ARTIFICIAL INTELLIGENCE AND LAW: DO WE NEED A THOUGHTFUL RECONSIDERATION?

EMAD ABDEL RAHIM DAHIYAT[1]

DEPARTMENT OF PRIVATE LAW, UNITED ARAB EMIRATES UNIVERSITY,
AL-AIN, UAE

Artificial Intelligence ("AI") plays a significant role in improving our quality of life in different areas. However, the emergence of this technology may give rise to considerable difficulties at a conceptual level especially regarding the fit between law and technology. This paper thus explores the main doctrinal questions posed by the advent of AI in the areas of electronic commerce and autonomous vehicles. Furthermore, this paper critically examines the main approaches that have been advanced to deal with intelligent software programs. It also proposes the gradual approach as a way of overcoming the difficulties of such programs. It does so by adopting different standards of responsibility depending on whether the action is completed autonomously by an unattended software, or whether it is completed automatically by an attended software. This paper, however, is not intended to provide the definitive answer to all questions and challenges in this regard, but rather to identify the main components, and provide some perspectives on how to deal with such issues.

Keywords: software agent, liability, legal personality, autonomous vehicles, artificial intelligence

1. Professor of Computer Law, Chairman of Private Law Department, United Arab Emirates University. Corresponding author email: e.aldehyyat@uaeu.ac.ae.

INTRODUCTION

Artificial Intelligence ("AI") is no longer just an imaginary dream or science fiction, but a well-established fact with many successful applications that mimic or even surpass human intelligence. AI applications are capable of independent action rather than merely following instructions. They further exhibit high levels of autonomy, reasoning, and learning capacity through which they can take the initiative, interact with their environment, and even make decisions based on their self-modified instructions without any human intervention.

It is only a matter of time before commerce will lie mainly in the hands of software systems, which have recently begun to play a significant role in the negotiation, formation and execution of electronic contracts on the Internet. An obvious example of such systems is an intelligent software agent which operates without

human review or involvement. By operating in such a way, software agents raise questions as to what they are, what their real role in the contractual process is, how they differ from other programs, what their status should be in the eyes of the law, and how best to consider them.

Another example is in public transport, where AI technology allows the vehicle to perform all safety-critical driving functions and monitor the conditions on the road so that the human driver only needs to provide destination or navigation input. This is not referring to driver-assistance technology that merely contains features to avoid or mitigate collisions such as crash warning systems, adaptive cruise control, lane-keeping systems, and parking assistance systems. Instead, it refers to autonomous technology, which transfers driving tasks to the vehicle instead of the driver.

It is perhaps time to seriously consider fitting intelligent software technology into our legal and social structure, since there will be an increasing reliance on such technology in the near future. Over the next few years, many disputes and litigation may arise examining the validity of electronic contracts concluded via such technology, and determining the liability for intelligent programs' actions. Therefore, we are faced with an urgent need to find convincing answers to new questions derived from the emergence of such technology. Instead of traditional jurisprudential thinking, it is desirable now to think outside the box and address what should follow when software becomes capable of generating actions autonomously. Let us now investigate whether the advent of AI represents a real problem that demands a conceptual reconsideration of the legal nature of such technology, or whether it is simply a question of explanation that needs only some kind of clarification.

I. INTELLIGENT SOFTWARE AGENTS IN E-COMMERCE

The emergence of intelligent software agents has given rise to strong suspicions and debates on what such agents are, what their real role in the contractual process is, how they differ from other programs, what their status should be in the eyes of the law, and how best to consider them. This section will be divided into two parts: the first part will discuss the legal difficulties in the case of contracts which intelligent software agents concluded autonomously, while the second part will deal with practical problems arising with respect to liability issues. Studying contract and liability issues is not the ultimate purpose of this section.

Rather, the main goal is to explore a number of legal difficulties that accompany contracting through software agents, discover how such agents differ from other software applications, and how such differences are legally relevant.

A. Intelligent software agents and contracts

Although electronic contracts *per se* exhibit more similarities to their non-electronic counterparts than differences, the emergence of intelligent software agents—which operate not only automatically but autonomously—has generated questions as to whether agent-generated contracts are legally binding contracts, and whether the basic conditions of a valid contract are fulfilled in cases where such agents are used on one or both sides. Furthermore, the involvement of intelligent agents gives rise to many difficulties especially with regard to the attribution of liability for the actions of such agents.

The first condition for the creation of a valid contract is that there must be at least two parties, a promisor and a promisee, and the relevant parties involved in the transaction must have the legal capacity to undertake legally binding contracts.[2] Such capacity, however, is generally conferred only to natural or legal persons, and not to objects, devices or tools. Without the necessary juristic capacity and legal personality, an actor is not considered a distinct party to a contract and cannot be subject to liability for an act. It would be equally difficult to describe those who lack such capacity and personality as autonomous subjects, endowed with legal subjectivity and separate patrimony. Under the current law, computers and software applications are not legal persons and have no subjectivity in any legal sense. There are no special legal rules yet that recognize them as having the capacity to contract on their own or on behalf of others, or that provide them with separate patrimony to respond to any liability. This simply means that they have no juridical standing in the eyes of the law. Such being the case, computers cannot be viewed as distinct parties to a contract or even as agents of other involved parties, but only as the instrumentalities of the persons who use them. This implies that computers cannot be considered separately from their users, and that computer-based contracts can only be upheld on the grounds of the user's capacity.

Such conclusions pose no legal difficulties with regard to neutral software applications that only act according to their built-

2. JOHN CIBINIC JR. ET AL., FORMATION OF CONTRACT 364 (4th ed. 2011).

in knowledge and predetermined criteria without any degree of autonomy, intelligence, consciousness, and self-determination, and without any ability to deal with unprecedented situations. Examples include EDI systems, vending machines, mechanical parking attendants, automated teller machines, and other ordinary software programs. Such applications are unable to autonomously initiate offers or alter the terms and conditions of the transaction. This is why their steps are always predictable, predefined and repeated each time they operate, and this is also why they often perform very simple and routine tasks which users could, subject to time constraints, carry out themselves. Very often, human users review and control the functions, processes, and actions of such applications and have the opportunity to confirm or reject any transactions. In other words, one way or another, the decision to initiate an offer or to accept or reject an offer occurs through the exercise of human choice and discretion. In this regard, the user does not see what is going on, but he knows, or ought to know, what is happening or what is going to happen.

The transactional process of such applications usually involves minimal or no negotiation, and may even involve no mental interaction at all. This occurs because such applications do not exercise discretion or independent judgment and lack any ability to negotiate[3] or to make autonomous decisions and significant choices. With these applications, transactions often take place within a closed or semi-open environment, either between parties who are well known to each other—in which case offline negotiation takes place in advance,[4]—or between parties who may have had no previous contact. In this case, the interaction will only be at a physical level, not at a "mental" level.[5] However, the inability of such applications to interact at a "mental" level or to exercise discretion does not prevent the creation of a contract.[6] In all cases, the contractual capacity of such applications is superfluous to the

3. *See, e.g.*, Thornton v. Shoe Lane Parking Ltd [1971] 2 Q.B. 163 at 169 (articulating that the plaintiff may "protest to the machine, even swear at it" but, he says significantly, " . . . it will remain unmoved").

4. A good example here is EDI transactions, which occur under the guidance of so-called trading partner agreements entered into at the beginning of the commercial relationship.

5. Physical interaction involves usually very simple and repeated tasks such as distributing the goods when receiving the money (as in the case of vending machines), or opening the barrier and producing a ticket (as in the case of the car park's machine).

6. Such inability can have no effect on the validity of the contract. *See, e.g.*, McCaughn v. American Meter Co, 67 F. 2d. 148, 149 (3d Cir. 1933) (holding that a slot-vending machine which automatically fulfilled the functions of selling, product delivery and price collection was capable of concluding a contract even without the direct intervention and the continuous supervision of human users).

transaction and what is essential is that the relevant parties themselves have the requisite capacity. This implies that the software applications' activities are deemed to be those of the proprietor, and that the application is no more than a facilitator used to extend the reach of interaction between the contracting parties. Hence, the implication is that there is no need for the law to give separate consideration to such applications.

As mentioned above, such a conclusion is not controversial in either theory or application with respect to ordinary software programs, or even with respect to the first generation of electronic agents that display a very limited degree of autonomy, mobility, and intelligence. The position, however, may be different when so-called intelligent software agents are used since such agents exhibit such high levels of sophistication, autonomy, mobility, and intelligence that the application of the above conclusion is less convincing. Because of their cognitive, reactive and proactive processes, such agents do not only respond to specific pre-set situations within their built-in knowledge and rules, but are also capable of unplanned behavior and quick responses to novel and unexpected situations.

Unlike other software applications, intelligent software agents play a crucial role in electronic contract formation. Their functions extend from simple tasks, such as searching, comparing prices, recommending products or services, filtering news lists, etc., to the more complicated tasks, including brokering, representing buyers and sellers, buying and selling goods on behalf of users, engaging in auctions,[7] negotiating the price of and eventually effecting payment of goods,[8] etc. Furthermore, intelligent software agents are expected to be able in the near future to initiate contractual offers autonomously, alter the terms of the transaction, and even generate novel terms and conditions, some or all of which might not be contemplated by the human parties to the contract. In doing so, such agents do not only assist human users through the various stages of a contract, but they also proceed to the next stage and act

7. The advanced generation of intelligent software agents is even able to submit a bid in an auction, which requires some form of evaluation and judgment as to the optimal price to bid. Electronic auction systems nowadays are structured to interact with intelligent agent systems. A good example here is the worldwide E-Bay system. Examples can also be found at yahoo.com, where a bidding agent places bids on the user's behalf at the lowest possible increments. For further details, see TURBAN ET AL., ELECTRONIC COMMERCE 2004: A MANAGERIAL PERSPECTIVE, Ch. 11 (3rd ed. 2003).

8. For example, the MIT group has developed Tete-a-Tete (T@T) a program that enables electronic agents to negotiate not only price, but several other terms of a transaction, including warranties, shipping, service, returns, and payment options.

more like initiators or intermediaries than instruments or facilitators.

In fact, one point of legal significance in the intelligent agent technology is that one or both of the human parties to the contract will often have no conscious role in entering into, and sometimes performing the contract. With such modern technology, it is highly likely that human users will not know exactly what computing operations or strategies their software agents will perform, nor the features of the environments in which their software agents are going to operate in. This immediately raises the question of how can we deem the human user employing an agent a party to a contract, where the human user lacks personal knowledge of the existence of the communication and contract. Furthermore, how can we still regard such an agent as a simple automatic extension of that user?

At this point it may be helpful to monitor the implications of the use of agent technology and to differentiate between software agents according to their functions and their degree of autonomy, sophistication, mobility, and intelligence. It must, however, be borne in mind that several issues are unavoidably conflated in this regard, such as the ascription of intentionality or conscious phenomenal experience, the possibility of exercising free will by software agents, the issue of whether or not such agents need to understand contractual terms before the contract can be binding, and so on. In this debate, many approaches have been suggested in order to give intelligent software agents the required legal capacity to enter into enforceable contracts, such as ascribing legal personhood to an intelligent agent, establishing that such an agent is a legal agent acting on behalf of its user, or deciding that the agent is merely a communication tool for transmitting the user's consent.

When we deal with the issues of "subjectivity" and legal "capacity," it would perhaps be helpful to consider a need to grant these artifacts some kind of subjectivity. If so, the question becomes whether intelligent software agents possess the minimum requirements necessary for them to be granted a minimum level of legal capacity. This of course does not mean providing the agents with full subjectivity, but rather, it would mean providing them with at least a minimum level that would allow for the legal and economic advantages of limited liability and the continuation of legal capacity. It is also necessary to examine how intelligent software agents differ from other software applications and then to consider how such differences are legally relevant. This requires a detailed inquiry into the characteristics of such agents and their locations, roles, functions, and the extent to which human users

have control over their agents or other technical aspects of the contractual process. Only after doing so may we decide which legal rules need to be enacted, modified, or developed so as to address the emergence of intelligent software agents.

B. Intelligent software agents and liability

The absence of human review and the fact that one or both human parties to the contract will often have no conscious role in the electronic contracting process, can easily lead to errors in computer-mediated environments. Imagine, for example, that a software agent, in the course of searching and roaming the Internet, infringes the rights of others (such as a copyright or privacy right), performs illegal transactions, or operates without the user's authorization and sells rather than buys certain shares. Further suppose that this agent, while gathering information, corrupts a third party's database, or causes the server to crash. In these cases, who is held liable for the damage the agent caused? Should the law automatically deem the human user responsible for causing such harm?

As a general principle, legal responsibility is assigned when it appears that someone, in some way, was the actual cause of damage. The law is also likely to be invoked when control is not exercised or when no warning of the danger is given by those who could have controlled the outcome of events, forecasted the results, and warned of the dangerous outcome. Although there are several parties who might be potential perpetrators of a wrongful act (such as the programmer, a third party, or the supplier, etc.), most responsibility analyses focus on the human users or legal entities on behalf of which these systems are operated, and adopt the legal fiction that anything issued from these systems is considered issued by the natural or legal persons who use them.

Such responsibility analyses, which place full responsibility on the shoulders of the user, arouse little or no difficulties when they are applied to neutral software applications or the first generations of software agents that have very limited autonomy. With these applications, the software agents, subject to a set of tacit or complicit user-specified constraints, simply browse the Internet to complete an acceptable deal. In this case, the strategic directions, final purposes, contractual details, and general conditions of the agent's actions are all pre-determined by their users or programmers. It can thus be said that the software applications' abilities to understand and react will be limited more by the ideas, instructions, aspirations, and efforts expended by the creator or

user than by the state of the art.[9] This is why the applications' actions cannot be considered the source of new entitlements and liabilities, and why they should never be the scapegoats for blame taken away from the individual user. Under this position, any activity involving such applications would be construed as simple transmission. If things go wrong, the damage can often be traced to human mistakes either in programming or parameterization. It can then be argued that we should apply the same rules relating to the guardianship of things and objects, to such applications.

The matter, however, might take a different turn. The idea of custody might seem inappropriate with respect to the advanced generations of intelligent software agents where the harm is a function of many factors. These advanced generation agents have some degree of control over their internal states, and their actions are, to a certain extent, determined by their own experiences, which makes it so that no one can either know the full context of the program or forecast their behavior in all possible circumstances. Even programmers involved in the construction of such agents will be incapable of either writing instructions to handle all circumstances optimally or determining the pattern of their mechanism of action over the long term.

In many cases, it may be impossible or extremely difficult for a consumer who knows little about the unique nature of the intelligent agent to determine exactly where the fault lies, identify the source of the negligence responsible for the defect, and establish whether this negligence is in the design of the system, in the operation of the system, or in the output of the system. Even if such identification is possible, the undesired outcomes of such systems might not be due to a defect in the code or in the input values and configuration, but might be due to the peculiar nature of intelligent agents—which have the ability to operate autonomously, modify their own code, and even generate new instructions. In such cases, it would require a very imaginative approach to consider such agents as mere transmission tools such as telephones or fax machines, or to classify every error as a transmission error. On the other hand, making the software company liable for all software failures might discourage it from attempting new and innovative projects, and might encourage consumers to behave recklessly if they know that the company would be solely liable in all cases. If that were the case, software companies would likely go bankrupt.

9. James Raymond Davis, *On Self-Enforcing Contracts, The Right to Hack, and Willfully Ignorant Agents*, 13 BERKELEY TECH. L.J., 1145, 1148 (1998).

The position is further complicated by the fact that different persons have contributed different skills to the development of a software agent and the standard of care differs from one contributor to the next. The questions that need to be addressed in this regard are: what attitude should the law take towards dealings with automated devices that interact with us, and which, to some extent, function independently of their owners. Can the law hold a human user responsible for damages they could not have prevented or foreseen? Could an intelligent agent be responsible? And if so, how could it answer for the damages caused to other third parties? Questions such as these cannot sensibly be answered without an analysis of all the issues associated with Software Intelligence.

II. ARTIFICIAL INTELLIGENCE IN THE AREA OF AUTONOMOUS VEHICLES

Many automobile companies, including Tesla,[10] Volvo,[11] Uber,[12] Google,[13] and Mercedes-Benz,[14] have already developed and tested their own versions of driverless cars, and hope to commercially release a line of fully autonomous vehicles within a few years. According to a recent forecast from HIS Automotive, there will be nearly twenty-one million autonomous vehicles on the world's roads by 2035.[15] Let us now explore the main theories of liability and their potential application in the context of autonomous vehicles. We will also further discuss doctrinal questions posed by the advent of autonomous vehicle technologies,

10. *Full Self Driving Capability*, TESLA, https://www.tesla.com/autopilot (describing the self-driving features, hardware, and processing equipment available on all new Tesla models).

11. *Why Autonomous Cars?*, VOLVO, https://www.volvocars.com/en-kw/own/own-and-enjoy/autonomous-driving [https://perma.cc/E72D-5XJW] (predicting its first "unsupervised autonomous vehicles" will be available by the year 2021).

12. David Shepardson, *Uber Unveils Next-Generation Volvo Self-Driving Car*, REUTERS (June 12, 2019), https://www.reuters.com/article/us-uber-selfdriving/uber-to-unveil-next-generation-volvo-self-driving-car-idUSKCN1TD1GO [https://perma.cc/RGJ3-5C6B].

13. Alex Davies, *Google's Self-Driving Car Company Is Finally Here*, WIRED (Dec. 13, 2016), https://www.wired.com/2016/12/google-self-driving-car-waymo/ [https://perma.cc/Z53V-C5HU] (describing the evolution of Google's in-house autonomous car project and its transformation into the stand-alone venture, Waymo).

14. Clifford Atiyeh, *BMW and Mercedes-Benz Will Stop Fighting and Join Forces to Make a Better Autonomous Car*, CAR AND DRIVER (Feb. 28, 2019), https://www.caranddriver.com/news/a26572775/bmw-and-mercedes-benz-joint-autonomous-car/ [https://perma.cc/4BS2-BRCR] (explaining that both Mercedes-Benz and BMW predict they will have fully-autonomous vehicles on the road within six years).

15. *Same-Day Analysis*, IHS MARKIT (June 7, 2016), https://ihsmarkit.com/country-industry-forecasting.html?ID=10659115737 [https://perma.cc/2W2V-QYP6].

and investigate whether or not our traditional legal framework is still sufficient to realistically answer such questions.

A. Overview

It is crucial to note that self-driving vehicles will not necessarily display the same degree or level of sophistication, autonomy, and intelligence. While the fourth and fifth generations of self-driving vehicles are capable of driving safely and making proper decisions, first, second, and third generations of such vehicles exhibit a very limited level of sophistication and lack significant ability to make autonomous decisions according to their own experiences.[16] Nevertheless, it is expected that artificial intelligence technology will, in the very near future, progress to the next level, where vehicles will gain more intelligence and autonomy and be active initiators and decision makers rather than merely assistants or facilitators.

However, the transition from human-driven vehicles to fully autonomous vehicles will not be easy. On one hand, autonomous vehicle technology has not yet arrived at the perfect level of reliability and autonomy to guarantee a safe assignment of responsibility to them. Autonomous vehicles, like any other internet-enabled devices, are potentially vulnerable to cyberattacks from hackers. At the same time, autonomous vehicles suffer from the problem of "many hands." There are a number of different parties involved in the manufacturing, developing, designing, programming, and using of such advanced vehicles, and hence there are several parties who are potential perpetrators of a wrongful act. This can make it quite difficult to determine precisely to whom the wrongful act should be attributed. Is it the manufacturer of the vehicle, the software programmer, the network provider, or the owner of the vehicle? It is still unclear how the law will handle autonomous vehicles and their potential effects on public roads. Furthermore, it is questionable whether existing laws are sufficient to address questions of liability in the context of autonomous

16. *See generally* U.S. DEP'T OF TRANSP., NAT'L HIGHWAY TRAFFIC SAFETY ADMIN, FEDERAL AUTOMATED VEHICLES POLICY: ACCELERATING THE NEXT REVOLUTION IN ROADWAY SAFETY, (Sept. 2016), https://www.transportation.gov/sites/dot.gov/files/docs/AV%20policy%20guidance%20PDF.pdf [https://perma.cc/GF4Y-3PFV] (releasing the National Highway Traffic Safety Administration (NHTSA) classification system, which partitions vehicle automation into five levels, ranging from level 0, "no automation," to level 5, "full self-driving automation").

vehicles, especially as the technology behind such vehicles evolves or is commonly used on a larger scale.[17]

It seems then that the main challenge is striking a balance between consumer rights and expectations on the one hand, and the need to encourage innovation on the other. If the law fails to recognize the autonomous aspects of self-driving technology or if it attributes liability without contemplating the role of different involved parties, then we will not arrive at an efficient outcome. Moreover, placing full responsibility on one party might hinder the development of autonomous vehicles and induce others to behave recklessly on the ground that the other party will be held liable for any accident, whether or not they exercise an appropriate level of care. In contrast, providing absolute protection from liability may eliminate the incentives for manufacturers of autonomous vehicle technology to make their products as safe as possible to prevent liability.[18]

Despite the benefits of this socially desirable technology, the legal response to self-driving technology is still limited to the testing process. In the United Kingdom, the Department for Transport has developed a Code of Practice to specifically regulate the testing of driverless and automated vehicles on public roads.[19] According to the Code of Practice, responsibility for safely testing these vehicles always rests with those organizing the testing.[20] Published by the Department for Transport in February 2015, the Pathway to Driverless Cars report concluded that "[r]eal-world testing of automated technologies is possible in the UK today, [provided] a test driver is present, and takes responsibility for the safe operation of the vehicle; and that the vehicle can be used compatibly with road traffic law."[21]

In the United States, many states are considering regulations related to the testing and operation of autonomous vehicle

17. Jessica S. Brodsky, *Autonomous Vehicle Regulation: How an Uncertain Legal Landscape May Hit the Brakes on Self-Driving Cars*, 31 BERKELEY TECH. L.J. 851, 859 (2016).

18. Gary E. Marchant & Rachel A. Lindor, *The Coming Collision between Autonomous Vehicles and the Liability System*, 52 SANTA CLARA L. REV. 1321, 1337 (2012).

19. DEP'T FOR TRANSP., THE PATHWAY TO DRIVERLESS CARS: A CODE OF PRACTICE FOR TESTING (2015), https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/446316/pathway-driverless-cars.pdf [https://perma.cc/ZD9A-F67F].

20. *Id.* at 7.

21. DEP'T FOR TRANSP., THE PATHWAY TO DRIVERLESS CARS: SUMMARY REPORT AND ACTION PLAN 8 (2015).

technology.[22] Such regulations focus mainly on enabling testing of autonomous vehicles by defining autonomous vehicle technology, determining the conditions for the testing, and establishing minimum safety standards to govern the deployment of such vehicles.[23] Legislation in California and Nevada, for example, allow self-driving vehicles to operate on roads so long as a human driver is sitting behind the wheel on alert, and require the adoption of safety standards to ensure the safe operation and testing of such vehicles.[24] Other states, however, permit autonomous vehicles to operate without human operators in the vehicle.[25] Washington, for example, issued an executive order in June 2017 specifying certain requirements for vehicles operated without human operators in the vehicle.[26]

At the international level, a first step toward recognizing self-driving technology was attained on March 23, 2016 with the entry of amendments to the Vienna Convention on Road Traffic.[27] The amendments allow control of the vehicle to be transferred to the vehicle system, provided that this system does not breach the United Nations' vehicle regulations or can be overridden or disabled by the driver.[28]

B. Legal liability for autonomous vehicle accidents

Autonomous vehicles are posing new concerns regarding how liability should be attributed for deaths and injuries caused by such

22. *Autonomous Vehicles | Self-Driving Vehicles Enacted Legislation*, NAT'L CONF. OF ST. LEGISLATURES (Feb. 18, 2020), https://www.ncsl.org/research/transportation/autonomous-vehicles-self-driving-vehicles-enacted-legislation.aspx [https://perma.cc/Q4D3-HMDM] (outlining that Alabama, Colorado, Florida, Louisiana, Michigan, Nebraska, New York, Pennsylvania, Utah, and Washington have all passed legislation related to the testing or operation of self-driving vehicles).

23. *See* Jeremy Carp, *Autonomous Vehicles: Problems and Principles for Future Regulation*, 4 UNIV. PA. J. L. & PUB. AFF. 81, 97–98 (2018).

24. *See* 13 CAL. CODE REG. §§ 227-227.54 (2018); NEV. REV. STAT. § 482A.070 (2017).

25. *See, e.g.*, Brendan Farrington, *Florida Law Lets Autonomous Vehicles Drive Without Humans*, ABC 7 WWSB (June 13, 2019), https://www.mysuncoast.com/2019/06/13/florida-law-lets-autonomous-vehicles-drive-without-humans/ [https://perma.cc/6DF8-NVS5].

26. Wash. Exec. Order No. 17-02 (June 7, 2017), https://www.governor.wa.gov/sites/default/files/exe_order/17-02AutonomouVehicles.pdf [https://perma.cc/52BS-8X4Y].

27. Press Release, United Nations Economic Commission for Europe, UNECE Paves the Way for Automated Driving by Updating UN International Convention (Mar. 23, 2016), https://www.unece.org/info/media/presscurrent-press-h/transport/2016/unece-paves-the-way-for-automated-driving-by-updating-un-international-convention/doc.html [https://perma.cc/SL2A-4PLA].

28. *Id.*

vehicles, and who is to bear the risk associated with unintended consequences. With these kinds of vehicles, the undesired outcomes or accidents might not be due to a defect in the programming or parameterisation of the vehicles, but because of the intervention of third parties or the peculiar nature of self-driving vehicles, which have the ability to react autonomously to their environments.

In a conventional vehicle crash, the damage can often be traced to human mistake or vehicle malfunction or defect.[29] This is why most responsibility analyses for ordinary vehicles or semi-autonomous vehicles with limited autonomy focus on the human driver.[30] The matter, however, might take a different turn, and the idea of custody might seem inappropriate in respect to the fully autonomous vehicle, where the harm is a function of many factors.

Among key concerns is the question of how to ensure that an autonomous car is making ethical decisions about how and when to take actions to avoid accidents. For instance, if a vehicle has the choice to either hit an elderly person, a dog, a pregnant woman, a child, a group of pedestrians or crash into a wall injuring or possibly killing its passengers, how can it make that split-second decision to avoid endangering human lives? Such debate shows the need to enact rules for algorithmic transparency and accountability. This is perhaps why the Association for Computing Machinery issued in 2017 a "Statement on Algorithmic Transparency and Accountability,"[31] and the Institute of Electrical and Electronics Engineers released at the end of 2016 a draft for public discussion

29. *See* U.S. DEP'T OF TRANSP., NAT'L HIGHWAY TRAFFIC SAFETY ADMIN., DOT HS 812 115, CRITICAL REASONS FOR CRASHES INVESTIGATED IN THE NATIONAL MOTOR VEHICLE CRASH CAUSATION SURVEY (2015) (outlining that in study of motor vehicle crashes, the study attributed 94 percent of car accidents caused by drivers' failures).

30. *Man Killed in Tesla 'Autopilot' Crash got Numerous Warnings: Report*, CNBC (June 20, 2017, 7:25 AM), https://www.cnbc.com/2017/06/20/man-killed-in-tesla-autopilot-crash-got-numerous-warnings-report.html [https://perma.cc/5993-6CL7] (last updated June 20, 2017, 10:10 AM) (in 2016, a man was killed in Florida when his 2015 Tesla Model S car with the Autopilot system crashed into a truck that was crossing the road in front of his car. Tesla Company has attempted to apportion blame onto the driver by arguing that the system provided the driver with sufficient messages during the trip warning him to put his hands on the wheel. The official investigation found that there was no defect in the Autopilot system and that the driver was not paying attention to the road).

31. Ass'n for Computing Machinery Pub. Policy Council, Statement on Algorithmic Transparency and Accountability (Jan. 12, 2017), https://www.acm.org/articles/bulletins/2017/january/usacm-statement-algorithmic-accountability [https://perma.cc/P8SS-M7KN].

titled "Ethically Aligned Design," emphasizing the need for accountability to avoid confusion or fear within the general public.[32]

From a legal perspective, autonomous vehicles are not yet recognized as legal persons able to be held liable for their actions or to be sued separately from their owners.[33] This implies that injuries stemming from their accidents would automatically be attributed to the natural or legal persons who utilize them, even if these vehicles were operating autonomously beyond the control of such persons. As long as self-driving vehicles continue to be merely sophisticated automata, their roles will not be considered separately, and such vehicles will continue to be governed by the legal discipline which deals with objects, and not by that which deals with subjects. Courts around the world have a tradition of attributing the actions of automated technologies to the user, and considering the user as the locus of liability, even if he or she is unaware of the operations of the automated machines.[34] This attitude, however, can be criticized not only because it ignores the active role and autonomous ability of some intelligent software applications, but also because it has the potential to produce harsh results on the users who would be bound to any actions or mistakes made by their software machines. Holding someone responsible is unfair as long as the individual has not caused harm, nor could have prevented or foreseen it. It is also unreasonable to expect a person, who is unable to appreciate the extent of a risk or to take steps to avoid its occurrence, to control the uncontrollable.

The question of liability in the case of using autonomous vehicles is far from simple. As self-driving technologies continue to develop and gain self-awareness, the legal claims related to these technologies will also rise in the not too distant future. This suggests that a different rule of attribution may need to be developed to contemplate both the human actor and the autonomous program. It may also be necessary for the law to

32. *See* INST. OF ELECTRICAL AND ELECTRONICS ENGINEERS, ETHICALLY ALIGNED DESIGN 8 (2016), http://cn.ieee.org/files/EAD_FINAL.pdf [https://perma.cc/UX5N-7M8H].

33. Shawn Bayern, *The Implications of Modern Business-Entity Law for the Regulation of Autonomous Systems*, 19 STAN. TECH. L. REV. 93, 95 (2015).

34. *See* Brouse v. United States, 83 F. Supp. 373, 374 (N.D. Ohio 1949) (attributing error to the pilot rather than the design of the autopilot feature of the plane, with the judge opining that, "[t]he obligation of those in charge of a plane under robot control to keep a proper and constant lookout is unavoidable."); *see also* Register.com, Inc. v. Verio, Inc., 356 F.3d 393 (2d Cir. 2004) (holding Verio liable for breach of contract because of the actions of the search robot); *see also* State Farm Mut. Auto. Ins. v. Bockhorst, 453 F.2d 533, 536–37 (10th Cir. 1972) (ruling that the insurance company was bound by the mistaken actions of its computer).

seriously consider the role every party plays in producing the action in question, and then use different standards of responsibility depending on whether an unattended software completes the action autonomously, or whether an attended software completes the action automatically.

We can distinguish three basic types of liability rules that might be applied in the case of damages stemming from accidents involving self-driving cars:

Strict liability: Strict liability is usually applied in connection with dangerous products and ultra-hazardous activities.[35] It is based on the relationship between risk and confidence and on the assumption that whoever uses such products would quite simply bear the risk of their mistakes and pay the cost of confidence in the binding nature of their outcomes.[36] According to this rule, liability will be attributed whenever damage occurs without having to demonstrate fault or address whether the damage could have been expected or avoided. This implies that liability could be allocated to the owners in light of the risks they accept through using the autonomous vehicle, regardless of whether the vehicle's operations were intended, authorized, or controllable, and also regardless of whether the owner was physically present in the vehicle at the time of the accident. Similarly, a vehicle manufacturer can be held responsible if the vehicle malfunctions or departs from its intended design, even if the manufacturer exercised all possible care in the preparation, marketing, and sale of the vehicle.

In order to avoid the undesired effects of strict liability and balance the various interests of involved parties, this rule is accompanied by the principles of reasonableness and good faith which may play a role in ensuring that such a rule will only be applied to the extent necessary to protect society without any extensiveness or exaggeration.[37] It is, however, important to recognize that a software vehicle's dangerousness does not reveal itself clearly, and hence it will be unfair and even commercially unreasonable, to hold the manufacturer or owner bound by

35. RESTATEMENT (SECOND) OF TORTS §§ 519–524A(1) (AM. LAW INST. 1977) ("One who carries on an abnormally dangerous activity is subject to liability for harm to the person, land or chattels of another resulting from the activity, although he has exercised the utmost care to prevent the harm.").

36. *Id.* at cmt. d.

37. Some courts have used the reasoning of a good faith defense indirectly without identifying that defense or its parameters. *See, e.g.*, State v. Williams, 115 N.E.2d 36**,** 42 (Ohio Ct. App. 1952) (truck driver with cargo of undersized fish was not guilty if he did not know about the cargo and it would have been impracticable and unreasonable to require an inspection). For more information, see Laurie L. Levenson, *Good Faith Defenses: Reshaping Strict Liability Crimes*, 78 CORNELL L. REV. 401 (1993).

unexpected reactions just because it was theoretically, or remotely, possible that the vehicle would produce them. This may deter people from purchasing autonomous vehicles and discourage companies from developing self-driving technologies.

While the application of this strict rule looks acceptable and convincing in relation to dangerous products that have the possibility to produce serious physical injuries or death, the same cannot be said with regard to self-driving vehicles that are mainly used to reduce deaths and injuries from road accidents. In fact, the application of strict liability in the context of autonomous vehicles can be subjected to many criticisms. First, the rule fails to recognize the inherent characteristics of such vehicles, which may act beyond our control. Instead of accommodating such characteristics (such as intelligence, autonomy, reactivity, proactivity, adaptivity, etc.), this rule deals with these vehicles as if they are mere conventional cars that human drivers are able to fully control whenever they want. Second, this rule places a heavy burden on the shoulders of users by making them liable for any damage caused by their vehicles, whether or not they exercised an appropriate level of care—resulting in an unwillingness to use self-driving technology. This rule can be justified if such vehicles exist in a vacuum in which the user is the only director and instructor and only the user plays a role in forming a vehicle's reaction. But is that the reality of driverless vehicles?

Negligence liability: Negligence liability may be defined as the failure to do something that a reasonably prudent person would have done under the same circumstances, or doing something that a prudent and reasonable person would not have done under the same or similar circumstances.[38] From this definition, it is clear that the standard of care which is used to determine whether or not there is negligence is that of the hypothetical reasonable person. Whether or not the owner of an autonomous vehicle has reached the required standard in any given case is a question of fact, whose answer depends greatly on what a reasonable person would have done under similar circumstances. According to this rule, the owner is only liable when the damage occurred due to his negligence and failure to exercise a reasonable level of care in using and controlling

38. *See* Blyth v. Birmingham Waterworks Co. (1856) 11 Exch 781, 156 ER 1047 (Alderson J) (defining negligence as "the omission to do something which a reasonable man, guided upon those considerations which ordinarily regulate the conduct of human affairs would do, or doing something which a prudent and reasonable man would not do").

his vehicle.[39] Unless the injured party demonstrates the existence of negligence on the part of the owner, he will not be compensated and the owner will then be absolved of the liability.[40] Similarly, the manufacturer is only liable when the accident occurs in an autonomous vehicle as a result of a flaw or shortcoming in its systems[41] or because of the failure to provide adequate instructions or warnings.[42] The manufacturer can also be held liable if a product fails to be of sufficient quality or if the manufacturer provides misleading information about the vehicle's capabilities.[43]

When human drivers have no direct control over the functioning of the fully autonomous vehicle, it is highly likely that autonomous vehicle technologies will shift the responsibility from the driver to the vehicle manufacturer. The application of negligence liability in cases involving fully autonomous vehicles can, however, be criticized because it is sometimes extremely

39. This may apply if the owner was careless in choosing or instructing his vehicle, if he used his vehicle in an inappropriate environment which clearly conflicted with the instructions of use, if he negligently failed to control the vehicle or was not diligent or serious in eliminating the unsuitable circumstances which increased the likelihood of the occurrence of damage.

40. Proving negligence is required in most accident claims. However, it is difficult for victim's to prove causation and obtain the evidence needed to establish liability. This implies that the defendant can simply escape liability by demonstrating that he or she has not violated the legal standard of care. *See, e.g.*, Ford Motor Co. v. Rushford, 868 N.E.2d 806, 810 (Ind. 2007) ("To prevail on a claim of negligence, a plaintiff is required to prove: (1) a duty owed by the defendant to the plaintiff; (2) a breach of that duty by the defendant; and (3) an injury to the plaintiff proximately caused by the breach."); *see also* Ryan Abbot, *The Reasonable Computer: Disrupting the Paradigm of Tort Liability*, 86 GEO. WASH. L. REV. 1, 12 (2018).

41. Colon Ex Rel. Molina v. Bic USA, Inc., 199 F. Supp. 2d 53, 85 (S.D.N.Y. 2001) ("[T]he plaintiff must show that a specific product unit was defective as a result of 'some mishap in the manufacturing process itself, improper workmanship, or because defective materials were used in construction,' and that the defect was the cause of plaintiff's injury.").

42. RESTATEMENT (THIRD) OF PRODUCTS LIABILITY § (c) (AM. LAW INST. 1998). There is, however, no clear consensus on when warnings are required. Nor is there a clear test as to when existing warnings are adequate. NIDHI KALRA, JAMES ANDERSON & MARTIN WACHS, LIABILITY AND REGULATION OF AUTONOMOUS VEHICLE TECHNOLOGIES 31 (2009).

43. Manufacturers could face misrepresentation or false advertising claims if they provide false or misleading impressions about their vehicles' capabilities, which consumers reasonably rely on, leading to unrealistic expectations and harm. However, there is still a difficulty in determining what is misleading information and what is not. The terms such as such as "full self-driving capability," "Autopilot," and "ProPilot" can be constructed and interpreted broadly, and the sophistication degree of autonomous vehicles is still not completely clear. For instance, in 2016 the German transport minister asked Tesla Motors to refrain from using the term "autopilot" in its advertising, as doing so may mislead drivers into believing that that they do not need to pay attention to the road. But Tesla defended its use of the term, arguing that it should be understood by its analogy to airplanes. *See* Alex Hern, *Germany Calls on Tesla to Drop Autopilot Branding*, THE GUARDIAN (Oct. 17, 2016), https://www.theguardian.com/technology/2016/oct/17/germany-calls-on-tesla-to-drop-autopilot-branding [https://perma.cc/V5QV-CK8N].

difficult, if not impossible, for the injured party who might know little to nothing about the peculiar complexity of AI, to prove and identify the source of negligence. By following this rule, we might then reach the position in which no party would be legally responsible for incidents involving autonomous vehicles, and the injured party would thus become entirely unprotected. The position is further complicated by the fact that different persons contribute different skills to the development of self-driving technology and that the standard of care differs from one contributor to the next. It is then difficult, in most circumstances, to show a chain of causation back to the programmer, owner, or any other party involved.[44] Furthermore, this rule can also be criticized because it has the potential to produce different conclusions regarding similar facts and situations. This is because this rule does not determine specific degrees or levels of negligence, but leaves such issues to be determined by reasonableness standards which differ from jurisdiction to jurisdiction. By doing so, this rule not only creates a kind of uncertainty and insecurity among the involved parties, but also unduly increases legal costs.

Vicarious liability: The most common example of vicarious liability is when an employer is held liable for the torts and acts of employees who commit these torts or acts within the course and scope of their employment. Under vicarious liability, the employer might not be liable for the acts of the employee, if these acts are wholly unconnected with the course of employment.[45] This implies that an employee in some cases might be responsible and sued for acts alleged to have occurred outside the course and scope of employment. In such cases, the employee will be asked not only to answer for the damage caused to other third parties, but also to meet the demands of the employer, who might claim contribution in order to recover compensation for the injured party. The fact that an employee might exclusively be held liable to pay damages, raises the question of whether it is still possible to apply the principles of

44. It may be difficult to prove this chain of causation because autonomous vehicles are not programmed to justify and explain their decisions and actions in detail. It could also be because the vehicle algorithm is somewhat inexplicable or not likely to be fully understood by the human user. What makes the matter further complicated is the fact that autonomous vehicles might act according to their own experience and self-created instructions and then commit mistakes independently of those who use them. For more information, see Jason Millar & Ian Kerr, *Delegation, Relinquishment, and Responsibility: The Prospect of Expert Robots*, *in* ROBOT LAW 126 (1st ed. 2016).

45. *See* Beard v. London General Omnibus Co [1900] 2 QB 530; *see also* Twine v. Bean's Express Ltd. [1946] 1 All ER 202.

vicarious liability in the case of accidents caused by autonomous vehicles.

There are still substantial differences between human employees and self-driving vehicles which might prevent an analogy from being drawn between them for the purpose of applying the principles of vicarious liability. A human employee enjoys legal personality and juristic capacity, and is employed under an agreed upon contract of service, in consideration of a wage or other remuneration.[46] An autonomous vehicle, on the other hand, lacks such legal personality or capacity, which enables it to contract on its own or to provide a required consent to any contract of service. Moreover, unlike human employees who have separate patrimonies distinct from their employers, the vehicle by itself has no patrimony or personal assets and thus it cannot compensate the victim for the damage it causes. This means that any liability will practically fall back on the user of the vehicle whether or not the user authorized the acts of the vehicle or whether the acts occurred in the course of the user's businesses. If that is the case, does it still make sense to consider the application of vicarious liability in the cases of accidents caused by autonomous vehicles?

The law, as it stands now, does not comprehensively regulate autonomous vehicles, and still treats them like other automated machines, ordinary tools, or other conventional vehicles.[47] By following this approach, the law fails to take into consideration the autonomous aspects of such vehicles, nor does it seriously consider

46. *See* Ready Mixed Concrete Ltd v. Minister of Pensions and National Insurance [1968] 2 QB 497 at 498 (holding that three conditions must be fulfilled for a contract of service to exist: first, the servant agrees, in consideration of a wage or other remuneration, to provide his own work and skill in the performance of some service for his master; secondly, he agrees, expressly or impliedly, that in the performance of that service he will be subject to the other's control in a sufficient degree to make that other master; thirdly, the other provisions of the contract are consistent with its being a contract of service).

47. The legal response to the advent of an autonomous vehicle is still limited to the testing, registration, and licensing of the vehicle and is not qualitative enough to accommodate the innovative aspects of its technology. However, the lack of technology-specific regulations does not imply that autonomous vehicles are unregulated at the state and federal levels. In fact, autonomous vehicles are subject to the same rules as traditional vehicles. *See, e.g.*, NAT'L HIGHWAY TRAFFIC SAFETY ADMIN., FEDERAL AUTOMATED VEHICLES POLICY 5 (2016) [hereinafter NHTSA POLICY 2016] ("[I]f a vehicle is compliant within the existing FMVSS regulatory framework and maintains a conventional vehicle design, there is currently no specific federal legal barrier to . . . being offered for sale."); *see also* CONN. GEN. STAT. § 13a-260(d)(1); D.C. CODE § 50-2352(2); Exec. Order No. 572 § 4 (Mass. 2016) (requiring that a natural person holding a valid license be seated in the control seat of the vehicle while in operation [and be] prepared to take immediate control of the vehicle if necessary). For more information, see Jeremy Carp, *supra* note 23, at 92–97.

their function, characteristics, or even the possible parties and factors involved in autonomous vehicle crashes. It seems that the current law is still based on the principle of technology neutrality, which is no longer accurate.[48]

This suggests that a different rule of attribution may need to be developed to address the emergence of self-driving technology in a manner that strikes a balance between the interests of consumers and producers of autonomous vehicles, as well as the practical and ethical considerations of such technology. This necessitates the full contemplation of the role of different parties and explicit recognition of machine-made mistakes. It would be useful if the law examined the extent to which the human driver has knowledge, accessibility, and control over the actions of the vehicle system, and whether or not the driver participated in some way in determining or choosing the degree of autonomy, mobility, and intelligence of that system. When attributing liability, the law must also differentiate between semi-autonomous vehicles that act automatically within the realm of human control, and fully autonomous vehicles that operate outside the realm of human control. This implies that we should not always hold the human driver responsible for the actions of the vehicle regardless of the circumstances of the accident, but we should instead consider and examine the role of other parties such as the manufacturers, road designers, service centers, pedestrians, programmers, Internet providers, etc.

In an environment where many parties are unwilling to take full responsibility for the actions of their intelligent software, or are unable to account for the unexpected events that happen on a road trip, insurance companies might play a role in making the distribution of liability more realistic and smooth. However, such a role or task will not be easy. On the one hand, insuring the risk

48. The principle of technological neutrality is intended to provide functionally equivalent treatment of electronic and traditional situations, irrespective of the medium used. This principle is based on the grounds that regulations should not be drafted in technological silos, and that technologies are not just 'tools' standing ready to serve the purposes of their users. That is to say, law should be independent of any particular technology, neither favoring nor discriminating against specific technologies as they emerge and evolve. Therefore, the fundamental rules should be the same online as off-line, and the same regulatory principles should apply regardless of the technology used. However, the application of the principle of technology neutrality might not always be desirable, especially when the matter relates to AI technology, which allows software to operate not only according to its user's instructions, but also according to its own experience and self-modified or self-created instructions. For further discussion about this principle, see Brad Greenberg, *Rethinking Technology Neutrality*, 100 MINN. L. REV. 1495 (2016); Winston Maxwell & Marc Bourreau, *Technology Neutrality in Internet, Telecoms and Data Protection Regulation*, 1 COMPUTER AND TELECOMMS. L. REV. (2014).

posed by the use of autonomous vehicles still faces difficulties in checking, assessing, or analyzing the operations of such vehicles. On the other hand, creating an insurance scheme or building a control system may be too expensive to justify seeking insurance for the risk posed by using self-driving technologies. In order to overcome this difficulty, better coordination and cooperation amongst all parties involved is essential.

Unless we understand the "why" of the technical failure, and unless we recognize the commercial demands of different business models, we cannot provide an appropriate legal framework to cope with the rapid pace of change and development in AI technologies. This implies that the law should take the lead and intervene in the early stages, rather than wait for the technological outcomes and react once society has been affected. This, however, necessitates further insight into the software development process so that the question of who is going to decide how an autonomous vehicle will operate and consequently what users will be able to do, will be discussed collectively by all the relevant parties of society involved in such technology (legislators, consumers, programmers, businesses). On the other hand, it is essential that the product development process take legal requirements into account during the initial stages of product design so that resulting outcomes reflect the existing legal framework pertaining to traffic safety.

Let us now examine the main legal approaches that have been advanced to deal with AI technology. Following this examination, the section will propose the gradual approach as a way of overcoming the difficulties of such technology by adopting different standards of responsibility depending on whether an unattended software completes the action autonomously, or whether an attended software completes the action automatically.

III. TOWARD A LEGAL DEFINITION OF SOFTWARE AGENTS: IS A CONCEPTUAL RETHINK IMPERATIVE?

This section discusses the main theories proposed to deal with intelligent software applications and their increasing role in our lives. As we will see, such theories can be classified into two main categories. The first category includes radical solutions that propose ascribing legal personality or some kind of subjectivity to software agents and treating them the same as human agents, while the second category includes the traditional solutions that consider software agents mere passive communication tools. Let us now discuss such theories and suggestions, and see whether or not there are any alternative mechanisms that might offer a convincing

answer to the questions generated by the emergence of intelligent software agents.

A. Ascribing legal personality to software agents

The first solution is to recognize electronic agents as legal persons and develop a theory of liability on that basis. This solution has been suggested by some scholars who consider that conferring legal personality to software agents brings with it the advantages of limited liability and the continuation of legal capacity—especially when such agents are self-modifying and acting according to their own experience, or when they have an autonomous capacity to act in some extra-legal manner.[49] By following this approach, computers would be subject to liability for their actions, to some extent, just as a natural person would, and would not be endowed with an unlimited power to bind their users. Moreover, this approach deals with an intelligent agent as a legal person who is capable of entering into contracts either as a principal or as an agent, and does not deviate from the traditional principles of contract law that require the will of a person and not of the artifact.[50]

This approach, however, leaves many questions unanswered. If we consider a software agent a legal person, then the creditors of the software agent could only sue that agent in order to receive compensation. Are we really ready for such a scenario? Are electronic agents currently autonomous enough to deserve and warrant legal rights? Which kind of responsibility and contractual commitment will be supported by the electronic agent itself if something goes wrong? Is it possible to consider a software agent as acting in good or bad faith? Can we consider this agent and its user as separate and distinct parties? At what point or degree of autonomy and sophistication would an agent acquire personality? The answers to the above questions are not straightforward. Several philosophical, legal, and technical issues are—unavoidably—conflated in this regard.

Conferring legal personality on software agents also raises the problem of how difficult it can be to identify the agent and

49. *See generally* Francisco Andrade et al., *Contracting Agents: Legal Personality and Representation*, 15 ARTIFICIAL INTELLIGENCE L. 357 (2007); Curtis E. A. Karnow, *The Encrypted Self: Fleshing out the Rights of Electronic Personalities*, 13 J. MARSHALL J. COMPUTER & INFO. L. 1 (1994); Lawrence B. Solum, *Legal Personhood for Artificial Intelligences*, 70 N.C. L. REV 1231 (1992).

50. *See* Tom Allen & Robin Widdison, *Can Computers Make Contracts?*, 9 HARV. J.L. & TECH. 26, 29–30 (1996).

determine whether it coincides with the hardware or with the software. The position will be further complicated in the case where the agent is distributed in more than one site but acts separately. Electronic agents can collaboratively multiply into undistinguishable copies,[51] so it should come as no surprise that they can become unrecognizable, which makes it difficult to determine which agent causes the actual damage. In many cases, no surgery can separate these linked agents from one another, or isolate them from viruses or environmental factors. For that reason, it is difficult to distinguish between outcomes that result from an agent's actions and those that result from viruses or an electronic environment.

What makes the matter more difficult is the fact that such an agent does not have an established physical location, and it may, at any time, disappear without any apparent reason other than its unreliable nature. Those who adopt this approach have justified their opinion by claiming that a system which achieves self-consciousness—that is to say which is able to learn, perceive their environment and to decide autonomously—is entitled to be treated as a legal person whose autonomous acts are considered separately.[52] However, this claim is problematic, since it is not at all certain that intelligent computers can achieve self-consciousness, or a strong self-image. Even if we accept that they can perceive their environment and respond in a timely fashion to changes that occur within it, it was not obvious until now that reactivity is a valid test for the entitlement to legal personality or consciousness. 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. This is particularly true if we contemplate the current and historical examples which show that lacking or achieving consciousness is not an essential factor in conferring or denying legal personality. For example, humans temporarily lacking consciousness (e.g. in comas or asleep) are not denied legal personality on that basis. Furthermore, companies are a clear example that lacking consciousness is not a reason for denying legal personality.[53]

51. *See* Curtis E. A. Karnow, *Liability for Distributed Artificial Intelligences*, 11 BERKELEY TECH. L.J. 147, 200 (1996).

52. Solum, *supra* note 49, at 1231.

53. *See* Jiahong Chen & Paul Burgess, *The Boundaries of Legal Personhood: How Spontaneous Intelligence Can Problematise Differences Between Humans, Artificial Intelligence, Companies and Animals*, 27 ARTIFICIAL INTELLIGENCE & L. 73, 76 (2019) (comparing AI and a corporation).

Historically, fully conscious humans—such as children, married women, and slaves were considered non-persons—while some legal systems have considered temples, spirits, and idols as legal persons.[54]

On the other hand, this approach requires that software agents be insured for the purpose of satisfying legal judgments. But the question here is what is the point in ascribing legal personality to such agents if the user bears all risk of loss. In this case the user would be responsible for procuring insurance simply because such agents lack personal assets. Similarly, if the attribution of a personality to intelligent software agents protects users by limiting their liability, how can such protection exist in light of the fact that such agents neither have personal assets nor are they regarded as an asset themselves. Even if such agents are provided with a patrimony, such patrimony will not change the matter and will not set limits on the user's liability since the user will pay the patrimony and will be responsible for paying any additional compensation if the patrimony is insufficient to satisfy a judgment.[55] This practically means that the user will ultimately bear all risk of loss. That being the case, does the attribution of patrimony to a software agent make any sense?

Advocates of this solution justify this approach by claiming that conferring legal personality on software agents is the ideal way of holding transactions arranged by such agents enforceable. They estimate that by ascribing legal personality to electronic agents, these agents would have the legal capacity to enter contracts. This justification misses the fact that having legal personality does not mean that a person is automatically capable of entering into a contract. Even if all of the conditions of a contract between two persons exist, the contract may nevertheless lack legal effect if one or both parties lack capacity to contract.[56] This is particularly so if we remember that some kinds of categories are accorded legal personality by the law but they nonetheless seem to lack the full capacity to contract and may need to act through agents to exercise their legal capacities. Good examples here are infants, persons in comas, and the mentally incapacitated.

54. *See generally* Jane Nosworthy, *The Koko Dilemma: A Challenge to Legal Personality*, 2 S. CROSS U. L. REV. 1, 10 (1998); S. M. Solaiman, *Legal Personality of Robots, Corporations, Idols and Chimpanzees: A Quest for Legitimacy*, 25 ARTIFICIAL INTELLIGENCE L. 155, 159 (2017).

55. Giusella Finocchiaro, *The Conclusion of the Electronic Contract Through "Software Agents" a False Legal Problem? Brief Considerations*, 19 COMPUTER L. & SECURITY SEC. REP. 20, 22 (2003).

56. GEOFFREY CHEVALIER ET AL., CHESHIRE AND FIFOOT'S LAW OF CONTRACT, 536–61 (16th ed. 2012).

We need then to recognize that the existence of personality does not necessarily imply the existence of capacity. Legal competence should be dealt with as a relevant but separate factor. It can thus be said that ascribing legal personality would not necessarily solve all of the problems of legal and contractual capacity.

B. The application of agency law

The second solution is to apply agency law to intelligent software agents and develop a theory for electronic contracting and liability on that basis. This solution has been mentioned by a number of authors as a means of dealing with the problems that the emergence of intelligent software agents have created in the world of e-commerce.[57] The advocates of this solution estimate that when the computers' communication is based upon pre-programmed instructions, and when these computers possess the cognitive capability to capture the unique goals of the user and act accordingly on the user's behalf, then it is time to recognize that computers may serve as agents. As agents, they should be treated in the same manner as the law treats the human agent with some exceptions where the electronic nature imposes additional requirements.

Besides securing the enforceability of computer-generated agreements, this solution can be used to set limits on the liability of the person using an electronic agent. Instead of conferring upon a software agent an absolute power to bind its user in all circumstances, the user, according to this solution, will not be held liable in cases where an electronic agent has exceeded its authority.[58] Consequently, just as we are not liable for the unauthorized actions of a human agent, users would be absolved of liability for intelligent software agents' unauthorized actions.

The advocates for this solution have taken into account that although intelligent software agents have the power to affect legal positions of persons, and produce rights and duties through their activities—the law does not yet recognize them as legal persons and consequently the subjects of rights and duties. Which is why

57. *See generally* John P. Fischer, *Computers as Agents: A Proposed Approach to Revised U.C.C. Article 2*, 72 IND. L.J. 545, 557 (1997); Ian Kerr, *Spirits in the Material World: Intelligent Agents As Intermediaries in Electronic Commerce*, 22 DALHOUSIE L. J. 189, 239 (1999); Suzanne Smed, *Intelligent Software Agents and Agency Law*, 14 SANTA CLARA COMPUTER & HIGH TECH. L.J. 503, 504 (1998).

58. Ian Kerr, *supra* note 57, at 244.

advocates of this solution insist that it is necessary to include electronic agents within the set of rules that form the external aspect of agency. That is to say that the only aspects of agency law that should be applied to electronic agents are the aspects that deal with agents as agents; the aspects of agency law that deal with agents as persons or humans are irrelevant.[59] Fischer, for example, notes that:

> The principles of agency extended to computers in the agency paradigm are only those that deal with agents *as agents*, that is, as entities doing the will of a human principal. The aspects of agency law that deal with agents *as persons,* that is, rules setting out the duties of agent to principal or of principal to agent which involve consideration of other aspects of an agent than its capacity to do the bidding of another, have been intentionally omitted from the agency paradigm …"[60]

Similarly, Kerr claims that " … the only aspects of agency law relevant to electronic commerce are those that pertain to the relationship between the person who initiates an electronic device and third parties who transact with that person through the device."[61]

This solution, however, could be subject to the following objections:

1. Autonomously vs. automatically

Agency, as restrictively defined in the Restatement of Agency, "is the fiduciary relation which results from the manifestation of consent by one person to another that the other shall act on his behalf and subject to his control, and consent by the other so to act."[62] This definition clearly displays that the essence of the agency is that agents should act on behalf of their principal and subject to their control. In other words, the agent has to perform as instructed, communicate all of the necessary information available, and comply with reasonable instructions given by the principal. Let us try to examine whether this definition successfully applies to electronic agents, whether their behavior could be completely

59. Fischer, *supra* note 57, at 558.
60. *Id*. at 570.
61. Kerr, *supra* note 57, at 241.
62. RESTATEMENT (SECOND) OF AGENCY § 1 (AM. LAW. INST. 1958).

determined by human users, and whether that definition can encompass unforeseen, unintended, or unauthorized actions.

First, we need to recognize that the advanced generation of intelligent software agents can learn from their experience, modify their code and instructions, and even create new instructions and directions. Furthermore, they participate well in fixing the contents of the transactions, and they conclude the purchase, in some cases, without any prior user intervention. Hence, they do more than human agents. It is extremely difficult, if not impossible, to accurately predict all contexts in which the electronic agent will operate, or to precisely forecast what data will form that agent at the time of the action, response or performance.[63] In most cases, users do not know in advance where their electronic agents perform their work, or the other systems and modules with which these agents will interact.[64] Moreover, there is no tangible connection between intelligent software agents and their users. This means that intelligent software agents have the ability to roam the Internet, and perform their tasks while the user is disconnected, logged out, or away from a Web interaction. This is why users often not only lack knowledge of the precise terms of agreements that are generated by their agents, but they are also completely unaware that these agreements are being made. That being the case, can we conclude that intelligent software agents serve the same function as human agents?

The advocates of this solution thus undoubtedly ignore the independence of the advanced generations of such agents, and confuse the concept of automation and that of autonomy. There is also confusion between intelligent software programs that represent a free and non-standardized format, and other ordinary software programs that operate in a restricted and standardized environment. This confusion seems clear and apparent, for example, Fischer supports this solution and justifies his situation by claiming that,

63. For more information, see Sam N. Lehman-Wilzig, *Frankenstein Unbound: Towards a Legal Definition of Artificial Intelligence*, 13 FUTURES 442, 446 (1981).

64. See, for example, AuctionBot where the user specifies a number of parameters, and after that, it is up to the agent to manage the auction, monitor the price change, interact with other bidding agents, and compete autonomously in the marketplace for the best bids. Unlike popular online auction sites such as eBay's Auction-Web, which requires consumers to manage their own negotiation strategies over an extended period of time, AuctionBot autonomously performs tasks that require immediate response to events with no delay while its user is away from a Web interaction. For more information, see Robert H. Guttman et al., *Agent-Mediated Electronic Commerce: A Survey*, 13 KNOWLEDGE ENG. REV. 147, (1998).

> Computer agents, however, have no independent existence outside of their capacity as agents. They perform precisely as instructed by the principal, and do nothing when not following programmed instructions. Short of a systems error, a computer does not deviate, either intentionally or unintentionally, from the instructions given it by the principal. Indeed, the accuracy of computers, and their ability to follow directions precisely, makes them arguably better suited to the role of agent, in the limited circumstances posited here, than humans.[65]

If we accept Fischer's conclusion, and if the intelligent agents perform precisely as instructed by the principal, then there will never be any problem, and there is no need at all to apply agency principles. But is this truly the reality? Should an intelligent software agent just follow instructions in all cases without any sort of autonomous or creative discretion? Is this the real point of intelligent agent technology? If so, then there is no need to use intelligent agents because any program will sufficiently serve this function.

2. Liability vs. inability

According to agency law, an agent will be held responsible in cases of exceeding authority, defective performance, failure to exercise discretion, or acting in a wholly unreasonable manner.[66] If we try to apply this principle to software agents, many of the questions repeat, especially those regarding how an electronic agent can be charged with any loss or depreciation in value resulting from exceeding its authority. Can an electronic agent really answer for damages, and meet other responsibility demands? Answering these questions is not an easy task, especially if we remember that the law does not yet recognize electronic agents as capable legal persons, and so these agents do not enjoy the consequences of legal personality such as financial autonomy, and patrimonial rights. What is the point then in declaring electronic agents liable if they lack personal assets and cannot be sued? Furthermore, this solution will be useless to the third party as long as electronic agents cannot incur liability in a material way. It

65. Fischer, *supra* note 57, at 558.

66. JEFFREY F. BEATTY & SUSAN SAMUELSON, THE LEGAL ENVIRONMENT AND BUSINESS LAW Ch.38 (10th ed. 2012) (ebook).

appears thus to be very doubtful whether the analogy can actually be drawn in that regard without further legal bases.

Before intelligent software technology reaches a level of technical sophistication where the software agents become reliable, able to consent, and bear legal consequences from their acts, the application of this solution will be solving a problem by creating other problems. Several doctrinal, practical, and technical difficulties that face the application of such a solution illustrate these issues. At a doctrinal level, electronic agents do not at present enjoy legal personality and capacity, and hence it will be very difficult to describe the electronic agent as a distinct party who can consent to agency contract or act as a legally binding agent on behalf of its user.[67] The doubt then arises with respect to the ability of the agent to provide consent or fulfill duties and fiduciary obligations to the principal (such as the duty of loyalty, the duty of obedience, etc.). At a practical level, the doubt arises in relation to the ability of such agents to be blamed, held responsible for damages, and meet liability demands. The liability would then fall back on the user in all cases, even if the agents malfunction, fail to perform the required task, exceed their authority, or act in an unknown, unforeseen, or unintended manner. The user will have no recourse against such agents and then one can wonder if the notion of agency still presents any interest in this regard. Technically, this solution does not contemplate or recognize the inherent unreliability of electronic agents, nor the dynamic nature of the digital environment in which such agents communicate and perform their tasks. It can be said that this solution ignores the potential sources of the problem and deals with the matter as if there exists direct communication, previous relations, and a trading partner (interchange) agreement between the user and the third party—without accounting for other involved parties and factors such as network providers, administrators of electronic shopping malls, programmers, server owners, the environment, viruses, etc.

C. Electronic agents as mere communication tools

The third solution is to consider electronic agents as mere communication tools or conduits by which their owners or users can express their own will and conduct their business. Here, electronic agents are treated as a passive implement or extension of the relevant human traders, where all agent actions are regarded as

67. *See* Emad Abdel Rahim Dahiyat, *Towards New Recognition of Liability in the Digital World: Should We Be More Creative?*, 19 INT'L J.L. INFO. TECH. 224, 226 (2011).

coming directly from the person owning, controlling, or instructing them. Accordingly, there is no need for the law to give separate consideration to electronic agents or to consider them as distinct contracting parties. The promoters of this solution have clarified that legal problems relating to the conclusion of contracts through computers are not new, and that the lack of direct human intervention does not represent a phenomenon that demands regulatory innovation.[68] According to such advocates, attributing all the consequences and activities of electronic agents to their users will give human users a strong incentive to carefully choose, operate, and monitor their electronic agents.[69] Consequently, this will play a role in producing efficient outcomes and promoting proper practices towards a more satisfactory and safe electronic environment.

Furthermore, the advocates of this solution justify their opinion by arguing that the party that assents to the means also assents to the consequences. The parties are not bound because they wanted the contractual contents, but because they chose to delegate to their electronic agent the contracts in their name.[70] This argument has the potential to produce unfair outcomes by being unnecessarily harsh on the party using such agents.[71] The application of this solution will place full responsibility on the shoulders of the user or owner of an electronic agent, even in the case where an electronic agent malfunctions, is tampered with, or performs unintended operations.

IV. THE GRADUAL APPROACH: A PROPOSAL FOR A NEW SOLUTION

A. Critiques of previous solutions

As noted above, previously proposed solutions fail to provide a proper answer to the challenges derived from the advent of electronic agents in the world of electronic commerce. There are common points of interest that can be considered as areas of weakness in such approaches. The first common point is that all previous approaches are based on the attitude of "one size fits all." Such approaches attempt to address the issues posed by the

68. Giusella Finocchiaro, *The Conclusion of the Electronic Contract Through "Software Agents" a False Legal Problem? Brief Considerations*, 19 COMPUTER L. & SEC. REP. 20, 23 (2003).

69. *Id.* at 22.

70. *See* Giovanni Sartor, *Cognitive Automata and the Law: Electronic Contracting and the Intentionality of Software Agents*, 17 ARTIFICIAL INTELLIGENCE & L. 253 (2009).

71. *See* Allen & Widdison, *supra* note 50, at 46.

technology of software agents without accounting for the location, functions, and roles of such agents, or the different kinds of electronic agent generations endowed with different levels and degrees of autonomy, mobility, reactivity, intelligence, and sophistication. By following this line of thinking, all previous solutions have addressed software agents as belonging to the same category—as legal persons or nothing. Such extreme dealing is by itself a conceptual mistake that can lead to confusion between the concept of autonomy and automation, and a divorce between legal theory and technological practice.

Another common area of weakness is the lack of reference to the programming and development process of an agent, and the precedent relationship between the developers/programmers and the users/owners. The solutions do not mention any other relationships that could precede or follow the development process. Such approaches not only turn a blind eye to the issue of how risks should be structured in the electronic environment, but also ignore agent conditions and requirements that should exist before any commercialization. This creates insecurity and uncertainty, opening doors to a variety of individual arrangements that conflict with the global nature of electronic commerce. Furthermore, such approaches do not address relationships between parties involved in this technology, such as; the relationships between the owner of a running agent and the owner of the agent platform, the owner of a running agent and the administrators of electronic shopping malls, or the user and the website that offers the agent. Failing to address these relationships leads to confusion and mistrust in electronic commerce via electronic agents. What complicates the position further is that involved parties do not conduct business and transact face-to-face, which implies that the existence of individual arrangements and agreements to solve problems is in fact very difficult, if not impossible.

Moreover, the previous solutions do not seriously account for the environments in which agents operate and the role such environments may play in creating unauthorized actions. Directly or indirectly, the solutions simply attribute the electronic agent's actions to its user alone, without any investigation into whether the user has knowledge, accessibility, and control over the actions of the agent, and without accounting for other parties' involvement in creating the agent's reactions. Overall, the solutions have failed to address the unique characteristics of intelligent agents and the active role that the environment and other relevant parties play in the electronic commerce process. By doing so, such solutions have commonly threatened the balance between the various, and often

contradictory, interests of the involved parties on the one hand and the commercial, technical, and legal considerations on the other.

In order to avoid a divorce between legal theory and technological practice, and in order to translate solutions successfully into law, it is necessary to recognize the unique characteristics of electronic agents and provide for the possibility that an autonomous electronic agent might operate in a manner unknown, unforeseen, or unauthorized by the user. New solutions, therefore, must be based on a deep understanding of different aspects of such technology and must take into account the environment as a part of a problem. Furthermore, new solutions must clarify the relationship between electronic agents, programmers, users, and third parties without treating electronic agents as if they were either legal persons or nothing.

B. A new solution

Our solution is based on a gradual approach that considers the kinds and generations of electronic agents with different levels and degrees of sophistication, as well as a number of parties and factors that play an active role in agent-based commerce. This gradual approach recognizes the series of relationships and stages, which precede, accompany, and follow the development process of software agents. In addition, the proposed solution aims to determine the optimal legal status of agents, guarantee the validity and enforceability of electronic agent transactions, and limit the liability of those who employ electronic agents in their business.

In order to achieve such goals, this solution necessitates that we first differentiate between electronic agents and other software applications, and even between electronic agents themselves. Then, we must consider how such differences are legally relevant. Only after doing so, may we determine how the law should treat such agents, and how risk should be structured in intelligent agent e-contracting. Although there are different designations and various kinds of electronic agents, we can differentiate between three fundamental generations or groups.

1. The first generation

The first generation of electronic agents include agents exhibiting a very limited degree of autonomy, mobility, and intelligence, and only operating automatically, not autonomously,

according to their programming and user instructions.[72] For this reason, the users often predict the agent's actions and outcomes. The agents perform simple and secondary tasks in the contractual process such as; searching the Web for product details, comparing prices, recommending products based on user preferences, and sometimes making offers to purchase according to pre-programmed parameters and within pre-defined limits. In other words, these agents mainly compare the user requirements and preferences with the product's features.[73] In such cases, the human user remains in control of the contractual process and decides whether to select the product or merchant, or to confirm or reject a transaction.

According to the gradual approach, it is reasonable to consider such agents as sophisticated communication tools charged with transmitting the will of their users, extending the reach of interaction between parties, and presenting another mode by which natural persons can conduct their business. Based on this analysis, such agents should be classified as having no legal capacity. Without legal capacity, agent-based contracts can only then be upheld on the grounds of the user's capacity, with no need to separately consider such agents or regard them as distinct parties. This means that anything emanating from such agents would be attributed to the natural or legal person using them.

However, according to this proposed solution, such attribution should not be absolute, but creative enough to contemplate and provide for exceptional situations, technical errors in programming, and subsequent intervention from the administrators of the platform, internet service providers, third parties, or other parties. At the same time, this proposed solution establishes that reasonableness requirements must be introduced and imported. The proposed solution must allow for a certain margin of flexibility in situations where the faith of the third party is not legitimate, where it is unreasonable to believe that the user consented to the agent behavior or transaction, or where the user knows or has

72. Such agents usually act as shopping assistants or information searchers. They assist their users in deciding what to buy and where to buy, by searching for best prices and offers that satisfy the users' specific needs. *See* Amy Greenwald & Jeffrey O. Kephart, *Shopbots and Pricebots*, *in* AGENT MEDIATED ELECTRONIC COMMERCE II: TOWARDS NEXT-GENERATION AGENT-BASED ELECTRONIC COMMERCE SYSTEMS (2000); *see also* SYED RAHMAN & ROBERT J. BIGNALL, INTERNET COMMERCE AND SOFTWARE AGENTS: CASES, TECHNOLOGIES AND OPPORTUNITIES 76 (2001).

73. A good example here is Bargain Finder, which is regarded as the first Internet shopping agent. This system assists users interested in music compact discs to find the desired CD. First, the user provides it with the name of a specific music CD and then the agent provides the user with the price at a number of online music stores, and with the necessary hyperlinks to order. *See* Bruce Krulwich, *Information Integration Agents: BargainFinder and NewsFinder*, AAAI TECHNICAL REPORT 72, 73–74 (1996).

reason to know that the software agent is not working properly. The users might recover their loss from the programmer/supplier of the agent if technical faults in programming or supply caused the damage—which can be determined by contract or product liability laws.

2. The second generation

Unlike first generation agents, second generation intelligent software agents exhibit a considerable degree of autonomy, mobility, and intelligence. Second generation agents are provided with reasoning and the ability to make decisions based on their built-in knowledge, user instructions, and their own experience and cognitive state.[74] Such agents operate in open, remote, and complex networks, and are usually located on external servers, not on user computers, placing them outside the human users' full control.[75] While second generation agents' abilities introduce a whole new set of advantages and opportunities, they also bring with them a number of challenges and uncertainties concerning how the law should treat such agents. Considering such agents as passive tools is unrealistic and unnecessarily harsh on the user, and will unavoidably produce unreasonable results at a practical, commercial, and legal level. However, agent technology has not yet progressed to such an extent, where ascribing legal personality to electronic agents is desirable. This leads us to an imperious need to analyze the matter in a different way.

An alternative solution would create companies for online trading,[76] using electronic agents in their course of business. The companies would be accompanied by a highly qualified team with wide knowledge of the technical aspects of electronic agents. The company would fulfill all legal requirements, especially those regarding capital. Under this approach, electronic agents would act in the name of the company, and their relevant location would be the domicile of that company.

74. *See generally* STUART J. RUSSELL & PETER NORVIG, ARTIFICIAL INTELLIGENCE: A MODERN APPROACH (3d ed. 2009) (providing a comprehensive explanation of the theory and practice of artificial intelligence).

75. A good example here is Kasbah, which is a multi-agent system created by the MIT Media Lab. The software agents proactively seek out potential buyers or sellers and autonomously negotiate with them and then conclude the contract autonomously on behalf of their users. For more details, see Alexandros Moukas, et al., *Agent-Mediated Electronic Commerce: An MIT Media Laboratory Perspective*, 4 INT. J. ELECTRONIC COMMERCE 5 (2000).

76. This could open doors to the creation of a new type of "hybrid" personality consisting of a human and software agent operating in tandem. For further discussion, see Allen & Widdison, *supra* note 50, at 40.

Establishing a distinct registry system of electronic agents may be too expensive to justify itself, but creating a system as a part of a company's registry would be practical and economical. The online register would allow counterparties to check the soundness of the agent and steer their decision to conclude the contract. This solution is more realistic because it may be easier to accept that a company has personality, intention, and other subjective states more apparent than that of an electronic agent alone.

This solution also precludes the possibility that an electronic agent will have unlimited power to bind its user financially. By following this approach, the human owner will not bear all risk of loss alone, but will indirectly meet the demands of responsibility through contribution and shares in the company. The greatest benefit of trading through such companies is the concept of limited liability. Shareholders are under no obligation to the company using AI technology, or its creditors—beyond their obligations on the par value of their shares. Furthermore, this solution would serve the purpose of guaranteeing the enforceability and validity of the electronic contract. Companies' agent based contracts could be upheld because companies already have the requisite personality and legal capacity.

Until software agents reach a reliable level, where it becomes appropriate to personify them, we are faced with an urgent need to handle the difficulties arising from the absence of human understanding and awareness in the contractual process. One of the most effective solutions in this case is developing technical and legal standards for regulating agent-based contracts, and increasing user information and awareness in the contractual process involving intelligent agents. This solution is based on a combination of legal and technical standards balancing the need to keep the minimum level of human review and awareness, with the need to protect the key features of software agents (e.g. autonomy, flexibility, dynamism, speed). With respect to the legal standard-setting, this solution provides that specific international trade organizations or professional associations can draft specific terms and conditions for the contract involving the intelligent agents. The terms must address everything relating to online contract formation, limitation of liability, warranties, the legal aspects of delegation, attribution of risk, etc. These standard terms must clearly set out the possible outcomes, and contemplate electronically generated mistakes, viruses, and errors in the operating system, server, and agent host. It is a good idea to include a guarantee or insurance clause to protect against any harm to the

network, third parties, user, or agent platform.[77] These terms and conditions, especially those relating to consumer statutory rights[78] or any exclusion or limitation of liability, need to be reasonable,[79] conspicuous, and properly drafted and displayed[80] in order to be enforceable. This necessitates that such terms must seriously consider the relevant legal restrictions and the necessity of striking a balance between the rights and obligations of the involved parties.

However, one might argue that adopting such a solution where parties draft their own standard terms could lead to a "counter-offer" or the "battle of forms," which could ultimately lead to no contract formation. This is especially true when the standards and conditions of the vendor differ significantly from those of the buyers, and when each party insists that their own standard terms must prevail. That being so, can we say that this solution will allow electronic commerce to flourish? What if the interaction occurs between two software agents from different vendors or different research projects, and the language of such terms, key codes, concepts and signals are different? How can we ensure that such agents will interpret and understand the meaning of such terms correctly and uniformly?

We recognize that this solution is unable to circumnavigate all shallows, but it attempts to make the main sea-lanes more reliable. In other words, this solution is not intended to provide the final answer to all problematic questions posed by the emergence of intelligent software agents, but is designed to provide temporary

77. Insurance companies could play a vital role in making the distribution of liability more realistic and smooth. However, agent insurance may not be appropriate for lay users or consumers who rarely, irregularly, or unknowingly use software agents to conduct simple or routine transactions without knowledge of the particular agents they are using. This gives priority to the solution based on creating online companies that regularly use software intelligence technology to transact business.

78. This includes, for example, the right to cancel the contract under certain conditions.

79. *See, e.g.*, St Albans City and Dist. Council v. Int'l Computs. Ltd. [1995] 4 All ER 481 (relating to a faulty computer program designed by ICL for St. Albans which resulted in a loss of nearly £1 million. In this case, the court held that ICL's standard terms and conditions limiting its liability to £100,000 was unreasonable, considering that ICL had the resources to remedy the damage as well as an insurance policy of £50 million); *see also* Salvage Ass'n v. CAP Financial Services [1995] FSR 654 (holding that a clause limiting liability to £25,000 was unreasonable due to the fact that the supplier, who can obtain insurance far more easily and cheaply than the customer, already had insurance covering up to £5 million in damages).

80. *See* Interfoto Picture Library Ltd. v. Stiletto Visual Programmes Ltd. [1987] QB 433 (holding that the vendor had a duty to draw attention to particularly surprising or onerous terms using boldface type or a separately attached note); *see also* Microstar v. Formgen, 942 F. Supp. 1312 (S.D. Cal.1996) (admonishing the merchant for placing the restrictive terms in a separate, non-cross-referenced file that the customer did not necessarily have to review).

relief until such agents reach a more reliable and autonomous level whereby law begins to regard them, rather than their users, as the source of the relevant action. However, it should be noted that the internet environment does not usually involve parties in a continuing relationship. Rather, it often involves interactions between parties who have never met.[81] It might also involve other interactions between two or more software systems without any human intervention at all. Thus, the paradigm of face-to-face communication, whereby the parties first establish the terms by which they agree to conduct online business, is not common in this environment. Nevertheless, to avoid a "counter-offer" or a "battle of forms," we propose that specific international trade organizations or professional associations such as the United Nations Centre for Trade Facilitation and Electronic Business (UN/CEFACT), the International Chamber of Commerce (ICC), or the Organisation for Economic Co-operation and Development (OECD), draft such standard terms and model contracts. This will not only contribute to establishing harmonized treatment of e-commerce and developing a unified code of practice for web traders, but will also avoid some of the problematic questions as to how the agent might react to non-standard situations, while encouraging good business practice.

The other aspect of this solution is the technical standard-setting which relates to the structure of the website and the procedures, mechanisms, and protocols that should be incorporated into the agent's programming. The website must provide parties with full access to the terms of the contract, and provide sufficient notice of the existence of software agents and the possibility that such agents might be downloaded automatically without human action. Although the main issue with intelligent software agents is not solely due to the terms of their use on websites, the existence of clear terms and guidelines may play a role in raising consumer confidence in agent-based commerce—creating transparency and informing consumers of the probable risks associated with using such technology. To this end, the law ought to require the home page of every website that offers intelligent software agents to effectively notify a user of the existence of such agents, their functional capabilities, and how to use them. The law is also advised to compel such websites to include clear terms that clarify the rights and obligations of users, the scope of their liability, and

81. MICHAEL CHISSICK & ALISTAIR KELMAN, ELECTRONIC COMMERCE: LAW AND PRACTICE 67 (3d ed. 2002).

whether or not the agents employed have been assessed for risk, or audited for compliance with relevant consumer protection laws.

On the other hand, it is essential to keep human users informed throughout the contractual process, and to notify or provide them with information about the relevant events and contractual terms either immediately or shortly after the conclusion of the contract.[82] However, this does not mean that the user finalizes the web-based contracts or that the validity of such contracts is conditional upon the user's knowledge, verification and confirmation. What is intended, is to provide the user with an opportunity to review the contract and correct any errors within a short time. This will not only comport with the speed of online commerce, but will also allow for effective error handling before the contract activates—preventing difficult or expensive issues that may appear later on.

Moreover, the above mechanism aligns with contract law and the general tendency of courts to highlight the importance of human consensus in the case of web-based contracts. Despite the differences between the internet and physical world communication methods, traditional rules of contract law—especially those relating to intention and assent—will still apply to online commerce.[83] According to our solution, it is advisable to program software agents to record and store contractual processes, orders, and initial parameterization. Such mechanisms will not only be used for the purpose of evidence in order to identify the source of the problem and attribute the responsibility accordingly, but also to observe the common errors in the process, develop appropriate measures to manage errors, and avoid their occurrence in the future.

3. The advanced and future generation

It is clear that the role of intelligent software agents in e-commerce will increase rapidly in the near future. Research in AI and machine learning will convert software agents from mere facilitators or simple mediators to initiators and decision makers with increased autonomy and responsibility. Once intelligent

82. Software agents should be designed in a way that allows them to send an acknowledgement of receipt or to register relevant contractual events, and then forward and report these events back to the user. Such notification can be performed through e-mail, Short Message Services (SMS), linking, or other web features and data transmission.

83. *See, e.g.*, Specht v. Netscape Communications Corp., 150 F. Supp. 2d 585 (S.D.N.Y 2001), *aff'd*, 306 F.3d 17 (2nd Cir. 2002) (holding that a software license agreement's terms were too obscure, making the contract unbinding).

software agents reach a reliable level—where it becomes possible to precisely identify and regard them rather than their users as the source of the relevant action—then it would become desirable to provide them with a legal personality. At that time, intelligent software agents could be considered distinct and separate parties, with patrimonial rights and under the influence of principles of agency law. However, this is a future solution.

Meanwhile, we shall give due attention to the relationships and stages, which precede, accompany, and follow the development process of software agents. According to this approach, the relationships between different involved parties should be clarified from the beginning. The user should have knowledge of the probabilities of failure associated with the given agent. Software developers should inform the user of the scope, function, and the suitable environment of the electronic agent, and warn of potential risks. Similarly, an agent's creator needs to understand where the agent works and the other systems and agents with which it might interact. A user is obliged in this regard to inform the creator of whether the agent will be used for an unusual purpose or for a specific sector of the electronic market.

Losses and lawsuits are less likely when software companies outline in contracts the reasonable care that they will use in the software development, the certain degree of sophistication that agents require, and which party will bear the potential risks. This approach advises consumers and providers to stipulate in the contract which liability rules will apply in the event of an electronic agent's failure. These contractual terms are subject to judicial review in order to reduce the imbalance between contracting parties. On the other hand, establishing technical and legal standards before any commercialization, plays an active role in preventing unreliable software on the market. Establishing technical criteria by law is not easy. For this reason, it is important for computer scientists to play a role in the legislative process, and it is essential for legal requirements to be taken into account during the programming/development process. Reciprocal cooperation should be established between law and technology since technology cannot legalize technology, and legal rules alone cannot deal comprehensively with technical issues.

We must take the reasonable expectations as well as the commercial considerations into account. This implies that we should not always hold the users responsible for the actions of their electronic agents regardless of the circumstances of the transaction, but we should instead consider and examine the role of other factors and parties in forming an agent's action.

CONCLUSION

AI technology challenges traditional legal concepts. Such technology generates different kinds of risks: commercial, technical, and legal. These risk categories are not self-contained; commercial and technical risks may contain legal aspects, and legal risks may combine technical and commercial elements. In order to adequately accommodate the autonomous aspects of intelligent software technology, it is necessary to recognize the unique characteristics of software agents and provide for the possibility that an autonomous software agent might operate in a manner unknown, unforeseen, or unauthorized by the user. This implies that our solutions must be based on a deep understanding of different aspects of such technology, and must also take into account the digital environment as a part of a problem.

In this paper, it was argued that intelligent software agents have not yet reached a level of sophistication and reliability at which it becomes desirable to issue a final judgment on what their permanent legal status should be, or to treat them as legal persons who are capable of entering into contracts as either principals or agents. At the same time, it is no longer convincing to deal with software agents as if they function in a vacuum without contemplating their levels of complexity or the interrelationships between the different parties involved in the agent technology. It is also difficult to apply the agency paradigm without considering the legal status of the software agent. That being the case, it becomes urgent to look for creative approaches that protect the user from the unlimited power of his software agent without losing the advantages of flexibility, autonomy, and intelligence that enable the agent to act more efficiently within the digital environment of the internet.

We have to simply think outside of the box without considering software agents as if they were either legal persons or nothing. We have to creatively and imaginatively reform our concepts in accordance with the relevant technical, practical, and commercial considerations whilst remembering that agents display varying degrees of sophistication.