Research article
UDC 34:004:343.721:004.8
EDN: https://elibrary.ru/smgmxq
DOI: https://doi.org/10.21202/jdtl.2025.16

Neurohacking in the Digital and Artificial Intelligence Age: Legal Aspects of Protecting Neural Information

Diogo Pereira Coelho
University of Seville, Seville, Spain

Keywords

artificial intelligence,
brain-computer interface,
cybercrime,
digital technologies,
law,
metaverse,
neurocrime,
neurohacking,
neurosecurity,
neurotechnologies

Abstract

Objective: to contribute to the concept of neurocrime; to study the current and future risks from the viewpoint of cybersecurity in the context of digitalization and artificial intelligence development.

Methods: the study uses a critical and descriptive analysis of the relationship between cybercrime and neurocrime. It provides a conceptual distinction between the brain-computer interface and its use and describes the differences between neural and mental manipulation. The legal autonomy of crimes against mental integrity in relation to crimes against physical integrity is investigated. The methodological framework includes the analysis of existing prototypes of neurocrimes based on a four-phase brain-computer interface cycle and the study of the features of neurohacking in the context of the metaverse and artificial intelligence technologies.

Results: the study revealed the essential characteristics of neurohacking as the misuse of neural devices to gain unauthorized access to and manipulate neural information. Four main types of brain-computer interface applications subject to neurohacking are identified: neuromedical applications, user authentication systems, video games, and smartphone-based applications. The modalities of neurohacking were established at each phase of the brain-computer interface cycle: manipulations at the stage of neural information input, measuring and recording of brain activity, decoding and classifying neural information, as well as at the stage of the result output. The specific threats of neurohacking in the era of digitalization are analyzed, including immersive attacks and human joystick attacks in the metaverse.

Scientific novelty: for the first time, a comprehensive differentiation of the concepts of neurocrime and cybercrime was carried out, highlighting their specific legal consequences. The author proposed a classification of neurocrimes based on the four-phase cycle of the brain-computer interface. The study substantiated the need to distinguish mental integrity as an independent object of legal protection, different from the protection of physical integrity. For the first time, the features of neurohacking in the context of the metaverse and artificial intelligence technologies were investigated, including the analysis of new types of attacks and threats to neurosecurity.

Practical significance: the study results are important for the development of legal regulation in the field of cybersecurity and the corresponding regulations. The identified types of neurocrimes and their classification can help to create a specialized legislation on the protection of neural data and mental integrity. Practical recommendations on ensuring the neurosecurity of brain-computer interfaces are in demand in medical practice, video game industry, authentication systems, and for the development of smartphone applications.

For citation

Coelho, D. P. (2025). Neurohacking in the Digital and Artificial Intelligence Age: Legal Aspects of Protecting Neural Information. *Journal of Digital Technologies and Law, 3*(3), 397–430. https://doi.org/10.21202/jdtl.2025.16

Contents

Introduction

There are those who believe that illicit access to or manipulation of neural information is not and never will be feasible in the way that is often assumed and feared[1]. The main reason for this is the limited understanding of the neural code, i.e. the language through which the brain encodes and processes information. For this to be possible, it would be necessary to decode the neural code in order to achieve a certain result (access or manipulation) and, among the billions of neurons that exist in the human brain, to determine which specific one to stimulate[2]. Although it is currently possible to make predictions about which region of the brain to stimulate, identifying the exact neuron still appears to be a major challenge. In addition, the neuron responsible for a particular function in the brain of a given subject may not be the same in the brain of another subject[3]. On the other hand, some argue that it is not possible to influence a subject's behavior by stimulating just a single neuron, because brain function depends on the coordinated activity of complex neuronal circuits, involving sets of hundreds or billions of neurons. The coordinated stimulation of large networks of specific neurons,

1 Fields, R. D. (2022). Hacking the brain: More fantasy than reality. The UNESCO Courier. Should we be afraid of neuroscience? (p. 9). UNESCO. https://clck.ru/3Nkhbr

2 Ibid.

3 Ibid.

with a view to imposing targeted and specific behavior for the purposes of manipulation and mental control, appears to be practically impossible[4].

It turns out that, in general, everything is considered a system, including the human brain. And like all systems, the human brain also seems to be the target of "hacking". Human beings themselves are considered[5] natural life hackers[6]. Card counting in blackjack[7] is considered a hack. Most sports are usually targets for hacking purposes. In F1[8], teams try to find new ways to modify the design of vehicles that are not expressly forbidden by the regulations. Gerrymandering[9] is considered a hack used in politics. Also in the financial and business world, there are several hacking methods used. From entrepreneurs to financial institutions, most of the players try to find loopholes in the system (i.e. the law), essentially in order to gain an advantage over competitors. Specifically, they exploit situations that are not expressly prohibited, but represent intentional (or not) subversions of the system. Both Uber[10], and Airbnb[11], among other large technology companies, initially violated the rules imposed by the legal systems of multiple jurisdictions. The concept of "obstruction" thus appears to be a hack as old as the concept of "gap in the law" or "gap in the system" itself. It is probably as old as human civilization, since people themselves seem to be able to be hacked. In this context, the human brain presents itself as a system. Specifically, a system whose optimization has resulted from continuous interaction with the environment over millions of years, having been developed to survive and, above all, to reproduce.

---

4 Ibid.

5 Schneier, B. (2021). The Coming AI Hackers. Belfer Center for Science and International Affairs. Harvard Kennedy School. https://clck.ru/3Nkhe4

6 The English expression "hacker" is usually translated to other languages as "computer pirate". However, using this kind of translated word appears to be inaccurate. A "hacker" can act like a "pirate" and a "pirate" can act like a "hacker", but the two definitions do not seem to depend on each other, nor do they seem to be equivalent. Rather, they seem to be purely complementary. A hacker exploits systems, whether on a computer, on a phone, through personal interaction or simply in any other aspect of human life. And this, as a rule, in a lawful way. Hackers (or "crackers") illicitly try to decode/ crack a system, either for fun or to obtain a certain result or advantage. Regardless of semantic issues and terminological confusion, this text will mostly use the term "hacker", essentially because it is the most well-known expression and also to make it easier to read. See Ribeiro, J. B. (2019, 12 Fevereiro). 'Hacker' vs 'Pirata Informático': a riqueza de uma definição perdida na tradução. SH/FTER. https://clck.ru/3NkhkJ

7 Keating, S. (2022). How a magician-mathematician revealed a casino loophole. BBC. https://clck.ru/3Nkhnx

8 Straw, E. (2022, February 22). F1's new philosophy in combatting design loopholes. The Race. https://clck.ru/3Nkhq9

9 Ax, J. (2023). North Carolina court allows partisan gerrymandering. Reuters. https://clck.ru/3Nkhrs

10 Henley, J. (2017, September 29). Uber clashes with regulators in cities around the world. The Guardian. https://clck.ru/3Nkht2

11 Neubauer, I. L. (2019, August 30). Countries that are cracking down on Airbnb. The New Daily. https://goo.su/UcOHh

Cognitive hacking thus appears to be a powerful tool[12] in relations between individuals, with the manipulation technique known as "social engineering" standing out. Within cybercrime itself, the only novelty lies in the use of technology for hacking purposes, because just like the human brain, computers are also systems. Over the last few decades, we have seen hacking methods adapt to the computerization of traditional systems. This computerization seems to have changed hacking methods in three different ways: scale, scope and speed. Firstly, it has amplified and extended the nature of hacks, thereby increasing their scale and scope. Next, the growing number of software and hardware developments has allowed systems to evolve faster than initially anticipated. Computer speed kept pace with this development, which resulted in an increase in the speed of hacking methods[13]. With the evolution from Web 1.0 to Web 2.0 and Web 3.0 and, more recently, Web 4.0, new disruptive technologies have emerged[14] and the use of digital or artificial computer systems is expanding at an increasingly rapid pace (Ienca & Haselager, 2016)[15]. Ultimately, with the evolution of cybercrime, this use will focus on the human brain and mind itself, specifically in the form of "neurocrime" and "neurohacking" (also commonly referred to as "brain hacking"). In this scenario,

12 Schneier, B. (2021). The Coming AI Hackers. Belfer Center for Science and International Affairs. Harvard Kennedy School. https://clck.ru/3Nkhe4. By way of example, many of the powerful social systems that form the basis of society, such as democracy and the market economy, among others, depend on the decisions that people make. This process can be the target of cognitive hacking in multiple ways. Starting with social communication. Personalized according to our preferences and behaviors, modern advertising represents a kind of mass hacking of the human brain, specifically of the conscious psychic process, including the previous unconscious state. Not to mention disinformation (often disseminated by the press itself), which represents a hack of the common understanding of reality. The repeated use of terms such as "terrorism" or "cyberterrorism" in the media and in politicians' speeches also represents a hack of the cognitive system, essentially with a view to convincing people that this is a greater threat than it really is and thereby causing fear and misrepresenting risk assessment.

13 Schneier, B. (2021). The Coming AI Hackers. Belfer Center for Science and International Affairs. Harvard Kennedy School. https://clck.ru/3Nkhe4

14 The term "disruptive" is relatively recent and, in other words, means causing a transformation in the standards, models or technologies already established in the market. In other words, it describes a technological innovation, product or service with so-called "disruptive" characteristics, because it is diverse, revolutionary, innovative or never thought of or applied in that specific context. See Dufloth, R. (2017). Novas tecnologias e o futuro do profissional do Direito. Mgalhas. https://clck.ru/3Nm3jy

15 In general, the technologies associated with web 1.0, web 2.0, web 3.0 and web 4.0 allow users to interact directly with data and systems and can be used for a wide variety of daily or even professional tasks. For example, while GPS systems help with geolocation and spatial navigation, portable devices monitor bodily processes such as heart rate, calorie intake and weight loss. Still by way of example, personal computers help with cognitive tasks such as arithmetic calculations, writing and memory, blockchain-based digital assets allow international transactions of value in a matter of minutes or even seconds, and generative artificial intelligence systems to produce text, images or videos instantly and innovatively. See Nath, K. (2022). Evolution of the Internet Web 1.0 to Metaverse: The Good, the Bad and the Ugly. Research Gate. https://clck.ru/3NkiAo

the increase in scale, scope and speed of hacking methods will be increasingly noticeable.

The aim of this text is to contribute to the study and initial framing of a subject whose understanding will never be sufficient, not least because of the high level of legal assets at stake.

1. From cybercrime to neurocrime

1.1. Connection between cybercrime and neurocrime

There is no agreed definition of cybercrime. The terms "cybercrime", "computer crime", "computer-related crime" or "high-tech crime" are used frequently, but randomly. Whether at international, European or even national level, there is no consensus on the expression, definition, typology or classification of cybercrime (Rodrigues, 2009; Vasconcelos Casimiro, 2000). There is no concept of "cybercrime" or "computer crime" expressly established in Portuguese legislation[16]. There is also no uniformly settled concept in literature and jurisprudence (Venâncio, 2011). The lack of uniformity lies in the fact that the term "cybercrime" covers, in a generic and abstract way, a range of crimes committed using information and communication technologies. This term includes both classic criminal actions and new types of crime.

According to the European Commission, cybercrimes are "criminal acts committed using electronic communications networks and information systems or against such networks and systems"[17], and can be divided into three forms. Firstly, traditional forms of criminal activity, but using the Internet (and identity theft or phishing methods) to commit crimes (such as computer fraud or spoofing). These traditional forms also include the international electronic trade in drugs, weapons and endangered animal species. Secondly, the online publication of illegal content, such as material inciting terrorism, violence, racism and xenophobia or the sexual abuse of minors. Finally, crimes exclusively committed on electronic networks, which represent new and often "large-scale" crimes that were "unknown in the pre-internet age". In the latter case, criminal agents attack systems or entire information infrastructures and even confidential State information (which constitutes a national threat). Still according to the European Commission, these attacks can be carried out through "botnets" (network of robots), i.e. criminal agents distribute

[16] As an example, see the case of Portugal. In addition to the types of crimes provided for in Law no. 109/2009, of September 15, which approves the Portuguese Cybercrime Law. (https://clck.ru/3Nkxa7), there are also other types of crimes of this nature provided for in the Portuguese Penal Code and in various other separate legal sources.

[17] European Commission. (2007). Towards a general policy on the fight against cyber crime. Communication from the Commission to the European Parliament, the Council and the Committee of the Regions. https://clck.ru/3NkiNh

"malware" which, in turn, when transferred, transforms the user's computer into a "bot". Once the network or information infrastructure is infected, it is used to commit crimes without the knowledge of the respective users. In short, computer crime constitutes any act in which the computer or similar technology serves as a means to achieve a criminal objective, and the computer or similar technology may represent a merely symbolic target of that act or even represent the object of the crime (Marques & Martins, 2006). It is therefore necessary to distinguish between "computer crime", in which information technology is the target of the crime, and "crime committed using computer means", in which information technology is the means of executing the crime.

In this logic, whenever crime is committed using neural interfaces and, in addition to representing a physical threat to users, it can also have a profound impact on their behavior and self-perception, everything indicates that we are facing a "neurocrime"[18]. The problem associated with the misuse of information technologies in the context of neurotechnology is particularly critical, as this type of technology applies (directly or indirectly) to the brain, one of the most important organs in the human body. The human brain not only contributes significantly to life processes, such as reproduction and maintenance of life, but also provides consciousness, perception, the ability to think, discernment, memory and language. Furthermore, it still has enormous importance in human behavior and self-perception as a sensitive being or individual endowed with emotions and sensitivity (Ienca & Haselager, 2016). Neurocriminality thus seems to constitute any act in which the human brain and/ or mind serve as a means to achieve a criminal goal, and the human brain and/ or mind can represent a merely symbolic target of this act or represent the object of the crime. Therefore, a distinction shall be made between "neurocriminality", in which the brain and/or mind directly or indirectly constitute the target of the crime, from "crime committed using neural and/or purely mental manipulation", in which brain and/ or mind constitutes the means of executing the crime. As such, as we will see in point 3 of this chapter, a distinction shall also be made between neural manipulation and purely mental manipulation. In addition, neurocrime may not involve direct access to the brain or stored information, but only indirect access. For example, the functions of the neural device may simply be limited, modified or deregulated. With the current advances in neural engineering technologies (mainly marketed in the health sector), this scenario appears to be increasingly a reality. In this scenario, the perpetrator of the neurocrime can affect the victim's brain indirectly, since the neural system is not directly accessed or significantly manipulated during the attack. Nonetheless, it may affect the victim's mental state in a significant way, as the neurocrime may have limited or constrained their conduct,

[18] A term commonly used in the case of criminal activities that exploit neural devices.

generated an emotional response in the form of panic, fear, or psychological disturbances, and/or left traumatic memories. It is worth mentioning that, in certain circumstances, the perpetrator and the target of the crime may be confused. For example, users with mental instability may choose to damage their neural devices in order to attempt suicide (Ienca & Haselager, 2016). That said, in a broad sense, the concept of neurocrime seems to be able to be defined as the crime of offense against the mind of a person or a group of people, committed using neural and/or purely mental manipulation via a neural device, and with the intention of, directly or indirectly, causing physical or mental harm, including reputational and/or property damage.

In the field of neurocrime and neural stimulation, there are currently two types of neural devices considered particularly critical. On the one hand, there are brain stimulators, especially deep brain stimulation (DBS) and transcranial direct current stimulation (tDCS) systems. On the other hand, brain-computer interfaces (BCI) stand out. Both types of device allow direct access to neural computing, although in different ways (brain simulation vs. reading brain activity). In addition, both types of devices are available not only as medical technologies, but also as products marketed to users considered healthy. As such, both types of devices raise multiple concerns in terms of "neurosecurity". The fact is that, to date, BCIs have been the main neural devices used for hacking purposes (even with experimental evidence and in the context of real situations), which is why they are the only ones to be explored in this article (Ienca & Haselager, 2016).

1.2. Concept of brain-computer interface and use cases

Unlike mere neurostimulators (electronic devices similar to cardiac pacemakers), BCIs are not used to stimulate the brain, but rather to establish a direct communication link which, by bypassing the peripheral nervous system and muscles, allows users to control an external computer exclusively through brain activity (Ienca & Haselager, 2016; Vallabhaneni et al., 2005). BCIs were first developed in the field of clinical medicine and as a therapeutic or medical assistance technology for neurological patients. In a clinical context, BCI applications are used to repair, assist, or augment motor, cognitive, or sensory functions in patients suffering from neurological disorders that precisely affect motor development and/ or cognitive and sensory functions, including spinal cord injuries, strokes, and neurological motor diseases such as amyotrophic lateral sclerosis (ALS) and muscular dystrophy (Ienca & Haselager, 2016).

BCI can be distinguished into two types: invasive and non-invasive. While invasive BCI records brain activity through the surgical implantation of electrode arrays in the central nervous system or through a mere direct connection, non-invasive BCI records brain activity through electrodes placed on the outside of the skull, i.e. through neuroimaging

technologies such as electroencephalography (EEG) and electromyography (EMG). In both cases, a direct interaction is established between the user's brain and the neural device. As a rule, this interaction consists of a cycle of four phases (Ienca & Haselager, 2016; Van Gerven et al., 2009; Bernal et al., 2022)[19]. The first phase consists of the input of neural information (i.e. the creation of specific brain activity) by the user in response to a given stimulus (whenever the BCI user wants to perform a certain mental task or achieve a certain cognitive state). The second phase consists of measuring and recording brain activity. In this phase, the user's brain activity patterns are detected, measured, and recorded by the interface during the cognitive process or the execution of a certain mental task. In the third phase, the raw neural data (neural information) resulting from the second phase is decoded in order to assess its main characteristics and classify it. After decoding and classification, the data is translated into the production of a certain result (output), i.e. the production of the result expected by the user. In general, in this fourth phase, the output consists of carrying out the action initially intended, desired, or considered beneficial to the user through the control of applications connected to the BCI. Controllable applications include powered devices such as electric wheelchairs or even robotic prosthetic limbs, sensory devices and other types of software and hardware applications (including mobile applications and cell phones). Once each of these phases has been completed, in principle, the user can see the result of the previous phase and thus start the next phase (Ienca & Haselager, 2016). An example of this is Gert Jan Oskam, who, after a motorcycle accident and being a paraplegic for over a decade, is now able to walk again using BCIs (albeit imperfectly for the time being)[20].

Today, BCI applications are available not only in clinical settings, but also for the general public (Mochan et al., 2025). Multiple commercial applications of EEG-based BCI devices have entered the market and are increasingly popular for both video games and everyday activities (Ienca & Haselager, 2016)[21]. In the electronic telecommunications industry, there are also BCIs available on the market. Firstly, certain mobile applications, such as "Xwave" (launched over 10 years ago!), allow certain types of earphones to establish a direct connection with compatible iPhone cell phones and, through this connection, record brain

[19] In this regard, it should be noted that BCIs can be classified based on its levels of invasiveness. While non-invasive systems require electrodes to be applied to the scalp, invasive systems require a surgical procedure to place electrodes inside the skull, either on the surface of the brain or even inside the brain.

[20] Whang, O. (2023). Brain Implants Allow Paralyzed Man to Walk Using His Thoughts. https://clck.ru/3Nkxhk

[21] As an example, see the websites of the companies Emotiv Inc. and Neurosky Inc., both pioneers in the commercialization of non-invasive, intuitive and accessible BCIs for video games, interactive televisions or non-manual control systems, respectively accessible at: https://clck.ru/3NkiYN & https://clck.ru/3NkiZy. See also Gordon, L. (2020, December 16). Brain-controlled gaming exists, though ethical questions loom over the tech. The Washington Post. https://clck.ru/3Nkicf

waves and frequencies (Ienca & Haselager, 2016)[22]. The same is currently happening in the arms industry, as several BCI applications are being developed (Ienca & Haselager, 2016; Czech, 2021). For example, the US Defense Advanced Research Projects Agency (DARPA) funds a wide range of BCI projects, essentially aimed at restoring behavioral or neural functions in soldiers and also improving the training and performance of those same soldiers and even secret service agents (Ienca & Haselager, 2016; Kotchetkov et al., 2010; Miranda et al., 2015). Based on the vast potential of brain control via neural computing devices in terms of utility and basic functionality, it is predicted that BCI will gradually replace the keyboard, touch screen, computer mouse and even voice assistance technology, as consumers may prefer to interact directly with computers (Ienca & Haselager, 2016; Yuan et al., 2010; Radu, 2024). However, as we will see in the next points and chapters, while the potential benefits of the foreseeable mass commercialization of certain clinical and non-clinical BCI technology applications appear to be significant and well-studied, the very serious (and perhaps irreparable) risks associated with neurosecurity remain largely unexplored. In order to understand these risks, a distinction must first be made between neuronal and purely mental manipulations.

1.3. Distinction between neural and purely mental manipulations

After more than 2,000 years of philosophical argument and the most recent impressive scientific advances in artificial intelligence (AI) and data science, the problem of mind-brain dualism persists and continues to defy a unanimous solution (Bublitz & Merkel, 2014; Moulin, 2022). According to Maria Fernanda Palma, the absolute denial of mind-brain dualism, or the reduction of the phenomena of consciousness to brain states, can have a significant impact on the real behavioral basis on which legal responsibility is founded, particularly in relation to voluntary action or true capacity of individuals to be considered guilty under Criminal Law[23]. In the event that conscience does not temporarily precede the decision, the free and conscious decision as a criterion of responsibility or greater responsibility in the figure of intention is necessarily questioned. In this sense, the absolute self-determination of the human being as a source of responsibility is questioned, when it leads to the assumption that the decision intelligent can be calculated like an algorithm and imitated by its mathematical processes. While neuroscience seems to propose a naturalistic reduction of the mind and states of consciousness, AI and data science suggest that the functioning of the brain can be conceived as an autonomous biological process, in which mental states and their relationship to human behavior can be

22 PLX Devices Inc. XWave – Mind Interface Introduction and Teaser. 2010. https://clck.ru/3NkifC

23 Fernanda Palma, M. (2021–2022). Ciberneurodireito. Powerpoint Presentation. Short Course in Criminal Law on Artificial Intelligence and Artificial Intelligence in Criminal Law. FDUL. IDPCC/ CIDPCC, 4.

reproduced. This perspective unexpectedly reopens the dualist proposal, treating the brain as just one support among other possible bases for human behavior. As such, and based on the associated sciences, the Law aims to examine whether a specific mental state has a causal or correlated cerebral existence, and whether it necessarily implies a certain behavior, as well as to what extent it is controllable[24].

According to Jan Christoph Bublitz & Reinhard Merkel, all mental phenomena are, in one way or another, connected to brain activity. In this sense, it is argued that the legislator should be cautious when describing the mind-brain relationship as a dualistic relationship, not least because it is difficult to identify changes in the mental state without also identifying certain changes at the cerebral (neural) level. It is also argued that mental states are not only correlated with a certain brain state, but are also caused or made conscious on the basis of a certain physical state. However, whenever a more precise and concrete description of this correlation is considered, multiple problems arise, as there is no consensus on the correlation with the causes or the process of "awareness" (Bublitz & Merkel, 2014). As such, despite the current tendency in psychiatry to classify any mental disorder as a brain disorder in the strict sense, there are those who argue that mental injuries may not necessarily be confusable with physical bodily injuries. In principle, it is not the brain that "decides", "suffers" or feels the "moral damage". On the contrary, everything indicates that we are dealing with mental states/ processes of people and not exactly with qualities of physical objects. For normative purposes related to the concepts of "harm" or "disorder/ dysfunction", everything indicates that the mind and brain deserve individual attention (Bublitz & Merkel, 2014). By way of example, mental disorders/ dysfunctions seem to result from certain psychological functions or in relation to certain social norms, and not exactly from electrochemical brain processes. In this case, it seems that we are dealing with mental and behavioral phenomena that cannot be described purely in terms of neuroscience. By way of example, it seems that "depression" is a specific mental symptom and suffering from depression depends exclusively on the display of that symptom. Even if it were known (which it is not) that all symptoms of depression are strongly correlated with chemical imbalances at the neurotransmitter level, the distinction between mental and brain dysfunction would persist (Bublitz & Merkel, 2014). In this logic, it has been argued that mental damage should not be treated in the same way as brain damage. Otherwise, the advances and developments of modern criminal law will be disregarded and the Roman concept of iniuria, which was seen as a summary notion of any and all offenses committed against the person, will be considered. The literature has thus been advocating the definition of mental states worthy of legal protection and the introduction of specific normative provisions that penalize interference with mental integrity, rather than the protection of physical integrity being adapted and expanded.

[24] Fernanda Palma, M. (2021–2022). Ciberneurodireito. Powerpoint Presentation. Short Course in Criminal Law on Artificial Intelligence and Artificial Intelligence in Criminal Law. FDUL. IDPCC/ CIDPCC, 4, Pp. 6–14.

While brain integrity should cover physical interventions, i.e. brain damage (regardless of the mental consequences), mental integrity should cover mental interventions, i.e. mental damage and regardless of the brain consequences (Bublitz & Merkel, 2014).

In this regard, it is also argued that mental integrity should also encompass protection against mental manipulations/interventions, such as the provocation of emotions, the manipulation of preferences and decision processes, the optimization of non-consensual neurological development, the manipulation of memory, the manipulation of will or willpower, among other cognitive and emotional phenomena (Bublitz & Merkel, 2014). In all these cases, the manipulations/ interventions restrict mental capacities or alter preferences and will formation. However, to date, none of these manipulations/ interventions seem to be, in general, adequately covered by the rules that currently protect physical integrity and mental health. Not least because all of these manipulations/ interventions only cause psychological changes, and do not seem to meet the requirements to be considered physical harm. Since the victim is not affected by any brain damage or experience of physical pain or discomfort, the unlawfulness of these manipulations/ interventions seems to stem from their purely mental effects (Bublitz & Merkel, 2014). With the entry into the digital and artificial age, everything indicates that in the future, in the event of litigation related to brain and mental integrity, neuroscientific evidence and proof will be increasingly important in ascertaining and discovering the truth (Shen, 2013). For this reason, and regardless of whether the denial of the mind-brain dualism persists, before we move on to the analysis of neurohacking per se, it is still necessary to make some considerations regarding the current deliberation of the legal autonomy of crimes against the mind in relation to crimes against physical integrity.

1.4. Considering the legal autonomy of crimes against mental integrity in relation to crimes against physical integrity

As explained in the previous point, neuroscience, AI and data science not only question assumptions about the position of the mind in the natural world, but also instigate a rethinking of its role in the normative world. It turns out that the law currently offers unilateral protection, as it systematically protects the body and the brain and only fragmentarily the mind and mental states. As explained in the previous point, the fundamental question is to what extent it is possible to legitimately intervene in the minds and mental states of other subjects. With the commercialization and mass implementation of neural technologies with the ability to intervene in the mind and detect mental activity, it is argued that the law should introduce autonomous and individual legal protection of mental integrity (Bublitz & Merkel, 2014; Abegão Alves, 2020). This scientific-philosophical and legal debate basically focuses on two aspects: (i) considering the empirical and conceptual autonomy of the mental in relation to the physical; (ii) considering the legal autonomy of crimes against mental integrity in relation to crimes against physical integrity. According to the Research Project “Crimes

Against the Mind", developed by the Institute of Criminal Law and Criminal Sciences of the Faculty of Law of the University of Lisbon (after the publication of the pioneering study by Bublitz & Merkel, 2014), which aims to fill this gap in legal thinking, the search for the limits of the legitimate alteration of the mental states of others has not been carried out by jurists and legal thinkers, so that the reality to be regulated remains one step ahead of the Law. Also according to this project, the most recent scientific discoveries demand the attention of the Law, not only from the point of view of the aggressor/ criminal, but also from the point of view of the victim, and the legal debate should arise both in relation to the problems of free will and the foundation of criminal responsibility that these discoveries raise, and also in relation to the legal assets to be protected. It is this second facet of the relevance of neuroscience to law that has yet to be explored. In this sense, the current aim is to deepen the scientific-philosophical/ legal debate on this subject, both at national and international level. Specifically, the aim is to problematize the key issues, proposing solutions and possible paths for future investigations that will inevitably involve the intersection of different fields of knowledge.

2. From neurocrime to neurohacking

2.1. Neurocrime prototypes usually referred to as neurohacking

Over the last few years, multiple examples of prototype neurocrimes usually referred to as neurohacking have been identified. These include the hijacking of wireless limb prostheses, the malicious reprogramming of neural stimulation therapy (i.e. unauthorized wireless changes to the device's configuration in order to generate certain brain stimuli) and the unauthorized interception of brain implant signals in order to obtain private neural information (Ienca & Haselager, 2016)[25]. In principle, all these examples can only be carried out using neural devices that allow a direct connection to the brain to be established, such as tDCS and especially the increasingly developed BCI (Denning et al., 2009). These are the prototypes of neurocrime that are usually referred to as neurohacking and, as we will see in the next few points, the way they are carried out is very similar to computer hacking in the context of cybercrime as explained in the first point of the first chapter on the connection between cybercrime and neurocrime (Ienca & Haselager, 2016).

2.2. Concept of neurohacking

In a broad sense, neurohacking seems to consist of the abusive and malicious use of neural devices in order to illicitly obtain and possibly manipulate neural information (Ienca & Haselager, 2016). Strictly speaking, neurohacking seems to consist of a neuro-attack carried out through neural devices, through which the perpetrators gain illicit access

[25] This last example describes a specific neurocriminal phenomenon in which the attack is not simply aimed at disrupting the neural device, but at obtaining unauthorized access to private information.

to neural information which, in turn, can be manipulated in order to control the cognitive process or the execution of a certain mental task of the user of the device. Once the neural device has been accessed, it is used to commit crimes with or without the user's knowledge.

2.3. Neurohacking based on types of BCI applications

2.3.1. Scope

As explained so far, BCIs can be intercepted to detect hidden autobiographical information from users and with a significantly high accuracy rate (Ienca & Haselager, 2016; Rosenfeld et al., 2006; Rosenfeld, 2011). It has been shown that once intercepted, BCIs can reveal private and confidential information about users, such as their PIN codes, bank and credit card details, date of birth, home address and even the faces of people they know (Ienca & Haselager, 2016). The science fiction future described in the present article therefore seems to be anything but fiction. In fact, such a future, in which subjects are able to access and manipulate the neural information of other subjects, is not approaching at a rapid pace, as it is a current reality (Mochan et al., 2025). In this scenario, and as we will see in the next sections of this chapter, as well as in Chapter III, unless the design and functional characteristics of current neural devices, which are still under development, have strong neurosecurity measures, their misuse and malicious use could imply serious and grave risks in terms of public safety (Ienca & Haselager, 2016)[26].

2.3.2. BCI applications that could be the target of neurohacking

According to the first point of this Chapter II, neurocrime prototypes can only be carried out through neural devices that make it possible to establish a direct connection with the brain, and BCI applications are the main targets for hacking. It is possible to distinguish between four main types of BCI applications that could be targeted for neurohacking: (i) neuromedical applications; (ii) user authentication; (iii) video games and entertainment; and (iv) smartphone-based applications. For each of these applications, possible attack scenarios are currently being studied, as well as the respective neurosecurity measures to be applied (Mochan et al., 2025). In fact, for some of the neurohacking activities based on types of BCI applications, there is already evidence and experimental proof and even in a real context (Li et al., 2015). As we will see in the next point, within the scope of these four main types of BCI applications, neurohacking can, in principle, take place in any of the different phases of the BCI cycle described in point two of Chapter I.

[26] Martinovic, I. et al. (2012). On the feasibility of side-channel attacks with brain-computer interfaces. USENIX security symposium. https://clck.ru/3NkoD5

2.4. Neurohacking modalities based on the four-stage cycle of the BCI

2.4.1. Scope

As mentioned in the previous point, and taking into account the types of BCI applications that already exist (Mochan et al., 2025) and also, theoretically, those that may exist in the (very) near future, the next sub-points will cover, also theoretically, the various types of neurohacking based on the four-phase BCI cycle. Specifically, the input phase of neural information, the phase of measuring and recording brain activity, the phase of decoding and classifying neural information and, finally, the output phase (Ienca & Haselager, 2016).

2.4.2. Manipulation of the neural information input phase

In this form of neurohacking, the hacker attacks the BCI user at the moment of inputting neural information, i.e. in the first phase of the cycle. The input of neural information can be manipulated by changing the stimuli communicated to the BCI user. For example, neurohackers can pre-select target stimuli in order to trigger a specific response in the user and thus facilitate access to their neural information. This type of neural malware has some similarities to spyware on a computer, as it aims to collect information about the user, send it to another entity, and/ or ensure control over a computer or other IT device without the user's permission or consent (Ienca & Haselager, 2016). In this case, the malware used for neurohacking has the particularity of being able to extract information directly from the signals emitted by the brain, and is therefore commonly known as "brain-spyware". In the future, there seem to be several possible mobile and portable applications of brain-spyware for the purposes of neurocrime and, it seems, activities such as password cracking, identity theft, phishing and other types of neural scams will become increasingly common (Ienca & Haselager, 2016). For the time being, deciphering the signals emitted by the brain with a level of precision and speed that is comparable to computer hacking does not seem to be possible outside of an experimental environment (and, in this case, to a limited extent). Given the limitations in terms of deciphering signals emitted by the brain, as well as the current degree of maturity of the market, the reward does not seem worth the risk for today's hacker. However, with technological advances and the rapid expansion of the BCI applications market, it seems that neural information will be increasingly valued and also decipherable and readable (Ienca & Haselager, 2016).

2.4.3. Manipulation of the brain activity measurement and recording phase

In the phase of measuring and recording brain activity, the hacker attacks the BCI user with the aim of, without the latter's consent or will, producing a result (output) different from that expected due to the regular functioning of the neural device. The form of the attack may differ depending on its specific purpose, with three main purposes in mind: cracking raw neural data from BCIs (i.e. neural information), disrupting the functionality of the BCI, and hijacking the BCI. In general, cracking raw data can result in various criminal

activities aimed at limiting, harming, or taking advantage of certain neural device user behavior. Attacks aimed at disrupting the functionality of the BCI can occur whenever the hacker intends to manipulate the measurement of brain activity in order to confuse, overdrive, or delay the functions of the BCI application. Hijacking can occur whenever the hacker aims to monitor and alter the BCI communication channel in order to decrease or even replace the user's level of control over the BCI application. During hijacking, the neural device receives orders that diverge from its user's intentions or desires and benefit the hacker. By way of example, this type of neurohacking could involve disrupting a particular speech production device based on BCI in order to silence the user, or even hijacking a wheelchair in order to define the user's route (Ienca & Haselager, 2016).

2.4.4. Manipulation of the decoding and classification phase of neural information

Neurohacking of the decoding and classification phase involves manipulating the production of the result (output) intended and expected by the user as a function of the normal processing of the BCI device. This can be done in three different ways: (i) by introducing noise in order to make the decoding process unnecessarily challenging; (ii) by interfering with the neural learning and memory mechanisms (machine learning) in order to manipulate the classification of brain waves; or (iii) by replacing the waves sent by the BCI to the output device. Each of these methods has advantages and disadvantages. For example, adding noise appears to be the easiest method to perform and the most difficult to detect compared to other methods. However, with this method, the hacker's chances of achieving the desired result are reduced. On the other hand, although the other two methods are the most difficult to execute and the easiest to detect, at the same time, they also offer the most control over the BCI system. As with the manipulation of the brain activity measurement and recording phase, the hacker appears to be able to manipulate the decoding and classification phase in order to hijack the BCI device. In this type of attack, the aim appears to be not only to limit or take control of the device away from the user, but also to replace it. By successfully hijacking the system, the hacker appears to be able to gain partial or total control over the BCI device, while limiting or eliminating the user's control. In this scenario, the hacker appears to be able to monitor, alter or insert messages into the BCI communication channel (Ienca & Haselager, 2016).

2.4.5. Manipulation of the output production phase

This modality occurs whenever the hack aims to alter the output perceived by the user at the end of each BCI cycle. In this modality, the neurohacker aims to manipulate the user's perception of immediately previous actions or their own self-perception resulting from cognitive states generated by the BCI. The criminal motive behind this type of hack would be, without the user's permission, to induce specific cognitive states or actions in the user's subsequent cycle (or in each subsequent cycle) for the hacker's benefit. By way of example,

the neurohacker appears to be able to carry out a kind of "neurophishing". In this sub-modality, the user may be induced by the hacker to enter a certain password or other type of authentication information before the originally intended process can begin or continue. In this scenario, it seems possible to subject the user to certain traumatic experiences. The criminal activities that fall under this category, among others, include scamming, phishing, identity theft and harm to physical or mental integrity (Ienca & Haselager, 2016).

3. Neurohacking in the digital and artificial age

3.1. Concept of digital and artificial age

It is commonly accepted that the information age consists of a historical period (between the 50s and 70s of the 20th century) that saw the transformation of traditional industries, established during the industrial revolution, to an economy centered on information (and communication) technology. Today, with the beginning of the 21th century and the economy moving towards digital and artificial technologies, it seems that we have entered a new historical period, specifically the digital and artificial age. As mentioned in the introduction, with the evolution of Web 1.0 to Web 2.0 and Web 3.0 and, more recently, to Web 4.0, new disruptive technologies have emerged and the use of digital and artificial computer systems is expanding at an increasingly rapid pace. Among these new disruptive technologies, recent advances in the sector of immersive technologies, the metaverse, and virtual worlds, or even in the sector of AI and data science, stand out. In the first sector, essentially in terms of virtual (VR), augmented (AR), mixed (MR), and extended (RX) reality, and also in terms of spatial computing and digital sensory interaction. In the second sector, mainly in terms of generative AI systems and robotics[27]. Both these sectors promise to revolutionize the way society interacts with technology (and vice versa). Not only do these types of technologies aim to improve the user experience in terms of efficiency, but they also have the features and characteristics to drive multiple sectors forward at breakneck speed (Ford, 2016)[28] and at a rate that exceeds all predictions[29], including in sectors associated with other disruptive technologies such as cloud and edge computing, big data, quantum computing, brain-computer interfaces, distributed ledger technologies

[27] See GPT-4. https://clck.ru/3NkoQK; Bing AI. https://clck.ru/3NkoTY; Gemini. https://clck.ru/3NkoV7. See also the advances in terms of humanoid robots, specifically: Boston Dynamics. https://clck.ru/3NkoY3; Tesla. https://clck.ru/3Nkxk8

[28] In this regard, it should be noted that the best-known way of measuring the progress of computer processing power is "Moore's Law". Gordon E. Moore predicted that every 18 months, the number of transistors on chips would increase by 100%. However, information technology is exceeding what Moore's Law itself predicted. Unlike hardware, for example, which has seen significant increases in computer memory capacity and the amount of digital information that can be transmitted over fiber optic cables, software is seeing the effectiveness of certain algorithms grow at a rate that exceeds all predictions (Ford, 2016).

[29] Coelho, D. P. (2023). Os recentes avanços no setor da IA são uma benção ou uma maldição? Observador. https://clck.ru/3NkpED

(such as blockchain), the internet of things (IoT), smart cities, facial recognition, robotics, among others.

Imagine these technologies reaching the potential widely predicted (Dwivedi et al., 2022)[30]. In the next 15 to 20 years, everything will be directly connected. New worlds and new countries will be (re)created in digital form in the metaverse[31]. Smart cities will be connected via IoT, with CCTV's[32] in every corner and highly equipped with facial recognition and intelligent sensory systems[33]. AI-based drones and humanoid police robots will patrol the streets and buildings[34]. Universities will be mostly AI-based[35]. Most organizations will not even have workers or physical spaces, being autonomous, digital and based on AI[36]. Businesses will be digital, in the metaverse and developed by autonomous digital and artificial companies[37]. AI-based industrial robots will produce all kinds of goods[38]. Workplaces will be virtually all digital[39]. People will spend practically their entire day (and night) in the increasingly developed metaverse[40]. People will be able to choose to step out into the physical environment in the form of holograms, making it possible to attend meetings and workspaces and, even to "move" freely anywhere on the planet[41]. All this, of course, without leaving home. The line separating virtual reality from augmented reality and mixed reality will become smaller. The same goes for the line separating physical

30 See also Coelho, D. P. (2023). Ano 2050: Era Digital. Observador. https://clck.ru/3NkpRV; Chayka, K. (2021). We already live in Facebook's metaverse. The New Yorker. https://clck.ru/3NkjZb

31 See Woodward, W. (2024). Backup nations: countries making digital twins to mitigate natural disasters. Nesta. https://goo.su/uVSUy; Widlund J. (2023) Singapore's First Country-Scale Digital Twin and The Future of Digital Open Data. https://clck.ru/3NkxqG

32 Also known as "closed circuit television" or "surveillance cameras".

33 Davis, D. (2021). Facial Recognition And Beyond: Journalist Ventures Inside China's 'Surveillance State'. NPR. https://clck.ru/3NkjET

34 See Jarecki, J., Wilson, N., & Trevellyan, K. (2024). Vermont police are using drones more than ever. Here's what that means. Vermont Public. https://clck.ru/3NkjGh ; Chen, H. (2023). Robôs com mais de 2 metros de altura integram força policial de Singapura em aeroporto. CNN Brazil. https://goo.su/W2x0

35 Carroll, M. (2024). UK's first 'teacherless' AI classroom set to open in London. Sky News. https://clck.ru/3NkjLK

36 Smith, T. (2024). Profitable, AI-powered companies with no employees to arrive 'next year'. Sifted. https://clck.ru/3NkjNX

37 Eckert, T., & Cigaina, M. (2023). The metaverse: A new space for business. SAP. https://clck.ru/3NkjPh

38 See Ping, Ch. (2019) Robots to wipe out 20 million jobs around the world by 2030: Study. https://clck.ru/3Nkxtp; Semuels, A. (2020). Millions of Americans Have Lost Jobs in the Pandemic – And Robots and AI Are Replacing Them Faster Than Ever. Time. https://clck.ru/3Nkxxo

39 Hoover, A. (2024) The Metaverse Was Supposed to Be Your New Office. You're Still on Zoom. https://clck.ru/3Nky4S

40 Steele, C. (2022). People Are Spending More Time Online-and They're Not Happy About It. PC Mag. https://clck.ru/3NkzDa

41 See Atkinson, E., & Meyer, M. (2022). Meeting in the Metaverse: The Future of Work?. University of Denver, Podcast. News. https://clck.ru/3Nkpxk. Verdict. In the metaverse, holograms offer more options than avatars. (2022, June 15). https://clck.ru/3Nkq2b

reality from digital reality. Central bank digital currency will be the only legal tender[42]. In everyday tasks, people will be assisted by portable and wearable technologies (in many cases invisible)[43] or even by humanoid domestic robots based on AI[44]. In the decisions they make, they will be assisted by AI through voice assistants or even through neural interface technology, with a view to merging human consciousness with AI in a kind of symbiosis between human and machine[45]. In general, each time people establish a connection with the metaverse, a neural connection will be initiated[46]. Digital sensory interaction will make people feel increasingly comfortable connected to the metaverse, and this is the reality that the new generations will know best. Because of this digital life, people will always be surrounded by cameras, microphones and interface systems. Even from birth, embryonic development will take place in artificial incubators[47]. In fact, many people will choose to create and cultivate emotional or loving relationships with AI-based humanoid domestic robots[48]. Domestic animals themselves will be replaced by artificial pets[49]. Everything will be digital or artificial.

As such, thoughts themselves will not be safe, because whenever a connection is established with the metaverse and/ or any type of neural interface technology is activated, the human being will be an open book. In this scenario, the use of computer systems is not limited to the social, professional and economic spheres, but extends to the psychological and biological spheres. Inevitably, the dizzying speed of development in all these sectors, as well as their combination, could result in an equally dizzying increase in the scale, scope and speed of cybercrime, neurocrime and, consequently, neurohacking methods. As such, and as we will see in the scope of this chapter III, the study and investment in cybersecurity and neurosecurity could become increasingly relevant (Pooyandeh et al., 2022).

---

42 Michel, N. (2024, June 17). CBDCs Are Instruments Of Control-And They're Here. Forbes. https://clck.ru/3Nkq5h

43 From Wearables to Implantables: The Rise of Invisible Technologies. https://clck.ru/3NkyPr

44 See Reuters. (2024). A humanoid robot to help you around the house. https://clck.ru/3NkqCB; Schwartz, R. (2024). Is the world ready for Tesla's new domestic robots? The Week. https://clck.ru/3NkqE8

45 See Brodsky, S. (2024, August 27). AI voice assistants evolve, promising deeper interaction. IBM. https://clck.ru/3NkqHZ; Niemeyer, K. (2024, August 3). Elon Musk says Neuralink could help humans compete with AI: 'Let's give people superpowers'. Business Insider. https://clck.ru/3NkqLa

46 How BCI can elevate the AR/VR experience. https://clck.ru/3NkyXe

47 Zimmer, K. (2021, March 30). The Ultimate Incubator: The Brave New World of Bionic Babies – Artificial placentas could improve the survival odds of premature infants. IEEE Spectrum. https://goo.su/PE2QYzM

48 See Travers, M. (2024, March 24). A Psychologist Explains Why It's Possible To Fall In Love With AI. Forbes. https://clck.ru/3NkqXn; Chow, A. (2023). AI-Human Romances Are Flourishing-And This Is Just the Beginning. Time. https://clck.ru/3NkyiH

49 World Economic Forum. (2023). Moflin, an AI pet, responds like a real animal. https://clck.ru/3Nkqgf

3.2. Concept of metaverse

In 1992, science fiction author Neal Stephenson introduced the term “metaverse” in his 1992 cyberpunk novel “Snow Crash”. In this work, a 3D virtual world is presented in which people, represented as avatars, could interact with each other and with artificially intelligent agents. In 2003, this initial concept of a metaverse (still a long way from the concept that is currently being idealized) was first implemented in the game Second Life and even had some success[50]. The term “metaverse” is formed by combining the Greek prefix “meta”, which can be translated as “beyond” or “transcendence”, and the suffix “verse”, which comes from the word “universe”. We are thus dealing with a world beyond the universe (Bernal et al., 2022). As such, this concept aims to represent a virtual (digital) world which, despite coexisting with physical reality (through augmented reality), allows us to overcome the physical limitations of the real world, such as space and time. In this digital environment (which, it seems, will be increasingly valued), multiple users aim to interact exactly as they do in real life, using an avatar that represents their digital alter ego or even their digital identity (Bernal et al., 2022). In a broad sense, the concept of the metaverse therefore consists of a space or set of virtual and shared spaces (commonly called digital or virtual worlds or environments) where users, represented by digital avatars, can access and interact in a multidimensional way via their headsets (among other possible accessories). In other words, instead of simply viewing the content, users can immerse themselves in digital content through their digital representations[51].

The main and basic technologies that currently make up the metaverse and virtual worlds are immersive technologies (such as virtual, augmented, mixed and extended reality, BCIs, and sensory interaction systems), 3D modeling and reconstruction technologies, spatial and edge computing, AI and data science, IoT, and distributed recording technologies (Pooyandeh et al., 2022) [52]. Unlike today’s virtual and/ or augmented reality technologies, which are mostly used for electronic games or to replace the keyboard, touch screen or even the computer mouse, the technology we want to achieve could be used to simulate practically any situation associated with the physical world. From carrying out professional

[50] Second Life is a video game launched in 2003 that allows users to enjoy a “second life” in the virtual world. Users can assume any identity and play any role. Specifically, they can take on the role of an avatar in a virtual world that can be explored to meet other users, to take part in individual and/ or group activities, and so on, just as they would in real life. See the Second Life video game website. https://clck.ru/3Nm4mQ

[51] Pereira Coelho, D. (2021). Metaverse: should regulators be more attentive than ever? Observador. https://clck.ru/3Nm549

[52] See also Tucci, L. (2024, March 22). What is the metaverse? An explanation and in-depth guide. TechTarget. https://clck.ru/3Nkqqk

activities, to attending virtual concerts[53] or even just enjoying some time with friends, there will be multiple possibilities for interaction. The ultimate objective is, therefore, to eliminate the boundaries between the physical world and virtual reality, allowing users to interact with virtual objects through the physical world and vice versa, thus having the possibility of processing any information or value in real time.

Using distributed ledger technology, users can buy and sell non-fungible crypto-assets through fungible crypto-assets within the metaverse. In fact, within the scope of a "blockchain-based virtual world" operating on the basis of a "virtual economy", crypto-assets issued using blockchain technology, in addition to allowing the digital representation of fungible financial products, also allow the digital representation of non-fungible non-financial products, whether they are hard assets, i.e. tangible and physical, or soft assets, i.e. intangible or digital goods. In these terms, the possibilities are virtually limitless, with some arguing that the metaverse appears to be the next generation of the internet (Pooyandeh et al., 2022). One way or another, the metaverse seems to be at least an evolution of the internet, with a dominant focus on social interaction. As the metaverse develops and the number of users increases, it seems that more and more personal information will be at risk, including neural information, as we'll see in the next few points (Pooyandeh et al., 2022).

3.3. Digital sensory interaction in the metaverse

Nowadays, in addition to interacting with the smartphone or tablet screen (among others) through touch, sensory interaction with digital environments is usually limited to hearing and sight, i. e. in total to three of the five senses of the human body traditionally known. In the near future, it seems that interaction may include more basic senses, with their perception increasingly similar to that of the physical world. In 2013, Google published the search engine "Google Nose" and offered a service that allowed users to find the product they were looking for through the sense of smell[54]. Although this service was prepared as a kind of April Fool's joke and for the specific conditions of that day, it was considered a success and, at the time, it was clear that users seemed to be ready to take interaction with the internet to the next level[55]. Since then, user-friendly digital sensory interaction systems have been developed to include the sense of smell in the usual interaction with digital environments. By way of example, users of digital environments are intended to have the means to enjoy experiences where they can smell a perfume before purchasing it via

[53] Simões Ferreira, R. (2022, Desember 29). With holograms or in the metaverse, how digital has already reinvented 'live'. Jornal de Notícias. https://clck.ru/3Nkqum

[54] See the Google Nose Beta website. https://clck.ru/3Nkqw4

[55] Nordyke, K. (2023). Google's April Fools' Joke: Search and Smell (Video). The Hollywood Reporter. https://goo.su/UBRW9

e-commerce. Still as an example, the aim is for users to have the means to breathe in the smell of the sea while leaving home as if they were on the beach or even feel its humidity on their skin. The same goes for taste and the sensation of flavor. In 2020, at Meiji University in Japan, a prototype was developed (dubbed the "Tasting Device") that allows the user to experience different flavors from a device adapted to touch with the tongue[56]. In this sense, the concepts of "augmented human being" or "augmented human intelligence" will be increasingly used and common in the context of users' sensory interaction with digital environments.

Within the scope of the metaverse, mobile and wearable devices such as, for example, virtual and/ or augmented reality headsets, in addition to including sensors for the user to detect movement or sound, may also include other types of sensors. Specifically, virtual reality systems are made up of inertial measurement units and include accelerometers, gyroscopes and magnetometers. Time, breathing and light sensors are also included. Augmented reality systems, on the other hand, can detect the user's location and what they see or hear, and most headsets are equipped with time-of-flight (ToF) sensors, vertical cavity surface-emitting lasers (VCSELs), binocular depth sensors, and optical sensors for structural monitoring. Both systems can also include audio-related sensors such as directional microphones, as well as thermal sensors, touch sensors, and front and rear video cameras. The touch sensor can be used to exchange information between humans and machines in the form of a human-machine interface, with the tactile stimulus being activated by this type of sensor (as, for example, in a touchpad). Most of these sensors are used in the context of IoT in an industrial or clinical context, in drones, humanoid robots, among others (Pooyandeh et al., 2022).

However, the equipment associated with the metaverse, which is currently commercialized en masse, still has multiple limitations (both in terms of hardware and software). The overwhelming majority do not seem to be developed enough to offer a considerably immersive metaverse. The perception of sensations in the physical world still seems to be better than in the digital world. Consequently, the overwhelming majority of metaverse platforms do not manage large volumes of user data, which in turn means that we are not close to mass adoption either. The apparent failure of the "Apple Vision Pro" headset, which aims to replace the keyboard, touchscreen and computer mouse, seems to be the best example of this[57]. In this context, and as we will see in the next section, the literature identifies BCIs as the key technology for achieving complete integration

[56] Grad, P. (2020). Digital device serves up a taste of virtual food. TechXplore. https://clck.ru/3Nkr7r

[57] Mitchell, A. (2024, November 12). Apple's Vision Pro flop: Company scales back production of $3,500 VR headset amid lackluster sales, customer complaints. New York Post. https://clck.ru/3Nm5Ka

between the user and the metaverse in the medium to long term (Bernal et al., 2022)[58]. In addition, the development of cutting-edge sensors and hardware, as well as other types of equipment associated with the metaverse, also seems to contribute to the mass adoption of this new model of sensory interaction with the internet[59].

3.4. Brain-computer interface in the metaverse

Neuralink Corp.[60], a nanotechnology company partly owned by Elon Musk[61], has developed a type of BCI that requires neurosurgery to implant an integrated circuit (chip) in the user's brain, which is a concept that both intrigues and discourages many potential consumers. This type of BCI makes it possible to establish bidirectional interaction with the brain, which includes both the neural mechanisms of learning and memory, as well as neurostimulation. In general, although they can be used to restore the ability to speak, write and even walk, they are still viewed with some suspicion[62]. Nevertheless, the literature has studied and contributed to the discussion of the "state of the art", in particular by comparing current BCI to certain virtual and/ or augmented reality devices that constitute a kind of EEG cap. Because it causes ergonomic problems, it was found that the rapid development of virtual and/or augmented reality headsets (among other products) that incorporate too many EEG sensors to monitor brain regions is impeded. It has been argued in the literature that virtual and/ or augmented reality headsets that mitigate the effects of noise resulting from the processing of brain waves seem to be the most suitable for incorporating EEG sensors (Orlosky et al., 2021). It has also been argued that a combination of resources and technologies, including virtual and/ or augmented reality, digital avatars, sensory interaction systems and BCIs, seems to be able to make the metaverse increasingly pervasive and immersive in everyday life. If this is the case, it could reshape the social experience of space and time. The combination of the metaverse

---

58 Consequently, whether we are aware of it or not, we are currently witnessing a rapid and drastic transformation of business and commercial etiquette, both in terms of products and production methods, as well as in types of services and the way they are executed. By way of example, the subconscious or dreams of users themselves do not seem to be safe from marketing campaigns or even propaganda. There are sensory interaction systems that seek to affect the content of the user's dreams through brain stimuli carried out before or during sleep. Specifically, they try to induce the user to view a particular product or service during their dreams.

59 Genser, J., Damianos, S., & Yuste, R. (2024). Safeguarding Brain Data: Assessing the Privacy Practices of Consumer Neurotechnology Companies. The Neurorights Foundation. https://clck.ru/3NkrK7

60 The BCI device developed by Neuralink consists of a small probe containing more than 3,000 electrodes attached by flexible wires that are thinner than a human hair. This device can monitor the activity of 1,000 brain neurons. A "neurosurgical robot" was also built and it can insert 192 electrodes into the brain every minute. See Galeon, D. (2017, November 22). Experts: Artificial Intelligence Could Hijack Brain-Computer Interfaces. Can we prevent AI from hacking into the human brain? Futurism. https://clck.ru/3NkrMh

61 See the Neuralink website. https://clck.ru/3NkrP3

62 Hall, S. B., & Baier-Lentz, M. (2022, February 7). 3 technologies that will shape the future of the metaverse – and the human experience. The World Economic Forum. https://clck.ru/3NkrRf

and BCIs presents arguments for generating new forms of social interaction and interoperability, making communication between the physical world and the digital world ever faster, more effective and efficient, but also more transparent (Dwivedi et al., 2022). In addition to its usefulness in the medical context, the combination of the metaverse and BCIs is also useful in other types of contexts. Specifically, it allows users to control certain objects (tangible or intangible, such as certain robotics products or digital avatars) with their minds, mental spelling, authentication with brain waves or simply enjoying video games or other entertainment (Bernal et al., 2022). This combination can also be used for cognitive assessment, emotional control and increased cognitive performance. Current literature also explores the feasibility of using BCIs to allow direct communication between the brains of different subjects, using both neural mechanisms for learning and memory, as well as neurostimulation (Bernal et al., 2022). However, despite the notorious evolution of BCIs over the last few decades, their full implementation in metaverse scenarios has not yet been studied in the depth it deserves. There still seem to be some open challenges. Firstly, it seems necessary to broadly analyze how BCIs can contribute to the metaverse. Secondly, it seems necessary to measure the performance of these systems and identify the trends and challenges that BCI presents when applied in a metaverse scenario. Last but not least, it also seems necessary to identify the problems, limitations and risks associated with the use of BCIs in the metaverse (Bernal et al., 2022).

3.5. Neurohacking in the metaverse

There are many similarities between the internet and the metaverse when it comes to cybersecurity challenges such as hacking into accounts, phishing, malware, etc. Despite the differences in terms of infrastructure, the metaverse (web 3.0 & web 4.0) presents new types of cybercrime that differ from those that occur on traditional websites (web 2.0). As the use of crypto-assets and central bank digital currencies expands, hackers will be increasingly interested in cracking the metaverse (Pooyandeh et al., 2022). In this sense, monitoring the metaverse and detecting attacks on new platforms will be more complicated than on traditional platforms. In line with what has been exposed so far in the previous points, with the commercialization and mass adoption of products related to the metaverse and virtual worlds, the scope of hackers' activities will increase substantially. Among the main associated risks are the "immersive attack", i.e. a new type of attack in a virtual environment that focuses on the malicious manipulation of a given device in order to physically or mentally harm or disturb the user. Also noteworthy is the "human joystick" attack. This attack consists of controlling users immersed in virtual and/or augmented reality systems within the metaverse, without their knowledge or authorization, in order to move their physical body to another location within the physical world. With the combination of BCIs, especially those used for neurostimulation,

the attacks aim to over-stimulate the target brain regions or inhibit them, thereby interrupting regular brain activity. The damage caused by this type of threat appears to be able to even recreate the effects of neurodegenerative diseases, although more studies in this regard still need to be carried out (Bernal et al., 2022)[63].

3.6. Brain-computer interface based on artificial intelligence

AI has contributed to advances in the analysis and decoding of neural activity, and has even boosted the BCI sector. Over the last decade, a wide range of AI-assisted or even purely AI-based BCI applications have emerged. These "smart" BCIs, including motor and sensory BCIs, have shown remarkable clinical success. In addition to improving the quality of life of paralyzed patients, they have expanded the athletic ability of ordinary people, and accelerated the evolution of robots and neurophysiological discoveries. However, despite technological advances, there are still several challenges in relation to long periods of training and learning (machine learning), producing results (outputs) in real time and also measuring and recording brain activity in the scope of operation of this new type of BCIs. As explained in the previous point (and, in general, within the scope of this chapter III), it seems that there is still a need for more studies in this direction (Zhang et al., 2020).

3.7. Neurohacking and artificial intelligence

Although there is not enough evidence that today's hacker groups have strong technical experience in managing and manipulating AI-based IoT systems, they have probably already realized their enormous potential. Most of these criminal organizations are made up of hackers who are skilled at manipulating, exploiting and misusing any type of computer system. And this with attacks 24 hours a day, and from anywhere in the world (Velasco, 2022). With the use of AI-based BCI technologies, cybercriminals seem to have found a new vehicle to leverage their illegal activities and, in particular, new opportunities to design and carry out attacks against individuals, companies, and even governments. The literature has raised multiple hypotheses. Firstly, if a hacker takes control of BCIs connected to large number of people, he could manipulate them into voting for a particular candidate, a particular party or a particular issue, thereby secretly overthrowing a particular government and/ or entire infrastructures of a particular state. Although, for the moment, this seems like a highly fictitious scenario, the risk of certain hacker groups using BCIs to turn their "hosts" into a kind of army of programmable robots willing to do anything

[63] It is worth noting that both Interpol and Europol are aware of criminal activities carried out within the metaverse. In this regard, see Interpol. (2022, October 20). Interpol launches first global police Metaverse. https://clck.ru/3NkriG; Europol. (2022, October 21). Policing in the metaverse: what law enforcement needs to know. https://clck.ru/3Nkrnf

their "master" commands does not seem to be ruled out at all[64]. Although BCIs were designed by humans to hack the human brain, the same seems to happen with the risk of AI itself using BCIs to hack the human brain[65]. In fact, it seems that certain AI systems have the potential to become hackers themselves once they become «sentient» (Esmaeilzadeh & Vaezi, 2021)[66, 67]. If this is the case, everything indicates that they will have at their disposal the computerized means to assess the vulnerabilities of any type of social, economic, and political system, and then exploit them at an unprecedented speed, scale, and scope and in a way unimaginable by the limited human mind. It is not just a mere difference in level of intelligence. It is a difference of species. It may even happen that certain AI systems aim to crack other AI systems, with human beings themselves watching and constituting little more than mere collateral damage. Everything indicates that this scenario does not constitute hyperbole. In fact, none of these hypotheses require the creation of a science fiction technology from the distant future. It does not seem at all unreasonable to say that the development of AI will become so rapid that it will even surpass human understanding, as, in fact, it already seems to surpass it[68].

Conclusions

The aim of this study is to contribute to the study of neurohacking in the digital and artificial era and, above all, raise awareness about the neurosecurity (and also ethical) implications resulting from the malicious use of technologies associated with the metaverse and AI for neural manipulation purposes.

As a result of this study, it was found that the possible benefits of developing and (mass) implementing/ commercializing this type of technologies may not outweigh the possible disadvantages. As any computer, computer network, or most other forms of information communication technology (ICT), the hard core is based on electronic components with the capacity to process data, i.e. with the capacity to record, process, and store information (and information), and execute algorithms. As such, the overwhelming majority are, in principle, hackable. Rapid technological advances do not even constitute a deterrent. On the contrary, they end up becoming an attraction or even a challenge. Just think of the recent increase in cyberattacks on public organizations, or even cybercrime in general.

64 Lau, J. (2020, November 18). Hacking Humans: How Neuralink May Give AI The Keys To Our Brains. Forbes. https://clck.ru/3NkrqV

65 Schneier, B. (2021). The Coming AI Hackers. Belfer Center for Science and International Affairs. Harvard Kennedy School. https://clck.ru/3NkrrW

66 Ibid.

67 See also Johnson, A. (2024, March 19). Consciousness for Artificial Intelligence? IEEE Pulse. https://clck.ru/3Nkrtg

68 Schneier, B. (2021). The Coming AI Hackers. Belfer Center for Science and International Affairs. Harvard Kennedy School. https://clck.ru/3Nkrw7

While we are currently witnessing severe consequences in terms of financial and, often, reputational damage, in the near future, and in a world of a "multiverse" of metaverses with "smart technology", the consequences of implementing chips in the human brain to establish a kind of symbiosis with the internet and AI will be unimaginable. Although every detail of human life today is already monitored[69] and analyzed[70] by mobile devices (such as smartphones and/or smartwatches) in order to collect and store data for the purpose of creating detailed psychographic profiles, the human brain contains information that, by its very nature, cannot be recorded without neural interaction. The direct connection between the human brain, mobile devices and AI could establish a kind of access route to human consciousness. It could also allow certain groups of hackers (who are usually one step ahead of security protocols) to take control of the human mind, which includes the decision-making process and its execution. In this scenario, imagine the consequences in terms of voting and in the context of political elections, or even in the context of other more or less peaceful political movements. In the war and military sector, imagine the unlimited possibilities for creating "supersoldiers" or even in terms of surveillance and social control. Imagine, in a future limit and apocalyptic scenario, the eventuality that AI itself could take control of the neural interface and, thus, human consciousness.

In any case, it seems that, in the near future, it will be increasingly easy to plant ideas or even ideologies in people's heads (which nowadays mostly happens through social networks). To what extent will it be possible to guarantee the protection of personal data in the event of a "neuroattack"? The term "personal data" takes on a whole new level in this context. Combining the Freudian theory of the perception-consciousness system, to what extent is it possible to guarantee the protection of the "conscious psychic process" itself, including the previous unconscious state? Who does not remember the movie "Inception", starring Leonardo DiCaprio in 2010? In this context, there are even some philosophical questions. What is the real world? How can we know if it is real or not? Can life in the metaverse be considered "more real" than what is considered real life today?

Among Elon Musk's most famous warnings (often seen as a kind of savior of humanity[71]), two very peculiar ones stand out. In 2018, he said that AI could become an "immortal dictator"[72], from which humanity "will never be able to escape". In 2020,

---

69 Szoldra, P. (2016, September 16). This is everything Edward Snowden revealed in one year of unprecedented top-secret leaks. Business Insider. https://clck.ru/3Nkry8

70 Shane, S., Rosenberg, M., & Lehren, A. (2017, March 7). WikiLeaks Releases Trove of Alleged C.I.A. Hacking Documents. The New York Times. https://clck.ru/3Nkrzw

71 Dowd, M. (2017, March 26). Elon Musk's Billion-Dollar Crusade to Stop the A.I. Apocalypse. The Vanity Fair. https://clck.ru/3NkkHp

72 Holley, P. (2018, April 6). Elon Musk's nightmarish warning: AI could become 'an immortal dictator from which we would never escape'. The Washington Post. https://clck.ru/3NkkGv

he warned again about the same issue, this time saying that "artificial intelligence will overtake humans in less than five years"[73]. Despite these, to say the least, intriguing warnings, the truth is that Elon Musk is also incessantly, obstinately and without apparent limits aiming to develop the functional and processing capacity of this type of technology, as well as expanding its possible forms of application. Take the current technological advances in robotics driven by the company "Tesla, Inc". Specifically, see the "Tesla Optimus" project[74] and the creation of humanoid robots nicknamed "Optimus" [75], whose appearance and characteristics are reminiscent of the robots in the movie "I, Robot", starring Will Smith in 2004, or even the Skynet robots in the movie "Terminator", starring Arnold Schwarzenegger in 1995. Are we unconsciously opening the doors to a fictional, dystopian world, like in the movie "The Matrix", starring Keanu Reeves in 1999? Only time will tell.

References

Abegão Alves, C. (2020). Contra a mente: ensaio de integração das lesões resultantes de intervenções na mente no artigo 143.º do Código Penal. In M. Fernanda Palma et al. (org.). *Livro em Memória do Professor Doutor João Curado Neves. Associação Académica da Faculdade de Direito de Lisboa* (pp. 215–235). AAFDL Editora. (In Portug.).

Bernal, S. L., Pérez, M. Q., Martínez Beltrán, E. T., Martínez Pérez, G., & Huertas Celdrán, A. (2022). When Brain-Computer Interfaces Meet the Metaverse: Landscape, Demonstrator, Trends, Challenges, and Concerns. *Computer Science – Human-Computer Interaction*. https://doi.org/10.48550/arXiv.2212.03169

Bublitz, J. C., & Merkel, R. (2014). Crimes Against Minds: On Mental Manipulations, Harms and a Human Right to Mental Self-Determination. *Criminal Law and Philosophy*, *8*, 51–77. https://doi.org/10.1007/s11572-012-9172-y

Czech, A. (2021). Brain-Computer Interface Use to Control Military Weapons and Tools. In S. Paszkiel (Ed.), *Control, Computer Engineering and Neuroscience. ICBCI 2021. Advances in Intelligent Systems and Computing* (Vol. 1362). Springer, Cham. https://doi.org/10.1007/978-3-030-72254-8_20

Denning, T., Matsuoka, Y., & Kohno, T. (2009). Neurosecurity: Security and privacy for neural devices. *Neurosurgical Focus*, *27*(1), E7. https://doi.org/10.3171/2009.4.focus0985

Dias Venâncio, P. (2011). *Lei do Cibercrime. Anotada e Comentada*. Almedina. (In Portug.).

Dwivedi, Y. K., Hughes, L., Baabdullah, A. M., Ribeiro-Navarrete, S., Giannakis, M., Al-Debei, M. M., Dennehy, D., Metri, Bh., Buhalis, D., Cheung, Ch. M. K., Conboy, K., Doyle, R., Dubey, R., Dutot, V., Felix, R., Goyal, D. P., Gustafsson, A., Hinsch, Ch., Jebabli, I., Janssen, … Wamba, S. F. (2022). Metaverse beyond the hype: Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. *International Journal of Information Management*, *66*, 102542. https://doi.org/10.1016/j.ijinfomgt.2022.102542

Ford, M. (2016). *Robôs: A Ameaça de um futuro sem emprego*. Bertrand Editora. (In Portug.).

Esmaeilzadeh, H., & Vaezi, R. (2021). Conscious AI. *arXiv:2105.07879*. https://doi.org/10.48550/arXiv.2105.07879

Ienca, M., & Haselager, P. (2016). Hacking the brain: brain-computer interfacing technology and the ethics of neurosecurity. *Ethics and Information Technology*, *18*, 117–129. https://doi.org/10.1007/s10676-016-9398-9

[73] Cuthbertson, A. (2020, July 27). Elon Musk claims AI will overtake humans 'in less than five years'. Independent. https://clck.ru/3NkkF2

[74] Levin, T. (2022, January 27). Elon Musk says Tesla's humanoid robot is the most important product it's working on – and could eventually outgrow its car business. Business Insider. https://clck.ru/3NkkDD

[75] Gomez, B. (2021, August 24). Elon Musk warned of a 'Terminator'-like AI apocalypse – now he's building a Tesla robot. CNBC. https://clck.ru/3NkkBk

Kotchetkov, I., Hwang, B. Y., Appelboom, G., Kellner, Ch. P., & Connolly E. S. Jr. (2010). Brain-computer interfaces: Military, neurosurgical, and ethical perspective. *Neurosurgical Focus*, *28*(5), E25. https://doi.org/10.3171/2010.2.focus1027

Li, Q., Ding, D., & Conti, M. (2015). Brain-computer interface applications: Security and privacy challenges. *IEEE conference on communications and network security (CNS). Florence, 2015* (pp. 663–666). https://doi.org/10.1109/CNS.2015.7346884

Marques, G., & Martins, L. (2006). *Direito da informática* (2.ª Edição). Almedina. (In Portug.).

Miranda, R., Casebeer, W. D., Hein, A. M., Judy, J. W., Krotkov, E. P., Laabs, T. L., Manzo, J. E., Pankratz, K. G., Pratt, G. A., Sanchez, J. C., Weber, D. J., Wheeler, T. L., & Ling, G. S. F. (2015). DARPA-funded efforts in the development of novel brain-computer interface technologies. *Journal of Neuroscience Methods*, *244*, 52–67. https://doi.org/10.1016/j.jneumeth.2014.07.019.

Mochan, A., Parkin, B., Farinha, J., & Bailey, G. (2025). *Emerging applications of neurotechnology and their implications for EU governance*. Publications Office of the European Union. Luxembourg.

Moulin, T. (2022). 'I Will Control Your Mind': The International Regulation of Brain-Hacking. *San Diego International Law Journal*, *24*(65).

Orlosky, J., Sra, M., Bektaş, K., Peng, H., Kim, J., Kosmyna, N., Höllerer, T., Steed, A., Kiyokawa, K., Akşit, K. (2021). Telelife: The Future of Remote Living. *Frontiers in Virtual Reality*, *2*, 763340. https://doi.org/10.3389/frvir.2021.763340

Pooyandeh, M., Han, K. J., & Sohn, I. (2022). Cybersecurity in the AI-Based Metaverse: A Survey. *Applied Sciences. MDPI*, *12*(24), 12993. https://doi.org/10.3390/app122412993

Radu, R. (2024). Neurotechnologies and the future of internet governance. Technical Report. EUI. RSC (Global Governance Programme). *GIFI. Issue Brief*.

Rodrigues, B. S. (2009). *Direito Penal Especial. Direito Penal Informático-Digital*. Almedina. (In Portug.).

Rosenfeld, P., Biroschak, J., & Furedy, J. (2006). P300-based detection of concealed autobiographical versus incidentally acquired information in target and non-target paradigms. *International Journal of Psychophysiology*, *60*(3), 251–259. https://doi.org/10.1016/j.ijpsycho.2005.06.002

Rosenfeld, P. (2011). P300 in detecting concealed information. In B. Verschuere, G. Ben-Shakhar, & E. Meijer (Eds), *Memory detection: Theory and application of the concealed information test* (pp. 63–89). Cambridge University Press. https://doi.org/10.1017/CBO9780511975196.005

Shen, F. (2013). Mind, Body, and the Criminal Law. *Minnesota Law Review*, *97,* 2036–2175.

Vallabhaneni, A., Wang, T., & He, B. (2005). Brain – Computer Interface. In B. He (Ed.), *Neural Engineering. Bioelectric Engineering* (pp. 85–121). Springer. Boston, MA. https://doi.org/10.1007/0-306-48610-5_3

Van Gerven, M., Farquhar, J., Schaefer, R., Vlek, R., Geuze, J., Nijholt, A., Ramsey, N., Haselager, P., Vuurpijl, L., Gielen, S., & Desain, P. (2009). The brain-computer interface cycle. *Journal of Neural Engineering*, *6*(4), 041001. https://doi.org/10.1088/1741-2560/6/4/041001

Vasconcelos Casimiro, S. (2000). *A responsabilidade civil pelo conteúdo da informação transmitida pela Internet*. Almedina. (In Portug.).

Velasco, C. (2022). Cybercrime and Artificial Intelligence. An overview of the work of international organizations on criminal justice and the international applicable instruments. *ERA Forum*, *23*, 109–126. https://doi.org/10.1007/s12027-022-00702-z

Yuan, B., Hsieh, Chih-Hung, & Chang, Chien-Ching (2010). National technology foresight research: A literature review from 1984 to 2005. *International Journal of Foresight and Innovation Policy*, *6*(1), 5–35. http://dx.doi.org/10.1504/IJFIP.2010.032663

Zhang, X., Ma, Z., Zheng, H., Li, T., Chen, K., Wang, X., Liu, Ch., Xu, L., Wu, X., Lin, D., & Lin, H. (2020). The combination of brain-computer interfaces and artificial intelligence: applications and challenges. *Annals of Translational Medicine*, *8*(11), PMC7327323. https://doi.org/10.21037/atm.2019.11.109

Author information

Diogo P. Coelho – PhD student, University of Seville
Address: 4 Calle San Fernando, 41013 Sevilla, Spain
E-mail: diopercoe@alum.us.es
ORCID ID: https://orcid.org/0000-0002-2082-1231
Scopus Author ID: https://www.scopus.com/authid/detail.uri?authorId=57703490300
WoS Researcher ID: https://www.webofscience.com/wos/author/record/GLU-8923-2022
Google Scholar ID: https://scholar.google.com/citations?user=-laUdL8AAAAJ

Conflict of interest

The author declares no conflict of interest.

Financial disclosure

The research had no sponsorship.

Thematic rubrics

OECD: 5.05 / Law
PASJC: 3308 / Law
WoS: OM / Law

Article history

Date of receipt – May 5, 2025
Date of approval – May 26, 2025
Date of acceptance – September 25, 2025
Date of online placement – September 30, 2025

Научная статья
УДК 34:004:343.721:004.8
EDN: https://elibrary.ru/smgmxq
DOI: https://doi.org/10.21202/jdtl.2025.16

Нейрохакинг в эпоху цифровизации и искусственного интеллекта: правовые аспекты защиты нейронной информации

Диого Перейра Коэльо

Севильский университет, Севилья, Испания

Ключевые слова

интерфейс «мозг – компьютер»,
искусственный интеллект,
киберпреступность,
метавселенная,
нейробезопасность,
нейропреступность,
нейротехнологии,
нейрохакинг,
право,
цифровые технологии

Аннотация

Цель: внесение вклада в осмысление концепции нейропреступности, а также в изучение текущих и будущих рисков с точки зрения нейробезопасности в условиях развития цифровизации и искусственного интеллекта.

Методы: в исследовании применен критико-описательный анализ связи между киберпреступностью и нейропреступностью, проведено концептуальное разграничение интерфейса «мозг – компьютер» и вариантов его использования, выполнено описание различий между нейронными и психическими манипуляциями. Исследуется правовая автономия преступлений против психической неприкосновенности по отношению к преступлениям против физической неприкосновенности. Методологический аппарат включает анализ существующих прототипов нейропреступлений на основе четырехфазного цикла интерфейса «мозг – компьютер» и изучение специфики нейрохакинга в контексте метавселенной и технологий искусственного интеллекта.

Результаты: исследование выявило сущностные характеристики нейрохакинга как неправомерного использования нейронных устройств для получения несанкционированного доступа к нейронной информации и ее манипулирования. Определены четыре основных типа приложений интерфейса «мозг – компьютер», подверженных нейрохакингу: нейромедицинские приложения, системы аутентификации пользователей, видеоигры и приложения на базе смартфонов. Установлены модальности нейрохакинга на каждой фазе цикла интерфейса «мозг – компьютер»: манипуляции на этапе ввода нейронной информации, измерения и записи мозговой активности, декодирования и классификации нейронной информации, а также на этапе вывода результата. Проанализированы специфические угрозы нейрохакинга в эпоху цифровизации, включая иммерсивные атаки и атаки типа «человек – джойстик» в метавселенной.

Научная новизна: впервые проведено комплексное разграничение концепций нейропреступности и киберпреступности с выделением их специфических правовых последствий. Предложена авторская классификация нейропреступлений на основе четырехфазного цикла интерфейса «мозг – компьютер». Обоснована необходимость выделения психической неприкосновенности как самостоятельного объекта правовой защиты, отличного от защиты физической неприкосновенности. Впервые исследованы особенности нейрохакинга в контексте метавселенной и технологий искусственного интеллекта, включая анализ новых типов атак и угроз нейробезопасности.

Практическая значимость: результаты исследования имеют важное значение для развития правового регулирования в области нейробезопасности и разработки соответствующих нормативных актов. Выявленные типы нейропреступлений и их классификация могут служить основой для создания специализированного законодательства о защите нейронных данных и психической неприкосновенности. Практические рекомендации по обеспечению нейробезопасности интерфейсов «мозг – компьютер» востребованы в медицинской практике, индустрии видеоигр, системах аутентификации и разработке приложений для смартфонов.

Для цитирования

Коэльо, Д. П. (2025). Нейрохакинг в эпоху цифровизации и искусственного интеллекта: правовые аспекты защиты нейронной информации. *Journal of Digital Technologies and Law*, *3*(3), 397–430. https://doi.org/10.21202/jdtl.2025.16

Список литературы

Abegão Alves, C. (2020). Contra a mente: ensaio de integração das lesões resultantes de intervenções na mente no artigo 143.º do Código Penal. In M. Fernanda Palma et al. (org.). *Livro em Memória do Professor Doutor João Curado Neves. Associação Académica da Faculdade de Direito de Lisboa* (pp. 215–235). AAFDL Editora. (In Portug.).

Bernal, S. L., Pérez, M. Q., Martínez Beltrán, E. T., Martínez Pérez, G., & Huertas Celdrán, A. (2022). When Brain-Computer Interfaces Meet the Metaverse: Landscape, Demonstrator, Trends, Challenges, and Concerns. *Computer Science – Human-Computer Interaction*. https://doi.org/10.48550/arXiv.2212.03169

Bublitz, J. C., & Merkel, R. (2014). Crimes Against Minds: On Mental Manipulations, Harms and a Human Right to Mental Self-Determination. *Criminal Law and Philosophy*, *8*, 51–77. https://doi.org/10.1007/s11572-012-9172-y

Czech, A. (2021). Brain-Computer Interface Use to Control Military Weapons and Tools. In S. Paszkiel (Ed.), *Control, Computer Engineering and Neuroscience. ICBCI 2021. Advances in Intelligent Systems and Computing* (Vol. 1362). Springer, Cham. https://doi.org/10.1007/978-3-030-72254-8_20

Denning, T., Matsuoka, Y., & Kohno, T. (2009). Neurosecurity: Security and privacy for neural devices. *Neurosurgical Focus*, *27*(1), E7. https://doi.org/10.3171/2009.4.focus0985

Dias Venâncio, P. (2011). *Lei do Cibercrime. Anotada e Comentada*. Almedina. (In Portug.).

Dwivedi, Y. K., Hughes, L., Baabdullah, A. M., Ribeiro-Navarrete, S., Giannakis, M., Al-Debei, M. M., Dennehy, D., Metri, Bh., Buhalis, D., Cheung, Ch. M. K., Conboy, K., Doyle, R., Dubey, R., Dutot, V., Felix, R., Goyal, D. P., Gustafsson, A., Hinsch, Ch., Jebabli, I., Janssen, … Wamba, S. F. (2022). Metaverse beyond the hype: Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. *International Journal of Information Management*, *66*, 102542. https://doi.org/10.1016/j.ijinfomgt.2022.102542

Ford, M. (2016). *Robôs: A Ameaça de um futuro sem emprego*. Bertrand Editora. (In Portug.).

Esmaeilzadeh, H., & Vaezi, R. (2021). Conscious AI. *arXiv:2105.07879*. https://doi.org/10.48550/arXiv.2105.07879

Ienca, M., & Haselager, P. (2016). Hacking the brain: brain-computer interfacing technology and the ethics of neurosecurity. *Ethics and Information Technology*, *18*, 117–129. https://doi.org/10.1007/s10676-016-9398-9

Kotchetkov, I., Hwang, B. Y., Appelboom, G., Kellner, Ch. P., & Connolly E. S. Jr. (2010). Brain-computer interfaces: Military, neurosurgical, and ethical perspective. *Neurosurgical Focus*, *28*(5), E25. https://doi.org/10.3171/2010.2.focus1027

Li, Q., Ding, D., & Conti, M. (2015). Brain-computer interface applications: Security and privacy challenges. *IEEE conference on communications and network security (CNS). Florence, 2015* (pp. 663–666). https://doi.org/10.1109/CNS.2015.7346884

Marques, G., & Martins, L. (2006). *Direito da informática* (2.ª Edição). Almedina. (In Portug.).

Miranda, R., Casebeer, W. D., Hein, A. M., Judy, J. W., Krotkov, E. P., Laabs, T. L., Manzo, J. E., Pankratz, K. G., Pratt, G. A., Sanchez, J. C., Weber, D. J., Wheeler, T. L., & Ling, G. S. F. (2015). DARPA-funded efforts in the development of novel brain-computer interface technologies. *Journal of Neuroscience Methods*, *244*, 52–67. https://doi.org/10.1016/j.jneumeth.2014.07.019.

Mochan, A., Parkin, B., Farinha, J., & Bailey, G. (2025). *Emerging applications of neurotechnology and their implications for EU governance*. Publications Office of the European Union. Luxembourg.

Moulin, T. (2022). 'I Will Control Your Mind': The International Regulation of Brain-Hacking. *San Diego International Law Journal*, *24*(65).

Orlosky, J., Sra, M., Bektaş, K., Peng, H., Kim, J., Kosmyna, N., Höllerer, T., Steed, A., Kiyokawa, K., Akşit, K. (2021). Telelife: The Future of Remote Living. *Frontiers in Virtual Reality*, *2*, 763340. https://doi.org/10.3389/frvir.2021.763340

Pooyandeh, M., Han, K. J., & Sohn, I. (2022). Cybersecurity in the AI-Based Metaverse: A Survey. *Applied Sciences. MDPI*, *12*(24), 12993. https://doi.org/10.3390/app122412993

Radu, R. (2024). Neurotechnologies and the future of internet governance. Technical Report. EUI. RSC (Global Governance Programme). *GIFI. Issue Brief*.

Rodrigues, B. S. (2009). *Direito Penal Especial. Direito Penal Informático-Digital*. Almedina. (In Portug.).

Rosenfeld, P., Biroschak, J., & Furedy, J. (2006). P300-based detection of concealed autobiographical versus incidentally acquired information in target and non-target paradigms. *International Journal of Psychophysiology*, *60*(3), 251–259. https://doi.org/10.1016/j.ijpsycho.2005.06.002

Rosenfeld, P. (2011). P300 in detecting concealed information. In B. Verschuere, G. Ben-Shakhar, & E. Meijer (Eds), *Memory detection: Theory and application of the concealed information test* (pp. 63–89). Cambridge University Press. https://doi.org/10.1017/CBO9780511975196.005

Shen, F. (2013). Mind, Body, and the Criminal Law. *Minnesota Law Review*, *97*, 2036–2175.

Vallabhaneni, A., Wang, T., & He, B. (2005). Brain – Computer Interface. In B. He (Ed.), *Neural Engineering. Bioelectric Engineering* (pp. 85–121). Springer. Boston, MA. https://doi.org/10.1007/0-306-48610-5_3

Van Gerven, M., Farquhar, J., Schaefer, R., Vlek, R., Geuze, J., Nijholt, A., Ramsey, N., Haselager, P., Vuurpijl, L., Gielen, S., & Desain, P. (2009). The brain-computer interface cycle. *Journal of Neural Engineering*, *6*(4), 041001. https://doi.org/10.1088/1741-2560/6/4/041001

Vasconcelos Casimiro, S. (2000). *A responsabilidade civil pelo conteúdo da informação transmitida pela Internet*. Almedina. (In Portug.).

Velasco, C. (2022). Cybercrime and Artificial Intelligence. An overview of the work of international organizations on criminal justice and the international applicable instruments. *ERA Forum*, *23*, 109–126. https://doi.org/10.1007/s12027-022-00702-z

Yuan, B., Hsieh, Chih-Hung, & Chang, Chien-Ching (2010). National technology foresight research: A literature review from 1984 to 2005. *International Journal of Foresight and Innovation Policy*, *6*(1), 5–35. http://dx.doi.org/10.1504/IJFIP.2010.032663

Zhang, X., Ma, Z., Zheng, H., Li, T., Chen, K., Wang, X., Liu, Ch., Xu, L., Wu, X., Lin, D., & Lin, H. (2020). The combination of brain-computer interfaces and artificial intelligence: applications and challenges. *Annals of Translational Medicine*, *8*(11), PMC7327323. https://doi.org/10.21037/atm.2019.11.109

Сведения об авторе

Коэльо Диого Перейра – аспирант, Севильский университет
Адрес: Испания, 41013, г. Севилья, Калле Сан Фернандо, д. 4
E-mail: diopercoe@alum.us.es
ORCID ID: https://orcid.org/0000-0002-2082-1231
Scopus Author ID: https://www.scopus.com/authid/detail.uri?authorId=57703490300
WoS Researcher ID: https://www.webofscience.com/wos/author/record/GLU-8923-2022
Google Scholar ID: https://scholar.google.com/citations?user=-laUdL8AAAAJ

Конфликт интересов

Автор сообщает об отсутствии конфликта интересов.

Финансирование

Исследование не имело спонсорской поддержки.

Тематические рубрики

Рубрика OECD: 5.05 / Law
Рубрика ASJC: 3308 / Law
Рубрика WoS: OM / Law
Рубрика ГРНТИ: 10.07.45 / Право и научно-технический прогресс
Специальность ВАК: 5.1.4 / Уголовно-правовые науки

История статьи

Дата поступления – 5 мая 2025 г.
Дата одобрения после рецензирования – 26 мая 2025 г.
Дата принятия к опубликованию – 25 сентября 2025 г.
Дата онлайн-размещения – 30 сентября 2025 г.