

Universidades Lusíada

Aguiar, Filipa Maria de Almeida Marques Pais de,
1982-

Duarte, Fernanda

González, José A.R. Lorenzo, 1965-

**Can the AI-endowed machine be regarded as a
moral agent? Jusphilosophical considerations on
the scope of protection**

<http://hdl.handle.net/11067/7780>

<https://doi.org/10.34628/G2Y4-CE33>

Metadados

Data de Publicação	2024
Editor	Universidade Lusíada Editora
Tipo	article
Revisão de Pares	Não
Coleções	[ILID-CEJEIA] Polis, s. 2, n. 09 (Janeiro-Junho 2024)

Esta página foi gerada automaticamente em 2025-02-23T22:45:48Z com
informação proveniente do Repositório

Can the AI-endowed machine be regarded as a moral agent?

Jusphilosophical considerations on the scope of protection of neuro-rights

Pode a máquina dotada de IA ser considerada um agente moral?

Considerações jus-filosóficas sobre o âmbito de proteção dos neuro-direitos

Filipa Pais D'Aguiar

Doutora em Direito. Professora Auxiliar da Faculdade de Direito da Universidade Lusíada. Investigadora do Centro de Estudos Jurídicos, Económicos e Ambientais. Investigadora associada do Instituto Galego de Análise e Documentación Internacional.
Email: 13001224@lis.ulusiada.pt
Orcid: 0000-0003-1614-1521

Fernanda Duarte

Pesquisadora Associada da Universidade Federal Fluminense. Doutora em Direito. Coordenadora do Collaborative Research Network 1 of the Law and Society Association. Membro do Instituto Nacional de Pesquisa e Promoção de Direitos Humanos. Juíza Federal.
Email: fernanda_duarte@id.uff.br
Orcid: 0000-0003-2156-6946

José A.R.L. Gonzalez

Professor Catedrático. Doutor em Direito. Diretor da Faculdade de Direito da Universidade Lusíada de Lisboa. Investigador do Centro de Estudos Jurídicos, Económicos e Ambientais e Coordenador do Grupo 2.
Email: 13000234@edu.ulusiada.pt
Orcid: 0000-0001-5473-6708

DOL: <https://doi.org/10.34628/G2Y4-CE33>

This essay begins with a brief introduction that looks forward to placing the scope of the debate proposed, which, in a concise synthesis, intends to comprehensively explore the intersection between artificial intelligence (AI), ethics, law and neuro-rights, highlighting the current limitations of AI compared to human intelligence and the importance of robust ethical regulation to guide the development and implementation of AI.

To this end, firstly, it seeks to address several questions about artificial intelligence and its relationship with morality, neuro-rights, and ethics, especially in the context of legal and philosophical implications. For start, a distinction is made between the three types of AI learning: *supervised*, which requires a set of labelled data provided by humans; *unsupervised*, in which the machine learns autonomously without reference values; and *reinforced learning*, in which the machine learns by trial and error, based on rewards and punishments.

The nature of artificial intelligence is highlighted by the fact that it can quickly process large volumes of data, learn autonomously, and make decisions based on data patterns. However, AI does not possess consciousness, emotions, or true understanding like humans do. Its capabilities are limited to the algorithms and to the data that feed it¹. Among the

¹ This work does not consider the discussions about the so-called technological singularity and the problems that would arise from it, since it is still hypothetical, from the perspective of the technology available today. In this extent, we refer the entry “singularity” in the Encyclopaedia Britannica online: “Singularity, theoretical condition that could arrive in the near future when a synthesis of several powerful new technologies will radically change the realities in which we find ourselves in an unpredictable manner. Most notably, the singularity would involve computer programs becoming so advanced that artificial intelligence transcends human intelligence, potentially erasing the boundary between humanity and computers. Often, nanotechnology is included as one of the key technologies that will make the singularity happen.” (TOUMEY, Chris. “Singularity”. Encyclo-

limitations of AI are the inability to think or reason independently, the lack of sensitivity, feelings, and awareness, and the dependence on the data it has been trained with, which can lead to biased results if the data is biased. Secondly, the similarities between AI and human intelligence are explored, highlighting that the quality of AI lies in the large amount of data it can process (*Big Data*) and that it can be more efficient than humans in recognising patterns and making data-based decisions. In this sense, ethics and AI are discussed in the context of the emergence of *machine ethics* to deal with moral dilemmas, as exemplified by the *trolley problem*, and the need to develop Artificial Moral Agents (AMA) capable to make ethical decisions. As for ethical agents, they are categorized into four types: *Ethical Impact Agents*, that produce moral consequences without intention; *Im-*

pedia Britannica, 24 Jun. 2024, <https://www.britannica.com/technology/singularity-technology>).

implicit Ethical Agents, that act ethically due to their programming; *Explicit Ethical Agents*, that process and make moral decisions; and *Full Ethical Agents*, that possess consciousness, intentionality, and free will, characteristics that are currently exclusive to humans. The essay also discusses the protection of neural data and the distinction between neural and biometric data, such as Law No. 21.383 in Chile, which protects neuro-rights, recognising brain and psychological integrity as consequent dimensions of human dignity. Thirdly, regarding human rights, the essay highlights regulations such as the European Union's Artificial Intelligence Regulation, which prohibit AI systems that exploit human vulnerabilities or manipulate behaviours, and the importance of codes of conduct and ethical principles in AI, as recommended by UNESCO.

Finally, the essay addresses future challenges and pondering, emphasising the ongoing need for *jusphilosophical* and ethical questioning and debate on AI and the complexity and unpredictability of human behaviour that still poses significant challenges to AI.

This being the brief path drawn, we move on to the development of the research work that ultimately highlights the importance of robust ethical regulation to guide the development and use of AI, committed to the protection and promotion of human rights.

§ 1. Regardless of the version one embraces, *machine learning* is not a reality portrayed in a simple way. Foremost, because one may identify three types: *supervised learning*, *unsupervised learning* and, finally, *reinforcement learning*².

Supervised learning requires a dataset that encompasses examples, answers, and values that the machine must follow to achieve the required or at least expected result. The machine's previously introduced data is dissected together with the results obtained up to a certain moment in time. From the combined analysis, the machine will then, and only at that exact moment, be able to predict all the variables and get to conclusions based

on the previously collected and analysed data. These conclusions are far from the data that the human being originally provided to the machine, at the moment of its design, as they result from data that the machine, by its own skills, came to acquire.

Unsupervised learning means that the system is devoid of any kind of evaluative formulation provided by the human being. So, the only *input* is the instruction to collect very specific data. From then on, the machine's learning is completely autonomous considering its "freedom" to make its own value judgments.

Finally, in *reinforcement learning systems*, instead of providing the system with the needed input and output to produce results, the machine is only provided with information about the intended goal together with the set of allowed procedures. Through a trial/error path, the machine will then proceed to a completion process to achieve, on its own, the goal that was originally proposed.

§ 2. When we asked ChatGPT, on the 18th of October 2023, "*why artificial intelligence is intelligent?*", the answer was (adapted to Portuguese from Portugal, and now translated into English):

"The expression "artificial intelligence" (AI) refers to the ability of a system or machine to perform tasks that would normally require human intelligence. It is, however, important to clarify that artificial intelligence is not "intelligent" in the same way as humans are intelligent. Artificial intelligence is built on algorithms and mathematical models that allow it to process information, learn from data, and make decisions, but this is done in a way that deeply differs from human intelligence.

There are several reasons why artificial intelligence could be considered 'intelligent':

Advanced Data Processing: AI capable of processing very large volumes of data at a much higher speed than humans.

Machine learning: AI that allows the system to enhance its performance while conducting a specific task, using self-learning techniques and being exposed to more data.

Decision Making: AI algorithms that are designed to make decisions based previously known data and patterns, often allowing quick and accurate decision-making.

Adaptation to Change: AI systems capable of

adapting to new information and environmental changes by progressively adjusting its models.

Natural Language Processing: AI capable of understanding and generating text in a similar way to human communication, e.g., chatbots, virtual assistants, and automatic translators.

Nevertheless, it is important to consider, that (at least as we speak) artificial intelligence does not possess consciousness, emotions, or true understanding like human beings. It is limited by programming and training and does not have an inherent understanding of the world. Artificial intelligence is only as 'intelligent' as the data and algorithms that are provided to it."

§ 3. To what extent is artificial intelligence³ not (for now) comparable to human intelligence?

Firstly, AI is a set of algorithms and instructions designed to perform specific tasks. These algorithms and instructions are human designed. Therefore, ultimately, human beings decide what inputs and outputs the AI system⁴ will benefit from, how it will process the data, and what decisions it will make based on them.

This means that the AI is not capable of independent thought or reasoning. Instead, it simply follows the previously instructions inputted in it. For instance, an AI system designed to identify objects in images may be able to

3 One may argue whether it is truly "artificial". "In reality, what we call «artificial intelligence» today is neither artificial nor intelligent. The early AI systems were heavily dominated by rules and programs, so some talk of «artificiality» was at least justified. But those of today, including everyone's favourite, ChatGPT, draw their strength from the work of real humans: artists, musicians, programmers and writers whose creative and professional output is now appropriated in the name of saving civilisation. At best, this is «non-artificial intelligences»" (EVGENY MOROZOV, "The problem with artificial intelligence? It's neither artificial nor intelligent", in *The Guardian*, 30/03/2023, available at www.theguardian.com/commentisfree/2023/mar/30/artificial-intelligence-chatgpt-human-mind).

4 To settle the terminology, "artificial intelligence system" (AI-system) means a computer program developed using one or more of the techniques and approaches listed in Annex I, capable of creating outputs, such as content, predictions, recommendations or decisions, for a given set of human-defined purposes, which influence the environments with which it interacts" [Proposal for a Regulation of the European Parliament and of the Council establishing harmonised rules on artificial intelligence (Artificial Intelligence Act), COM(2021) 206 final].

2 Also in this sense, CHAHAL AYUSHI and GULIA PREETI, "Machine Learning and Deep Learning", in *International Journal of Innovative Technology and Exploring Engineering*, no. 8, 2019, pp. 4910-4914.

accurately find a cat or dog in them, but it does not perceive what these animals are, or what they mean.

Secondly, AI is not truly intelligent because, apart from sensitivity, it also lacks feelings and consciousness, here understood as devices that derive from the mind. Fundamentally, there is a lack of an organic substrate, an organism, and the necessary devices to generate homeostatic feelings, that are crucial in the process of formulating choices in human intelligence⁵. This reflects the “[...] *war between feeling and reason*.” (Damásio, 2020, p. 20)⁶ because “We live life feeling, or reasoning, or both, depending on what circumstances require from us.”⁷ (Damásio, 2020, p. 20). Therefore, AI will be unable to understand the *nuances* of human behaviour, to interpret social situations and to make the expected decisions. AI has no aptitude for this type of inference. It does not have, at least, a past life that allows it to do so. For instance, an AI system designed to detect emotions in facial expressions may be able to correctly identify that a person is smiling but is not capable to comprehend the reason they do so, or the context implied.

Thirdly, AI is limited by the data on which it has been trained. AI systems rely on large amounts of data to learn how to perform specific tasks. Hence, the quality of the provided data significantly affects the system's quality performance (though, the same reasoning may apply to human intelligence).

For instance, an AI system trained on a dataset that only includes images of light-skinned people may not be able to accurately identify dark-skinned people as it was not exposed to sufficient examples of this second category and did not learn to identify them accurately. Therefore, an AI system that has been trained with biased data will, necessarily, make biased decisions (thus leading, eventually, to the perpetuation of inequalities or biases). Consequently, it will not be as impartial and objective as one would expect. It will all rely on the quality and quantity of data provided

“Thus, the praised «human intuition» actually consists of «pattern recognition». Good drivers, bankers, and lawyers don't have magical intuitions about traffic, investment, or negotiation. Instead, by recognising recurring patterns, they identify and try to avoid careless pawns, inept borrowers, and wrongdoers.”

to the system. This explains why developing codes of conduct and observing the ethical principles of AI is so important, as it was previously highlighted by UNESCO, in 2022, in the Recommendations on AI Ethics, namely, regarding the adoption, by Member States, of core values and principles in line with the United Nations Charter, Human Rights, International Law and the Sustainable Development Goals, promoting the necessary changes and drafting of legislation⁸.

§ 4. We now ask, instead, to what extent does human artificial intelligence resemble artificial intelligence?

The quality and reliability of artificial intelligence relies on the wide amount of information it is provided with. The so-called *Big*

Data phenomenon comprises mathematical methods applied to large amounts of information to find correlations and to infer probabilities.

Bearing this in mind, one may concede that it is not necessary to develop computers with human-like intelligence but, on the contrary, that human may change their way of thinking to resemble AI systems.

“Causality is an important part of human thinking, particularly in science, but according to this view we do not need causality. Correlations are enough. For example, based on criminal data we can infer where crimes will occur, and use it to allocate police resources. We may even be able to predict crimes before they are committed, and thus prevent them”⁹.

However, though it is a possibility, the fact is that predictive policing systems, based solely on profiling or evaluating human beings' characteristics, have been ruled out by the European Union's Artificial Intelligence Regulation¹⁰. Systems that are likely to manipulate human behaviour and exploit human vulnerabilities are also prohibited, at least in the Europe¹¹. Ultimately, we are witnessing the advent of digital borders.

Regardless of the path forward, it is certain that the transmission of knowledge from human beings to computers – whether it operates through language or otherwise – raises serious problems. On the one hand, the human language¹² and natural languages¹³ issue. Human language, inseparable from “[...] evolutionary changes in the brain.”¹⁴, results from the “[...] existence of physiolo-

5 DAMÁSIO, António. *Sentir & Saber: a caminho da consciência*. 1.^a ed. Temas e Debates, Círculo de Leitores: Lisboa, 2020. P. 63 ss.

6 DAMÁSIO, António. *Sentir & Saber: a caminho da consciência*. *Idem*. P. 20. Author's italics.

7 DAMÁSIO, António. *Sentir & Saber: a caminho da consciência*. *Idem*. P. 20. Itálicos do autor.

8 UNESCO. *Recommendation on the Ethics of Artificial Intelligence*. 2022. Available at: <https://unesdoc.unesco.org/ark:/48223/pf0000381137>. Accessed on: 25 May 2023.

9 RAGNAR FJELLAND, *Why general artificial intelligence will not be realized, in Humanities and Social Sciences Communications*, 2020, p. 2.

10 PARLAMENTO EUROPEU. Direção-Geral da Comunicação. Regulamento Inteligência Artificial: Parlamento aprova legislação histórica. 2024. Disponível em: <https://www.europarl.europa.eu/news/pt/press-room/20240308IPR19015/regulamento-inteligencia-artificial-parlamento-aprova-legislacao-historica>. Acesso em: 4 Jun. 2024.

11 PARLAMENTO EUROPEU. Direção-Geral da Comunicação. Regulamento Inteligência Artificial: Parlamento aprova legislação histórica. *Idem*.

12 FROMKIN, Victoria; RODMAN, Robert. *Introdução à linguagem*. Trad. Isabel Casanova. Coimbra : Almedina, 1993. P. 28 ss.

13 BENVENISTE, Emile. *O homem na linguagem*. Trad. Isabel Maria Lucas Pascoal. Lisboa : Veja, 1992. P. 59 ss.

14 FROMKIN, Victoria; RODMAN, Robert. *Idem*, p. 30.

gical, anatomical and “mental” preconditions that allow linguistic development.”¹⁵element. Thus, the existence of an organic substratum, an organism, a brain is fundamental in the transmission of knowledge through human language as well as in the comprehension of the subtlety of concepts in natural languages¹⁶. Thus, the dilemma faced by AI systems design and development that allow a computer to simulate the performance of skilled experts (for instance, in making medical diagnoses), is that an important part of the specialised knowledge is obtained implicitly. If these experts try to express the knowledge they apply in their work, they usually regress to a lower level. The same happens even with simpler tasks. For instance, most humans are experts at walking. However, if each one tries to express the way one walks, one will certainly give a description that does not express the skills involved in walking¹⁷. Therefore, the transmission of knowledge by a human being, even in the most basic tasks, through language also means a consideration of the subtleties of one’s natural language and the need to attend to all its nuances and specificities, organic dimensions and choices based not exclusively on reason, but also as a result of homeostatic feelings. This will perhaps be a dimension that, for now, remains profoundly human and that still distinguishes natural intelligence from artificial intelligence¹⁸. On the other hand, “in recent decades, research carried out in areas such as neuroscience and behavioural economics has allowed scientists to *hack* humans and, in particular, to perceive greater understanding of their decision-making. It turns out that human choices

- of everything from food to mates - are not the result of some mysterious *freedom of choice*, but rather of billions of neurons acting to calculate probabilities within a fraction of a second. Ultimately, freedom is defined by reference to the inner dimension of the human being and is identified with the absence of external coercion. Therefore, “[...] When we say that we are free, [...], it means that our behaviours are determined by what happens inside us, in the brain, and that they are not forced from the outside.”¹⁹ (Rovelli, 2022, p. 67). Considering neuroscience, “We are what we decide.” (Sigman, 2015, p. 65) and when we choose,

“De manera imperceptible, como si cada alternativa decantase naturalmente, comparáramos el universo de opciones posibles en una balanza mental, lo ponderamos y finalmente decidimos (membrillo, por supuesto). Sobre esas alternativas ponemos en acción los circuitos mentales que conforman la maquinaria de la decisión.”²⁰ (Sigman, 2015, p. 65).

Thus, the praised «human intuition» actually consists of «pattern recognition». Good drivers, bankers, and lawyers don’t have magical intuitions about traffic, investment, or negotiation. Instead, by recognising recurring patterns, they identify and try to avoid careless pawns, inept borrowers, and wrongdoers. It was also possible to conclude that the biochemical algorithms of the human brain are far from perfect. They rely on heuristics, shortcuts, and outdated circuits more adapted to the savannah rather than the urban jungle. It is no coincidence that good drivers, bankers and lawyers sometimes make gross mistakes. This means that AI can outperform humans even in tasks that supposedly require “intuition.”. Particularly, AI may be better at jobs that require the use of intuitions about other people’s behaviour. Many kinds of jobs - such as driving a vehicle on a pedestrian-filled street, lending money to strangers, and negotiating a contract - require the ability to correctly assess other people’s emotions and desires.

While it was claimed that such emotions and desires would be based on an *immaterial spirit*, it seemed obvious that computers would never be able to replace human drivers, bankers, and lawyers. For how could a computer understand the divinely created human spirit? However, if these emotions and desires are in fact no more than biochemical algorithms, there is no reason why computers cannot decipher them and do so much better than any *Homo sapiens*.

The driver who anticipates the intentions of a pawn, the banker who estimates the credibility of a potential borrower, and the lawyer who assesses the surrounding environment at the negotiating table do not rely on paranormal knowledge. Instead, but unconsciously, their brains are recognising biochemical patterns by analysing facial expressions, tones of voice, hand movements and even body odours. An AI equipped with the right sensors could do all of this with much greater accuracy and reliability than a human.”²¹

Perhaps, therefore, there is no major difference between human intelligence and so-called artificial intelligence. In any case, considering neuroscience knowledge, it seems that artificial intelligence is, for the time being, limited by the lack of an organic substrate, of a human organism and by the absence of homeostatic feelings that are reflected in the human choice process²². Essentially, the choices of human intelligence result from an organic, natural substratum, where homeostatic feelings play a fundamental role in the process of choice, giving it some degree of vulnerability. Artificial intelligence’s choices have no natural origin or organic substratum, and its robust and efficient decision-making process is solely probabilistic, statistical, rational and based on mathematical formulas (Damásio, 2020, p. 251-256)²³. Thus, on the one hand, by seeking to eliminate the vulnerabilities and the typically human affective dimension from artificial intelligence systems, human beings, actually, ended up conditioning the creative potential and the level of intelligence of these systems (Damásio, 2020, p. 251-256).²⁴

15 FROMKIN, Victoria; RODMAN, Robert. *Idem*, p. 30.

16 BENVENISTE, Emile, *Idem*, p. 60. “[...] Natural languages [...], in fact, contain one of the most subtle riches of concepts and distinctions, and fulfill a variety of functions [...]. Everyone who knows how to speak a language has, without a doubt, an implicit mastery of these concepts and these subtleties.”

17 RAGNAR FJELLAND, “Why general artificial intelligence will not be realized”, *op. cit.*, p. 3.

18 PAIS D’AGUIAR, Filipa (2024) – *Direitos Humanos e Inteligência Artificial: principais dimensões jurídicas, éticas, sociais e culturais, no contexto europeu e transnacional*. In Carlos Ajjonso Leony Neto, Leandro de Matos Coutinho, Patricia Desirée de Mello (Coord.) – *Mentes digitais : do zero ao infinito : Homagem à Professora Fernanda Duarte*. Prejácio de Guilherme Calmou Nogueira da Gama. São Paulo : Editora GZ. ISBN 978-65-5813-116-8. 242 p..

19 ROVELLI, Carlo. *Sete breves lições de física*. Trad. Vasco Gato. 1.ª ed., 3.ª republ. Penguin Random House : Lisboa, 2022. P. 67.

20 SIGMAN, Mariano. *La vida secreta de la mente : nuestro cerebro cuando decidimos, sentimos y pensamos*. (Debates). 1.ª ed. Penguin Random House : Buenos Aires, Barcelona, 2015. P. 65.

21 YUVAL NOAH HARARI, “21 Lessons for the 21st Century”, in *Vintage Publishing*, 2019, p. 19 to 20.

22 PAIS D’AGUIAR, Filipa, *Idem*, p. 7.

23 DAMÁSIO, *Idem*, p. 251-256.

24 DAMÁSIO, *Idem*, p. 251 ss. Also, PAIS D’AGUIAR, Filipa, *Idem*, p. 7.

On the other hand, despite the possibility of identifying biochemical patterns and algorithms of the human brain, human decisions still remain free because they are the result of the laws of nature that act on each one's brain, that is, they are "[...] freely determined by the results of the transient and very rich interactions between the billions of neurons in our brain: they are free when it is the action of these neurons that determines them." (Rovelli, 2022, p. 67)²⁵. Therefore, the unpredictability of human behaviour still remains and persists, that is, its complexity makes prediction difficult to the extent that "[...] We have hundreds of billions of neurons in our brains, as many as the stars in a galaxy, and an even more astronomical number of connections and combinations in which they can be found." (Rovelli, 2022, p. 67-68)²⁶. In short, freedom of choice means being determined by the laws of nature that act individually on the brain of each human being²⁷. The standardisation of «intuition» itself becomes unpredictable, but not impossible. Perhaps, for this reason, due to the effective and real possibility of *hacking* the human brain, we have been witnessing some legislative initiatives (e.g., Brazil, Chile, EU) that consider the protection of neuro-rights, seeking to delimit the scope of protection of neural data. For instance, in Chile, on October 25, 2021, the protection of neuro-rights, or rights of the brain and mind, as a dimension of human dignity, was recognized with the approval of Law No. 21.383, which modifies Article 19(1) of the "[...] Constitución Política de la República, de la siguiente forma: [...] "El desarrollo científico y tecnológico estará al servicio de las personas y se llevará a cabo con respeto a la vida y a la integridad física y psíquica. La ley regulará los requisitos, condiciones y restricciones para su utilización en las personas, debiendo resguardar especialmente la actividad cerebral, así como la información proveniente de ella;" (Chile, 2021)²⁸.

Apart from the debate regarding the protec-

tion of neuro-rights framework, the Chilean initiative also stressed the need to distinguish between the concepts of neural data and biometric data and to delimit their scope of protection²⁹. Essentially, neural data are distinguished from biometric data through the object in which they materialise.³⁰ Biometric data is identified with organs and body tissues, while neural data is the result of the operations of neurons and independent of an organic support³¹. It was in this sense that the justification of the Projeto de Lei n.º 522/2022 of the Brazilian Chamber of Deputies, in 2022, defined neural data as "[...] any information obtained, directly or indirectly, from the activity of the central nervous system and whose access is carried out through brain-computer interfaces, or any other technology, invasive or non-invasive; [...]"³². Examples of these technologies are, for instance, neuroimaging techniques such as computed tomography, positron emission tomography (PET), single-photon emission computed tomography (SPECT), functional magnetic resonance imaging (fMRI) and electroencephalography (EEG).³³ element. Also, in the Brazilian's initiative view, there is a need to distinguish the concepts of neural data and biometric data, to delimit and substantiate their scope of protection. Thus, neural data configure "[...] a semantic, or language, feature with the brain. In this sense, this information is data belonging solely to the domain of the brain. The operating mechanisms of neurons transmit signals filled with information about the person's neurocognitive state. Therefore, because neural data is part of people's minds,

the holder must have the right to protect not only their privacy, but also their psychological integrity."³⁴

The multiple advantages of these technologies are not questioned. The question that arises is different. It is the possibility of overcoming "[...] a barrier that seemed insurmountable, i.e., the human consciousness."³⁵. Considering that these technologies allow data collection directly from the central nervous system, this imply the possibility of recording data from the subconscious and accessing, for instance, "[...] thoughts that would never be communicated or transformed into actions [...]"³⁶. Thus, the rigors of the previous frontier between ethical experience and legal experience, based on externalized and internalized conduct, also disappear³⁷. In Europe, for instance, the Artificial Intelligence Regulation of the 13th of March 2024, considering various levels of risk, removes, as stated above, artificial intelligence systems that are likely to exploit human vulnerabilities or manipulate human behaviour.³⁸ Although this initiative responds to the need to contemplate Artificial Intelligence systems in the light of Human Rights and the European fundamental values and principles, it does not exhaust or close the debate and the philosophical and ethical pondering³⁹ at a global level. For Dragos Tudorache, rapporteur of the Committee on

25 ROVELLI, Carlo. *Idem* 2022, p. 67.

26 ROVELLI, Carlo. *Idem* 2022, p. 67-68.

27 ROVELLI, Carlo. *Idem* 2022, p. 67.

28 CHILE. Ministerio de Ciencia, Tecnología, Conocimiento e Innovación (2021) – *Ley 21383*. Biblioteca del Congreso Nacional de Chile : Chile. Disponível em: <https://www.bcn.cl/leychile/navegar/imprimir?idNorma=1166983&idVersion=2021-10-25>. Acesso em: 05 Jun. 2024.

29 WAJNERMAN PAZ, A. Is Your Neural Data Part of Your Mind? Exploring the Conceptual Basis of Mental Privacy. *Minds & Machines* 32, 395–415 (2022). <https://doi.org/10.1007/s11023-021-09574-7>.

30 WAJNERMAN PAZ, A., *Idem*. Também, BRASIL. Câmara dos Deputados. *Projeto de Lei 522/22*. Portal da Câmara dos Deputados. 2022. Disponível em: <https://www.camara.leg.br/propostas-legislativas/2317524>. Acesso em 05 Jun. 2024.

31 WAJNERMAN PAZ, A., *Idem*. Também, BRASIL. Câmara dos Deputados. *Projeto de Lei 522/22*, p. 5.

32 Também, BRASIL. Câmara dos Deputados. *Projeto de Lei 522/22*, p. 1.

33 HALLINAN, D., P. Schütz, M. Friedewald and P. de Hert 2014. Neurodata and Neuroprivacy: Data Protection Outdated? *Surveillance & Society*. 12(1): 55-72. Available at: <http://www.surveillance-and-society.org> | ISSN: 1477-7487. Accessed on: 5 Jun. 2024. P. 58. Also, BRAZIL. Chamber of Deputies. *Bill 522/22*, p. 5-6.

34 Também, BRASIL. Câmara dos Deputados. *Projeto de Lei 522/22*, p. 5.

35 BRASIL. Câmara dos Deputados. *Projeto de Lei 522/22*, p. 4-5.

36 BRASIL. Câmara dos Deputados. *Projeto de Lei 522/22*, p. 4-5.

37 TEIXEIRA, António Braz (2006) - *Sentido e Valor do Direito : Introdução à Filosofia Jurídica*. 3.ª Ed. Lisboa : INCM. P. 141 ss.

38 PARLAMENTO EUROPEU. *Regulamento Inteligência Artificial, P9_TA(2024)0138*, Resolução legislativa do Parlamento Europeu, de 13 de março de 2024, sobre a proposta de regulamento do Parlamento Europeu e do Conselho que estabelece regras harmonizadas em matéria de inteligência artificial (Regulamento Inteligência Artificial) e altera determinados atos legislativos da união (COM(2021)0206 – C9-0146/2021 – 2021/0106(COD)). 2024. Disponível em: https://www.europarl.europa.eu/doceo/document/TA-9-2024-03-13_PT.html#sdocta2. Acesso em: 05 Jun. 2024. Par. (29).

39 *Vide*, o elenco dos princípios éticos a observar pelos sistemas de inteligência artificial e fornecedores, pela UNESCO. *Recommendation on the Ethics of Artificial Intelligence*. 2022. Disponível em: <https://unesdoc.unesco.org/ark:/48223/pf0000381137>. Acesso em: 25 Maio 2023.

Civil Liberties:

“The EU has achieved results. We have linked the concept of artificial intelligence to the fundamental values that form the basis of our societies. However, there is a lot of work ahead that goes beyond the Artificial Intelligence Act itself. AI will lead us to rethink the social contract at the heart of our democracies, our educational models, labour markets and the way we conduct war. The AI Act is a starting point for a new governance model built around technology. We must now focus on putting this law into practice.”⁴⁰ (PE, 2024, p. 4)

§ 5. The rise of AI has led to the arise of a so-called *machine ethics*.

Originally, it arises, mostly, from a question drawn from the well-known *trolley problem*: “The carriage of a *trolley* approaches a fork. Following the current track, it will run over and probably kill a group of five workers. If the driver diverts the carriage to the other branch, he will cause the death of a single worker. What should a person do under these circumstances? And, more importantly, what should an AI system do?”⁴¹.

Driverless vehicles put machines in a position to decide, in fractions of a second, in life-or-death contexts. As the complexity of road networks increases, so does the probability of dilemmas arising that are not very different from the example above or are even more acute. New technologies in the diverse fields of AI, genomics, and nanotechnology will combine in myriads of unpredictable ways to offer promises in everything, e.g., from increasing productivity to curing diseases. There is, therefore, a strong scholarship that suggests the need to incorporate artificial moral agents (AMAs)⁴² into these new technologies to ma-

nage their complexity. AMAs should be able to make decisions that protect privacy, fundamental rights, and individual freedom. Designing AMA's will not be easy, but it is certainly inevitable and indispensable.

To prevent the possibility of autonomous AI acting into ‘knaveish’ consequences, it will be essential to design actors whose decisions and actions are likely to be regarded as ‘benevolent’. However, it is not easy to define the kind of benevolent behaviour that autonomous systems should observe. Should a good multifunctional robot help a stranger even when this is likely to delay the fulfilment of tasks that it has been predetermined for by its owner/handler? At the borderline, should it disobey its owner/handler when all foreseeable options are capable of harming human beings⁴³?

A machine cannot be genuinely an ethical agent or even a genuinely autonomous agent. A true ethical actor acts intentionally, freely. Thus, the challenge facing AI only concerns the so-called artificial morality: the ability to, in other words, make artificial agents act as if they were (natural) moral agents and, to that extent, to endow them with the ability to separate «right» from «wrong».

It is not sufficient that a product is flawless to place one's trust in multi-purpose machines which operate daily without connection to their designer and which are programmed to respond flexibly in real or virtual environments. An autonomous system that causes damage cannot be morally censured, just as a toaster that catches fire is not subject to a judgment of disapproval. However, an auto-

nomous system must be able to assess the possible harmful consequences resulting from its actions and select them accordingly.

“A simple vacuum cleaner, like the *Roomba*, can already be confronted with morally significant situations. Unlike a conventional vacuum cleaner, *Roomba* is not directly commanded by a human. Therefore, to a certain extent it is autonomous. But even a system as primitive as this one, faces elementary moral challenges. For instance: should it vacuum and, consequently, kill a ladybug that comes its way, or should it go around or chase it away? And being a spider? Should it be extinguished or spared? (...).

Autonomous vehicles are a particularly delicate example because they face not only moral decisions, but also moral dilemmas. (...). Autonomous vehicles may face situations where injuring or even killing one or more people to save others may become unavoidable. Imagine that an autonomous vehicle is not able to stop in time and only has the option of bumping into one of two groups of people: on the one side, two elderly men, two elderly women and a puppy; on the other side, a young woman with two young children. If it hits the first group, the two women will die, the two men and the dog will be seriously injured. If it hits the second group, one of the children will die and the woman and the other child will be seriously injured.”⁴⁴

One may object that, actually, it is not the robot vacuum cleaner or the autonomous vehicle that, in these cases, make a moral decision, but rather, and foremost, its designers. However, the idea that *designers* may determine, in advance, all possible outcomes is misguided. The program is more advanced than its *designer*. Even if designers provide the system with all the needed algorithms it will operate with, they cannot anticipate every move. Instead, the system is equipped with a set of decision-making procedures that allow it to decide effectively on its own. Therefore, due to the unpredictability and lack of control of human actors, it makes sense to use the term [or rather, the expression] «artificial agent» for this kind of system.

On the one hand, when the concept of autonomy is used, it is necessary, however, to

moral agents which can take moral decisions and act on them. Artificial moral agents in this sense can be physically embodied robots as well as software agents or ‘bots’” (CATRIN MISSELHORN, “Artificial Moral Agents, Conceptual Issues and Ethical Controversy”, in *The Cambridge Handbook of Responsible Artificial Intelligence – Interdisciplinary Perspectives*, Part I – Foundations of Responsible AI, Cambridge University Press, 2022, p. 31).

43 Thus, implementing Asimov's first and second laws. The so-called “Three Laws of Robotics” (ISAAC ASIMOV, “Runaround”, in *I, Robot*, Gnome Press, 1950, p. 27) are: “First Law - a robot may not injure a human being or, through inaction, allow a human being to come to harm. Second Law - a robot must obey the orders given it by human beings, except where such orders would conflict with the First Law. Third Law - a robot must protect its own existence as long as such protection does not conflict with the First or Second Laws”.

40 PARLAMENTO EUROPEU. Direção-Geral da Comunicação. Regulamento Inteligência Artificial: Parlamento aprova legislação histórica. 2024 Disponível em: <https://www.europarl.europa.eu/news/pt/press-room/20240308IPR19015/regulamento-inteligencia-artificial-parlamento-aprova-legislacao-historica>. Acesso em: 14 Mar. 2024.

41 COLIN ALLEN, WENDELL WALLACH and IVA SMIT, “Why Machine Ethics?”, in *IEEE Intelligent System*, vol. 21, no. 4, 2006, p. 12.

42 “Artificial Intelligence (AI) has the aim to model or simulate human cognitive capacities. Artificial Morality is a sub-discipline of AI that explores whether and how artificial systems can be furnished with moral capacities. Its goal is to develop artificial

44 CATRIN MISSELHORN, “Artificial Moral Agents. Conceptual Issues and Ethical Controversy”, *op. cit.*, pp. 32 and 33.

distinguish the use of a stronger sense of autonomy – *free will* – or a weaker sense – *the ability to act without the need for the intervention of others*. On the other hand, “the processes running in AI systems cannot all be measured according to duties of care designed for human conduct, or not without adjustments”⁴⁵. Therefore, even when it is possible to maintain that the machine is bound by duties of care, the state of its existence will often demand its reformulation.

For now, any AI system will be autonomous in the, above referred, second sense. However, in terms of liability for damages, it is crucial to know whether, even so, it can be considered a moral agent (without truly being one).

§ 6. According to one possible classification, the so-called *moral or ethical agents* may be included to the following kinds⁴⁶:

a. *Ethical impact agents* are those that produce moral consequences without them being perceived as such, e.g., the digital clock that alerts its user to task fulfilment. Considering that the moral quality of these devices relies solely in the use that is made of them, it is questionable whether they should be regarded as actual agents.

b. *Implicit ethical agents* are those that act ethically as its internal functions tacitly promote ethical behaviour or, at least, prevent unethical behaviour, e.g., aircraft warning systems that trigger an alarm when they get too close to the ground or to another aircraft. Nevertheless, these are not truly autonomous, since their moral qualities are entirely due to the intervention of its *designers*.

c. *Explicit ethical agents*. In this case, unlike the previous two types, the agents can freely recognise and process relevant moral information and to produce moral decisions. They act not only according to moral guidelines, but also based on moral pondering. From an ethical point of view, this kind of systems can surpass a human being in the resolution of highly complex situations, such as the organisation of humanitarian aid in major

“And, at least out of necessity (given that “we want machines to treat us well”) there are undeniably AI-powered machines - such as autonomous vehicles - that should be conceived as artificial explicit ethical agents. Moreover, if one considers that animals can generate intentional actions, it certainly will not be hard to admit, likewise, that AI systems are also capable of performing action-creating effects with the purpose of obtaining them.”

disasters. Usually, in situations like these, a human being often struggles in properly processing information on the distribution of aid while computers, designed to deal with intricate problems that demand quick decisions, are generally much more competent.

d. *Full ethical agents* refer to those that are endowed with awareness, intentionality and *free will*. For now, as far as it is known, only

the adult human being is endowed with such qualities. Of course, there is a very strong scholarship that argues machines will never be able to cross the line – *the bright line* – that, in this sense, separates them from the human being. The ability to philosophise, for instance, also seems to remain exclusive to human beings⁴⁷. Not knowing the future, this degree assertiveness is, most likely, unwise.

Actually, it seems irrefutable that at least *explicit ethical agents* are comparable to moral beings (albeit *lato sensu*).⁴⁸ And, at least out of necessity (given that “we want machines to treat us well”)⁴⁹ there are undeniably AI-powered machines – such as autonomous vehicles – that should be conceived as *artificial explicit ethical agents*. Moreover, if one considers that animals can generate intentional actions⁵⁰, it

47 ORENSANZ, Martín. Inteligencia Artificial e Filosofía. *IECE Revista Digital*. Año VIII, 15 (Jul. 2023) 50-53. ISSN 2545-6326. Disponível em: <http://iece-argentina.weebly.com/>. Acesso em: 05 Jun. 2024. O autor, M.O., questiona o sistema de inteligência artificial, em 2023, quanto à capacidade de filosofar: “M. O.: ¿Creés que las inteligencias artificiales son capaces de filosofar? ChatGPT: Las inteligencias artificiales están diseñadas para procesar y analizar información de manera eficiente, pero actualmente carecen de la capacidad de razonamiento abstracto, intuición y experiencia subjetiva que son fundamentales en la filosofía humana. La filosofía implica explorar y reflexionar sobre preguntas fundamentales relacionadas con la existencia, la ética, la realidad, el conocimiento y otros temas complejos. Aunque las inteligencias artificiales pueden ayudar en la recopilación y análisis de información filosófica, así como en la generación de argumentos basados en datos, no poseen una comprensión profunda ni una perspectiva filosófica intrínseca. Su conocimiento se basa en patrones y datos previos y no en la experiencia personal o en la reflexión crítica. La filosofía humana implica una dimensión ética, emocional y subjetiva que aún no ha sido replicada en las inteligencias artificiales. Si bien es posible que en el futuro se desarrollen IA con capacidades más avanzadas, como la emulación de procesos cognitivos complejos, actualmente no existen IA que sean capaces de filosofar en el sentido completo y profundo en que los seres humanos lo hacen.”

48 “Machines cannot really possess mental states. They might, however, exhibit states that are functionally equivalent to mental states although they are not associated with phenomenal consciousness and have only derived intentionality mediated by their programmers and users. One might call such states quasi-beliefs, quasi-desires, etc.” (CATRIN MISSELHORN, “Artificial Moral Agents. Conceptual Issues and Ethical Controversy”, *op. cit.*, pp. 39).

49 J. H. MOOR, “The Nature, Importance, and Difficulty of Machine Ethics”, *op. cit.*, p. 21.

50 “Rats detect the causal relationship between their

45 *Liability for Artificial Intelligence and other emerging digital technologies*, European Commission, Expert Group on Liability and New Technologies New Technologies Formation, 2019, p. 23.

46 J. H. MOOR, “The Nature, Importance, and Difficulty of Machine Ethics”, in *IEEE Intelligent Systems*, vol. 21, no. 4, July-Aug. 2006, pp. 18 to 21.

certainly will not be hard to admit, likewise, that AI systems are also capable of performing action-creating effects with the purpose of obtaining them. This will always make AI systems closer to the human beings than, for instance, a legal person⁵¹. Although, in some cases, it may be endowed with a true social substrate, it still is, in its essence, only a form⁵².

actions and associated outcomes by psychological processes analogous to those mediating the acquisition of simple instrumental beliefs in humans, the content of which can be reported in the form of a judgment. To the extent that this is so, rats at least have some of the necessary processes for causal judgment" (CECILIA HEYES and ANTHONY DICKINSON, "The Intentionality of Animal Action", in *Mind & Language*, vol. 5, no. 1, 1990, p. 100. It is even conceivable to equate the collective action of human animals with that of non-human animals: "Collectively acting is widespread in the non-human animal world. Some of this collective action is merely distributive, but some of it is very likely also shared or joint, in the senses in which I have introduced those terms. This does not itself imply that collective intentionality has a corresponding range, but it does mean that the very same explanationist motivation for positing collective intentionality in human groups applies to non-human animals" (Robert A. Wilson, "Collective Intentionality in Non-Human Animals", in Marija Jankovic and Kirk Ludwig (ed.), *Routledge Handbook on Collective Intentionality*, New York, 2017, p. 430).

51 "The inability to pinpoint specific human responsibility for failure suggests that «the machine» or the network «system» should be blamed for damage it causes. The temptation to treat sophisticated intelligent agents as independent legal entities, thus absolving the humans involved, is powerful. The agents appear to be autonomous and «independent»; their pathological results are by definition unpredictable. No human will have done anything that specifically caused harm, and thus no one should be liable for it. Just as we are not liable for the consequences of a human agent's unforeseeable pathological actions, so too humans should be absolved of liability for the unforeseen results of machine intelligence's pathology" (CURTIS E. A. KARNOW, "Liability for Distributed Artificial Intelligences", *op. cit.*, p. 189).

52 Cf. v.g. Hugo Ramos Alves, "A pessoa coletiva entre a ficção e o realismo", *Revista de Direito Comercial, Liber Amicorum Pedro Pais de Vasconcelos*, 2021, pp. 619/620: "Se, por um lado, a fórmula da ficção apresenta uma excelente base para justificar a personificação, pois, em rigor, a pessoa coletiva apenas surge *ope legis*, na sequência do respetivo reconhecimento, esta conceção, per se, prima, formalmente, pela neutralidade. Assim, a tentativa de encontrar a realidade, seja pela via orgânica, seja pela via da fórmula neutra da idealidade jurídica, aponta para o facto de a pessoa coletiva ter uma dimensão fáctica. Em qualquer caso, os termos em que ambas as construções são contrapostas radica na circunstância de ambas serem incapazes de explicar satisfatoriamente

References

ALVES, Hugo Ramos. A pessoa coletiva entre a ficção e o realismo. *Revista de Direito Comercial, Liber Amicorum Pedro Pais de Vasconcelos*, 2021, pp. 619/620.

BENVENISTE, Emile. *O homem na linguagem*. Trad. Isabel Maria Lucas Pascoal. Lisboa : Veja, 1992.

BRASIL. Câmara dos Deputados. *Projeto de Lei 522/22*. Portal da Câmara dos Deputados. 2022. Disponível em: <https://www.camara.leg.br/propostas-legislativas/2317524>. Acesso em 05 Jun. 2024.

CATRIN MISSELMHORN, *Artificial Moral Agents, Conceptual Issues and Ethical Controversy*, in *The Cambridge Handbook of Responsible Artificial Intelligence – Interdisciplinary Perspectives*, Part I – Foundations of Responsible IA, Cambridge University Press, 2022, p. 31).

CATRIN MISSELMHORN, *Artificial Moral Agents. Conceptual Issues and Ethical Controversy*, *op. cit.*, pp. 32 e 33.

CATRIN MISSELMHORN, *Artificial Moral Agents. Conceptual Issues and Ethical Controversy*, *op. cit.*, pp. 39).

CECILIA HEYES e ANTHONY DICKINSON, *The Intentionality of Animal Action*, in *Mind & Language*, vol. 5, n.º 1, 1990, p. 100.

CHAHAL AYUSHI e GULIA PREETI, *Machine Learning and Deep Learning*, in *International Journal of Innovative Technology and Exploring Engineering*, n.º8, 2019, pp. 4910-4914.

CHILE. Ministerio de Ciencia, Tecnología, Conocimiento e Innovación (2021) – *Ley 21383*. Biblioteca del Congreso Nacional de Chile : Chile. Disponível em: <https://www.bcn.cl/leychile/>

a multiplicidade de situações enquadráveis na noção de pessoa coletiva. A isto acresce o facto de, na raiz, a alteridade imanente à personificação, no campo comercial, assentar na ideia de limitação de responsabilidade associada a uma determinada entidade legal: a sociedade, aspeto que demonstra não ser possível dissociar a pessoa coletiva do respetivo escopo e, acima de tudo, das finalidades subjacente, a que o ordenamento jurídico dá guarida".

[navegar/imprimir?idNorma=1166983&idVersao=2021-10-25](#). Acesso em: 05 Jun. 2024.

COLIN ALLEN, WENDELL WALLACH e IVA SMIT, *Why Machine Ethics?*, in *IEEE Intelligent System*, vol. 21, n.º 4, 2006, p. 12.

COMISSÃO EUROPEIA. *Liability for Artificial Intelligence and other emerging digital technologies*, European Commission, Expert Group on Liability and New Technologies New Technologies Formation, 2019, pág. 23.

CURTIS E. A. KARNOW, *Liability for Distributed Artificial Intelligences*, *op. cit.*, p. 189).

DAMÁSIO, António. *Sentir & Saber: a caminho da consciência*. 1.ª ed. Temas e Debates, Círculo de Leitores: Lisboa, 2020. P. 63 ss.

EVGENY MOROZOV, *The problem with artificial intelligence? It's neither artificial nor intelligent*, in *The Guardian*, 30/03/2023, disponível em www.theguardian.com/commentisfree/2023/mar/30/artificial-intelligence-chatgpt-human-mind.

FROMKIN, Victoria; RODMAN, Robert. *Introdução à linguagem*. Trad. Isabel Casanova. Coimbra : Almedina, 1993.

HALLINAN, D., P. Schütz, M. Friedewald and P. de Hert 2014. Neurodata and Neuroprivacy: Data Protection Outdated? *Surveillance & Society*. 12(1): 55-72. ISSN: 1477-7487. Acesso em: 5 Jun. 2024. Disponível em: <http://www-surveillance-and-society.org>

ISAAC ASIMOV, *Runaround*, in *I, Robot*, Gnome Press, 1950, p. 27.

J. H. MOOR, *The Nature, Importance, and Difficulty of Machine Ethics*, in *IEEE Intelligent Systems*, vol. 21, n.º 4, July-Aug. 2006, pp. 18 a 21.

J. H. MOOR, *The Nature, Importance, and Difficulty of Machine Ethics*, *op. cit.*, p. 21.

ORENSANZ, Martín. *Inteligencia Artificial e Filosofía. IECE Revista Digital*. Año VIII, 15 (Jul. 2023) 50-53. ISSN 2545-6326. Disponível em: <http://iece-argentina.weebly.com/>. Acesso em: 05 Jun. 2024.

PAIS D'AGUIAR, Filipa (2024) – *Direitos Humanos e Inteligência Artificial: principais dimensões jurídicas, éticas, sociais e culturais, no contexto europeu e transnacional*. In Carlos Ajjonso Leony Neto, Leandro de Matos Coutinho, Patricia Desirée de Mello (Coord.) – *Mentes digitais : do zero ao infinito: Homenagem à Professora Fernanda Duarte*. Prejácio de Guilherme Calmou Nogueira da Gama. São Paulo : Editora GZ. ISBN 978-65-5813-116-8. 242 p..

PARLAMENTO EUROPEU. [Proposta de Regulamento do Parlamento Europeu e do Conselho que estabelece regras harmonizadas em matéria de inteligência artificial (Regulamento Inteligência Artificial), COM(2021) 206 final].

PARLAMENTO EUROPEU. Direção-Geral da Comunicação. *Regulamento Inteligência Artificial: Parlamento aprova legislação histórica*. 2024. Disponível em: <https://www.europarl.europa.eu/news/pt/press-room/20240308IPR19015/regulamento-inteligencia-artificial-parlamento-aprova-legislacao-historica>. Acesso em: 4 Jun. 2024.

PARLAMENTO EUROPEU. *Regulamento Inteligência Artificial*, P9_TA(2024)0138, Resolução legislativa do Parlamento Europeu, de 13 de março de 2024, sobre a proposta de regulamento do Parlamento Europeu e do Conselho que estabelece regras harmonizadas em matéria de inteligência artificial (Regulamento Inteligência Artificial) e altera determinados atos legislativos da união (COM(2021)0206 – C9-0146/2021 – 2021/0106(COD)). 2024. Disponível em: https://www.europarl.europa.eu/doceo/document/TA-9-2024-03-13_PT.html#sdocta2. Acesso em: 05 Jun. 2024.

RAGNAR FJELLAND, *Why general artificial intelligence will not be realized*, in Humanities and Social Sciences Communications, 2020, p. 2.

Robert A. Wilson, *Collective Intentionality in Non-Human Animals*, in Marija Jankovic and Kirk Ludwig (ed.), *Routledge Handbook on Collective Intentionality*, New York, 2017, p. 430).

ROVELLI, Carlo. *Sete breves lições de física*. Trad. Vasco Gato. 1.^a ed., 3.^a republ. Penguin Random House: Lisboa, 2022.

SIGMAN, Mariano. *La vida secreta de la mente: nuestro cérebro cuando decidimos, sentimos y pensamos*. (Debates). 1.^a ed. Penguin Random House: Buenos Aires, Barcelona, 2015.

TOUMEY, Chris. “Singularity”. *Encyclopaedia Britannica*, 24 Jun. 2024. Disponível em: <https://www.britannica.com/technology/singularity-technology>. Acesso em: 12 Jul.2024

UNESCO. Recommendation on the Ethics of Artificial Intelligence. 2022. Disponível em: <https://unesdoc.unesco.org/ark:/48223/pf0000381137>. Acesso em: 25 Maio 2023.

WAJNERMAN PAZ, A. Is Your Neural Data Part of Your Mind? Exploring the Conceptual Basis of Mental Privacy. *Minds & Machines* 32, 395–415 (2022). <https://doi.org/10.1007/s11023-021-09574-7>. Acesso em: 05 Jun. 2024.

YUVAL NOAH HARARI, *21 Lessons for the 21st Century*, in Vintage Publishing, 2019, p. 19 a 20.

