REDES

Revista Eletrônica Direito e Sociedade

Canoas, v. 8, n. 2, 2020

Artigo

Recebido: 04.04.2019

Aprovado: 28.05.2019

Publicado: 01.07.2020

DOI http://dx.doi.org/10.18316/REDES.v8i2.5614

Critical considerations on Artificial Intelligence liability: e-personality propositions

Sthéfano Bruno Santos Divino

Universidade Federal de Lavras, Lavras, Minas Gerais, Brasil

http://orcid.org/0000-0002-9037-0405

Abstract: The present article intends to discuss the notion of liability of acts practiced by artificially intelligent entities. The first topic is the analysis and definition of the term artificial intelligence, while the second is the subject of discussion and theoretical assumptions about the *lato sensu* liability of such entities. In the end, it is shown that, although such entities are endowed with a certain degree of autonomy, subjective criteria for personal accountability are ignored. Thus, personally attributing the cause and liability of unlawful acts to them might render scientific-legal production unviable. As an alternative means, it is proposed the incision of an objective liability for the eventual resolution of disputes that may arise from this context. The present reasoning is anchored in the deductive and integrated research method, and in the legal analysis and reflection.

Keywords: Artificial Intelligence; New Technologies; E-personality; Liability.

Considerações críticas sobre responsabilidade de Inteligência Artificial: inferências à e-personality

Resumo: O presente artigo objetiva discutir a noção de responsabilidade pelos atos praticados por entidades inteligentes artificialmente. Incumbe ao primeiro tópico a análise e definição do termo inteligência artificial, enquanto o segundo pretende discutir e teorizar as diretrizes da responsabilidade em sentido amplo dessas entidades. Ao final, propõe-se que, apesar de tais entidades possuírem certo grau de autonomia, inexiste a presença de subjetividade em seus desígnios. Portanto, atribuir a causa e responsabilidade pessoalmente aos atos ilícitos cometidos por uma IA pode inviabilizar a produção científica-jurídica e legal. Como meio alternativo, propõe-se a incisão de uma responsabilidade objetiva para eventual resolução dos litígios que possam surgir a partir deste contexto. Ancora-se o presente raciocínio nos métodos de pesquisa dedutivo e integrado, bem como na hermenêutica legislativa.

Palavras-chave: Inteligência Artificial; Novas Tecnologias; E-personality; Responsabilidade.

Initial designs

The modern feeling by the technological nature that emerges in the twentieth century presupposes the triumph of new resonant identities in the contemporary scenario. The primary task is entirely new: how to include or accept these factual situations to incorporate them into the legal system in a way that helps us and helps us to seek discernment for a complete rationality.

Since Descartes, the body-mind relationship has acquired new intersections. A substantial dual posture was objected to inferring new concepts from physical and psychological reality. With the structuring and social orientation in purely market terms, one must evaluate and reconsider some positions hitherto worked by language as dogmas.

From this perspective, in the early 1950s we had an important writing in the field of philosophy of mind that served as the design for one of the greatest events we have at one and the same time: Alan Turing's seminal paper Computing Machinery and Intelligence published by Mind, one of the greatest existing philosophical vehicles. Turing's main observation was the proposition that machines, if constituted in certain aspects, can think like humans. From this point on, another important event occurred both in the technological scenario and in the legal scenario. John McCarthy (2007, p. 2-15), based on Turing's paper, in 1956 at the Darthmouth conference, coined the notion and concept of Artificial Intelligence (AI). At that time, the novelty was already surprising and only visible in the field of literary fiction, especially in the early 1950s, with Isaac Asimov (1976), creator of I, Robot. Since then, the concept and concept of Artificial Intelligence has been developed and worked on constantly, despite its disappearance in the 1980s, due to the lack of interest of the researchers. However, in the contemporary scenario, the presence of these entities is noticeable and presents challenges for legal science.

On the one hand, we have the classic institutes of liability, covering both civil and criminal aspects. On the other, we have the right struggle to incorporate new technologies and adapt them to these institutes. But is this the right position? Does the new suit the old? Would not law be better suited to new technologies? This continuous erosion of questions modifies and bring a perspective of coexistence between two sciences so different in their regulatory character. But the conflicting form does not exempt us from the analysis of the disputes brought by increasingly sophisticated electronic beams. And this is currently visible.

The performance of intelligent beings artificially covers a range of legal relationships, we bring some examples. In the intellectual area, we discuss about the authorship of the works produced by an IA. There are two patent cases: The Next Rembrandt and The Obvious Group Case. The first took place on April 5, 2016, when a group of Dutch museums and research institutions, in partnership with Microsoft, brought up a painting named by them as The Next Rembrandt. Rembrandt Harmenszoon van Rijn, a late and renowned painter, left several works, but this is not one of them, on the contrary, it is the result of the artistic work produced by an AI.

The intelligent entity artificially responsible for the elaboration of the work used two methods well known in computer science: machine learning, described by such scholars as the ability to acquire

and accumulate own experiences through the repetition and repetition of tasks through an algorithm (GOLDBERH; HOLLAND, 1988, p. 95-99); and deep learning, described as a method analogous to human brain functioning, in view of contextualizing the situations that were put to it and transcending them short of its initially stipulated program (ČERKA, GRIGIENĖ; SIRBIKYTĖ, 2015, p. 376-389) . Initially, artificial intelligence examined Rembrandt's entire work and work, pixel-by-pixel, through 3D scanned materials and high-resolution files. Hence, AI was able to fragment the author's entire work into objective statistical data. In approximately 500 hours, all features, geometric analysis, composition and painting materials, it was possible to accurately replicate depth and texture, shadows and light, contours and dimensions, to create The Next Rembrandt. So far there are no legal claims about copyright claim, but they may eventually arise.

The second patent case involved the Obvious group, consisting of three French students, Hugo Caselles-Dupré, Pierre Fautrel and Gauthier Vernier, with the aim of expanding and democratizing Artificial Intelligence through art. The group used an open source code written by the young Robbie Barrat, 19, who publishes his works on GitHub, to produce an algorithm capable of producing artistic pictures equivalent to The Next Rembrandt. One of the works made by the IA belonging to the group Obvious, called "Edmond de Belamy", was sold for $ 432,500. The problem is that those responsible for its elaboration did not give the due credits to Barrat, the programmer and initial developer of the code. The group does not deny the appropriation and use of Barrat's designs, however, until shortly before the sale was made, they quieted and did not divulge this fact. Barrat dissatisfied with what happened said in his social network "I had no idea what you were doing with it -" democratized "sounds like you were doing some open source project. Conveniently cutting out the part where I ask you to credit a few weeks later after I see you posting the images for the first time for sale "(BARRAT, 2018). The Obvious group was pressured and felt compelled to give credit to Barrat, since the computer community was designating him as a thief[1].

> We believe that the legal framework is not ready yet and that the technology is not advanced enough to grant the authorship of an artwork to a virtual person. An AI doesn't have an intention and is far from having one, as opposed to what we tend to see in science-fiction. We believe that the authorship should go to the entity holding the artistic approach (OBVIOUS, 2018).

Another patent case that we can highlight is the issue of civil liability and criminal liability of artificially intelligent entities. In 1950, Asimov drew up 3 laws for robotics[2] which, if followed, would tend to regulate all human-machine relations in an objective and rigid way. Under the glimpses of the complex and endless possibilities of human interactions (also with machines), Asimov's laws are problematic. (1) What would happen if a person ordered a robot to hurt a person for their good? (2) If the robot is in the police scope and the superior responsible for the operation determines the arrest of a subject and he resists,

[1] https://twitter.com/DrBeef_/status/1055360024548012033

[2] 1 – A robot may not injure a human being, or, through inaction, allow a human being to come to harm. 2 – A robot must obey the orders given it by human beings except where such orders would conflict with the First Law. 3 – a robot must protect its own existence as long as such protection does not conflict with the First or Second Laws (ASIMOV, 1976, p. 6)

how should he proceed without breaching the first law? (3) What would happen to the surgical-medical robot that has a request from the patient to pause its operation and then an order from the doctor in charge to proceed with the procedure because it is beneficial to the patient, should you fail the order? It should be noted that the above-mentioned laws referred only to robots.

How does this really apply? We have a patent case known as Robot Sophia. On October 25, 2017, Artificial Intelligence created by David Hanson, was the first robot to acquire citizenship, becoming citizen of Saudi Arabia (STONE, 2017). In addition, Sophia is also able to simulate human behaviors such as sense of humor and feelings. His discursive capacity also impresses. Sophia delivered a brief speech at the UN on humanitarian issues such as the lack of Internet access and the lack of electricity in much of the world (UNITED NATIONS, 2017).

Now it is asked: who would be responsible if Sophia committed a civil or criminal wrongdoing? Who would respond for the damage done? Could Sophia personally be designated as an offender and responsible for such acts? The most advanced dogmatic codes have no legal solution to such questions. We need to seek some answers in comparative studies in philosophy, especially in the mental area. This brief descriptive scenario is capable of demonstrating the impossibility of exhaustion of the main factual relations with legal reflexes in view of the infinite discipline that technology tends to relieve us and to bring in an experimental and definitive character.

Thus, the article intends to discuss, albeit in initial considerations, the responsibility of the artificial entities artificially. The participation of these inferences may be valuable for future conflicts to come. In no way do we intend to impose the discursive criterion. Even because this would be a paradox. It is intended to open doors to an indispensable and immature debate around the world, but which reveals distortions in the formative process of legal science itself and in the classic institutes of this system.

Thus, the first topic deals with a conceptual delimitation of the term artificial intelligence. The second deals with the essential elements for the consideration of the broad liability of these entities, as well as the indispensable propositions for the development of the personality and personality of these virtual communities together with the pre-existing social and juridical formations. Just as each of us is in a condition to find a place in the virtual to satisfy our own pretensions, such entities can, to a certain extent, strengthen the factual and juridical sense of a social composition based on the cooperation between man-machine and, perhaps, to create the necessary conditions for the restructuring of modern identity.

In the end, it is concluded that in view of the extreme extension of the area covered by AI's participation and the theoretical discussions and the complex experiences of recent years, the classic institutes of Law, especially the notion of liability, need to be updated to the factual definitions that embody our life. The relationship between the changes determined by information technologies and the changes in their concept require a form and possible reference to the necessary application and specific legal protection of the personal and instrumental sphere of those who are inserted in them. For that, deductive methodological reasoning and integrated research are adopted.

The definition of Artificial Intelligence: what is it?

The term Artificial Intelligence, as said earlier, was coined by John McCarthy in 1956 at the Darthmouth College Artificial Intelligence Conference: the next fifty years (MOOR, 2006, p. 87-91). McCarthy (2007, p. 2) design an AI as:

> It is the science and engineering of making intelligent machines, especially intelligent computer programs. It is related to the similar task of using computers to understand human intelligence, but AI does not have to confine itself to methods that are biologically observable.

Although McCarthy does not specify in detail, he leaves some conceptual gaps. What is intelligence, for the author? McCarthy's (2007, p. 2-3) concept of intelligence is tied to the notion of technology, describing it as the "computational part of the ability to achieve goals in the world." McCarthy assumes the impossibility of a conceptual and dogmatic construction of intelligence short of his ontological relation and equivalent to the human mind, but according to his propositions anything could be an entity endowed with intelligence, even machines as simple as thermostats, for the author, have beliefs. "- John Searle (2017, p. 34) questioned him," What beliefs does your thermostat have? "McCarthy replied," Thermostat has three beliefs – it's too hot here, it's too cold here and it's right here.

McCarthy's inspiration for formulating this concept derives from the theoretical guidelines of Turing and his seminal essay. However, the claims of the mathematical scientist and father of artificial intelligence seem at least confusing. There is no theoretical and philosophical precision in his propositions. This is how there are two philosophical strands intended to explain the ontology of an artificially intelligent entity, both grounded in philosophical specters in the mind. The first is called Strong AI, postulating the existence of reproduction of mental phenomena in machines in the same way as mental operations occurring in human brains. The second, described as Weak AI, or cautious, proposes only the simulation of the intentional phenomena and the causal power of the human brain in artificially intelligent beings. The difference between the two lies in the factual and conceptual scope. While the former advocates the possibility of reproduction, which consists of the production of mental content ontologically identical to the brain contents, the latter advocates the simulation as something abstract and unnatural, linking literally to the artificial term. Here, the computer acts only as an instrument to mediate the attitudes expressed in the facticity between the mental content, the functions of the mind and the phatic world. In the first case the computer is the mind itself.

If McCarthy's concept is verified, it is impossible to fit in any of these philosophical directives. For the author, since anything can be endowed with intelligence and the philosophical strands necessarily compare the intelligence of machines with cerebral mental phenomena, the notion of Strong IA or Weak IA in McCarthy becomes impaired and unintelligible.

The development and dissemination of AI as a technological innovation also affected the legal sector. Both the strong and weak conceptions of these beings have their followers. Hallevy, Yanisky-Ravid, and Velez-Hernandez designate characteristics and linking elements to the notion of identity of an AI. For the first author Hallevy (2010, p. 6) are five elements that designate what is an intelligent entity artificially: (1)

communication ability; (2) internal knowledge (of itself); (3) external knowledge (about the world); (4) goal driven behavior and (5) creativity. Shlomit Yanisky-Ravid and Luis Velez-Hernandez (2018, p. 7) transcend Hallevy's conceptual optics and list ten necessary and indispensable attributes for the identification and characterization of an artificially intelligent entity: (1) innovation; (2) autonomy[3]; (3) unpredictable; (4) independence; (5) rationality; (6) evolving and capable of learning ability; (7) efficiency; (8) accuracy; (9) goal orientated and (10) free will ability to make choices.

Although the number of elements inserted to identify and characterize an AI vary from author to author, apparently none of these fifteen items seems to take into account the ontology of AI itself. They are elements related to the observer himself. Behavioral and behavioral descriptions insufficient to designate the very nature of the thing. A square is a square because it has four sides. However, not everything that has four sides is a square. These characteristics can have consequences and consequences for the legal considerations that may arise later. Its ontology, for us, is from a computer program. An artificial intelligence is nothing more than a complex computer program based on algorithms. Their behavioral skills are acquired through objective standards made by their programmers through deep learning and machine learning. This, however, does not change its ontology. An artificially intelligent entity can simulate digestion, but it cannot be a stomach. In the same way, such a thing can simulate photosynthesis, but it can never be a plant. This, then, is not to be confused with reproduction. That is why AI's weak notion may be the most acceptable today.

Other jurists want to differentiate the AI from Robot concept. Calo, Froomkin and Kerr (2016, p. 1) postulate that robots are composed of "(1) some sort of sensor or input mechanism, without which there can be no stimulus to react to; (2) some controlling algorithm or other system that will govern the responses to the sensed data, and (3) some ability to respond in a way that is at least noticeable by the world outside the robot itself". Richards and Smart (2016, p. 11) elaborate their definition of a robot with substrates in a non-biological agent, treating it as an autonomous agent derived from a constructed system capable of presenting physical and mental activity, but which is not alive in the strictly biological sense.

The main difference between these subjects (if so we can designate them) is that the robot acts as a physical and intermediate receptacle between AI and the phatic world. Apparently, every AI can turn out to be a robot, but not every robot can be considered an AI. In the first case we can designate the robot Sophia itself. In the second, we can demonstrate existing and operative AI programs that are executed only in restricted computing environments, through software and hardware in a certain computer, such as the Uber automotive vehicles. Thus, even designating something as robot does not authorize us to confer artificial intelligence, since it depends on a complex algorithm elaborated by programmers to maintain an interaction between the software and the executing hardware of that AI to the robotic structure and its interaction with the world.

[3] "It means that, within certain limits, machines ought to be able to take "decisions" autonomously and independent of external (e.g., remote) control on how to proceed with a given task should new conditions arise unexpectedly". (NEELIE KROES, 2011, p. 357).

This conceptual divergence between robot and AI transcends the academic field and enters the legislative scenario as a novelty. The European Parliament's resolution of 16 February 2017, with recommendations to the Civil Law Commission on Robotics (2015/2103 (INL)) (EUROPEAN UNION, 2017), in its first paragraph, prefers not to differentiate AI robots, treating them only the robots. European legislation also characterizes these entities and they are endowed with: (1) acquisition of autonomy using sensors and / or data exchange with their environment (interconnectivity) and the exchange and analysis of these data; (2) self-learning ability with experience and interaction – resulting from machine and deep learning methods; (3) a minimal physical medium – receptacle to interact with the world; (4) adapting their behavior and actions to the environment – also a result of deep learning -; (5) and lack of life in the biological sense of the term (EUROPEAN UNION, 2017).

We will consider enough and adopt the differentiation of Calo, Froomkin and Kerr (2016, p. 1), since our focus is the responsibility of the intelligent beings artificially, and not only of the robots. You need software running such a program to designate it as intelligent. In this way, the term artificial intelligence is designated and conceptualized as the present topic has its objective satisfied. It is now necessary to understand the notion of liability prescribed by the classic institutes of law and to insert it in the technological conception to verify its compatibility and, if necessary, postulate some guidelines.

The liability of artificial intelligence entities: initial proposal for a future approach

The insertion of beings endowed with artificial intelligence in society is reality. Examples such as the Sophia robot are demonstrative of the numerous cases that are being programmed to be subsequently increased day by day. But how can one act if one of these entities commits an unlawful civil or criminal act and causes harm to others? Is there specific legislation to regulate? Is it possible to hold them accountable for the practice of such an offense? Are the classic institutes of civil liability and criminal liability sufficient for such an approach? These questions will be answered in this topic. And we will start with the criminal disposition.

The Criminal Liability

In order for a criminal agent to be criminally responsible for the practice of illicit crimes, criminal psychology sets objective parameters for the imposition of such ownership. Two elements must necessarily coexist. The first is the external element or indeed, translated into the agent's own criminal conduct. Such an element is known as *actus reus*. The second element, however, is the internal or mental design, of a subjective character, inferred by the knowledge or interaction vis-a-vis (in the face of) criminal. Here your appointment is made as *mens rea*. As stated, such elements must coexist with each other. If one of them is absent, the application of criminal responsibility becomes impossible (HALLEVY, 2010, p. 8).

Actus reus is responsible for the designation of actions or omissions. The participation of certain external elements such as occur, as in the case of obtaining specific results for certain conduct. On the contrary, *mens rea* may possess innumerable levels of mental elements. For Hallevy the most advanced

is knowledge, which may accompany a prior intention or a specific intention. In the lower brain levels, there is what we understand as guilt *stricto sensu*, which includes neglect, recklessness and malpractice. Hallevy's criminal conception brings these two basic criteria for the necessary attribution of criminal responsibility to a human being or any other type of entity, including legal entities. There are other detectable abilities in human action, but it is enough for the presence *of actus reus* and *mens rea* to hold it criminally responsible. And it exemplifies the author that a spider may be able to act but is unable to formulate or designate the requirement required by the *mens rea*. In the absence of this, the sting of the spider cannot be criminally held accountable. In the same sense, a parrot is able to repeat words it hears or has been taught, but is unable to configure the internal element of *mens rea* for the commission of offenses (HALLEVY, 2010, p. 10).

There are three theories on the "criminal liability of AI entities" subject. The first is known as the perpetration-by-another liability model. The second is designated as the natural-probable-consequence liability model. Finally, the direct liability model. The three theories are independent of each other. However, there may be situations that require a joint action between them to ensure effectiveness of criminal liability.

The first model (the perpetration-by-another liability model) postulates that an AI has no human attributes. Here AI is considered innocent. A machine is a machine, so it will never be human. However, Hallevy points out that the capabilities of an AI cannot be ignored. In this model, such capabilities, however, are not sufficient to generate indications of *actus reus* and *mens rea*. Therefore, an AI, according to this theoretical guideline, cannot be the author of an offense.

The most important locus and design in the perpetration-by-another liability model is the visualization of an intermediary for the commission of the wrongful act. One is attributed liability of AI, who does not have any mental capacity to a certain responsible person, guardian, curator or legal representative judicially designated for the conduct coming from that entity. As we are approaching the possible practices coming from an AI, who can we designate as perpetrator-via-another (the real criminal agent)? Apparently, there may be two: the first could be the programmer responsible for the elaboration of the software that gave rise to the AI; the second could be the end user who would control it. In the first case, the developer can build an algorithm and program an AI to commit criminal acts. In the second case, the end user can designate and determine that an AI commits such an offense described in the penal code. In any of these situations presented, according to this model, liability is attributed to the one who gave the order or to the one who programmed the intelligent entity artificially, since in this case it acts only as an instrument capable of practicing crimes. Therefore, for the perpetration-by-another liability model an intelligent entity artificially has no mental capacity or any other equivalent to human capabilities. So is our conception and theoretical line adopted (HALLEVY, 2010, p. 10-12).

The second model, called the natural-probable-consequence liability model, tries to distinguish the function for which the AI was initially programmed and what were the acts practiced later. It is necessary to verify the different conduct and the one considered illegal, as well as the actions of the programmers who have programmed it for activity X and, without their consent, AI practiced Y. A factual and quite discussed case

can be brought: accidents involving vehicles of Uber (KINGSTON, 2018). The multinational has developed a software and a system to operate exclusively as passenger transport inside its vehicles considered autonomous. However, in one such accident one person died. The intent of the company was not to kill someone; the AI apparently did not intend to do that. Thus, in this model, as the intention and presence of the *actus reus* and *mens rea* elements is verified, it is not possible to ascertain the possibility of criminal liability y for such an offense and to attribute it to the intelligent entity artificially. However, a person may be criminally liable if the wrongful act is a natural and probable consequence of the originally-practiced (natural-probable-consequence-liability) (HALLEVY, 2010, p. 15-17). The application of this model resembles the liability attributed to the co-authorship of criminal offenses, even though the participation occurred indirectly, or the one that induced or instigated someone to commit suicide or to help them do so.

> Natural-probable-consequence liability seems to be legally suitable for situations in which an AI entity committed an offense, while the programmer or user had no knowledge of it, had not intended it and had not participated in it. The natural probable-consequence liability model requires the programmer or user to be in a mental state of negligence, not more. Programmers or users are not required to know about any forthcoming commission of an offense as a result of their activity, but are required to know that such an offense is a natural, probable consequence of their actions (HALLEVY, 2010, p. 17).

In a sense, the natural-probable-consequence liability seems to be applicable in situations where an AI committed a criminal offense without the knowledge and involvement of its programmer/user. In this model, it is required that these people be acting in a mental state of neglect, not anymore. The participation or knowledge of the perpetrators of any future commission of an offense as a result of their activity is dispensable, but they are obliged to know that this offense is a natural and probable consequence of their actions (HALLEVY, 2010, p. 17).

Although Hallevy only inserts negligence into the natural-probable-consequence liability model, the complete inference of guilt *stricto sensu* seems to be more appropriate. The question is: why should the programmer or the end user be held responsible for an artificially intelligent entity that has acted negligently and not responsible for those who have acted with imprudence or malpractice that resulted in the practice of a criminal offense by such entity? It is necessary to analyze this questioning both in the cases of action and omission of all those participants involved in that legal situation. in this way, programmers and developers must foresee and necessarily have a probabilistic-statistical notion of the risks and possible acts that an AI may have in relation to the commission of certain crimes, and may be held liable on the basis of guilt *stricto sensu*, had not been so designated (HALLEVY, 2010, pp. 20-21).

The natural-probable-consequence liability presents a distinction from the previous model. There are two scenarios. The first occurs when an AI acts exclusively innocently, without the knowledge that its conduct is criminally punishable. Under such circumstances the liability of the actions of this entity is identical to the perpetration-by-another liability model. However, the second scenario happens when an AI does not act innocently. Here, in addition to the liability of the programmer/developer subsidized in the natural-probable-consequence liability, the AI must also be criminally liable for its attitudes (HALLEVY, 2010, p. 20-21).

Finally, the third and final model, the Direct Liability Model assumes that an AI is totally independent of its programmer or the end user who uses it; it focuses on the intelligent entity itself artificially. As the requirement to characterize a criminal offense is only the detection of *actus reus* and *mens rea*, no matter what other internal and external elements, for Hallevy nothing prevents these elements from being filled by an AI.

In Hallevy's guidelines, an AI algorithm may present numerous features and qualifications superior to those of an ordinary human being. However, such characteristics and qualifications are dispensable from consideration of criminal liability. When a person, whether physical or legal, meets the requirements of the external element and the internal element, there is the configuration of criminal liability. Thus, if an IA is able to meet and satisfy these requirements, which the author advocates compliance by such entities, there are no obstacles that prevent the characterization of liability to these entities (HALLEVY, 2010, p. 22-23).

The presence of the *actus reus* in a homicide committed by the robot Sophia is visibly verifiable. Just an analysis of the penal norm, which will describe the typical fact, as well as the event committed by AI. This facility is not present when analyzing the existence of *mens rea*. Because it is an internal element, AI entities at first cannot create it. Although this is a philosophical consideration dealt with later, Hallevy assumes the existence of different mental elements according to technological evolution. The greater the evolution, the greater the cognitive capacity of AI.

Hallevy (2010, p. 24) adopts the concept of knowledge as the reception and sensory understanding of factual data. With the technology of machine learning and deep learning, AI, according to the author, is very well equipped for this reception. From sensory sensing receptors, voice simulators, physical contact, touch, etc., are present in artificially intelligent entities, including the Sophia robot. It is through these receivers that data will be transmitted to the central processing units that will analyze the data through processes. For the author, these processes are not so different. In humans it is given through the five senses. Already in the cognitive scenario of an AI are advanced algorithms executed by software in a hardware that tend to simulate human brain processes.

In one aspect, trying to appear more realistic, abstracting from fiction, Hallevy infers that only certain processes can be simulated by AI. Strong feelings like love and hate are impassive of automated reproduction. However, for the author this is not due to the ontology of the mind, but because there is a technological limitation. And the same can be applied to criminal liability. The existence of knowledge of an AI on a criminal type already guarantees liability for its illicit acts. This theoretical guideline, however, does not exclude the first, as initially addressed. The creator of the AI or the user who used it as a tool may be taxed as co-perpetrators of the wrongdoing.

> When an AI entity establishes all elements of a specific offense, both external and internal, there is no reason to prevent imposition of criminal liability upon it for that offense. The criminal liability of an AI entity does not replace the criminal liability of the programmers or the users, if criminal liability is imposed on the programmers and/or users by any other legal path. Criminal liability is not to be divided, but rather, added. The criminal liability of the AI entity is imposed in addition to the criminal liability of the human programmer or user (HALLEVY, 2010, p. 29).

It is assumed that the Direct Liability Model can be applied without the presence of guilty *sctricto sensu, actus reus and mens rea* of programmers or of the person who used AI for illicit. Here, there is protection for humans. Only AI would be criminally responsible for the illicit acts arising from its autonomy. It is assumed that the mens rea will be filled through the computational cognition of machine learning and deep learning. Hallevy even postulates something surprising: starting from the premise that AI are formed by objective data and experiences, it may act in self-defense to defend its existence[4].

The Civil Liability

Alongside the criminal guidelines, some considerations in the field of civil liability must be woven. The Brazilian Civil Code brings the institute of civil liability expressed in two norms. The first is that inserted in art. 186, assigning liability to the person who commits an unlawful act by voluntary act or omission, negligence or recklessness, violating law and causing harm to others, even if of an intrinsically moral character. In this prism is the subjective liability. The purposes for verification of its occurrence are necessarily: unlawful act through an action, fault, damage and causal nexus. The second form of liability is the objective modality described in art. 927, sole paragraph. Unlike its subjective modality, a person who damages another by virtue of an unlawful act shall repair it irrespective of fault, in cases strictly designated by law, or when the activity normally developed by the person causing the harm implies, by its nature, risk for the rights of others.

The main consideration that should be given and what is relevant for the present theme is girdled in the analysis of the element of fault. Now, imagine that the robot Sophia was in a dialogue with Jose and during that conversation it inferred physical and verbal aggressions to Jose, causing to him physical and moral damages. How to verify the presence of guilt in Sophia's conduct to designate objective or subjective liability?

First of all, the configuration of subjective liability for action or omission stems from a voluntary conduct necessarily derived from a prior intentionality expressed in a culpable conduct. To exist the very conception of guilt the agent must have intentionality. One cannot verify the guilt in agents that do not have intentional phenomena. In this way, the conscientious being must be inserted into a particular background to understand the purposes of his action or omission to be legally held accountable. A classic example of this is the one prescribed in art. 932, I, of the Civil Code, which gives parents strict liability for the wrongful acts committed by their children. Thus, it designates the code by virtue of the relative or absolute incapacity of these agents, for it is incomplete the notion of discernment and mental phenomena to understand the

[4] Not only positive factual and mental elements might be attributed to AI entities. All relevant negative fault elements are attributable to AI entities. Most of these elements are expressed by the general defenses in criminal law; e.g., selfdefense, necessity, duress, intoxication, etc. For some of these defenses (justifications), there is no material difference between humans and AI entities, since they relate to a specific situation (in rem), regardless of the identity of the offender. For example, an AI entity serving under the local police force is given an order to arrest a person illegally. If the order is not manifestly illegal, the executer of the order is not criminally liable. In that case, there is no difference whether the executer is human or an AI entity. (HALEVY, 2010, p. 30).

reflexes arising from that conduct and, therefore, cannot be held responsible. Therefore, to provide subjective liability and assign it to the agent, the agent must have intentionality and understand the background that is inserted. As machines, as defined, cannot think and do not possess intentional phenomena, they do not contain the semantic meaning of the notion of guilt. The damage done by Robot Sophia may have been nothing more than a failure of previous programming by its developers/programmers, or even intentional act of these, but never of the Sophia entity. It is for this reason that the classic institute of civil liability, in its subjective modality, is incapable of being effective if applied in the actions or omissions of acts coming from Artificial Intelligence, since it is necessary the evaluation of guilt, content that does not exist in these entities.

On the contrary, the notion of objective civil liability, by dispensing with the analysis of guilt, can adequately frame the unlawful acts originating from Artificial Intelligence. This positioning is followed by European regulations, when postulating that robots cannot be held responsible for actions or omissions that cause damage to third parties. A human agent will be the specific responsible, such as the manufacturer, the operator, the owner or user, as well as the agent that could have predicted and avoided harmful behavior of the robot. This positioning is equated with liability for the product contained in arts. 12 to 17 of the Brazilian CDC, since these last subjects could be considered strictly responsible for the actions or omissions of a robot (EUROPEAN UNION, 2017).

ČERKA et al. refutes the possibility of assigning liability for the fact of the product of an AI to its agent programmer/developer. According to the authors, if AI has a system of self-learning such as machine or deep learning, which improves with experience and improves AI decisions, it is difficult to verify and prove where there was the mistake or the human error in its programming. The authors' solution is based on a theory attached to the theory of risk: the deep pocket theory. According to this theory, if a person is involved in dangerous activities that are profitable and useful to society, one must compensate the damages caused to the society by the profit obtained. It is assumed that all those involved in the creation, programming, distribution, execution or any contact with the AI that will provide the service to society is responsible for the damages of their conduits. Therefore, you must ensure your dangerous activities by requiring compulsory insurance of your civil liability.

> Also in some cases it would be difficult to apply the product liability case, because AI is a self-learning system that learns from its experience and can take autonomous decisions. Thus, for the plaintiff it would be difficult to prove an AI product defect and especially that the defect existed when AI left its manufacturer's or developer's hands. It is hard to believe that it is possible to draw the line between damages resulting from the AI will, i.e. derived from self-decision, and damages resulting from product defect; unless we would equate the independent decision-making (which is a distinctive AI feature) with a defect. [...]
>
> [...] Liability without fault is based on the theory of risk. The theory is based on the fact that a person carries out activities that he or she cannot fully control; therefore, a requirement to comply with the safety regulations would not be reasonable, because even if the person acted safely, the actual risk of damage would still remain. In this case, it would be useful to employ the "deep pocket" theory which is common in the US. The "deep pocket" theory is that a person engaged in dangerous activities that are profitable and useful to society should compensate for damage caused to the society from the profit gained. Whether the producer or programmer, the person with a "deep pocket" must guarantee his hazardous activities through the requirement of a compulsory insurance of his civil liability (ČERKA *et al,* 2015, p. 386).

It agrees in part with the authors. Regarding the impossibility of setting liability for the fact of the product, the authors bring the classic notion of guilt to the discussion. In Brazil, as the rule in the CDC is the objective liability it is not necessary to verify if there was a human error or misconception in the programming and the development of the artificial entity, since the notion of guilt is incompatible with AI. As much as the machine has an apprenticeship, it will be simulated. The mind is a biological product. The AI will acquire factual data and transform it into objective syntactic data. There is no behavioral reproduction, but only simulation. If an AI acquires the capacity to cook, for her, this is just a product because, apparently, she does not understand the semantic content of taste, hunger and satisfaction in eating what was done. Therefore, any damage arising from his conduct, falls within the possibility of arts. 12 to 25 of the CDC, in their respective modalities and rules.

Regarding the deep pocket theory, the authors present a plausible solution. The creation of a guarantee fund to compensate for any damage done by the IA is useful. However, this is apparently only possible in cases of strict liability, in which it waives the verdict of guilt and it is possible to attribute joint liability to any entity in the consumer chain. And in cases of subjective liability, when there is no relation of consumption and the artificial entity commits some damage, as in the hypothesis of robot Sophia? How to apply deep pocket theory? It is prized by the creation of an electronic personality. Some considerations must be made.

First, there is no confusion between the legal concept and the philosophical concept of person. The philosophical concept of person is set apart. It develops in another theoretical and ontological spectrum, because it refers exclusively to what we understand per person, in its natural sense – as being. Apparently, the notion of person as a linguistic animal in Taylor (2013) is adequate, but this complex argumentative development is not pertinent to the present study. The legal concept of a person, on the other hand, translates into personality as "a susceptibility to being the holder of legal rights and obligations" (DE CUPIS, 2008, p. 19).

Such confusion does not occur precisely because there is a legal difference between entities natural person and legal entity, each with its attributions and peculiarities.

Moreover, in this line, the notion of Artificial Intelligence does not seem to fit any of the two legal classifications of person, whether natural or juridical. While the former refers to human beings, endowed with consciousness and intentionality, the second refers to abstract entities, with their own description and historical composition. The insertion of AI into any of these categories would make this entity strange and dislocated, without a factual, legal and historical context. As AI is endowed with uniqueness in relation to preexisting legal categories, the regulation and insertion of an adequately adequate is viable: the electronic personality.

The electronic personality should be attributed exclusively to the intelligent beings artificially. It results in the creation of a legal regulation to suppress the verification of guilt in the occurrence of illegal acts that falls short of the legal hypotheses anticipated. That is, the electronic personality is a hypothesis to be inserted in the legal role of objective responsibilities. Its particularity is that it should be accompanied by a guarantee fund, as expressed in the deep pocket theory, to repair or compensate for the damage done to

another. In turn, it will not exclude CDC's possibilities of liability, as it will act in a complementary manner. The electronic personality, in principle, will be applicable in legal hypotheses in which there is no relation of consumption, since the CDC itself has rules to do so, setting it in arts. 12 to 25.

Another consideration concerns personality rights. These are understood as "legal faculties whose object are the various aspects of the subject's own person, as well as their extensions and projections" (FRANÇA, 1983, p. 37). It is not intelligible and feasible to carry out the extension of the rights of the personality of the natural person and the legal person to the intelligent beings artificially. The dogmatic construction of the rights of the personality is strictly linked to the philosophical notion of the person; person as being, and not mere legal description. Such a position is advocated by Perlingieri and Tepedino when they object to the extension of the personal rights of the natural person to the legal entity. And their arguments can also be used as coexisting with AI entities.

> It is possible to remove the misconception about the extension of human rights to legal persons. (...) Hence a dogmatic and unitary conception of subjectivity as a neutral fact. The value of the individual subject is, however, different from that of the legal entity. (...) Industrial secrecy, bank secrecy, etc. may also be partly guaranteed by law, but not on the basis of the general clause governing the protection of the human person. The attempt to justify banking secrecy with the protection of privacy should be denied. It expresses an existential value (respect for the privacy of the private life of the individual); an interest of the bank and / or the client (PERLINGIERI, 2002, p. 157-158)

> Still in reference to the subject in question, the general clause contained in art. 52 of the Civil Code, according to which "the protection of the rights of the personality applies to legal persons." The legislator was well advised not to grant the legal person rights informed by values inherent to the human person. The mechanism was limited to allow the application, by loan, of the technique of personality protection, and only to the extent applicable, to the protection of the legal entity. The latter, although endowed with the capacity to exercise rights, does not contain the justificatory elements (axiological foundation) of the protection of the personality, conceived as a legal good, the object of existential situations. This is how the text of art. 52 seems to recognize that the rights of the personality constitute a category dedicated to the defense and promotion of the human person. So much so that it does not assure to juridical persons the subjective rights of the personality, admitting, only, the extension of the technique of the rights of the personality for the protection of the legal person. (...) Strictly speaking, the fundamental logic of the rights of the personality is the protection of the dignity of the human person. Even so, probably for practical convenience, the coder intended to extend them to legal entities, which can not mean that the conception of personality rights is a neutral conceptual category, applicable indiscriminately to legal persons and to human persons (TEPEDINO, 2004, p. 55-56).

The origin of the rights of the personality has an affinity with the ontology of the philosophical concept of person, but not only. There must be the subjective character, which is not in the artificial entities because they do not have minds and do not have intentionality. Thus, the creation of a personality for Artificial Intelligence is basically a response to the insufficiency of the factual and legal adequacy of these entities in the verification and verification of subjective civil liability in legal relationships dispensed with the consumerist character. Making it objective, with the creation of a guarantee fund on behalf of the IA, according to the theory of the deep pocket, can bring greater security for those that relate to AI, because, in case of moral or material injury, there is no need to verify agent or programmer/creator.

A final consideration concerns the possibility of the IA conducting legal business. As manifestation of will is an indispensable element to the configuration and negotiation (PONTES DE MIRANDA, 2012), artificial entities cannot artificially figure as agents to externalize this will, since they are absent from intentionality. What may occur is a situation analogous to the existing business with a legal entity: in setting up the e-personality of the AI, its developer, programmer or other persons strictly related to its creation will automatically be considered representatives to carry out the business acts on its behalf, as well as act in their interests when in court. It will be of this representative that the volitional negotiation will be expressed, but never of the AI.

Philosophical considerations

And what is the philosophical justification for attributing personality to an intelligent entity artificially? Hallevy's arguments encourage enthusiasts and scholars of the area. The union between fiction and reality is clearly present in the authoritative personality of the author. But we must analyze it with caution. Under the philosophical approach, there is a deficiency in his work. First, the philosophical concept of mind is not equivalent to the philosophical concept of knowledge. Secondly, as characteristic of the mind we have subjectivity. Hallevy ignores this characteristic and defines it as something objective in its nature. The definition and the distinctive features are fit for different goals. But he cannot answer the questions: what causes a mind? Can a machine, as defined, think? Is it right to attribute Intelligence to an inanimate being? The supposed propositions that will guarantee a greater scientific rigor come from the analysis of two philosophical currents of the mind: the functionalism and biological naturalism.

So far, we have assumed the impossibility of holding artificial intelligence accountable for its acts committed by virtue of its lack of intentionality. But how can we sustain this position? At first, the philosophy of mind will give us some help, especially by confrontation between Turing's computational functionalism and Searle's Biological Naturalism. The first chain, called functionalism, advocates a view that mental states (events, properties, processes) are not identical to the brain states, nor are they dispositions of behavior; mental states are, first and foremost, functional states of an organism (SCHWARTZ, 2017, p. 182). Such functional states can be designated as that which produces or causes the behavior of an organism under specific conditions. It defines itself by what it does and its relations with other mental states. In short, it translates into a mental state capable of playing a causal role in the particular behavioral organization of a being. In functionalist idealization, this state is a function that receives an input stimulus, together with other mental states, generating a product called output, which depends on that input and on that set. Mental state tends to take an input and the entire mental state of the organism is to generate a product – and the products can be changes in the mental state of that organism. In this way, functionalism presupposes mental states as genuine internal states of the organism, not just behavioral redescriptions and dispositions for it, as advocated by the behaviorist chain (SCHWARTZ, 2017, p. 183-184).

In functionalism, according to Searle (2010, p. 345) mental states are identified by their functions and not by the way in which these functions take place in the brain. Searle contests this current, claiming it

to be unverifiable and unintelligible, for "mental states in question are intrinsic[5] and functions are always relative to the observer" (SEARLE, 2010, p. 345). "There are no two types of mental phenomena, the intrinsic ones and those related to an observer; there are, rather, attributions of mental predicates that do not attribute a mental phenomenon intrinsic to the subject of attribution" (SEARLE, 2010, p 341). "The assignment of a function to a system or an element of a system is always done in relation to an objective, purpose or purpose, and functions are never just the causes; are causes within the context of teleology (SEARLE, 2010, p. 345). In other words, since interiority is intrinsic to mental states, the functional states do not possess it, because they are bound exclusively to the observer. Therefore, they cannot be elements of mental states.

Putnam and Dennett gave a new feature to functionalism. They turned it into computational functionalism, or Turing's machine functionalism. With this incursion, a more robust and complete theory of mind is presented, equating mental processes with computational processes[6] (SEARLE, 2010, p. 347). According to their mental state guidelines, "they are in fact functional states, but not of any kind. They are, rather, logical states of a computer and therefore are intrinsic states, at least at the level of the computer program's description" (SEARLE, 2010, p. 347). Metaphorically, the mind can be equated to software while the brain is hardware.

The veracity of the Turing functionalist chain can be realized through a test. First of all, you need two rooms. Two rooms, is the first requirement. Each one will be occupied with different subjects: a human in one and a computer in another. Human agents elaborate and ask questions. Such agents interact with the two occupants of the rooms by asking them questions. Both the human agent and the computer should talk and give a written response. If the human agents responsible for the contact between the occupants of the room cannot distinguish which occupant is the human agent and which occupant is the computational agent, it will pass the test. Turing's assumption is that the computer will be indistinguishable from a human agent if its input and output are functionally indistinguishable from a normal person (SCHWARTZ, p. 67)

The Turing test has a trace of unintelligibility. Searle finds him and assumes the creation of the Chinese room to make him stand out. The hypothetical situation is the following: a room with a human agent inserted in it. This agent is locked in this environment, where there are several baskets with Chinese

[5] "Intrinsics" simply means that states and events actually exist in the mind / brain of agents; the attribution of these states and events must be understood literally, not as a force of expression nor as a synthesis of an assertion that describes a complex set of events and relations occurring outside the agent. " (Searle, 2010, pp. 122-123).

[6] A Turing machine can be viewed as nothing more than a finite system of instructions to perform simple operations on strings of symbols which constitute the "input". The instructions are gathered into "machine states", each of which is a finite sequence of instructions, and a master instruction, or state-switching function, which prescribes which sequence of instructions is to be followed given the input. Such a specification is obviously entirely neutral about how such operating and switching is to be accomplished, and hence a particular Turing machine can be "realized" in very different ways: by a mechanical tape-reading device, by simulation on a digital computer, or by "hand simulation", where the operations are performed by a person or persons following written instructions on "state" cards". (DENNETT, 1981, p. 257).

symbols. This agent does not understand any word in Chinese, does not even know what is written. However, you are provided with a manual with rules for manipulating your native language to produce Chinese symbols, with the aim of creating new words.

According to Searle, suppose that, after a while, the subject is so effective and so good at following the instructions to manipulate the Chinese symbols and the programmers are so good at writing the programs that, if observed from the external point of view, from the point from someone outside the room where the subject is locked, his answers to the questions initially asked are utterly indistinguishable from those of native Chinese speakers. So, just by looking at the individual's answers, no one can say that he does not understand and does not speak a word in Chinese[7].

For Searle the simple manipulation of symbols in the Chinese room is something purely formal and syntactic. Syntactic, in this case, means automatic. It does not resemble the semantic mental character. Nothing guarantees that the agent inside the room can speak Chinese. He only manipulates them according to the rules described in the manual that was granted to him. As in the Turing test, computer responses may or may not be similar to human responses depending on the degree of data entered into the machine. Here, it works exclusively in the field of syntax, not semantics. The difference between these elements lies in the contextualization of meaning in a specific background. According to Searle's biological naturalism, a digital computer, as defined, can even understand the meaning and concept of pain; including, probably, simulating it, such as photosynthesis; but he has no access to the content of pain and the content of photosynthesis. A machine does not feel anger, does not learn Chinese, and does not know how to derive the burden and semantic meaning of words short of its formal-syntactic programming. Thus, even if an intelligent entity is artificially developed and programmed using the computational cognitive methods of deep learning and machine learning, it will only simulate situations for which it was initially constructed, without understanding the real semantic meaning of its attitudes.

The ability of human programming and the malice of the individual in manipulating the symbols obeying the rules does not make possible the learning of the Chinese by the simple manipulation and obedience to the manual. From the point of view of an outside observer, by virtue of the completion of a formal computer program, that person behaves exactly as if he understands Chinese, but he does not even understand a single word. If the individual does not understand Chinese, no other computer can understand it because no digital computer, due to the simple execution of a program, has something that we do not have. From this we can deduce that computers do not have minds, are not intelligent and act exclusively in the syntactic formal scope, manipulating objective data to insert itself in social communication.

[7] "Suppose also that after a while I get so good at following the instructions for manipulating the Chinese symbols and the programmers get so good at writing the programs that from the external point of view that is, from the point of view of somebody outside the room in which I am locked -- my answers to the questions are absolutely indistinguishable from those of native Chinese speakers. Nobody just looking at my answers can tell that I don't speak a word of Chinese". (SEARLE, J 1980, p. 420)

In summary, Searle's elementary propositions (2017, p. 51-52) are:

(1) brains cause minds – the mental processes that make up the mind are caused by processes that occur within the brain.

(2) Syntax is not enough for semantics – here there is the articulation and distinction of what is purely formal and what has content.

(3) computer programs are entirely defined by their formal or syntactic structure.

(4) Minds have mental contents; specifically, have semantic content.

Searle's final assumption is that no computer program is by itself enough to create a mind to an electronic system. Because computer programs cannot be considered minds. The fact that there is similarity between the mental brain causal functions and a software or hardware is not enough to consider it. If Searle's assumptions are correct, Hallevy's guidelines are unintelligible. It is assumed that what causes minds must have causal powers equivalent to that of the human brain. Therefore, the realization and execution of a computational program in a constructed artifact endowed with mental states close to humans would not suffice to attribute it as a mental subject. Substantial equivalence to human brain powers is required.

There are ontological traits that distinguish and characterize the mind. Subjectivity is one of them. The theoretical development involved in this theme needs to take into consideration how the conception of consciousness needs to explain a set of processes that can eventually lead a person to a subjective state of sensitivity. There is also unity of mind. In a sense, experiences are not sparse in a person's head. They unite in a conscious place. Furthermore, there must be intentionality, understood as "that property of many states and mental events by which they are directed to, or about, objects and states of affairs in the world.

Other elements are indicated by Searle. Everyone has its due importance. But if we look at these elements under Hallevy's guidelines, we will find that his propositions are apparently false. First, then, if Searle is correct, knowledge has no elemental force to cause a mind. No matter what the technological state of science. The main and brain-causing element are brains.

Second, thoughts are about something. They must have meaning. Linked to the mind, thoughts have semantic content. They transcend the formal and syntactic structure of a computer program. There is the understanding, even if words cannot describe, the meaning of pain, to withdraw the life of someone, inscribed in a background. Machines do not even have the temporal sense to understand what is 1 month or 1 year, maybe homicide or illicit acts.

There is also a lack of *mens rea* to establish a criminal offense. Although *actus reus* is present, it is not enough to set up criminal liability. The *actus reus* is only the occurrence of the conduct described in the criminal type. A machine apparently feels nothing when it kills a person. She cannot speak of despair, sadness or euphoria. It has no intentionality. It only acts as programmed. And if you act differently, it may have been due to errors or misconceptions of your programmer. There is no

mind in this being. Minds have strict biological connection. And if we consider the term intelligence linked to the mind, it becomes a mistake to designate these beings as intelligent. The best designation for them is "computer program".

Finally, the directives of biological naturalism appear to be true, automatically distorting Hallevy's functionalism and assumptions. Therefore, it is believed that the most viable model is the perpetration-by-another liability model. The intelligent entity artificially cannot be held responsible for having no evidence of will. There is no mind. Responsibility should be assigned under those who have acted in the background. Programmers or end user. It presupposes a responsibility and the duty of surveillance of the human being under the intelligent entities artificially. Moreover, comparing a machine to a mentally handicapped person is at least naive to those who do, such as Hallevy. A person remains a person even with reduced mental abilities. Machine will always be considered a machine, even if the technology provides otherwise. The robot Sophia, in a few years, will continue being robot, although its system evolves drastically. Therefore, the false conception of intelligence tied to a fictitious mind cannot serve as a means to block the effective application of legal responsibility, whether in the civil or criminal sphere. Only those who really have *actus reus* and *mens rea* can be liable.

Final Considerations

Although new technologies bring new challenges to law, this institute must conform to other sciences to ensure more effective applicability. To assert the presence of mind in an entity designated as artificially intelligent is, in terms of the philosophy of mind, erroneous. Thus, the criterion of intelligence is intrinsically linked to the intentional phenomena of the mind. And this, as such, presents itself as a biological phenomenon acting in the linguistic semantic field, a fact distinct from the designs of artificial intelligence.

The problem initially proposed was the possibility of assigning responsibility to the intelligent entities artificially. It is evident that, in the field of law, especially in the civil area and in the criminal area, we encounter difficulties. The classic institutes of responsibility still have an intrinsic notion of guilt in their construction. Blame in the broad sense, encompassing both fault and guilt *stricto sensu*. So, if it is assumed that a robot equipped with artificial intelligence has no mind, it cannot be said that such an entity has the culpable animus. Thus, if viewed from the standpoint of classical notions of responsibility, an artificial intelligence cannot be held responsible for the absence of mental phenomena.

On the other hand, analysis from the point of view of objective responsibility brings with it important considerations, since it dispenses with the faulty analysis. Preference is given to this institute rather than its subjective mode. However, in order to avoid solipsistic embarrassments in certain judgments, it is proposed to create an electronic personality, translated into the hypothesis of legal insertion of an incision in the role of objective responsibility described in arts. 932 et seq, of Civil Code. The current wording of these devices is not enough to legally support the electronic personality. A broad interpretation and acceptance of the community is necessary to avenge and become applicable. However, the legal provision to regulate the subject is indispensable, since the judge's analysis in the concrete case to verify the presence of the

requirements of art. 927 would basically bring a subjective judgment. The electronic personality, therefore, would be the key to accountability of these entities, in conjunction with the deep pocket theory.

Anyway, all the insertions made here and all the ideas presented are the germ of countless and future discussions that can and will happen when approaching these subjects (if we can so designate them). It is intended to open the way and new doors to other writings and in no way consider this article as an end point, but only as a window pointed to an infinite horizon.

References

ASIMOV, Isaac. **Eu, Robô**. Trad. Luiz Horácio da Silva. São Paulo: Edibolso, 1976.

CALO, Ryan; FOOMKIN, Michael; KERR, Ian. **Robot law**. United Kingdom: Edward Elgar Pub, 2016.

ČERKA, Paulius; GRIGIENĖ, Jurgita; SIRBIKYTĖ, Gintarė. Liability for damages caused by Artificial Intelligence. **Computer Law & Security Review**, Maryland, v. 31, n. 3, p. 376-389, jun. 2015.

DE CUPIS, Adriano. **Os direitos da personalidade**. São Paulo: Quorum, 2008.

FRANÇA, Rubens Limongi. Direitos da personalidade: coordenadas fundamentais. **Revista dos Tribunais**, São Paulo, v. 72, n. 567, p. 09-16, jan. 1983.

GOLDBERG, David; HOLLAND, John. Genetic algorithms and machine learning. ***Machine learning***, Switzerland, v. 3, p. 95-99, 1988.

HALLEVY, Gabriel. The criminal liability of artificial intelligence entities - from science fiction to legal social control. **Akron Intellectual Property Journal**, Akron, v. 4, n. 2, p. 01-42, 2010.

KINGSTON, John. Artificial intelligence and legal liability. **International conference on innovative techniques and applications of artificial intelligence**. Cornell University Library. Fev. 2018. Disponível em: <https://arxiv.org/abs/1802.07782>. Acesso em: 16 dez. 2018.

MCCARTHY, John. **What is artificial intelligence?** Stanford University, p. 2-15, 2007. Disponível em:<http://www.formal.stanford.edu/jmc/whatisai.pdf>. Acesso em: 21 mar. 2019.

MOOR, James. The Dartmouth college artificial intelligence conference: the next fifty years. **AI Magazine**, Palo Alto, v. 27, n. 4, p. 87-91, 2006.

PERLINGIERI, Pietro. **Perfis do direito civil**: introdução ao direito civil constitucional. 2. ed. Trad. Maria Cristina de Cicco. Rio: Renovar, 2002.

PONTES DE MIRANDA, Francisco Cavalcanti. **Tratado de direito privado**: parte especial. negócios jurídicos. Representação. Conteúdo. Forma Prova. São Paulo: RT, 2012, t. 3.

RICHARDS, Neil; SMART, William. How should the law think about robots? In: CALO, Ryan; FOOMKIN, Michael; KERR, Ian. **Robot law**. United Kingdom: Edward Elgar Pub, 2016.

SCHWRTZ, Stephen. **Uma breve história da filosofia analítica:** de Russell a Rawls. Trad. Milton C. Mota. São Paulo: Edições Loyola, 2017.

SEARLE, John. **Consciência e linguagem.** Trad. Plinio Junqueira Smith. São Paulo: Martins Fontes, 2010.

SEARLE, John. **Intencionalidade.** Trad. Júlio Fischer e Tomás Rosa Bueno. São Paulo: Martins Fontes, 2002.

SEARLE, John. **Mente, *cérebro e* ciência.** Trad. Arthur Mourão. Lisboa: Edições 70, 2017.

SEARLE, John. Minds, brains and programs. **Behavioral and Brain Sciences**, Cambridge, v. 3, n. 3, p. 417-424, 1980.

STONE, Zara. Everything you need to know about Sophia, the world's first robot citizen. **Forbes**. nov, 7, 2017. Disponível em: <https://www.forbes.com/sites/zarastone/2017/11/07/everything-you-need-to-know-about-sophia-the-worlds-first-robot-citizen/#7b50fe9146fa>. Acesso em: 29 nov. 2018.

TAYLOR, Charles. **As fontes do self**: a construção da identidade moderna. Trad. Adail Ubirajara Sobral e Dinah de Abreu Azevedo. 4. ed. São Paulo: Loyola, 2013.

TEPEDINO, Gustavo. **Temas de direito civil**, 3. ed. Rio de Janeiro: Renovar, 2004, t. 1.

TURING, Alan Mathison. Computing machinery and intelligence. **Mind**, Oxford, v. 49, p. 433-460, 1950.

UNITED NATIONS. **Resolução do Parlamento Europeu, de 16 de fevereiro de 2017, com recomendações à Comissão de Direito Civil sobre Robótica (2015/2103/INL)**. 2017. Disponível em: <http://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//TEXT+TA+P8-TA-2017-0051+0+DOC+XML+V0//PT>. Acesso em: 25 fev. 2019.

UNITED NATIONS. **At UN, robot Sophia joins meeting on artificial intelligence and sustainable development.** United Nations. 11 out. 2017. Disponível em: <https://news.un.org/en/story/2017/10/568292-un-robot-sophia-joins-meeting-artificial-intelligence-and-sustainable>. Acesso em: 25 fev. 2019.

YANISKY-RAVID, Shlomit; VELEZ-HERNANDEZ, Luis Antonio. Copyrightability of artworks produced by creative robots and originality: the formality-objective model. **Minnesota Journal of Law, Science and Technology**, Minnesota, v. 19, n. 1, p. 01-53, 2018.