the errors as "specification errors" - "where the agent applies the rules the principal [erroneously] specifies",[142] "induction errors" – where the agent enters into a contract which was not authorized,[143] and "malfunction errors" – "which involve software or hardware problems whereby the principal's rules or parameters for the agent do not result in the intended outcome".[144] The authors then assert that the risk of specification errors would correctly fall on the user/principal, and the risk of induction and malfunction errors should initially rest with the user/principal or third party, who is in turn entitled to recovery against the operator for exercising the most control over the agent under the doctrine of *respondeat superior*.[145]

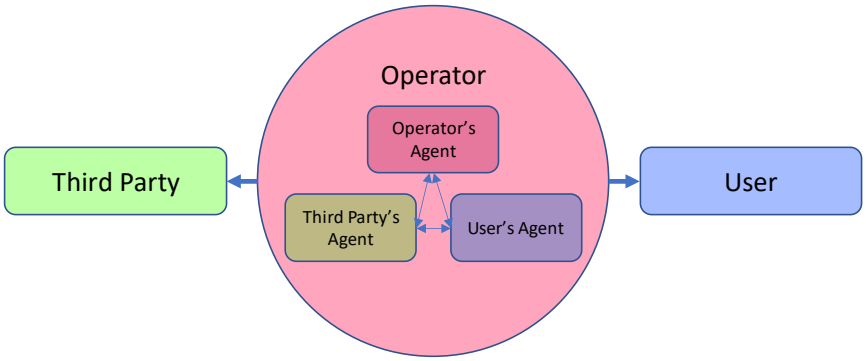

Figure 2: Modified figure from Chopra and White showing User, Operator and Third Party Agents in Multiagent Environment (ebay.com example)

29. Chopra and White contend that under the artificial agent as instrumentality theory, the user would be primarily liable for all three types of errors.[146] This result is only reached if one assumes that there is only *one* artificial agent operating in the circumstances. In reality, it would be more accurate, as Chopra and White themselves appear to concede,[147] that more than one agent is operating under the auspices of eBay. If one assumes that there is an artificial agent for the third party to conduct the auction, an artificial agent for the user to make bids at the auction, and an artificial agent for eBay to mediate and settle the bids and complete the transaction, there are actually *three* artificial agents, each

[141] ibid 48.
[142] ibid 46.
[143] ibid 49. At 46, the authors also describe this as "induct[ing] [into] a contract the principal *does* object to".
[144] ibid 46.
[145] ibid 49.
[146] ibid.
[147] ibid 44.

operating as the instrumentality of the respective party and interacting with one another, with eBay as the platform operator and overall controller for the artificial agents. In such a case, specification and induction errors will *prima facie* fall on the respective operators of the instrumentalities – be they the third party who erroneously described the item being auctioned or who failed to place restrictions on where the item being auctioned could be sold, or the user who erroneously programmed the wrong maximum bid price for the item or selected the wrong auctioneer. On the other hand, the risk of malfunction errors will clearly fall on eBay, since it is eBay who is the overall developer and operator for all the artificial agents. In this modified analysis, there is no need to resort to the characterization of the artificial agent as a legal agent to resolve issues of legal liability. Furthermore, this characterization correctly places the business risks on eBay to minimize all instances of malfunction errors, and puts eBay in the position to mediate between third party or the user as to their specification and induction errors.[148]

30. Therefore, in the overall analysis for tortious liability, the characterization of the artificial agent as a "legal agent" to enable a finding of legal liability appears unnecessary.

Conclusion

31. The law of agency underpins many commercial transactions and has facilitated the utility of legal entities such as the corporation. Increasingly, such entities are concluding transactions not through human intermediaries but through artificial agents. Many academic writers have thus postulated that artificial agents should be similarly ascribed the status of legal agents. However, the current state of AI technologies brings us nowhere near the concept of an autonomous artificial agent with the requisite level of rationality and intelligence to which we can accord the status of a legal agent. Furthermore, even if artificial agents achieve sentience and consciousness, unless legal personality is recognized or conferred, they cannot be legal agents. A piece of code, however sophisticated, cannot be a legal agent in the absence of recognition in law as a legal person. In addition, this recognition should be accompanied by some means to make such responsibility practically realizable. Corporations as legal persons achieve this through the requirement of a publicly disclosed share capital that is supported in certain circumstances by mandatory auditing, which enables a party dealing with a corporation to

[148] *See* eBay, User Agreement, <https://www.ebay.com/help/policies/member-behaviour-policies/user-agreement?id=4259#Returns> (accessed 29 June 2021); eBay, eBay Money Back Guarantee policy: Appeals and extensions, <https://www.ebay.com/help/policies/ebay-money-back-guarantee-policy/ebay-money-back-guarantee-policy?id=4210#section7> (accessed 29 June 2021).

decide whether to seek additional protection in the form of securities or assurances from third parties. In the absence of such a framework, many of the issues of contracting and liability in the immediate future can be resolved with reference to the treatment of artificial agents as instrumentalities of persons or legal entities. Such a treatment best accords with the functionalities of artificial agents and reposes the right duties and responsibilities on the human developers and operators of such artificial agents.

(9493 words with footnotes, excluding the abstract)