


Mental data protection and the GDPR

Marcello Ienca ^{1,*} and Gianclaudio Malgieri²

¹EPFL, College of Humanities (CDH), Lausanne, Switzerland

²EDHEC Business School, Augmented Law Institute, Roubaix, France

*Corresponding author. E-mail: marcello.ienca@epfl.ch

ABSTRACT

Although decoding the content of mental states is currently unachievable, technologies such as neural interfaces, affective computing systems, and digital behavioral technologies enable increasingly reliable statistical associations between certain data patterns and mental activities such as memories, intentions, and emotions. Furthermore, Artificial Intelligence enables the exploration of these activities not just retrospectively but also in a real-time and predictive manner. In this article, we introduce the notion of ‘mental data’, defined as any data that can be organized and processed to make inferences about the mental states of a person, including their cognitive, affective and conative states. Further, we analyze existing legal protections for mental data by considering the lawfulness of their processing in light of different legal bases and purposes, with special focus on the EU General Data Protection Regulation (GDPR). We argue that the GDPR is an adequate tool to mitigate risks related to mental data processing. However, we recommend that interpreters focus on processing characteristics, rather than merely on the category of data at issue. Finally, we call for a ‘Mental Data Protection Impact Assessment’, a specific data protection impact assessment designed to better assess and mitigate the risks to fundamental rights and freedoms associated with the processing of mental data.

KEYWORDS: Mental Data, Digital Mind, GDPR, Mental Privacy, Data Protection, Data Protection Impact Assessment

I. INTRODUCTION

In contemporary cognitive science, the human mind is typically described as the set of psychological faculties enabled by neural processes in the brain.¹ These include consciousness, imagination, perception, affection, thinking, judgement, language, and

1 Betty Pfefferbaum and Carol S. North, *Mental Health and the Covid-19 Pandemic*, 383 N. ENGL. J. MED. 510–512 (2020).

memory. Although characterized by a diversity of outlooks, the unifying theoretical commitment of cognitive science is that such mental faculties are constituted of information-bearing structures (sometimes called mental representations), which have informational content, therefore called mental content.^{2,3} However, the immense and sensitive value of this informational set is still not clear in legal terms. Accordingly, this article aims to understand what kind of legal protection ‘mental data’ have in the EU and whether the GDPR is an adequate tool of protection.

The urgency of this topic is clear: innovative data mining techniques, pervasive technologies, and the development of *emotion AI* demand a reflection on whether and how we should specifically protect the informational value of the digital mind and what is the state of the art in the EU legal framework.

While [Section 2](#) will focus on the technological challenges of the digital transformation for the human mind, [Section 3](#) will focus on the EU data protection framework, focusing in particular on the nature of mental data according to the GDPR, on the principle of lawfulness ([Section 3.A](#)) and on the risk assessment of mental data processing ([Section 3.B](#)), calling for a Mental Data Protection Impact Assessment (MDPIA) model.

II. DIGITAL TRANSFORMATION AND THE HUMAN MIND

In the last decade, the widespread adoption of smartphone-based mobile applications, wearable activity trackers, non-invasive neural interfaces in combination with the increased distribution of the Internet of Things (IoT) in both private and public spaces, has fueled a socio-technical trend known as the Quantified Self, ie the use of digital technology (broadly defined) for self-tracking purposes.⁴ Although the first generation of wearable devices and mobile tools could collect data, and provide insights only related to a small portion of human physiology and physical activity, chiefly mobility (eg daily steps and physical position), novel applications can now record a broader variety of human activities and underlying processes, including processes related to a person’s mental or psychological domain. This is due to a two-fold technological transformation.

First, self-quantification technologies have expanded in variety as to include data sources that could previously be collected exclusively via medical devices such as electroencephalography (EEG) and other neurotechnologies.^{5,6} This is possible mainly due to progress in the field of non-invasive brain-computer interfaces (BCIs). In recent years, BCIs and analogous neural interfaces have spillovered from the clinical and biomedical research domain onto the consumer technology market through a variety of personal and often direct-to-consumer applications. Second, smartphone-sensing

2 J. Garsd, *Business Is Booming for Therapy Apps, but What Really Works?* (Marketplace 15 April 2020) <https://www.marketplace.org/2020/04/15/covid-19-therapy-apps-mental-health/> (accessed Apr. 18, 2021).

3 Eran Klein and others, *Engineering the Brain: Ethical Issues and the Introduction of Neural Devices*, 45 HASTINGS CENTER REP., 26 (2015).

4 Melanie Swan, *The Quantified Self: Fundamental Disruption in Big Data Science and Biological Discovery*, 1 BIG DATA 85 (2013).

5 Marcello Ienca, Pim Haselager and Ezekiel J Emanuel, *Brain Leaks and Consumer Neurotechnology*, 36 NAT. BIOTECHNOL. 805 (2018).

6 Gabriella M. Harari and others, *Using Smartphones to Collect Behavioral Data in Psychological Science: Opportunities, Practical Considerations, and Challenges*, 11 PERSP. PSYCHOL. SCI. 838 (2016).

methods have improved in quality and reliability, now permitting a fine-grained, continuous and unobtrusive collection of non-neural psychologically, and socially relevant data such as speaking rates in conversation, tone of utterances, frequency of social interactions, ambient conversations, responses to cognitive tasks, 3D navigation tasks, sleep patterns, purchase preferences etc.⁷ This field of research is typically known as ‘digital phenotyping’.^{8,9} Third, advances in Artificial Intelligence (AI)-driven software, especially deep learning,¹⁰ are increasingly allowing us to derive insights about a person’s mental domain either from their brain data or from non-neural contextual information.¹¹ For example, smartphone apps can be used to infer a person’s cognitive status from their responses to gamified cognitive tasks such as 3D virtual navigation.¹² Convolutional neural networks (CNNs)—a type of network architecture for deep learning—have also proven effective to take in non-verbal cues from facial emotions and detect emotions from human facial images.^{13,14} A subfield of AI research called *emotion AI* (also known as *affective computing*) has emerged with the aim of studying and developing systems that are capable to detect, interpret, process, and simulate human affects and emotions.¹⁵ Although neurotechnologies such as BCIs can provide the informative basis for predictive inferences about mental processes from brain data (ie direct or indirect measures of brain structure, activity, or function), digital phenotyping, affective computing and other digital applications exploit non-neural contextual information such as behavioural and phenotypic data such as voice recordings, written text, and face images to make inferences about mental processes. It should also be noted that since the detection of affective information is highly dependent on collecting passive sensor data about physical states and behavior, emotion AI, and digital phenotyping are mutually intertwined.

The examples above attest that digital technology today can be used not only to measure relevant parameters of human anatomy and activity but also to gain exploratory information about mental faculties such as cognitive processes, personal preferences, and affective states. Furthermore, AI and big-data analytics potentially permit to explore these faculties not just retrospectively but also in real-time and in a predictive manner. When implemented in implantable BCIs, these AI features generate so-called ‘neuroadaptive technologies’, that is neuroinformatic

7 Ibid.

8 Jukka-Pekka Onnela and Scott L. Rauch, *Harnessing Smartphone-Based Digital Phenotyping to Enhance Behavioral and Mental Health*, 41 NEUROPSYCHOPHARMACOLOGY 1691 (2016).

9 Thomas R. Insel, *Digital Phenotyping: Technology for a New Science of Behavior*, 318 JAMA 1215 (2017).

10 Yann LeCun, Yoshua Bengio and Geoffrey Hinton, *Deep Learning*, 521 NATURE 436 (2015).

11 Marcello Ienca and Karolina Ignatiadis, *Artificial Intelligence in Clinical Neuroscience: Methodological and Ethical Challenges*, 11 AJOB NEUROSCI. 77 (2020).

12 S. Lawrence and others, *Face Recognition: A Convolutional Neural-Network Approach*, 8 IEEE TRANS. NEURAL NETW. 98 (1997).

13 Ibid.

14 Masakazu Matsugu and others, *Subject Independent Facial Expression Recognition with Robust Face Detection Using a Convolutional Neural Network*, 16 NEURAL NETW. 555 (2003).

15 Javier Marín-Morales et al., *Affective Computing in Virtual Reality: Emotion Recognition from Brain and Heartbeat Dynamics Using Wearable Sensors*, 8 SCI. REP. 13657 (2018).

systems that automatically adapt to the user’s mindset without requiring explicit instructions.¹⁶

These converging technological developments are increasingly enabling what can be defined *the digital mind*—namely the moment-by-moment quantification of the individual-level human mind using data from neural interfaces and other digital technologies—and a more intimate connection between minds and machines. Although several areas of cognitive science investigate informational structures of the mind and neuroscience provides increased evidence of their structural or functional realization in the brain, digital technologies are increasingly allowing us to grasp aspects of mental content from novel and various types of data sources.

II.A. Digital mind technologies and a definition of ‘mental data’

We define ‘digital mind technology’ any technology *for the exploration, analysis, and influence of mental data*. We define ‘mental data’ any data that can be organized and processed to infer the mental states of a person, including their cognitive, affective, and conative states. For the purposes of this study, we define ‘mental state’ any conglomeration of mental representations and propositional attitudes that corresponds to the experience of thinking, remembering, planning, perceiving, and feeling.

Types of mental data that appear increasingly suited to be explored, analyzed, or influenced using state-of-the-art digital tools include information related to emotions, memories, and intentions. Mental data can be generated from both neural and non-neural data. Inferring mental data from neural data involves a process of neural decoding, which typically occurs via reverse inference.¹⁷ This process, which has been often popularized under the misleading label of ‘mind reading’,^{18,19} generally involves establishing reliable statistical correlations between patterns of brain activity, function and structure, on the one hand, and mental information on the other hand. As we have seen, mental data can also be inferred from non-neural data sources such as behavioural and phenotypic data. Figure 1 clarifies this relationship.

Digital mind technologies can be useful to acquire information *from* and thereby provide assistive tools *for* people living with neurological and mental disorders. Although neural interfaces are increasingly used in neurology and neurorehabilitation, *digital mental health* is a growing field of research and clinical intervention based on the leveraging of digital technologies to improve people’s mental and psychological well-being.²⁰ Digital mind technologies have proven valuable in extending effective mental healthcare in a cost-effective manner and increasing the availability and accessibility

16 Thorsten O. Zander and others, *Neuroadaptive Technology Enables Implicit Cursor Control Based on Medial Prefrontal Cortex Activity*, 113 PROC. NATL. ACAD. SCI. USA 14898 (2016).

17 Russell A. Poldrack, *Inferring Mental States from Neuroimaging Data: From Reverse Inference to Large-Scale Decoding*, 72 NEURON 692 (2011).

18 Leo Kittay, *Admissibility of fMRI Lie Detection: The Cultural Bias Against “Mind Reading” Devices*, 72 BROOKLYN LAW REV. 1351 (2007) <https://brooklynworks.brooklaw.edu/blr/vol72/iss4/5>.

19 Matthias Gamer, *Mind Reading Using Neuroimaging: Is This the Future of Deception Detection?*, 19 EUR. PSYCHOL. 172 (2014).

20 David C. Mohr and others, *Accelerating Digital Mental Health Research From Early Design and Creation to Successful Implementation and Sustainment*, 19 J. MED. INTERNET RES. e153 (2017).

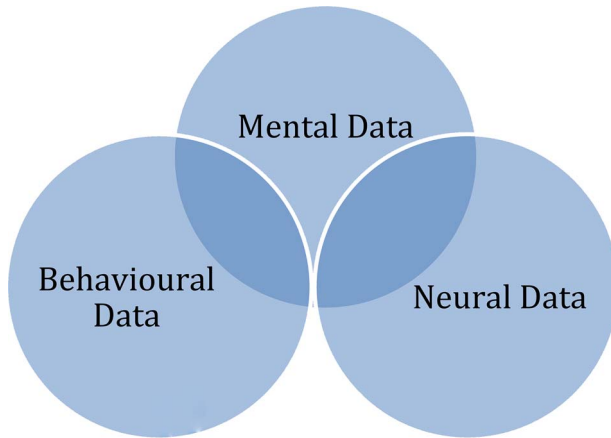


Figure 1. Relationship between neural, non-neural and mental data.

of mental health services.^{21,22} This digital shift in mental health is expected to be of paramount importance during and in the aftermath of the ongoing coronavirus disease (COVID-19) pandemic. The World Health Organization (WHO) has released expressions of concern and technical guidance about the psychological impact of the pandemic, with special focus on increased stress and anxiety. As many countries have introduced drastic containment measures such as social distancing and state-level lockdown, the WHO also expects a rise in levels of loneliness, depression, harmful alcohol and drug use, and self-harm or suicidal behavior.²³

These impacts are expected to outlast the pandemic. Although the ‘alarming’ mental health burden of the pandemic unfolds, experts called for enhanced monitoring of psychosocial needs and for the delivery of psychosocial support to mental health patients, health care providers, and the public. It has been observed, however, that monitoring psychosocial needs and delivering support through ‘direct patient encounters in clinical practice are greatly curtailed in this crisis by large-scale home confinement’.²⁴ Digital mind technologies hold promises for overcoming the limitations of physical care delivery, providing continuous remote monitoring of mental health parameters, and administering telemedical interventions. Since the beginning of the pandemic, an entire ecosystem of virtual therapy and mental health apps has been proliferating.²⁵

At the same time, this sector of bioelectronic and digital innovation raises unprecedented ethical and legal challenges. Compared to other digital measurements of the human body, the risks associated with mental data are likely of greater magnitude. The reason for that derives from the fact that the human mind governs cognitive

21 Chris Hollis et al., *Identifying Research Priorities for Digital Technology in Mental Health Care: Results of the James Lind Alliance Priority Setting Partnership*, 5 *THE LANCET. PSYCHIATRY* 845 (2018).

22 David C. Mohr, Heleen Riper and Stephen M. Schueller, *A Solution-Focused Research Approach to Achieve an Implementable Revolution in Digital Mental Health*, 75 *JAMA PSYCHIATRY* 113 (2018).

23 Mason Marks, *Artificial Intelligence Based Suicide Prediction*, 18 *YALE J. HEALTH POLICY, LAW, ETHICS* 98 (2019).

24 Pfefferbaum and North, *supra* note 1.

25 J. Garsd, *supra* note 2.

and affective phenomena such as consciousness, first-person subjectivity, memory, perception, emotions—all things that make us human.^{26,27} Furthermore, mental representations (and their underlying brain activity) are the closest psychological (and neurobiological) substrate of fundamental ethical-legal notions such as personal identity, personal autonomy, freedom of thought, mental integrity, and others.^{28–31} If mental information becomes unrestrictedly accessible by third parties, the very notions of personhood and personal autonomy degrade as the informational boundaries of the self-blur within the infosphere. Furthermore, mental data may encode private information about unexecuted behavior such as unuttered thoughts and intended action that are not otherwise accessible through simple behavioural observation.^{32,33} Therefore, if mental information becomes unrestrictedly accessible by third parties, then even this ultimate resort of private information may become observable from the outside world.

These inherent features of mental data raise meaningful ethical and legal implications. These implications are being currently addressed at several levels of governance such as technical and biosecurity standards, ethical guidelines and analogous soft law approaches,^{34,35} consumer protection regulation,³⁶ and international human rights law.³⁷ In this paper, we will focus on the governance of mental data from the perspective of data protection law. The reason for that stems from the fact that the inherent features of mental data described above raise the fundamental normative challenge of locating mental data within the current data protection landscape and defining the adequate conditions for their collection and processing. This focus on data protection should not be seen as mutually exclusive with the approaches described previously but rather as part of a multi-level approach to the governance of mental data.

Two caveats are important here for clarity and disambiguation purposes: Current technologies cannot yet decode mental information, that is, provide a detailed and causally robust account of the relation between certain data pattern and the semantic content of mental states. However, the technologies described above are already sufficiently sophisticated to establish statistically significant relations between certain patterns of neural, behavioural, or other data, on the one hand, and the actual

26 Klein and others, *supra* note 3.

27 Ralf J. Jox and others, *Disorders of Consciousness: Responding to Requests for Novel Diagnostic and Therapeutic Interventions*, 11 *THE LANCET. NEUROL.* 732 (2012).

28 Marcello Ienca and Roberto Andorno, *Towards New Human Rights in the Age of Neuroscience and Neurotechnology*, 13 *LIFE SCI., SOC. POLICY* 5 (2017).

29 JOSEPH J. FINS, *RIGHTS COME TO MIND: BRAIN INJURY, ETHICS, AND THE STRUGGLE FOR CONSCIOUSNESS* (Cambridge University Press, 2015) <https://www.cambridge.org/core/books/rights-come-to-mind/47C1316518BB222D79C7D5F6C8EED82A> (accessed Apr. 18, 2021).

30 Fabrice Jotterand, *Beyond Therapy and Enhancement: The Alteration of Human Nature*, 2 *NANOETHICS* 15 (2008).

31 Orsolya Friedrich, E. Racine, S. Steinert, J. Pömsl, R. J. Jox, *An Analysis of the Impact of Brain-Computer Interfaces on Autonomy*, *NEUROETHICS* 1 (2018).

32 Nita Farahany, *Searching Secrets*, 160 *UNIV. PENNSYLVANIA LAW REV.* 70 (2011).

33 Nita Farahany, *The Costs of Changing Our Minds*, 69 *EMORY LAW J.* 75 (2019).

34 Sara Goering and Rafael Yuste, *On the Necessity of Ethical Guidelines for Novel Neurotechnologies*, 167 *CELL* 882 (2016).

35 Ienca, Haselager and Emanuel, *supra* note 5.

36 Anna Wexler and Peter B. Reiner, *Oversight of Direct-to-Consumer Neurotechnologies*, 363 *SCIENCE* (New York, N.Y.) 234 (2019).

37 Ienca and Andorno, *supra* note 28; Rafael Yuste, Jared Genser and Stephanie Herrmann, 'It's Time for Neuro-Rights' 7.

occurrence of certain mental states. For example, assuming that a certain person X is experiencing fear or visually perceiving a human face, current technologies such as neurotechnology and affective computing are far from revealing the semantic content of such emotion of fear (ie what person X is afraid of) or the visual content of the associated perception (eg whose individual face person X is seeing). Furthermore, they are even more distant from revealing the phenomenology of such subjective experiences (ie what it feels like to have person X's experience of either fear or perceiving someone else's face). However, they are already sufficiently sophisticated to reveal from certain data patterns that person X is either experiencing fear or visually perceiving a human face. In addition, they current predictive models are sufficiently robust to generalize the inference that any time a certain pattern of data occurs, then a certain class of mental states (eg experiencing fear or seeing a human face) is likely to be involved. Although lacking detailed content, this type of information derived from mental data can still be sensitive, have severe privacy implications and generate novel thorny questions related to data protection.

A second caveat regards the relationship between mental data and neural data (also called 'brain data', which can be defined as direct measurements of (human) brain structure, function and activity. Authors in the field of neuroethics have long debated the nature and normative status of neural data. Some authors, such as Ienca et al. and Yuste et al., have argued that neural data are a particularly sensitive class of data because of their more direct causal link with mental processes, their greater elusiveness to conscious control, and their ability to predict a person's present and future health status and behavior.³⁸⁻⁴¹ In contrast, other authors such as Wexler have criticized this view in the light of the limited accuracy and reliability of currently available neurodevices.⁴² Mental data should be distinguished from brain data for two reasons. First, not all mental data are brain data as information about mental states and processes can be inferred also from non-neural data such as behavioural data. Vice versa, not all brain data are mental data as brain data can be processed to infer not only mental states but also basic brain anatomy and physiology, without revealing anything related to mental states and processes.

In this article we limit our focus to the EU GDPR, since it is one of the most advanced and comprehensive data protection laws in the world, having also an extraterritorial impact on other legal systems (See Article 3).⁴³

III. THE GDPR IMPLICATIONS: THE NATURE OF MENTAL DATA

The previously described technologies have huge implications on personal data protection and privacy of users. In this section, we are going to propose a first taxonomy

38 Ienca, Haselager and Emanuel, *supra* note 5.

39 Ienca and Andorno, *supra* note 28.

40 Ienca and Ignatiadis, *supra* note 11.

41 Rafael Yuste, et al., *Four Ethical Priorities for Neurotechnologies and AI*, 551 NATURE 159–163 (2017).

42 Anna Wexler, *Separating Neuroethics from Neurohype*, 37 NAT. BIOTECHNOL. 988, 988–990 (2019).

43 See Oreste Pollicino, *Data Protection and Freedom of Expression Beyond EU Borders: EU Judicial Perspectives in DATA PROTECTION BEYOND BORDERS: TRANSATLANTIC PERSPECTIVES ON EXTRATERRITORIALITY AND SOVEREIGNTY* (Federico Fabbrini, Edoardo Celeste and John Quinn eds., Hart Publishing 2020) <http://www.bloomsburycollections.com/book/data-protection-beyond-borders-transatlantic-perspectives-on-extraterritoriality-and-sovereignty> (accessed May 6, 2021).

of data protection issues related to these technological tools. In order to develop such a taxonomy, we should first wonder about the *nature* of data processed by digital mind technologies (personal data, special categories of personal data); then about the lawfulness of such processing activities (legal basis for processing personal data, sensitive data; safeguards to respect in case of automated profiling), and of their level of risk (relevant for accountability duties, like the implementation of safeguards throughout the GDPR and, more importantly, the Data Protection Impact Assessment).

Before analyzing these aspects, it is important to distinguish different variables that can highly influence the legal considerations of digital mind, in particular:

- (i) the (commercial or medical) *context* of the data processing;
- (ii) the (diagnostic, observational, or targeting) *purposes* of the processing;
- (iii) the *interests* in the data processing (public interests in diagnoses or data analyses; private interests in enhancing mental functioning or improving one's wellbeing; solely commercial interests in exploiting cognitive biases of consumers; etc.).

According to the definition of Article 4(1) GDPR, the related WP29 Guidelines⁴⁴, and the CJEU Cases (*Breyer*⁴⁵ and *Nowak*⁴⁶), data related to human brain and mind are always personal data if they allow to *single out* the data subject at stake.⁴⁷ However, one may wonder whether mental data are sufficient alone to be considered personal data, even without any additional identifiers to the concerned data subject.⁴⁸ Article 4(1) mentions a possible list of 'identifiers' and includes also 'one or more factors specific to the (...) *mental* (...) identity of that natural person'. The GDPR does not define mental identity: that wording seems quite obscure and even the EDPB has not clarified that concept. However, one might wonder whether 'mental data' are sufficient (even without any other identifier) to qualify as personal data. WP29 clarified that while some characteristics are so unique that someone can be identified with no effort, in general 'a combination of details on categorical level may also be pretty conclusive in some circumstances, particularly if one has access to additional information of some sort'.⁴⁹ In sum, we can preliminarily affirm that mental data, in combination with other data that allow to single out a data subject, are personal data. In addition, it has been noted that many types of neural data such as EEG and fMRI are uniquely related to an

44 Article 29 Working Party, Guidelines on the definition of personal data, 2014.

45 *Breyer v Germany* (C-582/14) CJEU (2016), [30].

46 *Peter Nowak v Data Protection Commissioner* (C-434/16) CJEU (2017).

47 Frederik J. Zuiderveen Borgesius, *Singling out People without Knowing Their Names—Behavioural Targeting, Pseudonymous Data, and the New Data Protection Regulation*, 32 *COMPUT. LAW SECUR. REV.* 256 (2016).

48 See, largely, Stephen Rainey and others, *Is the European Data Protection Regulation Sufficient to Deal with Emerging Data Concerns Relating to Neurotechnology?* J. *LAW BIOSCI.* (2020) <https://doi.org/10.1093/jlb/lsaa051> (accessed Feb. 24, 2021).

49 Article 29 Working Party, Guidelines on the definition of personal data, 2014, 13.

individual:^{50–52} accordingly, they might potentially be sufficient to identify a natural person (and so to qualify as personal data).

The subsequent legal issue to address is whether data related to human mind can be considered *special categories* of personal data or not. Article 9(1) states that special categories of data (hereinafter: sensitive data) include data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership, but also genetic data, biometric data for the purpose of uniquely identifying a natural person, data concerning health, or data concerning a natural person's sex life or sexual orientation.

As explained in the previous section, neurotechnologies and digital phenotyping tools (especially if combined with affective computing techniques) allow to collect and process exploratory information about mental states and cognitive or affective process, as well as information about a person's underlying neurophysiology and neuropathology. At the same time, in the age of big data and advanced analytics, these pieces of information can be also *inferred* rather than *observed* by data analytics based on retrospective data mining, pattern recognition and aggregation of multiple data sources or predictive analytics.

In order to classify the nature of these data, it is necessary to consider them separately. All information that may reveal a pathological mental status are sensitive data because they are in the definition of 'data concerning health'. In particular, Article 4(15) defines those data as 'personal data related to the physical or *mental health* of a natural person, including the provision of health care services, which reveal information about his or her health status'.⁵³ Recital 35 further clarifies that personal data concerning health should include 'all data pertaining to the health status of a data subject which reveal information relating to the *past, current or future* physical or *mental health status* of the data subject'.⁵⁴

This definition generally refers to 'mental health status', not exclusively to pathological statuses. Accordingly, we should infer that also information revealing a physiological mental status (ie the lack of any mental pathology) can be considered health data. In other terms, biological parameters that are usually necessary to infer mental illnesses are sensitive data even if in a specific context they do not reveal any illness, but just the correct functioning of brain physiology. Therefore, the definition of mental health should be considered extensively^{55,56} and should include also any form of cognitive processes and affective states of the data subject.

50 S. Yang and F. Deravi, *On the Usability of Electroencephalographic Signals for Biometric Recognition: A Survey*, 47 IEEE TRANS. HUM.-MACH. SYST. 958 (2017).

51 K. Aloui, A. A. Nait-Ali and M. Saber Naceur, *Using Brain Prints as New Biometric Feature for Human Recognition*, PATTERN RECOGN., ELSEVIER, (2017) <https://hal.archives-ouvertes.fr/hal-01681974> (accessed Apr. 18, 2021).

52 Rainey and others, *supra* note 48.

53 Italics added.

54 Italics added.

55 Gianclaudio Malgieri and Giovanni Comandé, *Sensitive-by-Distance: Quasi-Health Data in the Algorithmic Era* 26 INF. COMMUN. TECHNOL. LAW 229 (2017).

56 Giovanni Comandé and Giulia Schneider, *Regulatory Challenges of Data Mining Practices: The Case of the Never-Ending Lifecycles of "Health Data"*, 25 EUR. J. HEALTH LAW 284 (2018).

The analysis is more difficult for what concerns data that cannot reveal, not even in abstract, physiological conditions of the brain: eg non-pathological emotional information or information related to thoughts, preferences, or memories. Emotions are not per se ‘sensitive data’, but they might be if they are collected through emotion detection tools based on biometric tools such as facial recognition (‘biometric data for the purpose of uniquely identifying the data subject’).⁵⁷ On the contrary, it is difficult to consider emotion-related data detected through non-biometric methods (eg written text or voice records) as sensitive data. Similarly, affective or other mental states detected through or inferred from consumer-grade digital mind technologies such as consumer BCIs are unlikely to be considered sensitive data, unless those emotions can either be used even to infer the mental health status of the individual (eg digital biomarkers of neuropsychiatric disorder) or to reveal information about religious beliefs, political opinions, and sex life or sexual orientation. Accordingly, consumer neurotechnologies that claim to provide information about a person’s concentration and overall mental wellbeing are unlikely to be considered processors of sensitive data.

For what concerns data revealing information related to data subjects’ thoughts or memories, these data are *not* automatically sensitive data just because they refer to the ‘mental sphere’ of the subject. However, if—considering their content and the context and purpose of the data processing^{58,59}—it is likely that these kinds of data might reveal information about religious beliefs, political opinions, and sex life or sexual orientation, they are sensitive data under Article 9(1) GDPR.⁶⁰

Taking into account these last considerations, we observe that there is a clear conceptual and normative gap: even though most people would agree that mental data are the most intimate and sensitive information of the data subject, not all mental data are protected under the strict regime of sensitive data.

Rainey et al. have also emphasized this gap of protection, but for different reasons.⁶¹ In particular, they claim that the definition of ‘special categories of data’ in the GDPR (at Article 9(1)) is purpose-based: accordingly, if the initial declared purpose for processing those data is not related to healthcare (or to other sensitive purposes at Article 9(1)), those data cannot be considered sensitive regardless of their highly sensitive potentialities and eventual implications. We contend, as Article 9(1)⁶² and recital 51⁶³

57 Damian Clifford, *Citizen-Consumers in a Personalised Galaxy: Emotion Influenced Decision-Making, a True Path to the Dark Side?*, SSRN ELECTR. J. (2017), <https://www.ssrn.com/abstract=3037425> (accessed Dec. 16, 2018).

58 Paul Quinn and Gianclaudio Malgieri, *The Difficulty of Defining Sensitive Data—the Concept of Sensitive Data in the EU Data Protection Framework*, 22 GERMAN LAW J. 1583 (2021), forthcoming.

59 K. McCullagh, *Data Sensitivity: Proposals for Resolving the Conundrum*, 2 J. INT. COMMER. LAW TECHNOL. 190 (2007).

60 Rainey and others, *supra* note 48, 11.

61 *Ibid.* 14, 16, 17.

62 Article 9(1): ‘Processing of personal data *revealing* sensitive information (italics added). The reference to ‘revealing’ clearly refers to all potential implications and not merely to the purposes of data processing.

63 Recital 51: ‘Personal data which are, by their *nature*, particularly sensitive in relation to fundamental rights and freedoms merit specific protection as *the context of their processing* could create significant risks to the fundamental rights and freedoms’ (italics added). In this recital the focus is on ‘nature’ of data and their ‘context’, not on the purposes.

reveal and as it is largely discussed in literature,^{64–67} that the notion of sensitive data in the EU data protection field is mostly contextual and not purpose-based⁶⁸ (with the sole exception of ‘biometric data for the *purpose* of uniquely identifying a natural person’⁶⁹).

Consequently, we concur with Rainey et al. that there is a ‘gap’ of protection, but we disagree that this is due to the nature and definition of sensitive data. In contrast, we argue that this gap of protection stems from the fact that the list of sensitive data categories in the GDPR (health, biometric, genetic, political opinions, sexual orientations, etc.) is not comprehensive enough to include, eg ‘emotions’ or other ‘thoughts’ not related to health status, sexuality or political/religious beliefs. For example, the lawful and transparent processing of data about consumers’ emotions or moods on a social media platform would not be considered sensitive data processing if it is not possible to prove the (even just contextual and indirect) link between those emotions and sensitive areas (health, sexuality, and beliefs). It would be, thus, helpful if the GDPR could clarify that even indirect inferences, not strictly contextually related to the explicitly sensitive areas at Article 9(1) but anyway affecting the mental area could be considered sensitive.

III.A. The lawfulness of mental data processing

After this overview on the ‘nature’ of mental data, the subsequent problem is determining whether and when it is lawful to process these data under EU data protection rules.⁷⁰ To address this problem, it is necessary to understand if there is an appropriate lawful basis for processing mental data in a given context. The appropriate legal basis depends on the *nature* of the data at stake (if data are sensitive, the controller must comply not only with Article 6 lawful bases, but also with Article 9(2) lawfulness requirements). Before analyzing lawfulness conditions, it is important to remind that purposes should be specified, explicit and legitimate (Article 5(1)(b)). If the data controller processes mental data for, eg health self-monitoring purposes and then she uses those data for commercial purposes, she would commit a violation of the just mentioned purpose limitation principle.

In case mental data qualify as sensitive data (see the discussion in Section above), it is lawful to process them only if one of the lawfulness conditions under Article 9(2) of the GDPR can apply. In case the processing of these data has a *commercial* nature, the only possible legal basis is the explicit consent of the data subject according to

64 Yves Poullet and Jean-Marc Dinant, *Thoughts on Convention No. 108 for the Purposes of the Future Work of the Consultative Committee (T-PD)* 61, 43.

65 Quinn and Malgieri, *supra* note 58, 10–11.

66 Malgieri and Comandé, *supra* note 55.

67 Mccullagh, *supra* note 59.

68 According to the contextual approach in the GDPR, all personal data should be assessed against the background of the context that determines their processing, as determined by several contextual factors (eg the specific interests of the controller, the potential recipients of the data, the aims for which the data are collected, the conditions of the processing and its possible consequences for the persons involved). In contrast, the purpose-based approach essentially looks at the intention of the data controller and asks whether the controller intends to draw conclusions from the processing of particular data that could be regarded as being sensitive in nature.

69 Italics added.

70 For an early discussion on this topic, see also Dara Hallinan and others, *Neurodata and Neuroprivacy: Data Protection Outdated?*, 12 SURVEILL. SOC. 55 (2014).

Article 9(2)(a), but this consent should be informed and free as requested by Article 7. If the processing of mental data is based not only on the commercial interests of the data controller, but on the specifically expressed interests of the data subjects (self-monitoring, quantified self, exploration of mental activity, or cognitive training), it is more likely that data subjects' consent is free and, thus, in principle valid. However, research has shown that the collection and processing of mental data from consumer neurotechnology and digital phenotyping applications often occurs under weak consent regimes.⁷¹ This is due to the fact the Terms of Service of these digital tools are (i) rarely read by the users, (ii) typically uninformative about the whole data lifecycle and the specifications of data processing, and (iii) often based on presumed instead of affirmative consent. On the contrary, if the commercial nature of mental data processing is merely based on the interests of the data controller (eg in order to better micro-target the data subject through personalized ads), the data controller's burden to prove that consent was really free and informed seems more onerous. Actually, it is also possible that the data subject consents to mental data collection for a deliberate personal interest (say cognitive monitoring) but subsequently becomes subject to data processing activities (say microtargeting based on personalized cognitive or affective features) that are merely based on the interests of the data controller.

The nature of mental data processing can also be non-commercial. This is, for instance, the case of medical diagnosis, scientific research or other public interests. In case of mental data processed for *healthcare provision* reasons (diagnosis or therapy), Article 9(2)(h) allows such forms of processing without specific additional requirements.⁷²

In case of mental data processed for *scientific research*, Article 9(2)(j) allows such processing activities but under the specific condition that there is a 'Union or Member State law' authorizing it, and that it be 'proportionate to the aim pursued, respect the essence of the right to data protection and provide for suitable and specific measures to safeguard the fundamental rights and the interests of the data subject'. One might wonder whether such an intrusive inspection into the data subject (reaching her mental sphere) is 'proportionate' to that research. Most of research investigating mental data pertains to behavioural research, ie social experimenting (either online or in the physical world) in which emotions and other pieces of mental information of the data subject are gathered for psychological, social, or other kinds of research. In a preliminary opinion about scientific research, the EDPS stated that behavioural experiments are generally not in the scope of the research exemption in Article 9(2)(j) because they might lack an established ethical framework that would make it proportionate and justifiable under the GDPR.⁷³ In other terms, at least in the online (covert) behavioural experiments, the social and scientific benefits of such a research seems overridden by the detriment to the right to privacy and data protection of the research subjects.⁷⁴

It is open to debate whether this statement also applies to other research domains involving mental data, which currently lack an established ethical framework for ensur-

71 Ienca, Haselager and Emanuel, *supra* note 5.

72 Giulia Schneider, *Disentangling Health Data Networks: A Critical Analysis of Articles 9(2) and 89 GDPR*, 9 INT. DATA PRIVACY LAW 253 (2019).

73 EDPS, Preliminary Opinion on Scientific Research, 2020, 9 and 12.

74 Marks, *supra* note 23.

ing proportionality. For example, authors have argued that cognitive monitoring and self-administered neuromodulation via non-medical digital mind technologies might lack an established ethical framework that would make it proportionate and justifiable.^{75,76}

Furthermore, industry-funded biomedical studies also challenge the common normative distinction between research and other purposes. For example, social media giant Facebook has funded biomedical research on human subjects aimed at developing speech decoders that produce real-time decoding of speech in an interactive setting.⁷⁷ This research has important implications for patients with communication disability. However, it is possibly instrumental to Facebook's self-proclaimed commercial endeavor of creating a non-invasive BCI that would let customers type via brain activity.

As regards other possible public interests for which mental data can be processed, we could consider the case of mental data collected *within a legal procedure in a court*. Article 9(2)(f) allows sensitive data processing when the data processing is 'necessary for the establishment, exercise or defence of legal claims or whenever courts are acting in their judicial capacity'. We could consider, eg lie detector tools in criminal⁷⁸ or civil courts and in particular their capacity to read mental information in order to detect any omission or false statement of a witness or of an accused person.⁷⁹ In order to be under the scope of Article 9(2)(f) the data controller (the plaintiff, the defendant, and the prosecutor or the judge) should prove that lie detector tools are 'necessary' for the legal claims. Considering the level of sensitivity of such data, we might assume that the 'necessity' test in this case is particularly strict.⁸⁰ However, the second part of Article 9(2)(f) seems more lenient: 'whenever courts are acting in their judicial capacity'. This wording may even refer to the situation in which the judge autonomously orders the use of lie detector in a legal claim. However, it is important to notice that under the Charter of Fundamental Rights and the ECHR judges do not have an unfettered power: there should be a Union or National law (respecting the strict principles of Article 6 and 8 of the ECHR and of Articles 7, 8 47 and 48 of the EU Charter) eventually authorizing lie detector tools, with eventual additional safeguards.⁸¹

As for other forms of public interests, Article 9(2)(g) authorizes sensitive data processing, which 'is necessary for reasons of substantial public interest, on the basis of

75 See Ienca, Haselager and Emanuel, *supra* note 5.

76 Goering and Yuste, *supra* note 34.

77 David A. Moses and others, *Real-Time Decoding of Question-and-Answer Speech Dialogue Using Human Cortical Activity* 10 NAT. COMMUN. 3096 (2019).

78 Actually, criminal proceedings are regulated by Directive 2016/680 (Law Enforcement Directive). Article 10 of that Directive allows such a processing only 'where strictly necessary, subject to appropriate safeguards for the rights and freedoms of the data subject, and only: (a) where authorized by Union or Member State law; (b) to protect the vital interests of the data subject or of another natural person; or (c) where such processing relates to data which are manifestly made public by the data subject'. In case of mental data processing through lie detector tools, it seems difficult to overcome the strict necessity test, but in any case, there should be a law authorizing it (the application of one of the other options—b and c—seems quite unlikely to our case).

79 Marion Oswald, *Technologies in the Twilight Zone: Early Lie Detectors, Machine Learning and Reformist Legal Realism* (Social Science Research Network 2019) SSRN Scholarly Paper ID 3369586, <https://papers.ssrn.com/abstract=3369586> (accessed Jan. 21, 2021).

80 Information Commissioner Officer (ICO), *Special Categories of Data. What Are the Conditions for Processing?* (ICO, 2021) <https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/special-category-data/what-are-the-conditions-for-processing/> (accessed Apr. 18, 2021).

Union or Member State law which shall be proportionate to the aim pursued, respect the essence of the right to data protection and provide for suitable and specific measures to safeguard the fundamental rights and the interests of the data subject'. Here, there are many conditions: (i) a *substantial* public interest; (ii) a law; (iii) which is proportionate to the aim; (iv) which respects the essence of the right to data protection; and (v) which provide for suitable and specific safeguards to data subjects' fundamental rights. In abstract terms, it seems a very strict test for mental data processing: it is quite difficult to imagine a 'substantial' public interest for which it is necessary to detect mental data of individuals. Moreover, as in the reflections about scientific research, monitoring mental data seems rarely proportionate to an even hypothetical substantial public interest.

However, as discussed in the previous Section, the GDPR does not protect all mental data under the special regime of 'sensitive data'. Due to the narrowness of the sensitive data list at Article 9(1), any mental data that are not related (even just contextually and indirectly) to sensitive areas (health, sexuality, beliefs, etc.) cannot be protected under the strict rules of Article 9. For example, emotions or moods, desires or mental propensities would often be non-sensitive data.

For all mental data qualifying as *non-sensitive data*, the only lawfulness conditions to meet are at Article 6 GDPR: in addition to *consent* (for which, see the reflections above), the processing can be, eg carried out for a contract, a public interest or a legitimate interest. In case of *contract* (Article 6(1)(b)), it seems to be in the case of a data subject who is a consumer of a service for the monitoring of mental activities or brain data. As regards *legitimate interests*, under Article 6(1)(f) the processing must be 'necessary for the purposes of the legitimate interests pursued by the controller or by a third party, except where such interests are overridden by the interests or fundamental rights and freedoms of the data subject'. This wording states a balancing test containing the following requirements: the legitimacy of the interest of the data controller, the necessity of the data processing for such an interest, an analysis of the impact on the data subject, and the presence of eventual additional safeguards for the data subject. Even assuming that detecting mental data for, eg commercial reasons is a legitimate and necessary activity, the impact on the data subject (risks of vulnerability exploitation in commercial contexts, risks of manipulation and discrimination, unawareness of implication, etc.) seems however too burdensome to justify these interests.^{82,83} Next section will indeed explain why the impact of mental data processing on data subjects seems very considerable, even under the risk indexes of the GDPR.

The last possible lawful basis under which the data controller could process (non-sensitive) mental data is 'public interests' (Article 6(1)(e)). Although 'public interest' of Article 6(1)(e) seems a more lenient lawful basis if compared to the 'substantial public interest' of Article 9(2)(g), the data controller has an accountability duty to

81 Sjors Ligthart, et al. *Forensic Brain-Reading and Mental Privacy in European Human Rights Law: Foundations and Challenges*, 14 *NEUROETHICS* 191 203 (2020) <https://doi.org/10.1007/s12152-020-09438-4> (accessed Apr. 18, 2021).

82 Article 29 Working Party, *Opinion 06/2014 on the Notion of Legitimate Interest of the Data Controller under Article 7 of Directive 95/46/EC*, (2014).

83 Irene Kamara and Paul De Hert, *Understanding the Balancing Act behind the Legitimate Interest of the Controller Ground: A Pragmatic Approach*, in *THE CAMBRIDGE HANDBOOK OF CONSUMER PRIVACY* (Evan Selinger, Jules Polonetsky and Omer Tene eds., 1st edn, Cambridge University Press 2018) https://www.cambridge.org/core/product/identifier/9781316831960%23CN-bp-19/type/book_part (accessed Mar. 28, 2020).

prove the existence of public interests for detecting mental data of the data subject (such a public interest should be inferred, eg from existing laws or other legal sources) and to prove the necessity of such a processing for reaching those objectives in the public interest.⁸⁴

III.B. Mental data processing as high-risk processing: The DPIA measures

The previous two sections analyze the ‘nature’ of mental data under the GDPR and the lawfulness of their processing. As observed, many forms of mental data cannot be considered ‘sensitive’ due to the narrow list of sensitive data in Article 9 (eg emotions, moods or mental conditions not related to health, sexuality, political beliefs, etc.). Accordingly, many forms of mental data processing are not protected under the strict regime of Article 9(2), but under the more lawful grounds standards at Article 6: such processing can be based on consent, contract, legitimate interest, or public interest. Although these last lawful grounds have specific requirements and conditions, we argue that mental data—considering their sensitivity and implications—should be protected under higher standards (eg under the special regime of sensitive data at Article 9).

In fact, the GDPR offers other important accountability safeguards that could compensate for the lack of Article 9 protection for (some) mental data. We mention in particular the assessment and mitigation of data processing impact in Article 35 (the ‘Data Protection Impact Assessment’, hereafter: DPIA). Article 35 provides that for personal data processing at higher risks for fundamental rights and freedoms of data subjects, a specific impact assessment (where risks are assessed and mitigated) should be done prior to the data processing and on a regular basis whenever the level of the risk might change.⁸⁵ This impact assessment is different from the previously mentioned balancing test that should be conducted when processing data for ‘legitimate interests’ (Article 6(1)(f)): while the balancing test is just a provisional test based on the specific context at issue, the DPIA is a much more comprehensive assessment that should be documented in great detail and should include a description of the data processing, an analysis of necessity and proportionality, a specific description and assessment of possible ‘risks’ for data subjects with an accurate list of safeguards that should prevent or mitigate all those risks.⁸⁶ In addition, while the balancing test should be conducted only when the data processing is based on ‘legitimate interest’ lawful ground, the DPIA should be conducted for all forms of data processing ‘at high risk’.

If we consider the tremendous implications of processing mental data and the possible risks of that data processing, we can easily conclude that it is a high-risk data processing.⁸⁷ In more specific terms, Article 35(3) mentions three cases of high-risk data processing: (A) a systematic and extensive evaluation of personal aspects, which is

84 Paul Quinn, *Research under the GDPR—A Level Playing Field for Public and Private Sector Research?* 17 LIFE SCI., SOC. POLICY 4, 10 (2021).

85 Dariusz Kloza and others, *Data Protection Impact Assessments in the European Union: Complementing the New Legal Framework towards a More Robust Protection of Individuals* https://cris.vub.be/files/32009890/dpiala_b_pb2017_1_final.pdf.

86 About the difference between the balancing test and the DPIA, see Article 29 Working Party, Guidelines on Data Protection Impact Assessment, WP248 and Article 29 Working Party, Opinion 06/2014 on the notion of legitimate interests of the data controller under Article 7 of Directive 95/46/EC, 33.

87 See, e.g., Adam DI Kramer, Jamie E Guillory and Jeffrey T Hancock, *Experimental Evidence of Massive-Scale Emotional Contagion through Social Networks*, 111 PROC. NATL. ACAD. SCI. 8788 (2014).

based on automated processing, including profiling, and on which decisions are based that produce legal or similarly significant effects; (B) processing on a large scale of special categories of data or of personal data relating to criminal convictions; or (C) a systematic monitoring of a publicly accessible area on a large scale.

For the cases of mental data processing that we described in the previous sections, both conditions A and C might eventually apply. In particular, if the fully automated profiling based on mental data (emotions, sensory stimuli, cognitive characteristics, etc.) is aimed to take significant decisions for the data subject (eg lie detectors or crime predictions in cases of automated justice or in cases of border controls; micro-targeting for manipulative advertising on social media; and cognitive assessment in education, employment, for hiring purposes, for credit scoring; etc.) we are clearly in a situation of high risk under letter a). In this case, also Article 22 shall apply, implying a prohibition of such automated profiling unless an exception applies (explicit consent, Member State Law, and contractual necessity) together with suitable safeguards (at least the right to contest, the right to have a human intervention and to express one's view) and meaningful information about the logic, the significance and the envisaged effects of the data processing (under Articles 13(2)(g), 14(2)(h) and 15(1)(h)).^{88–91}

However, in many cases and business models described above, no significant individual 'decision' is taken after mental data collection. We can consider, eg the case of consumer apps to self-monitor brain activity or mental data processing for research purposes. In other cases, the eventual decision is not automated: eg in most cases of health-related mental data processing the medical decision is mediated by a human (a doctor). In all these cases, Article 22 does not apply, neither Article 35(3)(a) would apply.

In these cases, Article 35(3)(c) might often apply. Indeed, as discussed in [Section 3](#) above, most mental data can be considered special categories of data. Accordingly, if consumer apps, medical researchers or doctors process a large scale of such data (as often occurs), we are in a situation of high risk and the DPIA should be carried out.

However, in some situations, just few mental data are necessary for processing. In those cases, the definition of 'large scale' is disputable. In addition, as affirmed in [Section 3](#) there are many mental data that can be considered neither health data, nor related to other special categories of data at Article 9(1) (eg emotions; thoughts non-related to politics, religion, sexuality, etc.). Even in those cases there are clearly high potential risks for the data subjects (eg electoral or commercial mental manipulation,⁹² discrimination, loss of opportunities in several fields, identity theft, etc.).

88 Antoni Roig, *Safeguards for the Right Not to Be Subject to a Decision Based Solely on Automated Processing (Article 22 GDPR)* 8 EUR. J. LAW TECHNOL. <http://ejlt.org/article/view/570> (accessed Jan. 15, 2019).

89 Gianclaudio Malgieri, *Automated Decision-Making in the EU Member States: The Right to Explanation and Other "Suitable Safeguards" in the National Legislations*, COMPUT. LAW SECUR. REV. 105327 (2019).

90 Gianclaudio Malgieri and Giovanni Comandé, *Why a Right to Legibility of Automated Decision-Making Exists in the General Data Protection Regulation*, 7 INT. DATA PRIVACY LAW 243 (2017).

91 Maja Brkan, *Do Algorithms Rule the World? Algorithmic Decision-Making and Data Protection in the Framework of the GDPR and Beyond*, 91 INT. J. LAW INF. TECHNOL., 121 (2019) <https://academic.oup.com/ijlit/advance-article/doi/10.1093/ijlit/eay017/5288563> (accessed Apr. 24, 2019).

92 Daniel Susser, Beate Roessler and Helen Nissenbaum, *Online Manipulation: Hidden Influences in a Digital World* (Social Science Research Network, 2018) SSRN Scholarly Paper ID 3306006, <https://papers.ssrn.com/abstract=3306006> (accessed Feb. 28, 2019); Tal Z Zarsky, *Privacy and Manipulation in the Digital Age*, 20 THEORET. INQUIRIES LAW (2019) <http://www7.tau.ac.il/ojs/index.php/til/article/view/1612> (accessed Jan. 23, 2019).

Interestingly, the EDPB complemented the three high risks parameters at Article 35(3) with ten risk indexes:⁹³ where two of these indexes apply, the data processing should be considered at high risk and the DPIA should be done. Importantly, in the list of risk indexes, there are several that could apply to mental data processing in the research, commercial, and health-related fields. First of all, one of it is 'sensitive data processing'. The wording ('sensitive') is deliberately different from Article 9(1) that refers to 'special' categories of data. Indeed, as the EDPB specifies, in this category we should not only include data from Articles 9 and 10, but any 'data which may more generally be considered as increasing the possible risk to the rights and freedoms of individuals, such as electronic communication data, location data, financial data (that might be used for payment fraud)'. In addition, the EDPB also mentions, as an example, 'life-logging applications that may contain very personal information'. By analogy, we could easily conclude that also mental data should be considered 'sensitive' in this general meaning.

Among the other risk indexes, we find other relevant elements, such as 'evaluation or scoring' that might include all mental data processing aimed at assessing the subject even when there is no fully automated decision at the end of the processing. Another element is 'systematic monitoring': this might be the case of consumer apps based on brain monitoring.

In addition, the use of innovative technologies is also included in the risk indexes: mental data processing is actually often based on very innovative technologies. The EDPB explains that the use of such technology can involve novel forms of data collection and usage, possibly with a high risk to individuals' rights and freedoms. The reason why these processing should be considered at risk is that 'the personal and social consequences of the deployment of a new technology may be unknown'. An explicit example of this is IOT apps.

Another index that might be relevant for mental data processing is 'data concerning vulnerable data subjects'. The EDPB mentions some examples, eg 'employees', 'mental ill' persons, and 'patients' and explains that subjects' vulnerability is a risk index because of the higher power imbalance between the data controller and certain data subjects in certain moments. The examples of mental data processing that we referred to in the previous sections involve often patients (such as mentally ill persons), people with age-related cognitive decline and employees. However, in general, since data subjects' vulnerability has been defined as a transient characteristics and generally contextual effect of power imbalance,⁹⁴ we can easily include in this risk index also situations in which individual mental vulnerabilities of subjects are discovered, predicted, and exploited. This is typically the case with emotion-driven advertising⁹⁵, automated cognitive assessment (which could reveal digital biomarkers of cognitive decline), stress-management via neurofeedback etc.⁹⁶

93 European Data Protection Board, *Guidelines on Data Protection Impact Assessment (DPIA) and Determining Whether Processing Is "Likely to Result in a High Risk" for the Purposes of Regulation 2016/679* (2017).

94 Gianclaudio Malgieri and Jędrzej Niklas, *The Vulnerable Data Subject* 37 COMPUT. LAW SECUR. REV. (2020).

95 Robert Booth, *Facebook Reveals News Feed Experiment to Control Emotions*, THE GUARDIAN (29 June 2014) <https://www.theguardian.com/technology/2014/jun/29/facebook-users-emotions-news-feeds> (accessed June 2, 2019).

In all these cases, when two risk indexes can apply, mental data processing must be considered at high risk and the data controller is obliged to: describe the processing (including a description of the logic of the technology)⁹⁷, perform a balancing test based on necessity and proportionality of the data processing in relation to the purposes,⁹⁸ assessing the actual risks for fundamental rights and freedoms, and proposing suitable measures to address and mitigate those risks. This operation could imply an audit of the technological components of the processing (eg AI-driven processing) and a reconsideration of the algorithm in case some risks can be mitigated ‘by design’.⁹⁹ We call all these types of DPIA for mental data processing MDPIA.

In addition, also the Data Protection Agency could play a relevant role. Indeed, in case the data controller, after a DPIA, discovers that no adequate mitigations can be found for the existing risks, she can refer to the competent Data Protection Agency, that could give advice and recommendations (or even assign obligations) in a dialogue with the data controller.

In sum, even if we spot some ‘gaps’ in the existing protection of mental data under the GDPR framework, having a look at the broader picture of the GDPR principles and accountability duties we can find positive tools to limit abusive exploitation of mental data and better protect mental privacy of individuals.

IV. CONCLUSIONS

The impressive potentiality of AI applied to human mind of individuals (consumers, data subjects), in particular neurotechnologies and digital phenotyping tools (especially if combined with affective computing techniques), allow to collect and process large-scale and refined exploratory information about mental states and cognitive or affective process, as well as information about a person’s underlying neurophysiology and neuropathology. At the same time, these pieces of information can be also *inferred* rather than *observed* by data analytics based on retrospective data mining, pattern recognition and aggregation of multiple data, or predictive analytics.

Considering that the purposes, the outcomes and possible impacts of these different data processing activities are comparable, we claim to go beyond ‘neural data’ categories and analyze the data processing as a whole and the different and more diverse category of ‘mental data’, ie not only data directly derived *from* brain observation, but any data inferred directly or indirectly *about* mental states of a person, including their cognitive, affective, and conative states.

Once clarified in [Section 2](#) the topic of investigation, the purpose of this article was to analyze the existing legal protection for this broad category of ‘mental data’,

96 Anne Witt, *Excessive Data Collection as a Form of Anticompetitive Conduct—The German Facebook Case* (Social Science Research Network, 2020) SSRN Scholarly Paper ID 3671445, <https://papers.ssrn.com/abstract=3671445> (accessed Feb. 25, 2021).

97 Margot E Kaminski and Gianclaudio Malgieri, *Algorithmic Impact Assessments under the GDPR: Producing Multi-Layered Explanations*, INT. DATA PRIVACY LAW (2020) <https://doi.org/10.1093/idpl/ipaa020> (accessed Feb. 1, 2021).

98 Dariusz Kloza et al., *Data Protection Impact Assessment in the European Union: Developing a Template for a Report from the Assessment Process* (LawArXiv 2020) DPiLab Policy Brief, <https://osf.io/7qrfp> (accessed Dec. 1, 2020).

99 Lina Jasmontaite, et al., *Data Protection by Design and by Default*, 4 EUR. DATA PROTECT. LAW REV. 168 (2018).

even considering the meaningful risks for individuals' rights and freedoms. We limited our focus to the EU GDPR, since it is one of the most advanced and comprehensive data protection laws in the world, having also an extraterritorial impact on other legal systems (See Article 3).

The analysis in [Section 3](#) has therefore focused on: the nature of 'mental data' (personal, directly identifiable, special categories of data, etc.), lawfulness of their processing considering the different legal bases and purposes (commercial, healthcare, public, or private research) ([Section 3.A](#)), and compliance measures (particularly, considering risk assessment) ([Section 3.B](#)).

We concluded that, although the contextual definition of 'sensitive data' might appear inadequate to cover many examples of mental data (eg 'emotions' or other 'thoughts' not related to health status, sexuality or political/religious beliefs), the GDPR—through an extensive interpretation of 'risk' indexes as the EDPB proposes—seems to be an adequate tool to prevent or mitigate risks related to mental data processing. In sum, we recommend that interpreters and stakeholders should focus on the 'processing' characteristics, rather than merely on the 'category of data' at issue. That is why we considered 'mental data processing' as a whole and broader notion, rather than 'neural data'.

In particular, although mental data in some situations are not included in the *category* of 'sensitive' data under the GDPR, many *characteristics* of mental data processing (the profiling or scoring of individuals, the systematic monitoring of individuals, the use of innovative technologies, the presence of vulnerable individuals, etc.) might qualify as high-risk indicators and imply limitations and by-design safeguards to that data processing. In conclusion, we therefore call for 'Mental Data Protection Impact Assessment' (MDPIA), ie a specific DPIA procedure that can help to better assess and mitigate risks that mental data processing can bring to fundamental rights and freedom of individuals.

FUNDING

Marcello Ienca's work has been funded by the ERA-NET NEURON JTC 2020 "Ethical, Legal, and Social Aspects (ELSA)" project HYBRIDMIND (Swiss National Science Foundation's Grant Number: 32NE30_199436).

CONFLICT OF INTEREST

Marcello Ienca reports having served as ethics advisor to the Council of Europe, the OECD and Kernel.