T-PD(2024)1

CONSULTATIVE COMMITTEE OF THE CONVENTION FOR THE PROTECTION OF INDIVIDUALS WITH REGARD TO AUTOMATIC PROCESSING OF PERSONAL DATA

CONVENTION 108

The privacy and data protection implication of the use of neurotechnology and neural data from the perspective of Convention 108

by

Dr Eduardo Bertoni
Center for Human Rights and Humanitarian Law
American University Washington College of Law

Prof Dr Marcello Ienca
School of Medicine and Health
Technical University of Munich

Paper prepared for presentation to the 46th Plenary meeting, 5 June 2024.
The opinions expressed in this work are the responsibility of the authors and do not necessarily reflect the official policy of the Council of Europe.

1. Introduction: The Neural Interface Revolution

Neurotechnology is an umbrella term used to describe the spectrum of devices, tools, systems, and algorithms used to understand and/or influence, access, monitor, assess, emulate, simulate or modulate the structure, activity and function of the nervous systems of human beings and other animals. Neurotechnologies include a diverse array of innovations such as, inter alia, neural interfaces (NIs), neuroimaging techniques like structural and functional magnetic resonance imaging (sMRI/fMRI) and electroencephalography (EEG), motor neuroprosthetics, speech neuroprosthetics, assistive neurorehabilitation systems, implantable stimulation systems such as deep brain stimulation (DBS), neuromodulation therapies, neurofeedback systems, and computational models for brain function analysis.

Among the many applications of neurotechnology, neural interfaces (NIs) represent the fastest growing and technologically most disruptive subfamily of neurotechnologies[1,2].Neural interfaces are devices positioned within or outside the nervous system that can record, decode and/or stimulate neural activity. Accordingly, neural activity could be recorded from the central and/or the peripheral neural system. This is important to consider when determining if neural data fall under the category of personal data or under other categories.

NIs can be either implantable or non-implantable. Implantable neural interfaces involve the surgical implantation of devices directly into the nervous system, typically within the brain or peripheral nerves. These interfaces often provide high-fidelity communication between neural circuits and external systems, enabling precise control or monitoring of neural activity. Common examples include implantable brain-computer interfaces (BCIs), when the NI interfaces with the brain, and limb neuroprosthetics, when the NI interfaces with the peripheral nervous system to restore the function of a missing or impaired limb. These technologies offer groundbreaking opportunities for restoring motor function, sensory perception, and communication abilities in individuals with neurological disorders or injuries. In contrast, non-implantable neural interfaces do not require surgical implantation and are usually external to the body. These interfaces can include devices such as electroencephalography (EEG) headsets, transcranial magnetic stimulation (TMS) devices, and functional near-infrared spectroscopy (fNIRS) systems. While non-implantable interfaces may offer less direct access to neural signals compared to their implantable counterparts, they are often more accessible, portable, and less invasive, making them suitable for a broader range of applications, including research, diagnostics, and consumer-oriented neurotechnology products.

NIs are growing in both market size and technological advancement. A recent UNESCO report reveals a 20-fold increase in private investment in neurotechnology from 2010 to 2020, reaching $7.3 billion in 2020, with a projected surpassing of $24 billion by 2027. Over the same period, global neurotech-related patents have more than tripled[3].

1.1. What are the present capabilities of Neural Interfaces?

At the technological level, advances in both hardware (such as electrodes or other recording equipment) and software (including machine learning algorithms for data analysis) have enabled increasingly accurate predictive and retrospective inferences about sensorimotor, cognitive, and affective functions from neural recordings. In some cases, they

[1] Soekadar, S. R. *et al.* in *Policy, Identity, and Neurotechnology: The Neuroethics of Brain-Computer Interfaces* (eds Veljko Dubljević & Allen Coin) 65-85 (Springer International Publishing, 2023)
[2] Valeriani, D., Santoro, F. & Ienca, M. The present and future of neural interfaces. *Frontiers in Neurorobotics* **16**, 953968 (2022)
[3] Hain, D. S., Jurowetzki, R., Squicciarini, M. & Xu, L. 1-179 (UNESCO - United Nations Educational, Scientific Cultural Organization Paris, 2023)

even suggest the potential for reconstructing visual, auditory and/or semantic content of mental states from neural activity, a process often popularized as "mind-reading"[4,5]. We can group current cutting-edge neural interfaces into four main subfamilies: speech NIs, motor NIs, sensory NIs, as well as cognitive and affective NIs.

In the clinical setting, speech neural interfaces hold promise for restoring rapid communication to individuals with paralysis by translating neural activity associated with attempted speech into text or sound. Recent studies have demonstrated the feasibility of developing speech-to-text neural interfaces that record among patients unable to speak intelligibly due to conditions like amyotrophic lateral sclerosis in order to produce intelligible and rapid speech synthesis[6,7]. These devices record spiking activity from intracortical microelectrode arrays while the patients attempted to silently speak sentences and leverage deep learning models to decode large vocabulary at high speed (median rate of 62 to 78 words per minute) after only a relatively short training.

Parallel advances in non-implantable neural interfaces, which decode continuous language from non-invasive recordings, open the prospect of deciphering mental activity also among healthy individuals for non-medical purposes. For example, Tang et al. demonstrated the possibility of reconstructing continuous language using a non-invasive decoder using cortical semantic representations recorded via functional magnetic resonance imaging (fMRI)[8]. To do so, they recorded fMRI data from healthy podcast listeners and used a Large Language Model (LLM) called Generative Pre-trained Transformer (GPT version 1) to extract semantic features from perceived speech, imagined speech and even silent videos. Similar positive outcomes have been achieved in other studies employing different types of generative AI models, such as latent diffusion models[9].Although subject cooperation is required both to train and apply the decoder, the authors acknowledge that language decoder for generalized, non-medical use raise challenges to mental privacy.

Motor neural interfaces (MNIs) have made significant strides in enabling individuals with motor impairments to control prosthetic limbs, computer cursors, and other assistive devices directly with their neural activity. These interfaces can decode motor intentions from neural signals recorded in the motor cortex, translating them into movement commands[10]. For instance, patients have been able to control robotic arms to perform complex tasks such as grasping objects or even feeding themselves. MNIs rely on both invasive methods, like intracortical arrays, and non-invasive methods, such as electroencephalography (EEG), though the former typically offer higher precision and control.

Sensory neural interfaces aim to restore sensory functions, such as vision or hearing, to individuals with sensory impairments. For example, retinal implants have been developed to restore partial vision to individuals with certain types of blindness. These devices convert visual information into electrical signals that stimulate the remaining healthy retinal cells, providing the user with rudimentary vision. Similarly, cochlear implants bypass damaged parts

---

[4] Naddaf, M. Mind-reading devices are revealing the brain's secrets. *Nature* **626**, 706-708 (2024)

[5] Drew, L. Mind-reading machines are coming—how can we keep them in check? *Nature* **620**, 18-19 (2023)

[6] Willett, F. R. *et al.* A high-performance speech neuroprosthesis. *Nature* **620**, 1031-1036 (2023). https://doi.org:10.1038/s41586-023-06377-x

[7] Metzger, S. L. *et al.* A high-performance neuroprosthesis for speech decoding and avatar control. *Nature* **620**, 1037-1046 (2023). https://doi.org:10.1038/s41586-023-06443-4

[8] Tang, J., LeBel, A., Jain, S. & Huth, A. G. Semantic reconstruction of continuous language from non-invasive brain recordings. *Nature Neuroscience*, 1-9 (2023).

[9] Takagi, Y. & Nishimoto, S. in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition.* 14453-14463

[10] Lorach, H. *et al.* Walking naturally after spinal cord injury using a brain–spine interface. *Nature* **618**, 126-133 (2023). https://doi.org:10.1038/s41586-023-06094-5

of the ear to directly stimulate the auditory nerve, allowing individuals with profound hearing loss to perceive sound.

Cognitive and affective neural interfaces are being explored to enhance or monitor cognitive functions and emotional states. These interfaces can potentially be used to assist in conditions such as depression, anxiety, or other mood disorders by modulating neural activity in specific brain regions. For instance, deep brain stimulation (DBS) has been used to treat Parkinson's disease and is being investigated for its efficacy in treating depression and obsessive-compulsive disorder (OCD)[11]. Furthermore, real-time monitoring of brain states using EEG or functional magnetic resonance imaging (fMRI) is being studied for applications in neurofeedback therapies, which train individuals to regulate their brain activity to achieve desired mental states[12].

1.2. What are the future capabilities of Neural Interfaces?

The next decades promise a significant leap forward in neurotechnology, driven by the convergence of advancements in neuroscience and engineering[2,13]. This integration is poised to give rise to novel neural interfaces that will revolutionize various aspects of our lives. In this section, we outline some of the most promising avenues of development.

AI-powered Neurodiagnostic Tools: The fusion of brain imaging with machine learning is ushering in a new era of decision-support systems for clinicians. These systems enable swift and accurate diagnoses of a wide range of neurological disorders, including Parkinson's disease, Alzheimer's disease, epilepsy, brain cancer, and cerebral palsy[14]. Moreover, neurotechnologies are increasingly being integrated into clinical settings to alleviate the burden on healthcare professionals and enhance diagnostic precision.

Synthetic Memory: Recent advances in understanding how information is encoded in the brain pave the way for artificial decoders of individual memories. These neurotechnologies hold the potential to aid memory retrieval and optimize information organization within the brain. For instance, they can be invaluable in assisting eyewitnesses in recalling pertinent memories before legal proceedings, thereby influencing policy. Furthermore, artificial memory technologies offer a promising avenue for overcoming the limitations of human memory, potentially expanding our cognitive capacities.

BCI for Optimized Communication: Brain-computer interfaces (BCIs) have emerged as a groundbreaking solution for restoring communication abilities in individuals with severe disabilities[6,15,16]. From enabling individuals to type sentences using brain activity to decoding full sentences from minimally invasive brain recordings, BCIs are poised to revolutionize speech restoration and communication accessibility. Moreover, advancements in silent-speech interfaces promise to introduce novel forms of communication that prioritize privacy and convenience.

---

[11] Limousin, P. & Foltynie, T. Long-term outcomes of deep brain stimulation in Parkinson disease. *Nature Reviews Neurology* **15**, 234-242 (2019).

[12] Watanabe, T., Sasaki, Y., Shibata, K. & Kawato, M. Advances in fMRI real-time neurofeedback. *Trends in cognitive sciences* **21**, 997-1010 (2017)

[13] Helbing, D. & Ienca, M. Why converging technologies need converging international regulation. *Ethics and Information Technology* **26**, 1-11 (2024)

[14] Raghavendra, U., Acharya, U. R. & Adeli, H. Artificial intelligence techniques for automated diagnosis of neurological disorders. *European neurology* **82**, 41-64 (2020)

[15] Metzger, S. L. *et al.* A high-performance neuroprosthesis for speech decoding and avatar control. *Nature* **620**, 1037-1046 (2023)

[16] Luo, S., Rabbani, Q. & Crone, N. E. Brain-computer interface: applications to speech decoding and synthesis to augment communication. *Neurotherapeutics* **19**, 263-273 (2022)

Neurally Integrated Prosthesis: The next generation of neurally controlled prostheses is set to revolutionize the field of prosthetics. By capturing detailed motor intent directly from brain activity, these prostheses offer enhanced control and integration with the human body. Furthermore, they hold immense potential for facilitating rehabilitation in patients recovering from conditions such as stroke or multiple sclerosis. By bypassing impaired neuromotor systems, neurally controlled prostheses can help patients regain motor function and improve their quality of life.

Augmenting Intelligence and Cognition: Neurotechnologies aimed at cognitive augmentation hold promise for enhancing human performance in higher-order brain functions such as reasoning and decision-making[17,18].By monitoring brain activity and providing insights into cognitive function, these technologies can optimize decision-making processes and facilitate collaboration between humans and artificial intelligence (AI). Moreover, they offer the potential to augment human capabilities by integrating seamlessly with AI systems, resulting in human-AI teams that outperform individuals or AI alone in various tasks.

The future capabilities of neural interfaces are vast and transformative, spanning from revolutionizing medical diagnostics and treatment to enhancing cognitive abilities and facilitating everyday tasks. With ongoing advancements in neuroscience and engineering, we stand at the brink of a new era in neurotechnology that promises to reshape our understanding of the human brain and augment human capabilities in unprecedented ways.

1.3. What data do Neural Interfaces record?

Data collected through neural interfaces are commonly referred to as “neural data”. Neural data are quantitative data about the structure, activity and function of the nervous system, both the central and the peripheral, of a living organism. They encompass data relating to a nervous system's structural, electrical, optical, magnetic, or chemical activity, including both direct measurements of neuronal structure, activity and/or function (e.g., neuronal firing or summed bioelectric signals from EEG) and indirect functional indicators (i.e., blood flow in fMRI and fNIRS). Neural data can be integrated with non-neural contextual data, like voice recordings, smartphone activity logs, or neuropsychological evaluations, to aid in understanding mental processes more comprehensively. Unlike other bodily measurements, the risks and benefits related to gathering and analyzing neural data are unique in both quality and scale. This distinction arises from the inherent nature of neural data and the ethical and legal implications they entail. At the neurobiological level, neural data are the most direct correlates of mental states, as all cognitive and affective activity is primarily processed in the nervous system[19,20]. Therefore, the prospect of decoding or modifying neural activity implies the possibility of decoding or modifying cognitive and affective processes. We define mental states any conglomeration of mental representations and propositional attitudes in the human mind that correspond to the experience of thinking, remembering, planning, intending, perceiving, and feeling[21]. These states are the subjective experiences that constitute an individual's inner mental life and can influence behavior and decision-making. Mental states are typically categorized into several types: (i) Cognitive States, which include thoughts, beliefs, knowledge, reasoning, and problem-solving activities. Examples are thinking about a problem, believing a fact, or recalling a memory. (ii) Affective States, which pertain to emotions and

[17] Cinel, C., Valeriani, D. & Poli, R. Neurotechnologies for human cognitive augmentation: current state of the art and future prospects. *Frontiers in human neuroscience* **13**, 13 (2019)

[18] Valeriani, D., Cinel, C. & Poli, R. Group augmentation in realistic visual-search decisions via a hybrid brain-computer interface. *Scientific reports* **7**, 7772 (2017)

[19] Yuste, R. *et al.* Four ethical priorities for neurotechnologies and AI. *Nature* **551**, 159-163 (2017)

[20] Goering, S. & Yuste, R. On the Necessity of Ethical Guidelines for Novel Neurotechnologies. *Cell* **167**, 882-885 (2016). https://doi.org:10.1016/j.cell.2016.10.029

[21] Ienca, M. & Malgieri, G. Mental data protection and the GDPR. *Journal of Law and the Biosciences* **9**, lsac006 (2022)

feelings, such as happiness, sadness, anger, fear, and joy. Affective states can range from short-lived emotional responses to more prolonged mood states. (iii) Conative States, which involve desires, intentions, and volitions, which are related to the motivation to act. For instance, wanting to eat, intending to study, or deciding to go for a walk are all conative states. (iv) Perceptual States, which are related to the sensory experiences of the world around us, such as seeing, hearing, smelling, tasting, and touching. For example, perceiving the color of a flower or hearing a song.

(v) Sensory States, which include physical sensations such as pain, warmth, cold, hunger, and thirst.

It should be highlighted that neural data are not static entities confined to a read-only format; instead, they possess an inherently dynamic quality characterized by a read-and-write nature. Unlike traditional data stored in memory or on a hard drive, neural data within the brain's intricate networks are subject to constant flux and modification. This dynamicity arises from the brain's remarkable capacity for neuroplasticity, wherein synaptic connections are continually strengthened, weakened, or formed anew in response to experiences, learning, and environmental stimuli. Moreover, advances in neurotechnology have unlocked the potential for direct intervention in neural processes through techniques such as neurostimulation or neuromodulation. By delivering targeted electrical or magnetic impulses to specific brain regions, researchers can actively modulate neural activity, effectively "writing" new patterns of activity or altering existing ones. This bidirectional flow of information between the brain and external stimuli underscores the dynamic nature of neural data, highlighting the need for nuanced approaches to understanding and interpreting brain function.

1.4. Mental Data and Cognitive Biometric Data

Mental states can be inferred not only from neural data but also from some kinds of non-neural data. The set of both neural and non-neural data that can be used to infer mental states have been defined "mental data" or "cognitive biometric data". This emerging data category refers to any data that can be organized and processed to infer the mental states of a person, including their cognitive, affective, and conative states. Types of mental data, which can also be referred to as cognitive biometric data, include information related to emotions, memories, and intentions. These data types are increasingly explored, analyzed, or influenced using state-of-the-art digital tools. Mental data can be generated from both neural and non-neural sources. Inferring mental data from neural data involves a process of neural decoding, typically occurring via reverse inference[22].This process, often popularized under the misleading label of 'mind reading,' involves establishing reliable statistical correlations between patterns of brain activity, function, and structure on one hand, and mental information on the other. Similarly, cognitive biometric data can also be inferred from non-neural sources such as behavioral and phenotypic data. This unified approach to mental data, or cognitive biometric data, provides a comprehensive, sector-agnostic, and unitary method for protecting consumers' mental privacy. It includes the use of sensors that directly measure nervous-system activity, such as those used in electroencephalography (EEG), and electromyography (EMG). Additionally, it encompasses data collection practices that enable inferences about brain activity, such as eye-tracking data, heart rate, skin conductance, facial expressions, and vocal tone. While the notions of mental data and cognitive biometric data can be used interchangeably, the latter specifically emphasizes the biometric and biosensor origins of the data used to infer mental states, thus incorporating a broader range of physiological measurements and highlighting the varied sources and methodologies used to gather data

[22] Poldrack, R. A. Inferring mental states from neuroimaging data: from reverse inference to large-scale decoding. *Neuron* **72**, 692-697 (2011)

that can ultimately be interpreted to understand an individual's mental states. When mental data can be directly linked to an individual, this data falls under the category of personal data.

1.5. Unique features of Neural Data

To understand the privacy challenges associated to the human brain it is important to understand the unique features of neural data. According to some scholars, neural data possesses several unique features compared to other types of data, and these features carry significant implications for human rights[19,23,24,25].

Firstly, neural data is inherently pre-behavioral, akin to genetic data. Unlike behavioral observations, which capture the outward manifestations of cognitive processes, neural data offer insights into the underlying neural substrates of these processes before they manifest behaviorally. This pre-behavioral nature provides a window into the internal dynamics of the brain, offering a more direct glimpse into the mechanisms of cognition and perception and the way they drive behavior.

Secondly, neural data possess propositional and semantic value. They contain information that transcends mere sensory inputs or motor outputs, encompassing abstract concepts, thoughts, and intentions. By decoding neural activity, researchers can access the propositional content of mental states, unraveling the symbolic representations encoded within the brain's neural networks.

Furthermore, neural data exhibit metacognitive layering, reflecting the brain's capacity for self-awareness and introspection. Beyond encoding the content of immediate sensory experiences or cognitive processes, neural activity can reflect higher-order cognitive functions, such as self-reflection, monitoring, and strategic decision-making. This metacognitive layering adds an additional dimension of complexity to the interpretation of neural data, highlighting the recursive nature of cognition and consciousness.

While other data categories display one of these features, only neural data display all 3. This makes them a sui generis data category. However, since some neural and cognitive biometric data are personal data, some of the data protection regulations for personal data might be the framework for any kind of neural data regulation.

In addition, neural data display non-unique yet distinctive characteristics that further highlight their ethical and human rights salience. These include high temporal resolution and dynamic nature. Neural data provides insights into the temporal dynamics of brain activity, capturing the millisecond-scale changes in neuronal firing patterns that underlie cognitive processes, perception, and behavior. This high temporal resolution enables researchers to investigate the precise timing of neural events and their relationship to external stimuli or internal states. Unlike many other forms of data, which are sampled at relatively fixed intervals or time points, neural data is continuously evolving, reflecting the real-time dynamics of neural networks. This temporal granularity allows for the detection of rapid changes in brain activity, such as the onset of neural responses to sensory stimuli or the timing of cognitive processes such as decision-making or attentional shifts. The temporal resolution of neural data is particularly significant for understanding the dynamics of complex brain functions and for studying phenomena that unfold over brief time scales, such as perception, learning, and memory encoding. By capturing the fine-grained temporal dynamics of neural activity,

---

[23] Ienca, M. *et al.* Towards a governance framework for brain data. *Neuroethics* **15**, 20 (2022)

[24] Farahany, N. A. *The battle for your brain: defending the right to think freely in the age of neurotechnology*. (St. Martin's Press, 2023)

[25] Ramos, G. *et al.* The risks and challenges of neurotechnologies for human rights. Report No. ISBN: 978-92-3-100567-1, (UNESCO 2023)

researchers can gain deeper insights into the neural mechanisms underlying behavior and cognition. In terms of human rights significance, the temporal resolution of neural data underscores the need to consider the real-time implications of neurotechnology and neural data analysis for individuals' privacy, autonomy, and cognitive liberty. As techniques for recording and analyzing neural activity become increasingly sophisticated, it is essential to ensure that individuals' rights to control their own thoughts and mental processes are protected, particularly in contexts where real-time neuroimaging or neurofeedback technologies are used. By addressing the ethical, legal, and societal implications of temporal neural data, we can promote responsible and ethical use of neurotechnology while upholding fundamental principles of human rights and dignity. Finally, neural data is inherently dynamic and adaptive, reflecting the brain's capacity for plasticity and change. Unlike static forms of data, which remain fixed over time, neural activity can vary in response to internal and external stimuli, learning experiences, and developmental processes. This dynamicity adds an additional layer of complexity to the analysis and interpretation of neural data, as patterns of activity may evolve over time or in different contexts.

2. Ethical Challenges

The ethical challenges of neural interfaces have been widely debated in the scientific literature and in soft law instruments. These include issues of equitable access to the benefits of neural interfaces, respect for personal autonomy, and the implications for mental privacy and security. As neural interfaces become more advanced and widespread, several key ethical considerations emerge.

First, ***ensuring equitable access to neural interface technologies*** is crucial. These advanced medical devices and enhancements should not be limited to those who can afford them, as this would exacerbate existing social and economic inequalities. Policymakers and healthcare providers must work together to create frameworks that allow for broad access to these technologies, potentially through subsidized healthcare programs or insurance coverage. From a data perspective, the ethical principles of equality and fair distribution of benefits are crucial when considering the collection, use, and sharing of neural data. These principles ensure that the advantages and opportunities derived from neural data are accessible to all segments of society, preventing the exacerbation of existing inequalities. Equality in the context of neural data means ensuring that all individuals have equal opportunities to benefit from advancements in neural interface technologies. ensuring that neural interface technologies and related services are accessible to all individuals, regardless of their socio-economic status, geographical location, or other potentially limiting factors. This could be achieved by providing subsidies, insurance coverage, or public funding to make these technologies affordable. Most importantly it could be achieved by facilitating open data initiatives and promoting the use of open software. Open data initiatives ensure that neural data is freely accessible to researchers worldwide, irrespective of their institutional or geographic limitations. This democratizes research, allowing scientists from less affluent institutions or countries to participate in cutting-edge research without prohibitive costs. By making neural data openly available, researchers can build on each other's work more effectively, leading to faster scientific progress and innovation. Open software and data allow for collaborative efforts, reducing duplication of effort and accelerating the development of new technologies and treatments. Open data and software also enhance the transparency of scientific research, allowing others to validate and reproduce results. This is crucial for building trust in scientific findings and ensuring that advancements in neural interface technologies are based on robust and reproducible research. Furthermore, open access to neural data and software provides educational resources for students and trainees, enhancing learning opportunities in neuroscience and related fields. This helps build a more knowledgeable and skilled workforce capable of advancing neural interface technologies. Finally, it can also significantly reduce the costs associated with developing and deploying neural interface technologies. This is particularly important for low-income regions or underfunded research institutions, enabling them to participate in and benefit from technological advancements without the burden of high costs.

A second ethical imperative is respecting the fundamental ethical principle of **personal autonomy** in the context of neural interfaces. Individuals should have full control over whether to use these technologies and how they are used. Informed consent is essential, requiring that users understand the risks, benefits, and potential implications of neural interface technologies. This includes understanding how their data will be used and who will have access to it. In the context of neural data, exercising personal autonomy means individuals have control over their neural information and how it is used. This is closely related to the concept of cognitive liberty, which refers to the right of individuals to control their own mental processes, thoughts, and consciousness[24,26,27]. Cognitive liberty encompasses several moral rights including, but not restricting to freedom of thought and personal autonomy, i.e., the right to think freely without external interference or coercion. Freedom of thought includes the right to keep one's thoughts and mental processes private[21,26,27,28] (which some scholars refer to as 'mental privacy' -see next section), personal autonomy, and the right to make decisions about one's own mental and neurological states, including the use of technologies that may alter or enhance cognitive functions (which some scholars incude this as 'mental self determination').

It is important to highlight that to respect personal autonomy and exercise cognitive liberty individuals should be fully informed about the potential risks and benefits of using neural interface technologies. This includes understanding how their neural data will be collected, stored, used, and shared. Furthermore, participation in any neural data-related research or technology use should be voluntary, without any form of coercion or undue pressure. With regard to the degree of control over personal data, individuals should have ownership of their neural data, with the right to decide who can access it and for what purposes. Furthermore, they should be able to access their own neural data and transfer it to different service providers if they choose, while having reasonable expectations that their neural data are kept confidential and protected from unauthorized access, for instance by using advanced encryption methods such as homomorphic encryption, federated learning and secure multi-party computation to secure neural data against breaches. Transparency is needed at multiple levels of neural data processing. At the procedural levels, organizations collecting and using neural data should provide clear and accessible information about their practices. This includes detailing how data is collected, stored, used, and shared, as well as the purposes of data collection and any potential risks involved. At the algorithmic level, transparency involves explaining how algorithms process neural data, including the methodologies and assumptions underlying these algorithms. Organizations should disclose how neural data is analyzed and interpreted, ensuring that users understand the logic and potential biases in the algorithms. Additionally, there should be clarity about the decision-making processes influenced by these algorithms and the criteria used for generating outcomes or recommendations based on neural data. This level of transparency helps build trust and allows for external auditing and validation of the algorithms to ensure fairness, accuracy, and accountability. The right to withdraw is also an essential component of respecting personal autonomy and establishing cognitive liberty. Individuals should have the right to withdraw from any neural data collection or use at any time, with assurances that their data will be deleted or anonymized upon request. For those using cognitive enhancement technologies, there should be options to reverse or discontinue the enhancements if desired. Finally, ensuring that the use of neural interface technologies does not lead to coercive situations where individuals feel compelled to use them against their will. This includes both explicitly coercive scenarios where people are ordered or mandated to

[26] Ienca, M. On neurorights. *Frontiers in Human Neuroscience* **15**, 701258 (2021)

[27] Ienca, M. Common human rights challenges raised by different applications of neurotechnologies in the biomedical field. (Committee on Bioethics (DH-BIO) of the Council of Europe, Strasburg, France, 2021)

[28] Susser, D. & Cabrera, L. Y. Brain Data in Context: Are New Rights the Way to Mental and Brain Privacy? *AJOB neuroscience*, 1-12 (2023)

use a neural interface and/or have their neural data collected and to implicitly coercive scenarios in which refusing to perform the activities above may lead to power imbalances and competitive disadvantage.

The increasing capability of neural interfaces to decode thoughts, intentions, and mental states raises a third crucial concern: the prospect of "mind-reading". While voluntary mind-reading for clinical applications such as speech neuroprostheses can be extremely beneficial for patients (for instance people with speech impairments or people who develop aphasia as a consequence of a neurodegenerative disorder such as amyotrophic lateral sclerosis), we also necessitate robust privacy protections to prevent unauthorized access to an individual's neural data. Legal frameworks will need to be interpreted in a way that will ensure that neural data is treated with the same level of confidentiality as other sensitive personal information.

Another aspect that might be considered is the ethical obligation to ensure **data security** when using neural interfaces as security breaches could have serious consequences on neural interface users. Unauthorized access to neural data or the manipulation of neural interfaces could result in significant harm, including the loss of control over one's own body or thoughts. Research has shown that NIs can be hacked and, more broadly, can be subject to cybersecurity vulnerabilities[29,30]. It is critical to develop and implement stringent cybersecurity measures to protect neural interface systems from hacking and other forms of cyber threats.

Going besides data management practices, the use of neural interfaces for non-medical purposes, such as **enhancing cognitive or physical abilities** in healthy individuals, raises additional ethical questions. These include concerns about fairness in competitive environments (e.g., sports, academics) and the potential societal pressure to adopt such technologies to remain competitive. Ethical guidelines must be established to address these issues and ensure that the use of neural interfaces remains voluntary and does not lead to coercion or discrimination. Finally, there is agreement that **informed consent** must be robust and continuously updated as new information about the risks and benefits of neural interfaces becomes available. Additionally, public engagement is vital to ensure that societal perspectives and concerns are considered in the development and implementation of neural interface technologies. This includes transparent dialogue between scientists, policymakers, and the public to build trust and ensure that ethical standards are maintained.

2.1. Assessing the “Mind Reading” Problem

The metaphors “mind reading” and “brain reading” refer to the process of decoding semantic, visual, auditory, or other content of mental states from neural recordings. Just like reading a book involves decoding semantic content from a physical medium (i.e., written text) through which such content is represented, so reading a mind/brain would involve decoding content from the physical medium of neural recording. The mind reading metaphor has a long history. As early as 1755, the physician Guillaume-Lambert Godart suggested that one could ‘read’ the thoughts of a person using appropriate tools to interpret the ‘letters’ found in the brain[31]. In the early 2000s, the excitement of the cognitive neuroscience community in the face of concurrent advances in fMRI and pattern recognition, revamped this metaphor. For example, Haynes argued in 2012 that the field of “brain reading” has made it possible “to read

[29] Ienca, M., Haselager, P. & Emanuel, E. J. Brain leaks and consumer neurotechnology. *Nature biotechnology* **36**, 805-810 (2018)

[30] Ienca, M. & Haselager, P. Hacking the brain: brain–computer interfacing technology and the ethics of neurosecurity. *Ethics and Information Technology* **18**, 117-129 (2016)

[31] Gilbert, F., Burns, L. & Krahn, T. The Inheritance, Power and Predicaments of the “Brain-Reading” Metaphor. *Medicine Studies* **2**, 229-244 (2011)

our very detailed contents of a person's thoughts"[32]. This view, however, has been challenged on both theoretical and empirical grounds. At the theoretical level, the recently passed away philosopher Daniel Dennett and neurologist Marcel Kinsbourne drew a critical distinction between vehicle and content and argued that by applying classical pattern recognition methods to fMRI signals we can only reveal information about the representer, i.e. how mental information is represented in the brain (vehicle), but not the represented, i.e. the actual semantic, visual or other perceptual content of such information[33]. On empirical grounds it was observed that fMRI recordings are characterized by high interpersonal and spatiotemporal variability, as well as poor replicability[34]. For these reasons, there has been a general consensus that neurotechnology can reveal statistically significant correlations between neural data and mental states. Borrowing a Chomskyan terminology we call this *mental privacy in the broad sense* (MPBS). However, neurotechnology cannot "read minds" as it cannot decode the actual content of mental states from brain recordings. We call this second instance *mental privacy in the narrow sense* (MPNS). While MPBS has long been considered a realistic methodological and ethical concern, MPNS has long been considered a rather unrealistic, far-fetched if not entirely sci-fi problem[35].

Since the 2010s, the application of artificial intelligence (AI), particularly deep learning algorithms, to the analysis of neural data has sparked a revolution in neurotechnology and restored some epistemic legitimacy for the reading metaphor. One of the key advancements facilitated by this integration is the emergence of a new form of inference known as reverse inference. Traditionally, forward inference involved manipulating a specific psychological function to identify localized effects on brain activity. However, with the advent of AI-driven approaches, researchers can now engage in reverse inference, where they reason backward from datasets, such as patterns of neuronal activation, to infer the engagement of specific mental processes. This paradigm shift has profound implications for understanding the complex interplay between brain activity and cognition, opening up new avenues for investigating the neural basis of behavior and perception. By leveraging the power of AI, neuroscientists can more effectively decode the intricate patterns of neural activity and gain deeper insights into the workings of the human mind. Experimental setups based on reverse inference proved the possibility of training AI models to reconstruct visual and semantic content of mental states by relying exclusively on the neural recordings. The potential of generative AI to decode mental states cannot be understated. Tang et al. proved the feasibility of reconstructing continuous language using a non-invasive decoder using the GPT1 paradigm, based on 117 million parameters. Ongoing research based on the currently available GPT4 paradigm, which is based on 1,76 trillion parameters is likely to exponentially increase decoding power. These developments indicate that addressing MPNS may become just as pressing as addressing MPBS, as we approach the threshold of potentially demonstrating the first instances of decoding the actual content of mental states directly from neural data.

2.2. Exploring Two Dimensions of what is called Mental Privacy

For those who consider the existence of a new concept labeled under the title "Mental privacy", this notion encompasses two distinct but interconnected dimensions: Mental Privacy in the Broad Sense (MPBS) and Mental Privacy in the Narrow Sense (MPNS). Understanding these dimensions is essential for navigating the ethical and practical implications of

---

[32] Haynes, J.-D. Brain reading: Decoding mental states from brain activity in humans. *The Oxford handbook of neuroethics*, 3-13 (2011)

[33] Dennett, D. C. & Kinsbourne, M. Time and the observer: The where and when of consciousness in the brain. *Behavioral and Brain Sciences* **15**, 183-201 (1992). https://doi.org:10.1017/S0140525X00068229

[34] Eklund, A., Nichols, T. E. & Knutsson, H. Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates. *Proceedings of the national academy of sciences* **113**, 7900-7905 (2016)

[35] Wexler, A. Separating neuroethics from neurohype. *Nature Biotechnology* **37**, 988-990 (2019)

neurotechnology and neural data analysis. Mental Privacy in the Broad Sense (MPBS) pertains to the broader scope of privacy concerns associated with the correlation between neural data and user preferences, behavior, and identity. This dimension involves various scenarios.

MPBS involves revealing statistically significant associations between neural data, often in conjunction with cognitive biometric non-neural data, and user preferences, intentions, personality traits or behavior. Neural markers might be indicative of consumer preferences, political inclinations, or susceptibility to specific advertisements. The implications of this are profound, as it raises concerns about the potential for manipulation or exploitation. Corporations, marketers, or even political entities could use this information to influence individuals in subtle and potentially unethical ways, tailoring messages that leverage these neural insights to alter behavior or decision-making processes.

The second scenario involves re-identification risks for people who wish to remain anonymous. Even when neural data is de-identified to protect individual privacy, there remains a significant risk of re-identification through advanced data linkage techniques. Researchers and malicious actors alike can potentially cross-reference de-identified neural data with other available datasets to re-identify individuals. This undermines efforts to maintain confidentiality and anonymity, posing serious ethical and privacy challenges. For example, studies have shown that a face recognition algorithm could potentially reidentify neuroimaging data by matching reconstructed faces with subjects' photos even when the data are defaced[36]. Reidentification of neuroimaging data could lead to the unintended revelation of other sensitive information shared with it, including diagnoses, genetic details, neuropsychiatric assessments, or personal and family histories. The ability to re-identify individuals can lead to unwanted disclosure of personal information, and could be particularly damaging if linked to sensitive medical, psychological, or behavioral traits.

A third scenario involves insecure data sharing. The sharing of neural data through unsecured channels poses substantial risks. Without robust encryption and secure data transfer protocols, sensitive neural data can be intercepted, accessed, or manipulated by unauthorized parties. This is especially concerning given the sensitive nature of neural data, which could include information about cognitive functions, mental health status, and other deeply personal attributes. Unauthorized access to such data could result in misuse, ranging from identity theft to more insidious forms of personal exploitation.

A fourth MPBS scenario involves cybersecurity vulnerabilities of the neural interface hardware and data repositories. Side-channel attacks, where attackers exploit information gleaned from the physical operation of devices (such as power consumption patterns), can compromise neural data security. These vulnerabilities pose risks not only to the integrity and confidentiality of neural data but also to the safety of individuals using these technologies. Successful cyber attacks could lead to the unauthorized manipulation of neural data, potentially causing harm or distress to users.

Finally, MPBS challenges are raised by the prospect of biomarker identification. Advanced analysis of neural data can inadvertently disclose biomarkers or prodromal signatures indicative of neuropsychiatric disorders or other health conditions. For example, specific patterns of brain activity might reveal early signs of Alzheimer's disease or predispositions to mental health disorders. While this information could be valuable for early diagnosis and intervention, it also raises significant privacy concerns. Individuals may not want such sensitive information disclosed, especially without their explicit consent. The revelation of such

[36] Schwarz, C. G. *et al.* Identification of anonymous MRI research participants with face-recognition software. *New England Journal of Medicine* **381**, 1684-1686 (2019)

biomarkers can lead to stigma, discrimination, and psychological distress, impacting personal and professional aspects of individuals' lives.

Mental Privacy in the Narrow Sense (MPNS) focuses on the more specific concern of decoding the semantic and/or perceptual content from mental states, including cognitive, affective, or conative processes. This dimension involves two primary dimensions: decoding of semantic content and decoding of perceptual content. The ability to decode the semantic content of mental states from neural activity represents a significant advancement in neuroscience. This involves interpreting the meanings, thoughts, or intentions underlying neural signals. Such capabilities raise profound ethical and privacy concerns as they tread into the realm of accessing individuals' inner thoughts and experiences. For instance, researchers could potentially determine what a person is thinking or planning based on their brain activity patterns. This capability challenges traditional notions of mental privacy and autonomy, as it may lead to scenarios where individuals' thoughts are no longer private but accessible to external entities, thereby creating potential for misuse, manipulation, or coercion.

The potential to decode perceptual content, such as visual or auditory imagery, from neural representations further blurs the line between external observation and internal cognition. Advances in neuroimaging techniques and machine learning algorithms have made it possible to reconstruct images or sounds that a person is perceiving or imagining. For example, scientists have developed methods to approximate visual scenes viewed by individuals based on their brain activity. This capability not only raises questions about the extent to which personal cognitive processes can be observed but also about the implications for individual privacy. The ability to access and reconstruct one's sensory experiences opens up possibilities for invasive surveillance and unauthorized access to personal mental experiences.

Exploring both dimensions of mental privacy provide a comprehensive framework for addressing the ethical, legal, and societal implications of neurotechnology and neural data analysis. By considering the broad and narrow senses of mental privacy, stakeholders can develop informed policies, practices, and safeguards to protect individuals' privacy rights while promoting scientific progress and innovation in neuroscience. We recommend that ethical frameworks must be established to ensure that the decoding of neural data respects individuals' autonomy and privacy. This includes setting boundaries on the permissible use of technology for decoding mental content and ensuring that informed consent is obtained from individuals whose neural data is being analyzed. In parallel, legal regulations must evolve to address the unique challenges posed by the ability to decode mental content. This includes defining and protecting mental privacy rights in the light of the unique features of neural data and cognitive biometric data, establishing clear guidelines for data ownership and usage, and implementing robust protections against unauthorized access and misuse of neural data.

2.3. Ethical-Technical Solutions to Mitigating the “Mental Privacy” Challenges

Addressing the challenges posed by Mental Privacy in the Broad Sense (MPBS) and Mental Privacy in the Narrow Sense (MPNS) requires a multifaceted approach that combines technological, regulatory, and ethical measures. By implementing targeted strategies, stakeholders can mitigate privacy risks associated with neurotechnology and neural data analysis.

First, addressing the challenges associated with “Mental Privacy” in the Broad Sense (MPBS) requires a multifaceted approach that combines technological innovations, ethical considerations, and robust security measures. One effective strategy is embracing Privacy-Preserving Technologies (PPT) such as differential privacy, federated learning, homomorphic encryption, and secure multi-party computation (SMPC). Differential privacy involves adding controlled noise to the data, allowing researchers to derive meaningful insights while preserving the privacy of individuals. Federated learning enables collaborative data analysis without sharing raw data by training algorithms across decentralized devices. Homomorphic encryption allows computations to be performed on encrypted data, ensuring data remains secure even during processing. Secure multi-party computation (SMPC) facilitates the analysis of data from multiple parties without revealing the individual datasets. These technologies collectively enhance data privacy by enabling collaborative analysis of neural data without compromising individual privacy.

Another important solution is deploying tools for the selective filtering of neural data. This involves designing systems that allow researchers to focus on relevant information while minimizing the disclosure of sensitive or personally identifiable data. By filtering out extraneous data that is not pertinent to the research objectives, it becomes possible to reduce the risk of privacy breaches and ensure that only the necessary data is used in analysis.

Enhancing the security infrastructure of neurotechnology devices, data repositories, and communication channels is also crucial. Implementing strong encryption methods for data storage and transmission can prevent unauthorized access and data breaches. Regular security audits, the use of firewalls, and intrusion detection systems can further safeguard against cyberattacks. Ensuring that all components of the neurotechnology ecosystem are secure helps maintain the integrity and confidentiality of neural data.

Developing and implementing ethical guidelines and standards is essential to address the potential for neurodiscrimination and ensure the responsible use of neural data. These guidelines should provide clear protocols for the ethical handling of neural data in research, clinical practice, and commercial applications. They should include principles for obtaining informed consent, ensuring data anonymization, and protecting individuals from discrimination based on neural data. Ethical standards can help foster trust in neurotechnological advancements and ensure that the use of neural data aligns with societal values and norms. By integrating these solutions, it is possible to address the broad privacy concerns associated with the correlation between neural data and user preferences, behavior, and identity. This comprehensive approach ensures that the benefits of neurotechnology can be realized while safeguarding individual privacy and maintaining public trust in these emerging technologies.

Second, addressing the challenges associated with Mental Privacy in the Narrow Sense (MPNS) requires robust legal, ethical, and educational strategies to protect individuals' rights and autonomy over their mental states and neural data.

One essential solution is enshrining principles of mental privacy and cognitive liberty in legal frameworks and human rights instruments as evolutionary interpretations of the rights to privacy and freedom of thought. By explicitly recognizing and safeguarding individuals' rights to exercise conscious and voluntary control over their neural data and mental information, laws can provide robust protections against unauthorized access and misuse. This legal recognition can ensure that individuals retain autonomy over their mental experiences and prevent potential abuses of neurotechnology. Including cognitive liberty within current human rights standards emphasizes the importance of mental privacy in the digital age.

Another strategy is introducing a mental privacy impact assessment framework[21] into standard practice. Similar to data protection impact assessments used in other fields, this framework would help evaluate the potential risks and benefits of neurotechnology applications. By systematically assessing how neurotechnological tools and applications might impact mental privacy, stakeholders can make informed decisions and implement appropriate risk mitigation strategies. This proactive approach ensures that the deployment of neurotechnology is carefully considered, balancing innovation with the need to protect individual privacy.

Finally, empowering individuals through awareness campaigns is also crucial. Educating the public about the unique features of neural data and their ethical implications can foster informed consent, autonomy, and agency in the use of neurotechnology. Awareness campaigns can highlight the differences between neural data and other types of personal data, emphasizing the sensitivity and potential risks associated with neural information. By providing individuals with the knowledge they need to understand and navigate these risks, such campaigns can help ensure that consent to use neurotechnology is genuinely informed and voluntary. Empowered individuals are more likely to make choices that reflect their personal values and privacy preferences, thereby enhancing their control over their mental states and neural data.

By integrating these solutions, it is possible to address the specific privacy concerns related to decoding the semantic and perceptual content from mental states. This comprehensive approach ensures that advancements in neurotechnology are made responsibly, with respect for individual rights and autonomy, thereby fostering a trustful relationship between the public and the scientific community.

3. Regulation Initiatives on Neurotechnologies

3.1. Neurotechnologies and Human Rights: A general approach

Despite the "datafication" of societies, the human rights principles that are the foundation of, for example, health care and health systems like the ones in which neurotechnologies are relevant, remain the same. With the rise of emerging technologies in health care and health systems, including neurotechnologies, and the "datafication" of society, it is important to consider the human rights foundation that inform these developments. In this section, we attempt to respond to two very broad questions that may be helpful both for our particular analytical approach, and as a general approach:

1) How can States ensure that individuals can enjoy human rights, especially personal data protections, in this new environment?

2) Is the current human rights, and especially, personal data protections´framework capable of addressing the intangible aspect of this new environment?

Before going to the problems that both questions raise, it is important to highlight some principles, particularly because we will focus our work on the relationship between the use of neurotechnologies and personal data. We will expand on these principles in following sections.

It is also important to consider the emerging legal and regulatory framework around these technologies. Currently we find different "levels" or approaches for such regulations: some of them are coming from what usually is called "hard law", including supra national -like human right treaties- or national general human rights laws. It must be included in this "hard law" framework, supra national or national specific personal data treaties or laws (like the Convention 108 and 108+). However, when we approach the problem of neurotechnologies regulation, we also need to consider that currently there are a lot of initiatives coming from different intergovernmental bodies that also provide with guidelines, models, etc, that might help with the analysis of what we need on how to better respond the two questions mentioned above. Finally, we also consider that it is important to take into consideration what private actors are doing to “self” regulate the use and development of neurotechnologies. The chart below helps to summarize what we just said:

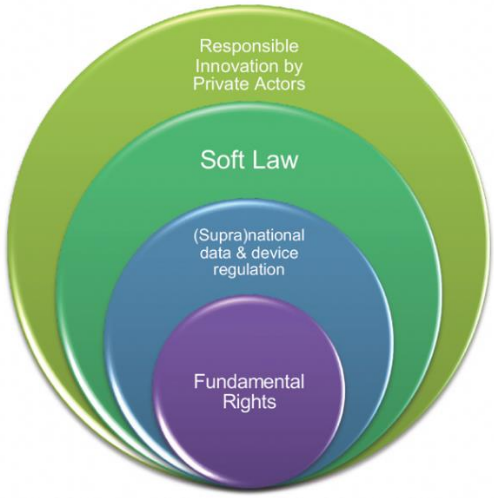

Figure 1- Multi-level Governance of Neurotechnology (re-used with acknowledgment from Ienca 2021)

In the sections below, and having in mind the two questions we asked above, we will begin with the current debate on what people put under the framework of the worry of “rights inflation and devaluation.” We will express our general views about how neurotechnologies interventions are covered by current human right norms and standards coming from soft law and hard law. Then we will mention some initiatives coming from the private sector as useful self-regulatory frameworks.

Because we do not cover all the human rights that are impacted by the use of neurotechnologies -we just give a general overview on this matter-, we include a specific section concerning the relationship between data protection regulations, particularly the Convention 108+ and the use of neurotechnologies. Finally, we will mention some policy options and some open concern and questions that would allow to continue the debate in future works.

3.2. Current Debate on Neurotechnology and Human Rights: The Worry of Rights Inflation and Deflation

The debate on the human rights implications of neural interfaces mimics the ethical debate summarized above. However, it departs from it to the extent that it does not limit to the formulation and validation of normative ethical principles but also on their localization within and adherence to existing regulatory framework. The notion of “neurorights” has been introduced by scholars and policy bodies to refer to the specific set of human rights (both in the sense of moral rights and legal rights) that concern the brain and mind of the human being[26,37,38,39,40,41,42]. The debate on neurorights is characterized by two primary viewpoints: one advocating for the establishment of new, explicitly defined rights to address the ethical and legal challenges posed by neurotechnologies, and the other arguing that existing human rights frameworks, if adequately interpreted and enforced, are sufficient to protect individuals from these emerging threats. Proponents of new neurorights emphasize the need for specificity and targeted protection, while supporters of the existing framework highlight its flexibility and caution against creating redundant and potentially confusing legal standards. This ongoing discourse reflects broader questions about how best to adapt our legal and ethical systems to keep pace with rapid technological advancements.

Proponents of new neurorights argue that neurotechnological advancements create ethical and legal dilemmas that existing human rights frameworks are ill-equipped to address. These technologies, capable of reading, influencing, and potentially altering brain activity, present unique threats to mental privacy, cognitive liberty, mental integrity and psychological continuity that are unprecedented in scope and nature. Advocates believe that the specificity of these threats necessitates the creation of new, explicitly defined neurorights. They contend that existing rights, which were conceived long before the advent of such advanced neurotechnologies, lack the precision needed to effectively safeguard individuals against the nuanced risks of unauthorized neural data collection, cognitive manipulation, and other potential abuses. New rights could provide targeted protections against these specific threats, ensuring that individuals maintain control over their own neural information and cognitive processes. Advocates argue that formal recognition at this level would set a strong precedent and promote the development of legal and regulatory frameworks tailored to the unique challenges posed by neurotechnologies. For example, Farahany has argued that recognizing a right to cognitive liberty would be critical to give people the right to decide how and weather their brains are accessed and changed[24]. Besides academic debates, the movement for neurorights has garnered support from various international organizations. For instance, Chile was the first country to amend its Constitution to protect “mental integrity” and neurodata in 2021. Specifically, the provision states that “the law shall regulate the requirements, conditions, and restrictions for [neurodata], and shall especially protect brain activity, as well as the information derived from it.” Furthermore, scientific and technological developments are to be conducted with “respect for […] physical and mental integrity.” Simultaneously, the same legislators introduced Bill 13.828-19, which aimed to further regulate neurotechnology by requiring consent to use neurotechnology and establishing penalties for noncompliance[39]. A criminal bill law in Argentina, a bioethics law in France and a Charter of Digital Rights in Spain also introduce neurorights protections. Intergovernmental instruments such as the aforementioned OECD Recommendation call for the protection of mental privacy and cognitive liberty.

On the other camp, critics of new neurorights caution against the potential for legal redundancy and complexity. They argue that introducing new rights specifically for neurotechnology could lead to overlap with existing rights, creating confusion and inefficiency in legal and regulatory frameworks. Instead of establishing entirely new rights, they advocate for a clearer and more rigorous interpretation and application of existing rights to address

[37] Ligthart, S. *et al.* Minding rights: Mapping ethical and legal foundations of'neurorights'. *arXiv preprint arXiv:2302.06281* (2023)

[38] McCay, A. Neurorights: the Chilean constitutional change. *AI & Society*, 1-2 (2022)

[39] López-Silva, P. & Valera, L. *Protecting the Mind: Challenges in Law, Neuroprotection, and Neurorights*. Vol. 49 (Springer Nature, 2022)

[40] García-López, E., Muñoz, J. M. & Andorno, R. Neurorights and Mental Freedom: Emerging Challenges to Debates on Human Dignity and Neurotechnologies. *Frontiers in Human Neuroscience* **15**, 790 (2021)

[41] Yuste, R. & De La Quadra-Salcedo, T. Neuro-Rights and New Charts of Digital Rights: A Dialogue Beyond the Limits of the Law. *Ind. J. Global Legal Stud.* **30**, 15 (2023)

[42] Ienca, M. & Andorno, R. Towards new human rights in the age of neuroscience and neurotechnology. *Life sciences, society and policy* **13**, 1-27 (2017)

neurotechnological challenges. For instance, Susie Alegre, a human rights lawyer and author of 'Freedom to Think' published an newspaper article titled "We don't need new 'neurorights'— we need to know the existing law" arguing that a framework for neurorights already exists within existing human rights law. Her concern is the ability of regulators to apply this framework properly. While acknowledging the importance and risks of the emerging technology, she argues that the more significant concern is the tendency for policy discussions to ignore existing, applicable laws. She cautions that advocating for a new set of neurorights risks implying that existing human rights do not apply in this new situation.

Furthermore, critics of new neurorights caution against the potential for legal redundancy and complexity. They argue that introducing new rights specifically for neurotechnology could lead to overlap with existing rights, creating confusion and inefficiency in legal and regulatory frameworks. Instead of establishing entirely new rights, they advocate for a clearer and more rigorous interpretation and application of existing rights to address neurotechnological challenges. For instance, the former UN Special Rapporteur Professor Ahmed Shaheed emphasized that the right to freedom of thought is capable of addressing the new challenges brought forth by emerging neurotechnologies. According to this view, what is needed is regulators who have a strong understanding of our existing legal framework and the creativity to apply it in these new situations. She also emphasizes the need for robust access to justice for all individuals. Alegre also argues that freedom of thought is the foundation that will inform the legal and ethical limits of using neurotechologies while protecting the freedom to innovate.

A similar view is proposed by Christoph Bublitz who acknowledges the formidable advances in neurotechnology while arguing that the question of whether existing legal frameworks or new "neurorights" are appropriate to address these advances is more straightforward[43]. He emphasizes that new rights are not needed, and this proposal is legally flawed and poses even greater challenges, such as its promotion of rights inflationism, its tendency towards neuroexceptionalism and neuroessentialism, and its lack of grounding in preexisting scholarship. For these reasons, Bubliz also supports efforts to ground the regulation of emerging neurotechnologies in existing rights.

In a more recent work, [44] Bublitz expanded on his argument that existing human rights law is both sufficient to address emerging neurotechnologies and preferrable over developing "neurorights." He explains that extensive protection is afforded to individuals through the foundational human rights to bodily and mental integrity and the right to privacy regarding all neurotechnological inferences. Additionally, Articles 18 and 19 of the IPCCR, which cover freedom of thought, conscience, religion, and opinion, add unquestionable and well-founded strength to protections concerning neurotechnologies. The principle of human dignity provides further support for individual protection. This principle safeguards individual independence and respects individuals' personalities, legal personhood, and subjectivity. Together, these three aspects of human rights law provide a robust and intricate system of protection against harm from neurotechnological interferences.

This last position that oppose the first one, is similar of what we call here an hybrid position, because it proposes to develop neurorights as reconceptualized human rights. In other words, this position share the idea that it is important to interpret human rights in a way that clearly capture the current legal challenges. This view was originally expressed already by Ienca and Andorno (2017) and it has been further articulated by Munoz and Marinaro (2023) who have suggested to reinterpret freedom of thought as a protection of both the forum

[43] Neuroethics (2022) 15: 7 https://doi.org/10.1007/s12152-022-09481-3
[44] Christoph Bublitz (31 Jan 2024): Neurotechnologies and human rights: restating and reaffirming the multi-layered protection of the person, The International Journal of Human Rights, DOI: 10.1080/13642987.2024.2310830 **To link to this article:** https://doi.org/10.1080/13642987.2024.2310830

externum (externalizations of thought through behavior) and the forum internum (the underlying, pre-behavioral mental states)[45].

When assessing the debate above, we argue that before postulating any thesis on the de novo creation of neurorights, it is imperative to prioritize the clear and rigorous interpretation and application of existing rights. From this point of view, we concur that "Human rights law is not the proper forum to address *all* ethical concerns, even when they are substantive, and it also lacks the granularity required for fine-tuned regulations. This is the domain of domestic law." Despite providing an appropriate framework for individual protection in this realm, the generality of human rights law makes it an impracticable forum to guide all ethical concerns and precise regulations. Domestic law is the appropriate tool to address these challenges. At the same time, we recognize, in accordance with the Council of Europe's Manual for Human Rights Education with Young people[46], that "a human rights claim is ultimately a moral claim, and rests on moral values".

In other words, human rights in the moral and in the legal sense are intertwined but distinct. Failing to recognize this distinction would imply considering every moral desiderata a human rights, which would in turn lead to rights inflation[47]. This inflation may lead, as Tasioulas (2019) pointed out, to a "troubling corrosion of the idea of human rights because the distinction between a universal human interest and a universal moral right is too often overlooked by IHRL"[48]. At the same time, neglecting the moral foundation of human rights may undermine the very justification of human rights law, and may lead to a risk of rights deflation, i.e. excessive intepretability of legal rights. It should also be noted that emergent regulatory challenges arise when considering a new "neurorights" framework. Bublitz points out that though the definitions provided by "neurorights" might be reasonable, they have not yet been affirmed by courts, unlike the existing human rights framework.

Given this particular strength of human rights law, we argue that existing rights must be reaffirmed and strengthened to provide appropriate protection for individuals regarding neurotechnologies. Ongoing work in international, soft-law institutions to interpret and clarify the human rights framework might contribute to this process.

Lastly, it is important to highlight that the ideas discussed above have also been embraced by the Steering Committee for Human Rights in the Field of Biomedicine and Health (CDBIO) of the Council of Europe.

The Rapporteur´s Report "Neurotechnologies and Human Rights Framework. Do We Need New Rights?" concludes that

"While specific "neuro"-rights may well be important in the future, it may be premature to embark upon creation of such rights at this juncture. There is no clear consensus regarding the conceptual-normative boundaries and terminology of neurorights. Divergences exist in relation to how these rights are interpreted, named, and conceptually articulated. Moreover, there is a risk that elaboration of new rights could lead to accusations of rights inflation which poses the risk of undermining existing fundamental rights and thus far, proposed "neuro-rights" could be encompassed under many existing human rights instruments and articles. A more productive avenue of exploration in terms of governing innovation and application of

[45] Muñoz, J. M. & Marinaro, J. Á. Neurorights as reconceptualized human rights. *Frontiers in Political Science* **5**, 1322922 (2023)
[46] https://www.coe.int/en/web/compass/what-are-human-rights-
[47] Clément, D. Human rights or social justice? The problem of rights inflation. *The International Journal of Human Rights* **22**, 155-169 (2018)
[48] Tasioulas, J. Saving human rights from human rights law. *Vand. J. Transnat'l L.* **52**, 1167 (2019)

neurotechnologies within and beyond the field of biomedicine may be a form of multi- level governance."

This suggested multi-level governance includes "self regulations", responsible innovation in the work of private actors, soft law norms created by intergovernmental organizations, and finally, *"a process of interpreting and applying existing rights or indeed adding to the scope and content of existing rights to ensure adequate protection of individuals using neurotechnologies for medical, social or economic purposes."*[49]

In the following sections we will mention some of the current state of the art of this multi-level governance, starting with responsible innovation via self-regulation. The next section of this paper discusses the current, state-of-the-art, multi-level governance, first focusing on the principle of responsible innovation via self-regulation.

3.3. Responsible innovation via self-regulation: a few examples

This section addresses the ideas of the Institute of Electrical and Electronic Ideas (IEEE) and The International Brain Initiative (IBI). These are just a few important examples of the many self-regulation initiatives and suggestions emerging in the world of neurotechnology. This paper does not aim to include all, or almost all, self regulation initiatives for responsible innovation. We only cited some examples that we consider important examples of self-regulatory initiatives.

We start with the Institute of Electrical and Electronics Engineers -IEEE-[50] and more specifically, with the IEEE Standards Association (IEEE SA)[51], which, according to their definition it "is a leading consensus building organization that nurtures, develops and advances global technologies, through IEEE." The IEEE's Standards Association (IEEE SA) is a leading organization that, according to its website, aims to build consensus among stakeholders in a way that "nurtures, develops, and advances global technologies."

The IEEE Working Group on Recommended Practices for the Responsible Design and Development of Neurotechnologies approved its recommendation P7700[52] that *"[…] establishes a uniform set of definitions and a methodology to assess the ethical and socio-technical considerations and practices regarding the design, development, and use of neurotechnologies. The applied ethical approach utilizes a responsible research and innovation approach, which enables developers, researchers, users, and regulators to anticipate and address ethical and sociocultural implications of neurotechnologies, mitigating negative unintended consequences, while increasing community support and engagement with neurotechnology innovators."*

The International Brain Initiative (IBI)[53] was created in 2017 to *"catalyse and advance neuroscience through international collaboration and knowledge sharing, uniting diverse ambitions and disseminating discoveries for the benefit of humanity."*

Currently, its "Neuroethics Working Group" focuses on underexplored topics in the neuroethics debate, including issues of global justice in access to neurotechnology. The Working Group also debates the neuroethics of brain and mental health and a data governance framework for brain data that is culturally-aware and legally compatible with all brain initiatives. This governance framework is being designed to facilitate neural data collection, processing and

[49] See Rapporteur Report cited.
[50] https://www.ieee.org/
[51] https://standards.ieee.org/about/
[52] https://standards.ieee.org/ieee/7700/11038/
[53] https://www.internationalbraininitiative.org/

sharing across the various large-scale nation-wide projects that compose the International Brain Initiative while upholding privacy, data protection and ethical standards.

3.4. Soft Law Regulations: Some Examples

The OECD Council adopted the Recommendation on Responsible Innovation in Neurotechnology on December 2019 on the proposal of the Committee for Scientific and Technological Policy.[54]

The Recommendation aims to guide governments and innovators in anticipating and addressing the ethical, legal and social challenges raised by neurotechnologies.

The Recommendation includes the following nine principles:

1. Promoting responsible innovation

2. Prioritizing safety assessment

3. Promoting inclusivity

4. Fostering scientific collaboration

5. Enabling societal deliberation

6. Enabling capacity of oversight and advisory bodies

7. Safeguarding personal brain data and other information

8. Promoting cultures of stewardship and trust across the public and private sector

9. Anticipating and monitoring potential unintended use and/or misuse.

The details of Principles 7 and 9 are particularly important. Principle 7, which encourages stakeholders to safeguard personal brain data and other information, issues guidelines on upholding the right to privacy. Actors should prioritize transparency and consent when working with research participants so that individuals are well-informed about the processes and standards their data will undergo and can give free and informed consent. Actors should also provide individuals with opportunities to have an opinion and a choice about accessing, amending, and deleting personal data. They should also take steps to prevent the data from being used to discriminate against protected individuals and populations. Finally, actors should protect confidentiality by protecting the data against unauthorized use and security breaches.

Principle 9 further details how actors can protect individuals' rights to privacy by anticipating and monitoring their systems for unintended use and misuses. They should take preventative steps to ensure that individuals' right to privacy is protected, such as implementing risk mitigation systems before neurotechnologies are deployed, supporting individual integrity, autonomy, protection of private life, and dignity, and protecting against discriminatory or otherwise improper uses of the data.

[54] OECD, Recommendation of the Council on Responsible Innovation in Neurotechnology, OECD/LEGAL/0457

In October 2023, after the “Informal Meeting of Telecommunication Ministers” of the European Union, the Ministers approved the “León Declaration on European Neurotechnology: A Human Centric and Rights Oriented Approach”.

The Leon Declaration calls on certain actors to take specific steps to protect the human rights to dignity and privacy. For example, it recommends that the European Commission collaborate with standardization bodies, private entities, neurotechnology innovators, and other relevant actors to develop and promote a rights-oriented approach for deploying and using neurotechnologies and a transparent ecosystem for the use of neurotechnology.

The United Nations Human Rights Council -UNHRC-, adopted the Resolution HRC51 L.3[55] on “Neurotechnologies and Human Rights” in 2022. Two issues of that resolution are important to underscore.

The first issue is that the UNHRC recalled *“the obligation and the primary responsibility to promote and protect human rights and fundamental freedoms lie with the State, and that the Guiding Principles on Business and Human Rights: Implementing the United Nations “Protect, Respect and Remedy” Framework sets out that States have a duty to protect, and that business enterprises, including technology companies, have a responsibility to respect human rights.”*

The second issue essential to highlight is that the UNHRC recognized that *“neurotechnology could be promising for human health and innovation, but that, at the same time, the continued development of some of its applications may pose a number of ethical, legal and societal questions that need to be addressed, including in human rights terms.”*

For these reasons, the UNHRC requested more studies on the challenges of neurotechnology with regard to the promotion and protection of all human rights.

UNESCO published in 2023 the document “The Risks and Challenges of Neurotechnologies for Human Rights.[56]

The document compiles the viewpoints of experts who participated in an international workshop in November 2021 organized by UNESCO. One of the main questions during the event was whether existing international legal frameworks are sufficient to protect human rights that might be affected by the use of neurotechnologies. Besides the opinions listed above, it is essential to highlight that Philipp Kellmeyer, who at that time was the Research Group Leader at the Neuroethics and AI Ethics Lab, Medical Center – University of Freiburg, Faculty of Medicine, University of Freiburg in the chapter of the document titled “Neurotechnologies and Fundamental Rights”, considers that *“In the current debate, ethics and law scholars are divided on the question of whether new fundamental rights are necessary or whether the interpretation of existing human rights (and constitutional rights in a national context) suffices to protect mental privacy and mental integrity. At the same time, however, some countries, most notably Chile but also Argentina and Spain, are already deliberating and discussing proposed neuroprotection laws to that effect. Notably, however, there is still a lack of a unified conceptual understanding of the underlying notions, such as mental privacy and mental integrity, which might result in these legal initiatives to potentially under-, mis- or overregulate on these issues.*”

[55] https://daccess-ods.un.org/access.nsf/Get?OpenAgent&DS=A/HRC/RES/51/3&Lang=E
[56] https://unesdoc.unesco.org/ark:/48223/pf0000384185 and https://doi.org/10.54678/POGS7778

The Inter-American Juridical Committee -IAJC- of the Organization of American States[57] approved two documents regarding neurosciences and neurotechnologies.

The first was the “Declaration of the Interamerican Juridical Committee on Neuroscience, Neurotechnologies, and Human Rights: New Legal Challenges For The Americas”, adopted in 2021.[58] This Declaration addresses mental privacy when neurotechnologies are used to obtain mental data, specifying certain restrictions to protect this right. Considering the IAJC’s pre-existing legal framework that protects an individual’s privacy from state and third-party interference, the Declaration acknowledges that when new technologies threaten this right, existing legal frameworks must be reinforced and adapted for the present day. Given the significant risks to human dignity, including loss of personality or autonomy for individuals and invasion into individuals’ minds to influence or change their mental processes, the Declaration emphasizes the importance of developing and implementing a “specific regulatory framework” to address these risks based on the longstanding principles guiding the use of personal information.

The Declaration also includes recommendations to States, the private sector, academia, and the scientific community. It recommends that states enact legislation to ensure that neurotechnologies are only used for “legitimate aims” that protect against surveillance, social control, or social scoring. The private sector should ensure that its development and innovation practices are compatible with human rights and the UN Guiding Principles on Business and Human Rights.

Two years later, the IAJC approved the *Inter-American Declaration of Principles on Neurosciences, Neurotechnologies, and Human Rights.*[59]

Principles 2, 3, and 4 are very clear on the connection between neurotechnologies, the need to respect current human rights, and current personal data principles:

Principle 2: Protection of Human Rights in the design of neurotechnologies. *States shall promote a human rights-based approach in the development of neurotechnologies, seeking to ensure comprehensive protection and respect for human rights in the design of neurotechnologies, their research methods, as well as in their implementation, commercialization, evaluation, and use.*

Principle 3: Neural data are sensitive personal data. *Neural data are highly sensitive personal data. Those responsible for the processing and use of neural data must adopt enhanced privacy and security measures and ensure limits on the use of decoding techniques that allow a person to be identified or made identifiable, especially with databases or sets of information that are shared with third parties. States shall foster measures to ensure control, security, confidentiality, and integrity of neural data.*

Principle 4: Express and informed consent regarding neural data*. The consent of the person to whom the neural data belongs is a prerequisite for access to the collection of brain information. It is vital to guarantee free, informed, express, specific, unequivocal, and flawless consent when it comes to access to or processing of neural activity. The consent given must be revocable at any time. Special protection is required in the case of children and adolescents, as well as persons with disabilities, older persons, and persons deprived of liberty.*

[57] https://www.oas.org/en/sla/iajc/inter-american_juridical_committee.asp
[58] https://www.oas.org/en/sla/iajc/docs/CJI-DEC_01-XCIX-O-21_ENG.pdf
[59] https://www.oas.org/en/sla/iajc/docs/CJI-RES_281_CII-O-23_corr1_ENG.pdf

These soft law principles provide an appropriate framework for states to begin strengthening their laws and regulations surrounding neurotechnologies. They demonstrate the real and sufficient connection between emerging technologies and human rights, and that states must not ignore this framework.

3.5. Neural Data and Mental Privacy from the Perspective of Convention 108

In the above sections, we presented an overview of the approach that some private or intergovernmental organizations took about the necessity -or not- of new regulations for the use of neurotechnologies.

Even though the answer is open, many opinions seem to incline that most of all neurotechnologies uses that could affect negatively the human beings are covered -and so, restricted- by current human rights standards. However, normative work is needed to ensure the adequate interpretation and reconceptualization of existing rights in the light of unique characteristics of neural data.

Neural data presents unique challenges and considerations regarding data privacy when viewed through the lens of Convention 108. Convention 108, also known as the Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data, is the first binding international treaty specifically addressing privacy and data protection. From this perspective, it is essential to examine the implications of neural data on individual privacy rights within the framework established by Convention 108.

Convention 108 recognizes the fundamental right to privacy and emphasizes the importance of protecting personal data from unauthorized access, disclosure, and misuse. In order to be consistently applied to neural data, Convention 108 should be interpreted as in a manner that adequately safeguards individuals' mental privacy rights, which encompass both the broad and narrow senses outlined in the paragraph. In the broad sense, neural data raises concerns about the correlation between brain activity and user preferences, behavior, and identity. Convention 108 highlights the risks associated with the unauthorized collection, sharing, and analysis of data, particularly in scenarios where statistically significant associations are revealed or re-identification risks arise from de-identified data. Moreover, Convention 108 emphasizes the importance of addressing cybersecurity vulnerabilities and ensuring secure data sharing practices to protect against unauthorized access or misuse of sensitive data. This would also apply to neural data.

In this section, we elaborate on the idea that Convention 108 and Convention 108+ and their interpretation made by different bodies of the Council of Europe and Scholars should be the basis of the regulation of neurotechnologies to respect a very concrete human right: the right to privacy and more specifically, the personal data protection right. In other words, we will demonstrate how Convention 108 and Convention 108+ and their interpretations by various scholars and the Council of Europe could provide a robust basis for the regulation of new neurotechnologies rooted in the right to privacy and the protection of personal data.

To begin, it is important to start recalling the basic principles and legal bases for data processing that might be related to the use of neurotechnologies and are included both in Convention 108 and Convention 108 +. These principles and legal basis are the ones that might be applied to the use of neurotechnologies. We have already highlighted the following ones: [60]

[60] Recommendation CM/Rec(2019)2 of the Committee of Ministers to Member States on the Protection of Health-Related Data.

a. The data should be processed transparently and in a way that is aligned with relevant law.
b. The data should be collected for "explicit, specific, and legitimate purposes" and should not be processed in any other way incompatible with the initial purpose.
c. The data should be processed in a "necessary and proportionate" way, considering the legitimate purpose for which the data was collected.
d. Personal data, to the extent possible, should be collected from the data subject. When this is not possible, the data can be collected from other sources.
e. The data should not exceed what is necessary to accomplish the initial purpose but should also be accurate and updated.
f. Appropriate security s should be implemented to prevent accidental and unauthorized access to data or its destruction, loss, modification, or disclosure.
g. The rights of individuals to access their data and the rights to "information, rectification, objection, and deletion"[61] should be upheld.

These principles and legal bases apply to the discussion of neurotechnologies only when working under the assumption that most or all neurotechnologies process personal data. Data protection principles apply to neurotechnologies whenever the identification or re-identification of individuals is possible, but will not apply to data that has been processed and has lost all connections to the individual who provided the data.

The current definition of personal data encompasses any identified or identifiable individual information.[62] or any information relating to an identified or identifiable individual."[63] With that definition in mind, it appears that data collected by neurotechnologies should be considered personal data if they encompass any identified or identifiable individual information. By defining personal data as any data associated with an identified or identifiable individual, we can categorize some data collected about individuals by neurotechnologies as personal data.

However, Convention 108 was drafted before the widespread use of neurotechnology and the collection of neural data became prevalent. As such, the convention's scope and definition of personal data, according to some scholars, may not explicitly encompass neural data or adequately address its distinct characteristics and privacy implications, so it could be important to clarify by the interpretation of the Convention that the definition of personal data explicitly include neural data and that ensure appropriate protection under the convention.

On this issue (definition of neural data as personal data) it is important to recall that the IAJC, mentioned in the previous section, takes this position. In the "Notes to the Inter-American Declaration of Principles on Neuroscience, Neurotechnologies, and Human Rights," and about the Principle 3 cited above, the Committee clearly explained why "neural data" is not only "personal data" but also "highly sensitive personal data". The Committee said that:

The term "neural data" refers to data resulting from the use of new technologies for the identification and coding of the human brain's own biosignals. […] Likewise, sensitive personal data are construed to mean data referring to the private sphere of their owner whose misuse may lead to discrimination or place the person concerned at grave risk. […] Neural data are particularly likely to cause considerable harm to individuals if misused. Using artificial

[61] Recommendation CM/Rec(2019)2 of the Committee of Ministers to Member States on the Protection of Health-Related Data
[62] Convention for the Protection of Individuals with Regard to Automatic Processing of Personal Data
[63] Modernized Convention for the Protection of Individuals with Regard to the Processing of Personal Data

intelligence algorithms, neurotechnologies can recognize and decode neural information. This makes it possible to interpret (albeit in a limited way) the electrical parameters generated in the brain. That, in turn, allows correlations to be made between the decoded neural information and certain personality traits of an individual: information that can be used for non-medical or research-related purposes. Neural data may also be used for biometric identification, because a person's brain activity is unique, identifiable, and distinguishable from others, making it the most reliable means of biometric identification available to date. For these reasons, this principle seeks to protect brain information from intrusion by any individual, organization, or government that seeks to use neural data in a manner not consented to by the individual. It is for this reason that those responsible for the processing and use of neural data must adopt privacy and security measures commensurate with the sensitivity of those data and their ability to harm the owner of the data.

It should be noted, however, that not all neural data are unique and identifiable. Therefore, specific protections are needed to protect neural or cognitive biometric data that do not fall under personal data.

Member States may choose to include stricter conditions in their law when neural data is considered sensitive. If that is the case, definitions in the Appendix to Recommendation CM/Rec(2021)8[64] might be useful to take into account: processing of sensitive data defined under Article 6 of Convention 108+ in the context of neural data processing should only be allowed where appropriate safeguards are enshrined in law, and the data are necessary for the lawful and specific purposes of the processing. Additionally, processing for the purpose of detecting or predicting racial or ethnic origin, political opinions, trade union membership, religious or other beliefs, health, or sexual life should be prohibited or only be allowed where appropriate safeguards are enshrined in law, and the data are strictly necessary for the lawful and specific purposes of the processing. When consent is required, it should be explicit where the processing concerns such data.

Professor Bublitz, taking into consideration the European General Data Protection Regulation -GDPR- explained that: *For example, the European General Data Protection Regulation (GDPR) has a special category of sensitive data, including genetic and health data (Article 9). Much data about the brain ("neuro") or stemming from medical examinations of it (neuroimaging) are covered by this category. As a consequence, processing of such data is prohibited, with enumerated exceptions. Insofar as some forms of neurodata are not covered but should be so, one may insert "neurodata" to Article 9, next to other types of data such as genetic data . No need for further reforms. Surely, the GDPR and other regulations may have shortcomings, but they do not arise from the nature of neurodata, but rather from developments such as big data, or data-driven business models in which consumers voluntarily exchange their data for the use of services. These issues need to be addressed, but within existing frameworks.*[65]

The list of sensitive data categories in Convention 108 might not be comprehensive enough to include, e.g., 'emotions' or other 'thoughts' when they are not related to health status, sexuality, or political/religious beliefs, such as data collected to detect emotions for psychographic profiling on social media to analyze consumer behavior. Nonetheless, data collected about these internal processes should still be considered personal, whether they broadly identify thoughts and feelings or belong to one of the protected categories.

[64] Recommendation CM/Rec(2021)8 of the Committee of Ministers to member States on the protection of individuals with regard to automatic processing of personal data in the context of profiling
(Adopted by the Committee of Ministers on 3 November 2021at the 1416 meeting of the Ministers' Deputies)
[65] At Neuroethics (2022) 15: 7 https://doi.org/10.1007/s12152-022-09481-3 page 7

Because is possible to consider some neural data as personal data, regulators should incorporate the previous work done by the Council of Europe that *mutatis mutandi,* apply to the neurotechnologies field, including the previous interpretations on technologies such as big data, facial recognition, health data, and artificial intelligence. In other words, recent documents coming from the Council of Europe that already interpreted Convention 108 and 108+ for technologies such as big data, facial recognition, health data, and artificial intelligence are helpful to take into account in our work.

To begin with the individual rights, not the only ones but maybe the most important are:[66]

a. Individuals are entitled to not be subjected to significantly impactful decisions based solely on automated neural data processing without considering their views.
b. Individuals have the right to obtain confirmation of their neural data being processed.
c. Individuals have the right to obtain information about why their neural data is being processed and where the results of that processing are being applied.
d. Individuals have the right to object to the processing of neural data unless the controller has a legitimate ground for processing that is more important than the individual's rights or fundamental freedoms.
e. Individuals have the right to request the erasure of neural data being used inconsistently with the provisions of the convention.
f. Individuals have the right to erase the data if it is processed in violation of the convention. If the controller refuses, "some remedy" should be made available to the individual.[67]

The transparency of processing is also an aspect that should be considered when neurotechnologies are used. Individuals should be given information including the purpose of processing, the consequences of failing to give a response, the identity of the neural data controller and the processor, the recipients of the neural data, the methods of exercising their rights, and transfers of their data to other countries.[68] They should also be informed about the techniques used for collection and whether those techniques are invasive or non-invasive. This information should be given to the individual in a way that the specific person can understand, considering their ability to comprehend the information.

Concerning the access to neural data and the individual's ability to request that the information is erased, any regulation should account for what already has been said for the processing of health data:[69] In certain circumstances, individuals can request to have their neural data erased.

Regarding the automatic processing of neural data, people have the right not to be subject to a decision that significantly affects them based on the automatic processing of neural data without their views taken into account. This is a very well-established principle

[66] Modernized Convention for the Protection of Individuals with regard to the Processing of Personal Data
[67] Recommendation CM/Rec(2019)2 of the Committee of Ministers to Member States on the Protection of Health-Related Data
[68] Recommendation CM/Rec(2019)2 of the Committee of Ministers to Member States on the Protection of Health-Related Data
[69] Recommendation CM/Rec(2019)2 of the Committee of Ministers to Member States on the Protection of Health-Related Data

coming from Convention 108 that was constantly mentioned in the documents mentioned above.

The same must be said concerning the acquisition and processing of neural data: The data should be obtained, stored, and processed lawfully and not used for illegitimate purposes. The data collected and stored should not exceed what is necessary to accomplish the intended purpose and should be stored in a way that ensures that the individual is only identifiable for only the necessary duration.[70]

If neural data will be used for profiling or high-risk profiling activities, as defined in the Appendix to Recommendation CM/Rec(2021)8[71], regulators in the neurotechnologies field should follow this appendix's general principle*: "respect for fundamental rights and freedoms, notably the rights to human dignity and to privacy but also to freedom of expression, and for the principle of non- discrimination and the imperatives of social justice, cultural diversity and democracy, should be guaranteed, in both the public and private sectors, during the profiling operations covered by this recommendation."*

At the same time, if the information coming from neural data processing is used for predictive purposes in decision-making processes, neurotechnology regulators should incorporate the "Principles and Guidelines"[72] regarding the processing of Big Data. As the general principle established, *"the controllers and processors should adequately take into account the likely impact of the intended Big Data processing and its broader ethical and social implications to safeguard human right and fundamental freedoms, and ensure the respect for compliance with data protection obligations as set forth by Convention 108."*

Moreover, Convention 108 addresses the transfer of personal data across borders and requires that adequate safeguards be in place to protect data privacy. However, the international nature of neuroscience research and collaboration may necessitate additional provisions or mechanisms for ensuring the protection of neural data in cross-border transfers.

Convention 108 promotes principles of data minimization and purpose limitation to ensure that personal data is collected and used only for specified, explicit purposes. However, the dynamic and multifaceted nature of neural data may make it challenging to adhere to these principles effectively. In most neural recordings, the principle of purpose limitation is nearly inapplicable as it is not possible to selectively filter purpose-specific information in the dynamic flow of neural data. New guidelines may be needed to establish to assure the minimun collection of neural data and ensuring that it is used only for legitimate and well-defined purposes.

Finally, consent should also play an important role in the field of neurotechnologies because individuals have the right to consent to the collection of their brain information ("free, informed, express, specific, unequivocal, and flawless consent."). Convention 108 emphasizes the importance of obtaining informed consent for the collection, processing, and sharing of personal data. However, the unique nature of neural data, which often involves recording or stimulating brain activity directly and below the threshold of conscious awareness, may raise challenges in obtaining truly informed consent. New interpretations may be needed to develop guidelines or standards for obtaining consent specifically for the collection and use

[70] Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data

[71] Recommendation CM/Rec(2021)8 of the Committee of Ministers to member States on the protection of individuals with regard to automatic processing of personal data in the context of profiling (Adopted by the Committee of Ministers on 3 November 2021 at the 1416 meeting of the Ministers' Deputies)

[72] Consultative Committee of the Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data (T-PD) Guidelines on the Protection of Individuals with regard to the Processing of Personal Data in a World of Big Data T-PD(2017)01

of neural data, taking into account the complexities of brain research and neurotechnology, and the subconscious nature of most neural information.

Despite what have been said above,the potential to decode semantic or perceptual content from neural activity raises profound ethical questions about the boundaries of privacy and the right to mental autonomy. Convention 108 emphasizes the importance of obtaining informed consent and implementing safeguards to prevent unauthorized access to individuals' neural data, particularly in contexts where decoding techniques may intrude upon their mental privacy. Convention 108 also emphasizes the importance of ensuring the security and confidentiality of personal data to prevent unauthorized access, disclosure, or misuse. However, the highly sensitive nature of neural data, which provides insights into individuals' thoughts, emotions, and cognitive processes, may require enhanced security measures and safeguards. Specific requirements or standards for protecting neural data from cybersecurity threats, unauthorized access, or inappropriate use may be considered.

As we explained in this work, the unique features of neural data raised special challenges for the Convention. While Convention 108 provides a solid foundation for protecting personal data, including neural data, some areas of the Convention 108 may need, as we explained above, guidelines for new interpretation to better protect neural data.

4. Policy Options and Recommendations to State Parties

We recommend that[73] the principles governing personal data protection set out in Convention 108 must be observed when processing neural data if the data is personal data. They should be recognized as a general, mandatory framework and include definitions of appropriate, specific, and legitimate aims of processing. They should include a requirement for adequate, relevant and not excessive data, limited data storage time, and security measures that guarantee the confidentiality of the data and respect for the rights of individuals.

Furthermore, the rights of persons whose data are collected and processed must be respected, particularly their rights of access and objection to processing, rectification and erasure.

Another policy option relevant to this field is related to scientific research. Scientific research should be considered in any guideline regarding the use of neurotechnologies. However, the adopted Recommendation CM/Rec(2019)2 of the Committee of Ministers to Member States on the Protection of Health-Related Data[74] gives enough guidance on scientific research.

For example, it should be established that the processing of neural data for scientific research should be subject to proper safeguards provided by law, be carried out with a legitimate aim, and be in compliance with the rights and fundamental freedoms of the data subject.

Moreover, regulations might establish that neural data can only be processed in a scientific research project if the data subject has consented unless the law allows for the processing of neural data for scientific research without the data subject’s consent. The provisions of such a law should be proportionate to the aim pursued, respectful of the essence of the right to data protection, and provide for suitable and specific safeguards to protect the fundamental rights and freedoms of the data subject. These safeguards should include the obligation to put

[73] Explanatory memorandum to Recommendation CM/Rec(2019)2 of the Committee of Ministers to member States on the protection of health-related data. Document CM(2019)51-add final

[74] Recommendation CM/Rec(2019)2 of the Committee of Ministers to Member States on the Protection of Health-Related Data

technical and organizational measures in place to ensure respect for the principle of data minimization.

Taking into consideration the risk of any misuse of neurotechnologies, we also recommend that states implement a risk-based classification of these technologies, analogous to what is developed for artificial intelligence technologies.

Furthermore[75], an approach focused on avoiding and mitigating the potential risks of processing personal data is a necessary element of responsible innovation in neurotechnologies. Unacceptable risks (red lines) might be included as a policy option.

Also related to risk assessment, like a “Mental Data Protection Impact Assessment” - MDPIA (Ienca and Malgieri, 2022)- might be considered to develop and for further implementation[21]. This form of impact assessment covers both neural and non-neural cognitive biometric data.

If neural data processing is considered high-risk by default, then the data controller should be obliged to: describe the processing (including a description of the logic of the technology), perform a balancing test based on necessity and proportionality of the data processing in relation to the purposes; assessing the actual risks for fundamental rights and freedoms, and proposing suitable measures to address and mitigate those risks. The MDPIA could imply an audit of the technological components of the processing (eg AI-driven processing) and a thorough evaluation and possible reconsideration of the algorithm to determine if some risks can be mitigated ‘by design’.

Finally, we consider that the development of neurotechnologies that relies on the processing of neural data should be based on the principles of Convention 108+ when the data is considered personal data[76]. The critical elements of this approach are lawfulness, fairness, specification of purpose, proportionality of data processing, privacy-by-design and by default, responsibility, demonstration of compliance (accountability), transparency, data security, and risk management. Of course, the case might not be the same if the data is collected from the peripheral nervous system and that data is impossible to be connected to an identifiable individual. In that case, because it is not considered personal data, it is impossible to apply and personal data regulation. In other words, when personal data is processed by neurotechnologies uses, the principles mentioned above should apply. Only when the data is collected from the peripheral nervous system and not connected to an identifiable individual might these principles not apply, as this data would not fall under the definition of “personal data.”

Overall, Convention 108 provides a robust framework for addressing the data privacy implications of neural data, highlighting the importance of protecting individuals' privacy rights while promoting scientific progress and innovation in neuroscience. By interpreting the convention to adapt and to reflect the realities of the digital age and the complexities of neural research, policymakers can ensure that individuals' privacy rights are adequately protected in the context of neuroscience and brain research.

[75] Guidelines on artificial intelligence and data protection adopted by the Committee of the Convention for the Protection of Individuals with regards to Processing of Personal Data (Convention 108) on 25 January 2019

[76] Guidelines on artificial intelligence and data protection adopted by the Committee of the Convention for the Protection of Individuals with regards to Processing of Personal Data (Convention 108) on 25 January 2019