Ciência da Informação *Express*

ISSN 2675-8717

https://doi.org/10.60144/v5i.2024.120

BRIEF CONSIDERATIONS ON ETHICS FOR ARTIFICIAL INTELLIGENCE NEUROTECHNOLOGIES

Breves considerações sobre ética para neurotecnologias de inteligência artificial

Sthéfano Divino
Doutor e Mestre em Direito Privado pela Pontifícia Universidade Católica de Minas. Professor Adjunto de Direito Civil do Curso de Direito da Universidade Federal de Lavras, Lavras, Minas Gerais, Brasil.
sthefanodivino@ufla.br
https://orcid.org/0000-0002-9037-0405

ABSTRACT

Introduction: implants and technological devices are being used to decode neural activity to move a prosthetic arm, control an avatar, and turn thoughts into text through an AI-based decoder. These situations are designed by Brain-computer Interface (BCI), one of the main AI-based neurotechnologies used to understand the brain and to improve people's welfare. In 2023, UNESCO already recognized its benefits but also revealed the potential ethical issues and problems, particularly with its use of non-invasive interventions. **Objective:** so, this essay aims to answer the following research question: Which Ethical standards can be designed and used to balance the person's rights with technological development to prevent vulnerability situations? **Method:** the methods used in this work is the bibliographic research plus the hermeneutic interpretation. **Results:** it proposes Ethical standards for protecting the rights of the vulnerable to ensure that these rights are respected. **Conclusions:** there is no need for the creation of a new neurorights. Privacy and intimacy can and will deal with all the issues of neurotechnologies. However, it is necessary to improve the protection of the owner's rights through strong ethical and governance standards.

keywords: Neurotechnology; Artificial intelligence; Ethics; Governance; Brain-computer interface.

RESUMO

Introdução: implantes e dispositivos tecnológicos estão sendo usados para decodificar a atividade neural para mover um braço protético, controlar um avatar e transformar pensamentos texto por meio de um decodificador baseado em IA. Essas situações são projetadas pela Interface Cérebro-Computador (ICC), uma das principais neurotecnologias baseadas em IA usadas para entender o cérebro e melhorar o bem-estar das pessoas. **Objetivo:** em 2023, a UNESCO já reconheceu seus benefícios, mas também revelou as possíveis questões e problemas éticos, especialmente com o uso de intervenções não invasivas. Portanto, este breve

ensaio tem como objetivo responder o seguinte problema de pesquisa: Quais padrões éticos podem ser propostos e usados para equilibrar os direitos da pessoa com o desenvolvimento tecnológico para evitar situações de vulnerabilidade? **Método:** os métodos utilizados neste trabalho são a pesquisa bibliográfica e o método hermenêutico. **Resultados:** propõe-se e analisa-se alguns princípios éticos para a proteção dos direitos dos pacientes vulneráveis objetivando garantir que esses direitos sejam observados. **Conclusões:** verifica-se que não há necessidade de criação de novos neurodireitos. Aparentemente, a privacidade e a intimidade e os outros princípios éticos e jurídicos podem e conseguem lidar com todas as questões atuais das neurotecnologias. Entretanto, é necessário aprimorar a proteção dos direitos do titular por meio de fortes padrões éticos e de governança.

Palavras-chave: Neurotecnologia; Inteligência artificial; Ética; Governança; Interface cérebro-computador.

1 INTRODUCTION

Due to the advances in Neurotechnology[1], companies can identify brain patterns related to specific behavior through AI[2] algorithms (SANI; PESARAN; SHANECHI, 2024). Implants and technological devices are being used to decode neural activity to move a prosthetic arm (Tam et al., 2019), control an avatar (GALIAUTDINOV; MKRTTCHIAN, 2020), and turn thoughts into text through an AI-based decoder (DEVLIN, 2023).Furthermore, the Neuralink's first human patient was able to control the mouse through thinking (REUTERS, 2024).

These uses/tools raise many legal and ethical questions. In 2023, UNESCO already recognized its benefits but also revealed the potential ethical issues and

[1] "Neurotechnology is defined as the assembly of methods and instruments that enable a direct connection of technical components with the nervous system. These technical components are electrodes, computers, or intelligent prostheses. They are meant to either record signals from the brain and "translate" them into technical control commands, or to manipulate brain activity by applying electrical or optical stimuli. Closed-loop interactions of readout and stimulation systems (control circuits) are subject of current research as well. In the following, we would like to offer some insight into the current state of basic and applied research, and possible clinical applications resulting from it. We will also address some of the ethical issues that emerge in the context of neurotechnology and describe some ongoing interdisciplinary research on brain-machine interfaces" (Muller; Rotter, 2017, p. 93).

[2] "We define AI as the study of agents that receive percepts from the environment and perform actions. Each such agent implements a function that maps percept sequences to actions, and we cover different ways to represent these functions, such as reactive agents, real-time planners, and decision-theoretic systems" (Russell; Norvig, 2010, p. VIII).

problems, particularly with its use of non-invasive interventions. Ethics[3] and biomedical Ethics[4] are some trending topics on Artificial Intelligence Neurotechologies, which “has the potential to solve many health issues, but it could also threaten human rights, freedom of thought and privacy” (AZOULAY, 2024). According to Gilbert and Russo (2024, p. 855), the issues discussed frequently include mental privacy (SHEN, 2013), mental freedom (BUBLITZ, 2016), and personhood (SOLUM, 2020).Furthermore, Yuste et al. (2017, p. 160) propose four areas of concern: privacy and consent; agency and identity; augmentation; and bias.Many questions arise about the use of AI in neuroscience. These include (Hildt; Laas; Sziron, 2020, p. 274):

> Data Concerns: Data management, data security, protection of personal data, surveillance, privacy, and informed consent. Algorithmic Bias and Discrimination: How to avoid bias and bias related problems? This points to questions of justice, equitable access to resources, and digital divide. Autonomy: When and how is AI autonomous, what are the characteristics of autonomous AI? How to develop rules for autonomous vehicles? Responsibility: Who is in control? Who is responsible or accountable for decisions made by AI? Questions relating to AI capabilities: Can AI ever be

[3] “[…] Ethics is the science that deals with conduct, in so far as this is considered as right or wrong, good or bad. A single term for conduct so considered is "moral conduct," or the "moral life”. […] The terms "ethics" and "ethical" are derived from a Greek word ethos which originally meant customs, usages, especially those belonging to some group as distinguished from another, and later came to mean disposition, character. They are thus like the Latin word "moral," from mores, or the German sittlich, from Sitten. As we shall see, it was in customs, "ethos," "mores," that the moral or ethical began to appear. For customs were not merely habitual ways of acting; they were ways approved by the group or society. To act contrary to the customs of the group brought severe disapproval. This might not be formulated in precisely ourterms—right and wrong, good and bad,—but the attitude was the same in essence. The terms ethical and moral as applied to the conduct of to-day imply of course a far more complex and advanced type of life than the old words "ethos" and "mores," just as economics deals with a more complex problem than "the management of a household," but the terms have a distinct value if they suggest the way in which the moral life had its beginning” (Dewey; Tufts, 2022, p. III and IV).

[4] First introduced in 1979 by Childress and Beauchamp (1994), there are four main Principles of Biomedical Ethics: The principle of respect for autonomy; Nonmaleficence; Beneficence; and Justice.
“The principle of respect for autonomy has a variety of interpretations. In clinical ethics it is usually understood as a right of an individual patient or research subject to decide about themselves according to their own principles, which gives them also responsibilities for the possible outcomes” (Rus; Groselj, 2021, p. 113).
“The principle of nonmaleficence requires from a physician to not create harm to a patient or a research subject. It is one of the fundamental principles in medical ethics since Hippocrates and it is known by the maxim “Primum non nocere”; first, do no harm” (Rus; Groselj, 2021, p. 113).
“The principle of beneficence defines a fundamental mission of healthcare providers—to contribute to the welfare of their patients (Rus; Groselj, 2021, p. 115).
“The principle of justice is, in general, defined by two concepts: equitability and distributive justice” (Rus; Groselj, 2021, p. 115).

> conscious or sentient? What would conscious or sentient AI imply? Values and morality: How to build in values and moral decision-making to AI? Are moral machines possible? Should robots be granted moral status or rights?

So, according to this background, this essay aims to answer the following research question: Which Ethical standards can be designed and used to balance the person's rights with technological development to prevent vulnerability situations?

2 AI NEUROTECHNOLOGIES FORMEDICAL TREATMENT AND ENHANCEMENT

As we saw, artificial intelligence neurotechnologies can be used in many ways. But, in the end, they aim for improvements in human life. According to Savulescu, Sandberg, and Kahane (2011, p. 8),

> Enhancements include different kinds of improvements: 1. Medical treatment of disease. 2. Increasing natural human potential – Increasing a person's own natural endowments of capabilities within the range typical of the species homo sapiens, e.g. raising a person's IQ from 100 to 140. 3. Superhuman enhancements (sometimes called posthuman or transhuman) – Increasing a person's capabilities beyond the range typical for the species homo sapiens, e.g. giving humans bat sonar or the capacity to read minds.

We believe that has no difference with this thinking about AI-neurotechnologies.[5]Thus, any ethical pattern must consider the need and adequacy

[5] "The incorporation of neural-based technologies into psychiatry offers novel means to use neural data in patient assessment and clinical diagnosis. However, an over-optimistic technologisation of neuroscientifically-informed psychiatry risks the conflation of technological and psychological norms. Neurotechnologies promise fast, efficient, broad psychiatric insights not readily available through conventional observation of patients. Recording and processing brain signals provides information from 'beneath the skull' that can be interpreted as an account of neural processing and that can provide a basis to evaluate general behaviour and functioning. But it ought not to be forgotten that the use of such technologies is part of a human practice of neuroscience informed psychiatry" (Rainey; Erden, 2020, p. 2439).
"From better diagnosis, prognosis, and prevention to more precise surgical procedures, AI has the potential to offer unique opportunities to enhance patient care and improve clinical practice overall" (Jotterand; Bosco, 2020, p. 2455).
"Clinical neuroscience is increasingly relying on the collection of large volumes of differently structured data and the use of intelligent algorithms for data analytics. In parallel, the ubiquitous collection of unconventional data sources (e.g. mobile health, digital phenotyping, consumer neurotechnology) is increasing the variety of data points. Big data analytics and approaches to Artificial Intelligence (AI) such as advanced machine learning are showing great potential to make sense of these larger and heterogeneous data flows. AI provides great opportunities for making new discoveries about the brain, improving current preventative and diagnostic models in both neurology and psychiatry and developing more effective assistive neurotechnologies. Concurrently, it raises many new methodological and ethical challenges. Given their transformative nature, it is still largely unclear how AI-driven approaches to the study of the human brain will meet adequate standards of scientific validity and affect normative instruments in neuroethics and research ethics" (Ienca; Ignatiadis, 2020, p. 77).

of AI for clinical use and medical treatment. First, if there is an efficient scientific method or medication for the clinical case, their use might be highly considered. This proposal focuses on the principles of beneficence and nonmaleficence due to the high uncertainty of AI technologies. Second, the use of AI neurotechnologies only should be recommended if there is any previous case attesting to his efficiency. If the technology is in test form, the physician should avoid his use and his recommendation. If there are no previous studies and tests, the appropriate method for its use in patients is not through clinical trials, but through research with human beings. Therefore, the ethical and legal rules are different. The proposal here focuses on clinical cases, not academic ones.

Thirdly, even if the technology is proven to be effective, its use depends exclusively on the patient's authorization. Therefore, based on a horizontal relationship between doctor and patient, the doctor should actively listen to the patient's preferences and respect them if they don't want to use the technology.

Finally, if the patient accepts the use of AI neurotechnology, the physician must inform them of all the risks of its use, as well as the possibilities of success or failure, and obtain their free and informed consent.

On the other side is using neurotechnologies for human enhancement[6]. It is understood that this is a slightly more sensitive relationship, as the patient seeks out the physician not for clinical purposes but to enhance an existing quality or skill.Therefore, greater care is required. The physician should only authorize the use of this type of technology once he has verified that: a) the risks are minimal; b) the patient is aware of the risks and issues that may happens during treatment; c) the patient has been adequately informed of the risks and has consciously assumed them; d) the benefits outweigh the risks and harms; e) there is no total or partial

[6] “The technological promise of “human enhancement” is on the one hand as old as Bacon’s The New Atlantis, and on the other hand, quite a recent concern in the ethics debate, especially after the development of recombinant DNA technologies suitable for genetic engineering. The directed and tailored modification of human genetic material in human individuals (germ line), pharmaceuticals and machines (prosthetic limbs and organs) that increase or improve physical, sensory and/or mental capabilities of humans, devices that establish a functional brain-computer interface, with bidirectional communication are only some examples” (Keskinbora, 2019, p. 280)

permanent reduction in limbs or biological function; and f) constant medical monitoring must be carried out.

Note that there is no substantial difference in clinical or enhancement use. While clinical use is needed, enhancement is also an option. However, both require all information and care to instruct the patient about the risks. The main difference relies on enhancement use should not cause any damage for the patient (non-maleficence) and only should be used and recommended as the last option. Now, we should see which ethical principles we should use and suggest for AI neurotechnologies

3 ETHICAL PRINCIPLES FOR ARTIFICIAL INTELLIGENCE NEUROTECHNOLOGIES

Although ethical standards are constantly being discussed today (Friedrich; Wolkenstein, 2021; Liao, 2020; Doya *et al*., 2022; Miśkiewicz, 2019), this discussion is not new (Wallach; 2011; Coenen, 2010; Bostrom, 2005). The authors demonstrate considerable concerns in various areas ranging from data protection, agency, intimacy, etc. According to Berger and Rossi (2021), AI Neurotechnologies must respect at least mental privacy, Human Agency and autonomy, Human identity, fairness, accuracy, transparency, Security, and Well-being. I will make some commentaries and discussions about each of them.

First, mental privacy is a concept hard to define. What is the difference between privacy and mental privacy? Does the law treat them differently? What is the object of protection for each of these rights? Is privacy mental only because it comes from the brain? (Arstila; Scott, 2011). Privacy can no longer be seen as the right to be alone (Warren; Brandeis, 1890), but as the power to control the information that enters and leaves a person's subjective sphere (Rodotà, 2004; 2005; 2009). Therefore, the right to privacy object of protection is the physical environment where the person is. Privacy protects the person from being violated externally by others. Take, for example, a person who is inside their home and is being watched by a stalker. Another example is a person who is on a bus and fiddling with their smartphone and a standing passenger looks at their chats. In both situations, the

right to privacy is violated because the subject has entered the physical/private sphere of another. Therefore, the right to privacy does not seem to be adequate to protect the holder of neurotechnologies using artificial intelligence.

The right to intimacy seems more appropriate for the subject under discussion. Firstly, because intimacy is subjective. It protects the subject's entire psychological apparatus and not just the physical environment, such as privacy. While privacy seems to be limited to the entry and exit of information from the physical environment, intimacy is concerned with protecting moments of vulnerability of its holder. Take, for example, a person who is at home and has their right to privacy protected. However, their bath and intimate moments are more subjective than simply being at home. Furthermore, an intimate relationship with a husband or wife is more private than simply being seen watching a stream in your TV room. Both behaviors are frowned upon, but realize that intimacy goes beyond the physical environment and violates the subjective aspects of the person.

Therefore, AI Neurotechnologies should not only restrict itself to the information that is taken from the environment, but also avoid collecting data, images, or using any other form of collection that is capable of violating the intimacy of its owner.

The second and third ethical principles are directly linked: Human Agency, and Human identity. Firstly, any device capable of emitting electrical waves and altering the user's behavior should be used with extreme caution. This cautious use is due to the possibility of affecting the patient's behavioral patterns. Therefore, in the interests of beneficence and non-maleficence, the use of neurotech AI should be suspended when the slightest change in the patient's behavior or identity has been verified. However, the physician may encounter a hard case: what to do if discontinuing the use of the technology causes the patient's clinical condition to worsen?In this case, the physician should inform the patient of the situation and allow them to make a free and informed choice: to continue with the treatment, with possible changes to their

agency and identity, or to suspend the treatment, with the suggestion of other clinical methods that are as effective as the technology used.[7]

The principle of fairness proposes that technologies are used in a fair, and equal way and that “systems do not present biases or discriminatory behavior that may harm certain social groups or commit moral violations” (Cortiz, 2023, p. 769). In other words, the use of AI neurotech should enable control over its use, as well as balancing the positive and negative effects where they exist.

Accuracy demands that the use of AI neurotech be precise for the right case. Off-label use should be avoided. Therefore, in clinical cases, this type of technology is recommended only when it has been developed for this purpose. It should be noted that accuracy in this case has an instrumental and useful meaning. It must be appropriate for the disease it is being used to treat. And it is possible. One example is the study by Metin et al. (2024) which demonstrated the efficiency of using Deep Learning to diagnose and treat treatment-resistant depression (TRD). As a result, “GoogleNet classified the healthy controls and non-TRD group with 88.43%, the healthy controls and TRD subjects with 89.73%, and the TRD and non-TRD group with 90.05% accuracy. The external validation accuracy for the TRD-non-TRD classification was 73.33%. Finally, the CAM analysis revealed that the TRD group contained dominant features in class detection of deep learning architecture in almost all electrodes” (Metin *et al*., 2024, p. 1). It is therefore recommended to use AI neurotech only when there is a high chance or probability that the results will be positive and when there is compatibility between the design and the purpose.

Transparency[8] is directly tied to the right to information in the doctor-patient relationship. Transparency is not only about how the AI works but also how the

[7] Furthermore, there is a serious concern about free will breach in AI technologies (Adomaitis; Grinbaum, 2024).

[8] “The origins of the transparency requirement in data protection law date to the 31st International Conference of Data Protection and Privacy Commissioners held in Madrid in November 2009, in which the importance of transparency to protect an individuals’ privacy was acknowledged. After being included in the proposal for the GDPR in 2012, the transparency principle made its way into the binding GDPR. Today, transparency is a core principle enshrined in Art. 5(1)(a) of the GDPR which states that personal data must be “processed lawfully, fairly and in a transparent manner in relation to the data subject,” thereby illustrating the close connection between transparency, lawfulness, and fairness. Art. 5(1)(a) of the GDPR, as the first of the core principles of data processing, is a “catch-all” provision, which is going to be typically called upon as a means of last resort if more concrete

physician instructs the patient about how it works. It must be seen from two perspectives. In the first situation, the doctor must inform the patient of all the known positive and negative effects. In addition, they should tell the patient how their life could improve or worsen its use. From the second perspective, the doctor must inform the patient about what data will be collected and how it will be processed by the doctor or the company responsible for developing the AI. It is a business relationship where the patient must understand what data is being collected and processed.Therefore, transparency brings with it explainability.[9] So, a neurotech AI is transparent when the user can identify how the inputs generated the outputs (Joyce *et al*., 2023).

Regarding Security, regulatory bodies should only authorize the use of AI technologies that are at least minimally acceptable as safe. Therefore, each country will establish safety standards, and Professional Councils will regulate the technical requirements for each professional. Security, then, is not just about device safety but also regulatory security, as they are directly connected. Thus, the use of neurotech AI is only authorized when it has been previously validated by competent government agencies. Furthermore, even if previously validated, we return to what has been stated: off-label use should be avoided.

Finally, regarding well-being, the ethical principle requires that the use of neurotech AI be aimed at improving the patient's quality of life. However, this improvement must be balanced with all other principles. Despite being the goal, improper use that violates other ethical principles cannot be justified to achieve this end. Thus, AI is a tool used in favor of humans, focusing on their subjectivity and aiming to enhance their living conditions.

principles are not applicable in a specific scenario. Failing to adhere to it can be punished with steep fines (cf. Art. 84 of the GDPR)" (Felzmann *et al*., 2019, p. 1).

[9] "Explainability methods can be deterministic or have gradations of stochasticity, and provide local (e.g. individual video frames) and/or global explanations (aggregate feature weights within a model). The methods may be algorithm-specific or agnostic, and be intrinsic to the model or require additional processing. The core methodologies of most interpretability approaches will nevertheless seem familiar to experimentalists: features undergo some class of iterative ablation or permutations, the change in performance after such manipulation is evaluated, and the manipulation's impact is succinctly summarized. How the permutation is performed, evaluated, and summarized differs between methods with varying relatedness to human processes and intuition for complex outcomes" (Goodwin, 2022, p. 102544)

4 CONCLUSIONS

It is noticeable that many ethical issues already present in the physician-patient relationship are replicated in the relationship where AI is used in neurotechnologies. However, new challenges arise from the possibility of intrusion into agency, identity, intimacy, and the collection of personal data. These are the main emerging problems in this relationship. Therefore, as suggested, it is clear that the use of neurotech AI must be approached with caution. It is not possible to prioritize one principle over another. All are important and must be upheld when this type of technology is used. Thus, the study conducted demonstrates the need for a deeper approach while respecting the existing legal and ethical premises to avoid proposing unnecessary theories or instruments. Throughout the approach, it is evident that the current ethical and legal system can handle the issues involving neurotech AI. There is no need to create a new ethical or legal framework, as the technology is merely a utilitarian extension of existing professions and issues.

REFERENCES

ADOMAITIS, Laurynas; GRINBAUM, Alexei. Neurotechnologies, ethics, and the limits of free will. **Integrative Psychological and Behavioral Science**, Nova York, p. 1-14, 2024. Available in: https://link.springer.com/article/10.1007/s12124-023-09782-3. Accessed in: 25 sept. 2024.

ARSTILA, Valtteri; SCOTT, Franklin. Brain reading and mental privacy. **Trames: A Journal of the Humanities and Social Sciences**, Tallinn, v. 15, n. 2, p. 204, 2011. Available in: https://www.kirj.ee/trames-2011-vol-15-issue-2/. Accessed in: 25 sept. 2024.

AZOULAY, Audrey. **The Ethics of Neurotechnology**: UNESCO appoints international expert group to prepare a new global standard. UNESCO. 2024. Available in: https://www.unesco.org/en/articles/ethics-neurotechnology-unesco-appoints-international-expert-group-prepare-new-global-standard. Accessed in: 11 sept. 2024.

BEAUCHAMP, Tom L.; CHILDRESS, James F. **Principles of biomedical ethics**. São Paulo: Edições Loyola, 1994.

BERGER, Sara; ROSSI, Francesca. The future of AI ethics and the role of neurotechnology. In: WORKSHOP ON ADVERSE IMPACTS AND COLLATERAL EFFECTS OF ARTIFICIAL INTELLIGENCE TECHNOLOGIES, 2021. **CEUR-WS**, 2021. Available in: https://research.ibm.com/publications/the-future-of-ai-ethics-and-the-role-of-neurotechnology. Accessed in: 11 sept. 2024.

BOSTROM, Nick. A history of transhumanist thought. **Journal of Evolution and Technology**, Hartford, v. 14, n. 1, 2005. Available at: https://jetpress.org/volume14/bostrom.pdf. Accessed on: 25 Sept. 2024.

BUBLITZ, Christoph. Moral enhancement and mental freedom. **Journal of Applied Philosophy**, Oxford, v. 33, n. 1, p. 88-106, 2016. Available at: https://onlinelibrary.wiley.com/doi/abs/10.1111/japp.12172. Accessed on: 25 Sept. 2024.

COENEN, Christopher. Deliberating visions: The case of human enhancement in the discourse on nanotechnology and convergence. In: **Governing Future Technologies: Nanotechnology and the Rise of an Assessment Regime**. Heidelberg: Springer, 2010. p. 73-87. Available at: https://link.springer.com/chapter/10.1007/978-90-481-2834-1_6. Accessed on: 25 Sept. 2024.

CORTIZ, Diogo. A narrative review of fairness and morality in neuroscience: insights to artificial intelligence. **AI and Ethics**, New York, v. 3, n. 3, p. 769-780, 2023. Available at: https://link.springer.com/article/10.1007/s43681-022-00147-6. Accessed on: 25 Sept. 2024.

DEVLIN, H. AI makes non-invasive mind-reading possible by turning thoughts into text. **The Guardian**, Londres, 2023. Available at: https://www.theguardian.com/technology/2023/may/01/ai-makes-non-invasive-mind-reading-possible-by-turning-thoughts-into-text. Accessed on: 11 Sept. 2024.

DEWEY, John; TUFTS, James Hayden. **Ethics**. Chicago: DigiCat, 2022.

DOYA, Kenji et al. Social impact and governance of AI and neurotechnologies. **Neural Networks**, Oxford, v. 152, p. 542-554, 2022.

FELZMANN, Heike et al. Transparency you can trust: Transparency requirements for artificial intelligence between legal norms and contextual concerns. **Big Data & Society**, Londres, v. 6, n. 1, p. 2053951719860542, 2019. Available at: https://journals.sagepub.com/doi/10.1177/2053951719860542. Accessed on: 25 Sept. 2024.

FRIEDRICH, Orsolya; WOLKENSTEIN, Andreas. Introduction: Ethical issues of neurotechnologies and artificial intelligence. In: FRIEDRICH, Orsolya;

WOLKENSTEIN, Andreas. **Clinical Neurotechnology Meets Artificial Intelligence: Philosophical, Ethical, Legal and Social Implications**, Nova York, p. 1-9, 2021. Available at: https://link.springer.com/chapter/10.1007/978-3-030-70942-7_1. Accessed on: 25 Sept. 2024.

GALIAUTDINOV, Rinat; MKRTTCHIAN, Vardan. Brain machine interface for avatar control and estimation for educational purposes based on neural AI plugs: theoretical and methodological aspects. *In:* GALIAUTDINOV, Rinat; MKRTTCHIAN, Vardan. **Avatar-Based Control, Estimation, Communications, and Development of Neuron Multi-Functional Technology Platforms**. Hershey: IGI Global, 2020. p. 294-316. Available at: https://www.igi-global.com/chapter/brain-machine-interface-for-avatar-control/244793. Accessed on: 25 Sept. 2024.

GILBERT, Frederic; RUSSO, Ingrid. Mind-reading in AI and neurotechnology: evaluating claims, hype, and ethical implications for neurorights. **AI and Ethics**, New York, p. 1-18, 2024. Available at: https://link.springer.com/article/10.1007/s43681-023-00195-x. Accessed on: 25 Sept. 2024.

GOODWIN, Nastacia L. et al. Toward the explainability, transparency, and universality of machine learning for behavioral classification in neuroscience. **Current Opinion in Neurobiology**, Oxford, v. 73, p. 102544, 2022. Available at: https://www.sciencedirect.com/science/article/pii/S095943882200084X. Accessed on: 25 Sept. 2024.

HILDT, Elisabeth; LAAS, Kelly; SZIRON, Monika. Shaping ethical futures in brain-based and artificial intelligence research. **Science and Engineering Ethics**, Dordrecht, v. 26, p. 2371-2379, 2020. Available at: https://link.springer.com/article/10.1007/s11948-020-00245-4. Accessed on: 25 Sept. 2024.

IENCA, Marcello; IGNATIADIS, Karolina. Artificial intelligence in clinical neuroscience: methodological and ethical challenges. **AJOB Neuroscience**, Abingdon, v. 11, n. 2, p. 77-87, 2020. Available at: https://www.tandfonline.com/doi/full/10.1080/21507740.2020.1736914. Accessed on: 25 Sept. 2024.

JOTTERAND, Fabrice; BOSCO, Clara. Keeping the “human in the loop” in the age of artificial intelligence: accompanying commentary for “correcting the brain?” by Rainey and Erden. **Science and Engineering Ethics**, Dordrecht, v. 26, n. 5, p. 2455-2460, 2020. Available at: https://link.springer.com/article/10.1007/s11948-019-00134-w. Accessed on: 25 Sept. 2024.

JOYCE, Dan W. et al. Explainable artificial intelligence for mental health through transparency and interpretability for understandability. **Digital Medicine**, Londres, v. 6, n. 1, p. 6, 2023. Available at: https://www.nature.com/articles/s41746-022-00735-3. Accessed on: 25 Sept. 2024.

KESKINBORA, Kadircan H. Medical ethics considerations on artificial intelligence. **Journal of Clinical Neuroscience**, Oxford, v. 64, p. 277-282, 2019. Available at: https://www.sciencedirect.com/science/article/abs/pii/S0967586819303210. Accessed on: 25 Sept. 2024.

LIAO, S. Matthew. A short introduction to the ethics of artificial intelligence. In: LIAO, S. Matthew (Ed.). **Ethics of Artificial Intelligence**. Oxford: Oxford University Press, 2020, p. 1-45. Available at: https://global.oup.com/academic/product/ethics-of-artificial-intelligence-9780190905033. Accessed on: 25 Sept. 2024.

METIN, Sinem Zeynep et al. Deep learning-based artificial intelligence can differentiate treatment-resistant and responsive depression cases with high accuracy. **Clinical EEG and Neuroscience**, Nova York, p. 15500594241273181, 2024. Available at: https://journals.sagepub.com/doi/abs/10.1177/15500594241273181. Accessed on: 25 Sept. 2024.

MIŚKIEWICZ, Julia. The merger of natural intelligence with artificial intelligence, with a focus on Neuralink company. **Virtual Economics**, Warsaw, v. 2, n. 3, p. 22-29, 2019. Available at: https://virtual-economics.eu/index.php/VE/article/view/24. Accessed on: 25 Sept. 2024.

MÜLLER, Oliver; ROTTER, Stefan. Neurotechnology: current developments and ethical issues. **Frontiers in Systems Neuroscience**, Lausanne, v. 11, p. 93, 2017. Available at: https://www.frontiersin.org/articles/10.3389/fnsys.2017.00093/full. Accessed on: 25 Sept. 2024.

RAINEY, Stephen; ERDEN, Yasemin J. Correcting the brain? The convergence of neuroscience, neurotechnology, psychiatry, and artificial intelligence. **Science and Engineering Ethics**, Dordrecht, v. 26, n. 5, p. 2439-2454, 2020. Available at: https://link.springer.com/article/10.1007/s11948-020-00254-3. Accessed on: 25 Sept. 2024.

REUTERS. Neuralink's first human patient able to control mouse through thinking, Musk says. **Reuters**, Londres, 2024. Available at: https://www.reuters.com/business/healthcare-pharmaceuticals/neuralinks-first-human-patient-able-control-mouse-through-thinking-musk-says-2024-02-20/. Accessed on: 11 Sept. 2024.

RODOTÀ, Stefano. Data protection as a fundamental right. In: **Reinventing Data Protection?** Dordrecht: Springer Netherlands, 2009. p. 77-82. Available at: https://link.springer.com/chapter/10.1007/978-1-4020-9498-9_5. Accessed on: 25 Sept. 2024.

RODOTÀ, Stefano. Persona, libertà, tecnologia - Note per una discussione. **Diritto & Questioni Pubbliche**, Palermo, v. 5, p. 25, 2005. Available at: http://www.dirittoequestionipubbliche.org/page/2005_n5.php. Accessed on: 25 Sept. 2024.

RODOTÀ, Stefano. Tra diritti fondamentali ed elasticità della normativa: il nuovo codice sulla privacy. **Europa e Diritto Privato**, Milão, v. 1, p. 1-11, 2004. Available at: https://www.iusexplorer.it/Riviste/DettaglioRivista?idRivista=119. Accessed on: 25 Sept. 2024.

RUS, Meta; GROSELJ, Urh. Ethics of vaccination in childhood—A framework based on the four principles of biomedical ethics. **Vaccines**, Basel, v. 9, n. 2, p. 113, 2021. Available at: https://www.mdpi.com/2076-393X/9/2/113. Accessed on: 25 Sept. 2024.

RUSSELL, S.; NORVIG, P. **Artificial Intelligence:** a modern approach. New Jersey: Pearson Education, 2010.

SANI, Omid G.; PESARAN, Bijan; SHANECHI, Maryam M. Dissociative and prioritized modeling of behaviorally relevant neural dynamics using recurrent neural networks. **Nature Neuroscience**, Londres, p. 1-13, 2024. Available at: https://www.nature.com/articles/s41593-023-01294-8. Accessed on: 25 Sept. 2024.

SAVULESCU, Julian; SANDBERG, Anders; KAHANE, Guy. Well-being and enhancement. In: SAVULESCU, J.; MEULEN, R.; KAHANE, G. (Eds.). **Enhancing Human Capacities**. West Sussex: Blackwell Publishing, 2011, p. 1-18. Available at: https://onlinelibrary.wiley.com/doi/book/10.1002/9781444393545. Accessed on: 25 Sept. 2024.

SHEN, Francis X. Neuroscience, mental privacy, and the law. **Harvard Journal of Law & Public Policy**, Cambridge, v. 36, p. 653, 2013. Available at: https://harvard-jlpp.com/vols-36-40/. Accessed on: 25 Sept. 2024.

SOLUM, Lawrence B. Legal personhood for artificial intelligences. In: SOLUM, Lawrence B. **Machine Ethics and Robot Ethics**. Abingdon: Routledge, 2020, p. 415-471. Available at: https://www.routledge.com/Machine-Ethics-and-Robot-Ethics/Anderson-Anderson/p/book/9781138585101. Accessed on: 25 Sept. 2024.

TAM, Wing-kin et al. Human motor decoding from neural signals: a review. **BMC Biomedical Engineering**, Londres, v. 1, p. 1-22, 2019. Available at:

https://bmcbiomedeng.biomedcentral.com/articles/10.1186/s42490-019-0005-5. Accessed on: 25 Sept. 2024.

YUSTE, Rafael et al. Four ethical priorities for neurotechnologies and AI. **Nature**, Londres, v. 551, n. 7679, p. 159-163, 2017. Available at: https://www.nature.com/articles/551159a. Accessed on: 25 Sept. 2024.

WALLACH, Wendell. From robots to techno sapiens: ethics, law and public policy in the development of robotics and neurotechnologies. **Law, Innovation and Technology**, Abingdon, v. 3, n. 2, p. 185-207, 2011. Available at: https://www.tandfonline.com/doi/abs/10.5235/17579961.3.2.185. Accessed on: 25 Sept. 2024.

NOTES AND CREDITS OF THE ARTICLE

• **Acknowledgments:** this article was developed following an invitation from the editors to explore the topic of artificial intelligence. The adopted approach reflects a critical analysis and a specialized contribution to foster discussions in the field, based on prior studies and experiences.

• **Funding:** not applicable.

• **Conflicts of Interest:** not applicable.

• **Ethical Approval:** not applicable.

• **Availability of Data and Materials:** not applicable.

• **Manuscript Published as Preprint:** not applicable.

• **Author Contributions:**

Contribution	Divino, D.
Study conception	x
Conceptualization	x
Methodology	x
Data collection / investigation	x
Data curation	x
Data analysis	x
Results discussion	x
Visualization (graphs, tables, etc.)	x
Original draft	x
Review and final editing	x

Ciência da Informação Express
ISSN 2675-8717

• **Publisher:** Federal University of Lavras (UFLA). The views expressed in this article are the responsibility of the authors and do not necessarily represent the opinion of the editors or the university.

• **Editor-in-Chief:** Nivaldo Calixto Ribeiro, Federal University of Lavras (UFLA).

• **History:**

Received: 16/09/2024	Accepted: 22/09/2024	Published: 02/10/2024

This form was prepared based on the best practices suggested by SciELO in its Open Science compliance form and on the Notes of Work forms from the scientific journals: Encontros Bibli, AtoZ: novas práticas em informação e conhecimento, and the Credit form of Revista Digital de Biblioteconomia e Ciência da Informação.