Parliamentary Assembly
Assemblée parlementaire

http://assembly.coe.int

COUNCIL OF EUROPE
CONSEIL DE L'EUROPE

Doc. 15147
24 September 2020

The brain-computer interface: new rights or new threats to fundamental freedoms?

Report[1]
Committee on Legal Affairs and Human Rights
Rapporteur: Mr Olivier BECHT, France, Alliance of Liberals and Democrats for Europe

Summary

The rapid progress made in neurotechnology in recent years has the potential to create increasingly effective brain-computer interfaces (BCI). Access to the neural processes that underlie conscious thought implies access to a level of the self that by definition cannot be consciously concealed or filtered. This risks profound violation of individual privacy and dignity, with the potential to subvert free will and breach the ultimate refuge of human freedom – the mind. Cognitive and sensory enhancement through BCI could create separate categories of human beings. Individual identity, agency and moral responsibility may be diminished through the merger of neurological and digital sensory experience and decision-making processes. Such outcomes could change the very nature of humanity and of human societies.

Democratic societies should ensure that basic ethical principles are respected. The huge potential benefits of neurotechnology are such that progress and innovation should not be stifled. Nevertheless, research should be steered away from foreseeably harmful or dangerous areas and towards positive applications.

The committee considers that a sensitive, calibrated approach to regulation of emerging neurotechnology, including BCI technology, is needed, encompassing both ethical frameworks and binding legal regulation. It therefore calls on member States, the relevant intergovernmental committees of the Council of Europe and the Committee of Ministers to take specific steps to this end.

---

[1] Reference to committee: Doc. 14814, Reference 4435 of 12 April 2019.

Contents **Page**

A. Draft resolution[2]

1. The Parliamentary Assembly notes the rapid progress made in neurotechnology in recent years, including the ability to record and directly stimulate neural activity, with the potential to create increasingly effective brain-computer interfaces (BCI). This progress has been driven by a combination of, improved understanding of the functioning of the brain, technical developments and the growing power of artificial intelligence systems. The ability to create a fully symbiotic connection between the human brain and digital computing systems, including the internet and artificial intelligence systems, remains a distant aspiration. Nevertheless, it is a goal that researchers and entrepreneurs are already pursuing and which many believe may eventually be achieved.

2. Neurotechnology, including BCI, is currently being developed and applied with a range of uses in mind. Amongst other things, huge sums are being invested in research to create new medical treatments for neurological and psychiatric disorders, such as direct control of robotic limbs, synthetic speech production, or the treatment of intractable mood disorders or post-traumatic stress disorder. Military and security establishments are researching neurotechology for use in intelligence, propaganda, interrogation, surveillance and combatants' performance enhancement. Private companies are researching the possible use of consumer devices to transform thoughts directly into typing; providing commercial lie-detection services based on brain scans; and selling direct-to-consumer neurotechnology devices, for example as computer gaming or wellness products. Researchers are exploring the development of 'neuromarketing' campaigns that would exploit subconscious preferences, and examining whether patterns of neural activity may be predictive of criminal recidivism.

3. Access to the neural processes that underlie conscious thought implies access to a level of the self that by definition cannot be consciously concealed or filtered. This risks profound violation of individual privacy and dignity, with the potential to subvert free will and breach the ultimate refuge of human freedom – the mind. Cognitive and sensory enhancement through BCI could create separate categories of human beings, the enhanced and the unenhanced, with enhancement available only to those with the necessary wealth and privilege, or used for repressive purposes. Individual identity, agency and moral responsibility may be diminished through the merger of neurological and digital sensory experience and decision-making processes. Such outcomes could change the very nature of humanity and of human societies.

4. Even if the more spectacular hypothetical applications of BCI remain speculative, the advances already made, and the resources being devoted to further research imply an urgent need for anticipation and precautionary regulation now. Democratic societies should ensure that basic ethical principles are respected. The huge potential benefits of neurotechnology, especially in the medical field, are such that progress and innovation should not be stifled. Nevertheless, research should be steered away from foreseeably harmful or dangerous areas and towards positive applications that do not threaten individual dignity, equality and liberty, which are the foundations also of democracy.

5. The Assembly considers that a sensitive, calibrated approach to regulation of emerging neurotechnology, including BCI technology, is needed, encompassing both ethical frameworks and binding legal regulation. It notes the similarities and connections between 'neuroethics' and bioethics, and the significance of artificial intelligence to the operation of BCI technology. It therefore welcomes the work already underway within the Council of Europe by the Committee on bioethics (DH-BIO) and the Ad hoc Committee on artificial intelligence (CAHAI). It further welcomes the work of other international organisations, notably the Organisation for Economic Cooperation and Development (OECD), which recently adopted a Recommendation on Responsible Innovation in Neurotechnology. The Assembly notes with interest developments such as those in Chile, where consideration is being given to constitutional amendment, legislation and other measures intended to protect human society from possible adverse consequences of neurotechnology.

6. The Assembly considers that the following ethical principles must be applied to the development and application of neurotechnology in general and BCI technology in particular:

   6.1. Beneficence and prevention of malign use. This technology should be developed and applied only for purposes that are consistent with respect for human rights and dignity. Research aimed at incompatible purposes should be prohibited. Special attention should be given to dual-use technology and technology developed for military or security purposes. New neurotechnology should be subjected to a prior human rights impact assessment before being put into use.

---

[2] Draft resolution unanimously adopted by the committee on 9 September 2020.

6.2. Safety and precaution. This technology should be safe for both the user and, in their intended or unintended consequences, society in general. Safety must be ensured before any new applications are put into use.

6.3. Privacy and confidentiality. At a minimum, information gathered by neurotechnological and BCI devices must be protected according to general principles of data protection. Consideration should also be given to protecting 'neurodata' as a special category, for example by analogy to prohibitions on commerce in human organs.

6.4. Capacity and autonomy. This technology should not be used against a subject's will or in a way that prevents the subject from freely taking further decisions about their continued use. Special care will be needed where such technology is used to treat chronic pain, drug dependency or other conditions where interruption of treatment could lead to discomfort or distress.

6.5. Human agency and responsibility. This technology should not prevent an individual from acting freely and being responsible for their actions. Human beings, acting freely according to their natural (as opposed to enhanced or symbiotic) consciousness, must remain the only decision-makers and the primary actors in society, especially in matters that may impact human rights and democratic processes.

6.6. Equity, integrity and inclusiveness. This technology should not create any form of privileged or superior status for their users; it should be implemented with respect for human equality and dignity, including of members of marginalised or vulnerable groups; and it should be made available as widely as possible, especially insofar as they are applied for medical purposes.

6.7. Ensuring public trust through transparency, consultation and education/awareness-raising. The implementation of new technologies, such as neurotechnology intended for use by individuals, will be best favoured and accepted if it takes place with the confidence of the public, in awareness of the benefits as well as the potential dangers.

7. The extent to which BCI technology may have the potential to change fundamentally the relationship between the individual's internal and subconscious self and the outside world implies unique and unprecedented threats to fundamental values of human rights and dignity. The Assembly notes with particular interest proposals to establish and provide legal protection for new human rights, sometimes referred to as 'neurorights'. These proposals are intended to fill the gaps in the existing human rights framework through which BCI technology might threaten enjoyment of currently protected rights and, beyond that, respect for fundamental human dignity. The rights in question have been expressed as cognitive liberty, mental privacy, mental integrity and psychological continuity.

8. The Assembly therefore calls on Council of Europe member States to:

8.1. establish ethical frameworks for research, development and application of neurotechnology, including BCI technology, taking into account the principles set out in paragraph 6 of the present resolution;

8.2. clearly define the limits of research, development and application of neurotechnology, including BCI technology, through specific legal frameworks that ensure effective respect and protection of human rights;

8.3. ensure that appropriate bodies exist for the oversight and regulation of research, development and application of neurotechnology, including BCI technology, so as to ensure effective implementation of the applicable ethical and legal frameworks;

8.4. consider the establishment and legal protection of new 'neurorights' as a particularly effective protection against possible risks posed by BCI technology.

9. As regards relevant work already underway within the Council of Europe, the Assembly:

9.1. encourages the DH-BIO to take an open and constructive approach to the question of new 'neurorights', including the possibility of assuring their protection under international law through an additional protocol to the Convention for the Protection of Human Rights and Fundamental Freedoms (ETS No. 005);

9.2. encourages the CAHAI to take account of the potential risks and opportunities arising from the application of artificial intelligence in the context of BCI systems and its particularly serious impact on human rights.

B. Draft recommendation[3]

1. The Assembly refers to its Resolution … (2020) entitled “The brain-computer interface: new rights or new threats to fundamental freedoms?” It recalls that this resolution was adopted as relevant work was ongoing within the Council of Europe by the Committee on bioethics (DH-BIO), concerning neurotechnology, and the Ad hoc Committee on artificial intelligence (CAHAI).

2. The Assembly therefore calls on the Committee of Ministers to:

   2.1. support work within the DH-BIO on human rights and neurotechnology, including by supplementing existing terms of reference to ensure that consideration is given to the possibility of protecting ‘neurorights’ through an additional protocol to the Convention for the Protection of Human Rights and Fundamental Freedoms (ETS No. 005);

   2.2. take into account the potentially unique and unprecedented impact on human rights of the use of artificial intelligence in connection with brain-computer interface (BCI) systems when assessing the feasibility of a legal framework for artificial intelligence.

---

[3] Draft recommendation unanimously adopted by the committee on 9 September 2020.

C. Explanatory memorandum by Mr Becht, Rapporteur

1. Introduction

1. The motion for a resolution (Doc 14814) underlying this report, which I tabled on 23 January 2019, was referred to the committee on 12 April 2019, following which the Committee appointed me as rapporteur on 29 May 2019.

2. This motion recalled that computers had always been seen as representing human progress, which is why the gradual connection to the human body of devices involving digital computing has occurred without much debate. Such devices now allow direct interface between the human brain and computers. On the one hand, this may have valuable medical applications, such as restoration of disabled people's ability to speak or manipulate physical objects; on the other, it could allow information to be read from, added to or deleted from the human brain – breaching a barrier to the ultimate refuge of human freedom: the mind.

3. Following the Committee's discussion of my introductory memorandum at its meeting in Berlin on 14-15 November 2019, I conducted a fact-finding visit to California, United States of America, from 25-27 February 2020. At the University of California, Berkeley, I met Professor Jack Gallant (with whom I discussed the latest developments in brain imaging technology, including its use to reconstruct visual images from neural activity), Professor Ehud Isacoff and Professor Stuart Russel; at the University of California, San Francisco, I met Josh Chartier, with whom I discussed the use of electrocorticography and 'neural network' machine learning algorithms to convert decoded neural activity into synthetic speech. Owen Phillips, CEO of start-up company BrainKey, told me of his plans to develop medical diagnostic tools based on artificial intelligence (AI) to analyse magnetic resonance imaging (MRI) scans of paying clients. José Carmena, CEO of start-up company Iota Biosciences, told me about his company's highly miniaturised medical diagnostic devices, sometimes described as 'neural dust'. At the University of Stanford, I spoke with Professor Hank Greely, Alix Roger and Daniel Palanker; Professor Oussama Khatib and Shameek Ganguly, with whom I discussed robotics and prosthetics; Professor EJ Chichilnisky, with whom I discussed artificial retinal implants and technological threats; and Professor Nick Melosh, with whom I discussed auditory prostheses. Professor Byron Yu of Carnegie Mellon University described the latest developments in neural recording and stimulation technology, including optical imaging. Alan Mardinly and Alex Feerst of Neuralink described their neural recording technology. With Luc Julia, Vice-President of Innovation at Samsung, I discussed consumer applications for interfacing with technology. With Greg Corrado, senior research director at Google, I discussed the use and possible abuses of technology in general and in health care in particular. I would like to thank all of these individuals for their time and contributions, and also to thank Emmanuel Lebrun-Damiens, Consul-General of France, San Francisco, and his colleagues for their invaluable assistance in organising my visit.

4. I had also intended to organise a hearing with experts at a meeting of the Committee but unfortunately this was made impossible by the Covid-19 pandemic. I would like to thank Dr Marcello Ienca, Chair of Bioethics, Health Ethics and Policy Lab, D-HEST, Swiss Federal Institute of Technology, ETH-Zurich, Dr Timothy Constandinou, Deputy Director, Centre for Bio-inspired Technology, Imperial College London, and Dr David Winickoff, Senior Policy Analyst, Organisation for Economic Co-operation and Development (OECD), for their willingness to participate in the planned hearing and for the written contributions to this report they provided instead.

2. Technology

5. The history of neurotechnology is relatively recent. It is just over a hundred years since the first electroencephalography (EEG) recording of electrical signals in the brain of an animal; the first human EEG was recorded in 1924. The first direct electrical stimulation of the human auditory system was conducted in 1957. In 1965, an EEG was used to compose music; by 1988, an EEG could be used to control a mobile robot. In 1997, the US Food and Drug Administration approved the use of deep brain stimulation (DBS, an invasive technology – see below) as a treatment for essential tremor and Parkinson's disease. Progress has accelerated over the past 20 years, in parallel with the exponential growth in computing power: in 2005, Matt Nagie became the first person to control an artificial hand using a brain-computer interface (BCI); in 2013, a BrainGate patient controlled a robot prosthetic limb via an array of micro-electrodes implanted into the brain; and in 2018, researchers at Berkeley created the world's smallest, most efficient implanted 'neural dust' wireless nerve stimulator. Leading actors in the technology industry are now working on commercial applications, including Elon Musk's Neuralink company and Facebook's Reality Labs.

6. The core technology in BCIs consists of two components: a device to record or stimulate brain activity; and a 'decoder' algorithm to extract information from the recorded activity or to create a signal to stimulate activity. The recording and stimulating technologies may be either non-invasive (remaining outside the skull) or invasive (introduced within the skull). Along with EEG, non-invasive recording technologies include functional magnetic resonance imaging (fMRI) and functional near-infrared spectroscopy (fNIRS), both of which record neural activity by measuring blood flow within the brain; non-invasive stimulating technologies include transcranial direct current stimulation (tDCS) and transcranial magnetic stimulation (TMS), both of which induce electrical current within the brain. Invasive BCI recording technologies include electrocorticography (ECoG) and cortical implants, which involve placing electrodes directly onto the cerebral cortex; at a more experimental level, 'neural dust' (wireless, battery-free miniature implants, fitted with sensors and stimulators and powered by ultrasound), 'neural lace' (tiny electrodes distributed along polymer 'threads' that are inserted – 'injected' – into the brain, as being developed by Neuralink) and 'neuropixels' (another type of multi-electrode array capable of accessing many different regions of the brain simultaneously) represent a further step in sophistication. Deep brain stimulation (DBS) and vagus nerve stimulation (VNS) are examples of invasive stimulating technologies.[4]

7. Impressive progress continues in all of these areas. For example, shortly before I arrived in California, the 'NexGen 7T MRI' brain scanner was installed at the Helen Wills Neuroscience Institute at UC Berkeley, 7T referring to the strength of the magnets (7 Teslas). The NexGen 7T MRI is an extremely expensive reportedly $13.4 million) and truly international (Siemens, a company headquartered in Germany, built the NexGen 7T MRI) endeavour, like much of the research in this area. The NexGen 7T MRI will have an imaging resolution of around 0.4 mm, which corresponds to the scale of neural column structures that respond to specific features of the sensory world. Whilst spatial resolution is thus becoming less of a limitation, MRI technology is still subject to significant temporal resolution limitations: whereas neurons fire at a rate of around 100 times per second, an MRI records at a rate of around once per second.

8. Despite the rapid recent progress and the variety of approaches currently being explored, there are still many significant constraints on development of more ambitious applications of BCI. Non-invasive approaches are unable to record detailed activity at cellular level and can only be used for simple binary-choice interfaces. At present, anything more sophisticated requires major surgery. Even these invasive technologies have serious limitations, however, including the degradation over time of the quality of recordings obtained via implants. Then there is the need to communicate data from within the skull: either a wire must be passed through the skull, or a wireless system used, which creates its own problems. The best current invasive multi-electrode array systems can record via up to a thousand channels, monitoring hundreds of neurons from a single area of the brain; but a more general purpose BCI would require sampling from tens, if not hundreds of thousands of sites, potentially across multiple areas of the brain. This introduces further computational and data analysis challenges. There are also engineering and surgical problems related to the manufacture and implantation of complex three-dimensional structures with integrated electronics. Above all, "We simply do not understand well enough the nature of distributed information representations and processing in the neocortex to be able to make more than a rudimentary estimate of what a particular sequence of activity might 'mean'."[5]

9. On the technological front, at least, some researchers anticipate significant progress in the coming years, on the basis of what they call 'neuralnanorobotics'. This would involve microscopic devices that would be introduced into the subject's bloodstream, which they would follow, where necessary crossing the blood-brain barrier, before locating and attaching to specific types of brain cell or cellular structure. These devices would then record and/or stimulate neural activity in order to "provide a non-destructive, real-time, secure, long-term, and virtually autonomous *in vivo* system that can realise the first functional BCI" – in other words, a system that would overcome the constraints mentioned above. "Such human BCI systems may dramatically alter human/machine communications, carrying the promise of significant human cognitive enhancement …"[6]

10. These proposals are not as far-fetched as they may seem. Electromagnetic nanoparticles have already been used to control intrinsic fields within mouse brains, and fluorescing carbon nanodots have been used to target and image specific cells in mouse brains. A human brain has been linked to the spinal cord of an anaesthetised rat, and a human brain has guided the movements of a cockroach along an S-shaped track via electrical stimulation of its antennae. Multiple brains have been connected in order to perform co-operative tasks: four connected rat brains were found to outperform a single brain in computational performance; and a

---

[4] For further information on different types of neurotechnology, see for example "iHuman: Blurring lines between mind and machine", the Royal Society, September 2019.

[5] Mitrasinovic et al, "Silicon Valley new focus on brain computer interface: hype or hope for new applications?", F1000Research 2018, 7:1327 (as updated 21 January 2019).

[6] Martins et al (2019) Human Brain/Cloud Interface, Front. Neurosci. 13:112.

three-human brain-to-brain interface (BBI), called ‘BrainNet’, allowed three human subjects to collaborate via non-invasive direct brain-to-brain communication (EEG and TMS) in order to take decisions. As the Royal Society paper says of imagined future neurotechnologies, “These things are a long way off but not impossible in some form. Think how distant and futuristic landing on the moon or the internet would have seemed in 1950, when few households even possessed a telephone …”

11. Progress in BCI technology has also been significantly driven by the rapid development of AI over the past decade or so. Analysis of brain images, notably those captured by MRI, by machine learning algorithms has contributed to our understanding of how the brain is functionally structured and our ability to decode neural activity in order to reconstruct the thought patterns that it represents. This has allowed researchers to reconstruct images from MRI scans of subjects watching movie trailers (as Professor Gallant and his team did as far back as 2011), or to convert brain signals captured by ECoG into synthetic speech (as Dr Chartier is doing – see further below). Just as in other fields as disparate as medical diagnostics or autonomous vehicles, the progress in AI has been instrumental. Further, general information on AI, including a description and an examination of the applicable ethical principles can be found in appendix to the present report.

3. Applications

12. As noted above, the earliest BCI applications used brain signals to control simple prosthetic and robotic devices.[7] As both sensory devices and computational power have developed, so have new, more sophisticated applications become possible.

13. With research on psychological disorders and geriatric neurological disorders, for example, attracting enormous amounts of funding, it is not surprising that much current research into neurotechnology focuses on medical applications. Josh Chartier at the University of California, San Francisco told me about the project he was working on to detect the patterns of neural activity associated with control of the vocal tract during speech. The resulting signals can then be decoded and used to generate ‘synthetic speech’. This approach gets around the enormous difficulty of identifying neural activity associated with specific words as such (as compared to the activity associated with the intention of speaking those words). The research is being conducted on severe epilepsy patients who had been fitted with a ‘neural net’ intended to monitor seizures, and who have volunteered to allow the device to be used also for this research. Such speech synthesising neurotechnology would only be used for persons in the most extreme situations, such as ‘locked-in syndrome’, since it is highly intrusive. Dr Chartier recognised the possibility that a technique which works with people who can still speak may not work for people who have for a long time been unable to speak, and who may no longer display the same detailed, consistent patterns of neural activity; although he did suggest that with practice, ‘locked-in’ syndrome patients may be able to ‘retrain’ their brains to produce patterns of activity that can be decoded and used to generate ‘synthetic speech’. The basic concept – deriving articulated language from neural activity – illustrates how increasing understanding of how the brain works, coupled with improving technology for reading the brain. The intrusive technology and the reliance on conscious mental effort to control the vocal tract may mean that some of the ethical concerns, in terms of privacy and mental integrity, associated with this approach are less prevalent.

14. José Carmena’s company, Iota Biosciences, is developing battery-free, ultrasonically powered bioelectronic devices known as ‘neural dust’, measuring only a few millimetres in length, that can interface directly with the central nervous system. These devices can gather precise data or directly stimulate nerves and could be used to diagnose and treat conditions from arthritis to cardiovascular disease. Iota aims to reduce the size of its devices to sub-millimetre level, which may eventually make possible their use in brain-computer interface technology.

15. Some companies are already supplying direct-to-consumer services or products. BrainKey, which was presented to me by the company CEO Owen Phillips, is assembling a database of MRI scans, mainly obtained from public databases (such as UK Biobank), with others from private clients. At present, Brainkey’s service is essentially descriptive/analytical, including statistics and a 3D-printed model of the individual’s brain. The intention is to develop AI diagnostic tools. Whilst Dr Phillips was against employers being allowed to require job applicants to undergo MRI scans that would then be analysed in an attempt to predict the individual’s future professional performance, he accepted that the technology could potentially be used to this end.

---

[7] See for example Chapin et al, “Real-time control of a robot arm using simultaneously recorded neurons in the motor cortex” (of a rat), Nat. Neurosci. 1999 Jul;2(7):664-670.

16. The Canadian company InteraXon Inc. sells a multi-sensor EEG-based device called Muse, intended to assist users during meditation practice by recording their neural activity (as well as their physical stillness, via an accelerometer, and their heart-rate, via an optical heart-rate monitor). An accompanying smartphone-based app gives users feedback on their practice and progress. Users are told that they retain full control over the EEG data generated when using their Muse device, but can opt in to a research programme that shares anonymised EEG (and other Muse sensor) data with "third parties involved in research related to improving the scientific understanding of the brain/body or to improving products and/or delivering better experiences and services."[8]

17. OpenBCI takes a different approach, selling open-source hardware (including EEGs, along with EMGs for sensing muscular activity and ECGs for heart function, as well as all sorts of associated components) that can be used with free, open-source software for various projects. OpenBCI's website states that "We work to harness the power of the open source movement to accelerate ethical innovation of human-computer interface technologies."[9] This can be seen as corresponding to the 'democratisation' model of neurotechnological development suggested by Dr Ienca (see further below).

18. Research funded by Facebook used ECoG technology to understand, on the basis of neural signals alone, what subjects were both hearing and intending to say. This was described as "illustrating the promise of neuroprosthetic speech systems for individuals who are unable to communicate";[10] or, as Facebook Vice-president Andrew Bosworth announced on Twitter, a "wearable device that lets people type by just imagining what they want to say" – which is perhaps more revealing of Facebook's commercial interest in this technology. Whilst Dr Chartier's research, for instance, has already given promising results using invasive ECoG technology, it is difficult to see how non-invasive 'wearable' devices could be comparably effective, in the foreseeable future at least.

19. Another area of commercial interest is the field of so-called 'neuromarketing', "a recent interdisciplinary field which crosses traditional boundaries between neuroscience, neuroeconomics and marketing research … primarily concerned with improving marketing strategies and promoting sales".[11] One study showed that subjects who preferred Pepsi Cola over Coca Cola during blind tasting had a strong response in one region of the brain (the ventral putamen); during unblind tasting, almost all subjects preferred Coke, with a particularly strong response in another part of the brain (the prefrontal cortex, which is linked to the sense of self – and which, in this context, seemed to 'over-ride' the taste-buds of those who had previously preferred Pepsi). Heightened understanding of this mechanism could allow advertisers to develop and test marketing strategies intended to engage sub-conscious preference mechanisms that are particularly responsive to image or branding, as opposed to a product's intrinsic qualities.

20. Neurotechnology and BCIs may have particular potential in the field of criminal and judicial proceedings. "[T]here are attempts to use neuroscience to develop objective methods for assessing what have been inherently subjective questions. Truth-telling and lie detection, mental capacity, pain, and memory reliability are useful areas of study for criminal and civil litigation. These attempts at mind reading are becoming theoretically possible."[12] Indeed, some of these technologies are already commercially available, even if they have not yet been accepted for use in the courtroom: for example, the companies No Lie MRI and Cephos both market products using fMRI as a basis for assessing an individual's truthfulness. Another study has suggested that activity in the anterior cingulate cortex region of the brain, which is associated with impulse control, was predictive of subsequent rearrest (recidivism) amongst offenders.[13] All of these developments would pose serious questions from the perspective of substantive rights and procedural guarantees.

21. Research on BCI neurotechnology is not only being done for commercial and civilian purposes; it is also being done with security and military goals in mind. The US Defense Advanced Research Projects Agency (DARPA) has been particularly active in this field. Much of its work has been primarily intended for

---

[8] Quoted text copied from the Muse End User License Agreement on 19 July 2020.

[9] Engineers at Netflix managed to combine, in a sense, these approaches by hacking a Muse, so that it could be used as a sort of hands-free remote control to scroll through a menu (using the accelerometer to detect head movements) and select a programme (using the EEG to detect neural activity, in this case concentration on the word "play").

[10] Moses et al, "Real-time decoding of question-and-answer speech dialogue using human cortical activity", Nature Communications, 2019. Other researchers have used the same technology to arrive at similar results, e.g. Salar et al, "Classification of Articular Movements and Movement Direction from Sensorimotor Cortex Activity", Scientific Reports/ Nature Research, October 2019.

[11] Ulman, Cakar and Yildiz, "Ethical Issues in Neuromarketing: 'I Consume, Therefore I am!'", Sci Eng Ethics (2015) 21:1271-1284.

[12] Goodnough and Tucker, "Law and Cognitive Neuroscience", Annu. Rev. Law Soc. Sci. 2010. 6:61–92.

[13] Aharoni et al, "Neuroprediction of future arrest", PNAS Early Edition, 2013.

general medical purposes, albeit with a focus on issues of particular relevance to the armed forces. The ‘revolutionising prosthetics’ and ‘reliable neural interface technology’ (RE-NET) programmes, for example, were intended to accelerate the development of BCI controlled prosthetics (neuroprosthetics); and the ‘reorganisation and plasticity to accelerate injury recovery’ (REPAIR) programme had the aim of restoring neural and behavioural function following neural injury or sensory deprivation.

22. Other DARPA-funded research may give rise to more complex ethical questions. The ’restorative encoding memory integration neural device’ programme (REMIND) developed technology that could either improve or impede a subject’s ability to record events in its memory. The ‘accelerated learning’ programme was intended to “revolutionize learning in the military environment” (for example, of rifle marksmanship). The ‘narrative networks’ (N2) programme aimed to detect brain activity associated with narrative influence, which could allow “faster and better communication of information in foreign information operations” and create BCI technologies that “close the loop between the storywriter and consumer, allowing neural responses to a narrative stimulus to dictate the story’s trajectory”. This could be used to create “optimal narratives tailored to a specific individual or group of people”. The ‘neurotechnology for intelligence analysts’ (NIA) programme was intended to develop new non-invasive BCI systems to increase the efficiency and productivity of imagery analysts by detecting neural responses to seeing ‘targets of interest’ on images. The ‘cognitive technology threat warning system’ (CT2WS) would use a similar approach in order to enhance the ability to detect and respond to threats during site security surveillance. DARPA has also been working on low-cost EEG headsets, with a view to “crowdsourcing data collection efforts for neuroimaging”. More recently, the ‘systems-based neurotechnology for emerging therapies’ (SUBNETS) programme hoped to produce new implantable technologies for neural recording and stimulation in order to treat neuropsychiatric and neurological conditions, including amongst armed forces’ veterans with service-related mental health problems. The ‘restoring active memory’ (RAM) programme aims at restoring memory in human patients with untreatable illnesses who suffer from memory deficits. Perhaps out of recognition of the potential ethical issues, DARPA is reported to have worked closely with the Food and Drug Administration on these latter two programmes.[14] Experts whom I met in California considered that the DARPA-funded projects were a long way away from achieving their stated goals, which were often deliberately expressed in extremely ambitious terms, in the knowledge that they went beyond what was currently possible and the expectation that at least some useful progress would nevertheless be made.

23. Elon Musk has said that Neuralink is intended to find a way to “achieve a sort of symbiosis with artificial intelligence” – a reflection of his stated belief that artificial intelligence is “our biggest existential threat”, “potentially more dangerous than nukes” (nuclear weapons). These concerns were echoed by the advocates of ‘neuralnanorobotics’ as a means of interfacing the human brain with the cloud: this “may be beneficial for humanity by assisting in the mitigation of the serious existential risks posed by the emergence of artificial general intelligence. [It would do this] by enabling the creation of an offsetting beneficial human augmentation technology …”[15]

24. Alan Mardinly, Neuralink’s research director, told me that Elon Musk’s ‘symbiosis’ vision was far away as a goal, with many technical and scientific unknowns to be addressed before it could be reached. The company’s immediate aim was to produce medical devices, notably to assist patients with spinal cord injuries. Since Neuralink’s intra-cortical ‘neural lace’ technology involved thousands of electrodes, arranged along ‘threads’ hanging from a 4x4 mm chip, it provided sufficient resolution to control devices. Mr Mardinly considered neural lace to be the best currently available technology for brain-computer interface, although its invasive nature meant that the range of applications would necessarily be narrow. Nevertheless, he considered it preferable to the existing industry-standard Utah electrode array: neural lace used flexible electrodes (as opposed to the rigid electrodes of the Utah array) and was thus both less disruptive of brain tissue and longer-lasting, since it did not cause localised scarring; and it offered greater resolution, with 8 to 16 electrodes per very narrow shank (as opposed to the Utah array’s single electrode on a relatively thick shank). In time, neural lace could be used to implant up to 100,000 electrodes, although an integrated, automated insertion system would be needed for this. To summarise, Neuralink aimed to produce technology that would be safe, stable and scalable. Neuralink was already working with the US Food and Drug Administration (FDA) on an ‘early feasibility study’; this was primarily concerned with safety issues, such as cybersecurity/ hackability, data/ privacy implications and input control issues, however, rather than ethics.

---

14. Miranda et al, “DARPA-funded efforts in the development of novel brain-computer interface technologies”, Journal of Neuroscience Methods 244 (2015) 52-67.
15. Martins et al, op. cit.

25. Looking further into the (possible) future, the advocates of ‘neuralnanorobotics’ envisage this technology, in conjunction with the global ‘cloud’ infrastructure, being used in many applications: fully immersive virtual reality that would be indistinguishable from reality; augmented reality, with information about the real world superimposed directly onto the retina; real-time auditory translation of foreign languages, or access to many forms of online information; and even to experience “fully immersive, real-time episodes of the lives of any willing human participant on the planet, via non-intrusive ‘Transparent Shadowing’”.

26. In his written contribution to this report, Dr Winnickoff set out the taxonomy of neurotechnologies below (NB I have slightly modified his proposal for the sake of simplicity and relevance). Whilst ‘brain-computer interface’ appears in relation to the primary function of particular technology, it evidently also encompasses reading and intervening/modulating functions. As this report has already described, BCI technology shows the potential to be used in all of the suggested spheres.

- Spheres of use, which include:
    - clinical/medical (neurology/neurosurgery, psychiatry, rehabilitation, pain medicine)
    - occupation (training, performance)
    - military (intelligence/interrogation, weapons/warfighter enhancement)
    - public (direct to consumer/do-it-yourself: education, wellness/lifestyle/entertainment)
- Primary function of the technology, including:
    - reading the brain (imaging, modelling/mapping)
    - intervening/modulating brain function
    - ‘engineering’ the brain (brain-computer interface, neuroprosthetics)
    - derivative (inspired by the brain, for example artificial neural networks)
- Health aims – to prevent, to restore, to replace, to augment…

27. As the Royal Society’s paper summarises:

> *“Implants, helmets, headbands or other devices could help us remember more, learn faster, make better decisions more quickly and solve problems, free from biases … Linking human brains to computers using the power of artificial intelligence could enable people to merge the decision-making and emotional intelligence of humans with the big data processing power of computers, creating a new and collaborative form of intelligence. People could become telepathic to some degree, able to speak not only without speaking but without words – through access to each other’s thoughts at a conceptual level … Not only thoughts, but sensory experiences, could be communicated from brain to brain … Mentally and physically enhanced military or police personnel could protect the public by being able to see more effectively in the dark, sense the presence of others and respond rapidly …”*

28. Nevertheless, it is still worth bearing in mind that current technology is very far from achieving almost any of these purposes, or at least not in a practically useful (or threatening) way. Without significant progress in our understanding of the structure and functioning of the brain, the more spectacular possibilities for BCI will remain purely speculative. Some experts have fundamental doubts about the potential of neurotechnology as a basis for interacting with computers. Greg Corrado of Google argued that the human nervous system had evolved to interact with the outside world using the natural senses, which would remain the primary means of interfacing with computers: in particular, the eyes and ears, to receive information, and the fingertips and vocal tracts, to impart information. Improvements in human-computer interface would thus occur on the computer side, such as natural speech processing (the ability of a computer to ‘understand’ human speech). In his view, the most promising possibilities for bioelectronic interface were direct stimulation of the muscular nerves and cochlear and retinal implants, to input information (stimulate), and the motor cortex to read information. In Dr Corrado’s view, the current state of BCI technology does not give rise to any ethical risks, a view that was shared by many whom I met in California. That said, its misuse could raise ethical concerns, for example if fMRI technology were used to predict individuals’ personalities, criminal propensities or professional capacities – which Dr Corrado considered to be akin to phrenology, the discredited theory whereby an individual’s personality traits and intellectual capacity could be determined by measuring physical features of the skull.

4. Concerns

29. The current state of BCI technology may not give rise to immediate concerns that manifestly exceed the scope of existing ethical frameworks (notably medical ethics and privacy/data protection regulations), but the pace of their evolution will go beyond the ethical concerns, and the breaking point beyond which current law becomes inadequate is close. The march of technological progress is unrelenting and some actors will inevitably seek to develop applications that do raise serious ethical issues, whether on the basis of new technologies or through the use of existing technologies. It is nothing more than simple precaution to anticipate what may happen in future and to set limits that steer research away from foreseeably harmful or dangerous areas and towards positive applications that do not threaten individual rights or democratic societies. Some of these ethical issues are outside the scope of the present report: for example, the fact that interfacing with a brain "inevitably changes the brain".[16] For present purposes, I will concentrate on the human rights-related issues, which are based in an underlying concern to ensure respect for human dignity and the essential characteristics of the human being as an autonomous moral agent.

30. Access to the neural processes that underlie conscious thought implies access to a level of the self that by definition cannot be consciously concealed or filtered – the ultimate violation of privacy. Even today's neuromarketing technology, combined with large-scale dissemination of targeted messages via advertising and social media, could have a devastating impact on freedom of choice and democratic processes. BCIs could be used to create advanced lie detectors that may be seen as so reliable that the information they obtain would be admitted in evidence in criminal proceedings. This may, however, violate the protection against self-incrimination: it would in effect be impossible for a suspect to decline to provide information to an interrogator. At the same time, BCIs could be used to create false memories, or amend or delete real ones, rendering human testimony unreliable as evidence.

31. 'Dual-use' research, of the type funded by DARPA (see above), has given rise to particular concerns. Dual-use refers to both the use of civilian technology for military, national security or police/judicial purposes, and the possibility of harmful misuse, including by non-State actors, of using otherwise beneficial technology. In this respect, it has been noted that "The reliance of these technologies on computation and information processing also makes them potentially vulnerable to cyber-attack",[17] whose potential consequences for the individual become ever more dangerous as BCI technology becomes more sophisticated and its ability to affect neural processes more powerful and specific.

32. There are also issues of access and fairness. Risks are already emerging that access to powerful cognitive enhancement for non-clinical purposes could be dependent on wealth, with some able to afford it and others not. This could create two categories of human being, the enhanced and the non-enhanced. These technologies are also such that decisions on what is possible and who should benefit could be left to the unregulated market, or be dictated by the interests of potentially autocratic governments.

33. Finally, "[o]n a more philosophical level, there are fears that widespread use of neural interfaces could lead to human decisions being directed by what some have called 'neuro-essentialism' – the perception that the brain is the defining essence of a person and that our choices can be reduced to a set of neurobiological processes, leaving no room for individual agency or moral responsibility …"[18] One might ask, what is the point of developing BCIs as a way of defending humanity against the potential risk of general AI, if in doing so we negate the defining qualities of individual human existence?

34. Professor Rafael Yuste of New York's Colombia University has articulated these concerns on behalf of the 'Morningside Group' of 25 experts working in and around the field of neurotechnology. In a 2019 speech to the Inter-Parliamentary Union meeting in Doha (Qatar), Professor Yuste described how possible misuse and a lack of regulation could lead to problems in the following five areas:[19]

– Personal identity: "the more we are connected to the net through brain computer interfaces or devices, the more we dilute our own identity, our self. The dependency that we are witnessing now on our devices is an appetizer of what's to come: as we increase the bandwidth of our connection to the net, by using non invasive brain computer interfaces, we will become increasingly dissolved in it."

---

[16] Amy L. Osborn and Bijan Pesaran, "Parsing learning in networks using brain-machine interfaces", Curr Opin Neurobiol. 2017 October; 46: 76-83.

[17] Ienca, Jotterand and Elger, "From Healthcare to Warfare and Reverse: How Should We Regulate Dual-Use Neurotechnology?", Neuron 97, January 17, 2018, 269-274.

[18] The Royal Society, op. cit.

[19] See https://www.ipu.org/sites/default/files/documents/yuste_speech_final_0.pdf.

- Free will: “if we use external algorithms and information to make decisions, we are relinquishing our own agency. Who is making the decision? … what will happen when we have a life GPS that advices us as to what we should be doing at any moment”.
- Mental privacy: “If brain data is accessible and can be deciphered, then our mental processes, our thoughts, will be accessible from the outside. Moreover, even thoughts we are not aware of, or subconscious, could be deciphered. We think brain data should be protected with the same legislative rigor as body organs. In fact, our brain data is an organ, not a physical organ, but a mental organ, and it should be forbidden to commerce with it, as it represents who we are.”
- Cognitive augmentation, including augmented learning that “could enable some groups of the society in some countries to augment their mental and physical abilities, by enabling them to access external algorithms and robotics for daily life. We think that guaranteeing the principle of justice in the development and deployment of these technologies should ensure equality of access and that the use of these technologies for military application should be severely regulated.”
- Protection against biases and discrimination, “since algorithms used in AI often have implicit biases so these technologies could inadvertently implant these biases into our brain processing. It will be terrible to undo our historical march towards equality and justice by spreading biases with the new technology.”

5. Responses

35. There is widespread agreement amongst researchers on the need for anticipatory regulatory action in relation to emerging neurotechnologies, including BCIs: what the authors of the Royal Society paper describe as an “’early and often’ approach”. Industry has also supported such an approach: Mark Chevillet, Research Director at Facebook Reality Labs, has observed that “We can’t anticipate or solve all of the ethical issues associated with this technology on our own. What we can do is recognize when the technology has advanced beyond what people know is possible, and make sure that information is delivered back to the community. Neuroethical design is one of our program’s key pillars — we want to be transparent about what we’re working on so that people can tell us their concerns about this technology.”[20] Others have argued for co-ordination between research in neuroethics and related fields, such as AI – the overlap between the two fields in terms of technology should be reflected in a common approach to ethical principles.[21] Some have argued in favour of a “neurosecurity framework … designed and implemented to maximize security across the whole translational continuum between scientific research and society (and reverse) … particularly sensitized to anticipate and promptly detect neurotechnology-specific threats, especially those that concern the mental dimension … [This framework] should include, at least, three main levels of safeguard: calibrated regulatory interventions, codes of ethical conduct, and awareness-raising activities.”[22]

36. In 2013, the Nuffield Council on Bioethics published a detailed report entitled “Novel neurotechnologies: intervening in the brain”. This report, which focuses primarily but not exclusively on medical applications or neurotechnology, notes that “The brain has a special status in human life that distinguishes it from other organs. Its healthy functioning plays a central role in the operation of our bodies, our capacities for autonomous agency, our conceptions of ourselves and our relationships with others – and thus in our abilities to lead fulfilling lives. This means that the novel neurotechnologies that we consider in this report, each of which intervenes in the brain, raise ethical and social concerns that are not raised to the same extent by other biomedical technologies.” The report proposes an ethical framework constructed in three stages:

- Foundational principles of beneficence and caution, arising from “a tension between need and uncertainty.” Whilst serious brain disorders cause severe suffering, there is an absence of (other) effective interventions. At the same time, the full benefits and risks of neurotechnology are also not yet fully understood due to their novelty and an incomplete understanding of how the brain itself works. “The special status of the brain therefore provides both a reason to exercise beneficence by intervening when injury or illness causes brain disorders, and a reason for caution when we are uncertain what the effects of doing so will be.”
- The implications of the principles of beneficence and caution should be examined against a set of five ‘interests’: safety, (unintended) impacts on privacy, autonomy (both in treatment-specific decisions and the wider context of patients’ lives), equity of access to new treatments and trust in novel technologies.

---

[20] “Imagining a new interface: Hands-free communication without saying a word”, Tech@facebook, 30 July 2019.
[21] Marcello Ienca, “Neuroethics Meets Artificial Intelligence”, theneuroethicsblog.com, 1 October 2019.
[22] Ienca, Jotterand and Elger, op. cit.

– The report also proposes three ‘virtues’ to guide the practice of actors in this area. These are inventiveness (both in innovation and in providing wider access), humility (acknowledging the limits of knowledge and our capacity to use technologies to alleviate brain disorders), and responsibility (through the use of robust research practices and refraining from exaggerated or premature claims for neurotechnologies).

37. Similar approaches are reflected in other proposals for ethical frameworks to regulate BCI and other neurotechnologies. For example, one group of researchers has proposed the following list of ‘Neuroethics Guiding Principles’:

i. Make assessing safety paramount;

ii. Anticipate special issues related to capacity, autonomy, and agency;

iii. Protect the privacy and confidentiality of neural data;

iv. Attend to possible malign uses of neuroscience tools and neurotechnologies;

v. Move neuroscience tools and neurotechnologies into medical or nonmedical uses with caution;

vi. Identify and address specific concerns of the public about the brain;

vii. Encourage public education and dialogue;

viii. Behave justly and share the benefits of neuroscience research and resulting technologies.[23]

38. The issue of ‘dual-use of cognitive technology’ (including BCI) was addressed by Marcelo Ienca in a 2018 article.[24] Noting that “cognitive technologies have the potential to accelerate technological innovation and provide significant benefit for individuals and societies”, Dr Ienca considers that “due to their dual-use potential, they can be coopted by State and non-State actors for non-benign purposes including cybercrime, cyberterrorism, cyberwarfare and mass surveillance. In the light of the recent global crisis of democracy, increased militarisation of the digital infosphere, and concurrent potentiation of cognitive technologies [CT], it is important to proactively design strategies that can mitigate emerging risks and align the future of CT with the basic principles of liberal democracy in free and open societies.”

39. Dr Ienca’s therefore proposes “the democratisation of CT”, adopting elements of the two extremes of governance and regulation, pure laissez-faire and strict regulation. ‘Democratisation’ shares with a strict regulatory approach an appreciation of the novelty of CT as a relatively recent and still immature field, lacking a consensus on core concepts or policies that would be needed to maximise the benefits whilst minimising the risks. It also recognises the magnitude of the opportunities and risks involved – “the potential of influencing human cognitive capabilities, hence determining a non-negligible effect on human cultural evolution and global equilibria” – and the fact that the novelty of CT means that “human societies are now at a historic juncture in which they can make proactive decisions on the type of co-existence they want to establish with these technologies. Privileging laissez-faire approaches at this stage of development would defer risk-management interventions to a time when cognitive technology is extensively developed and widely used, hence refractory to modification.” Democratisation would, however, accept the laissez-faire viewpoint that “over-regulation can (a) obliterate the benefits of cognitive technology for society at large, and, if managed by non-democratic or flawed democratic governments, (b) produce an undesirable concentration of power and control.”

40. This ‘democratisation’ of CT would be based on six normative principles – which are comparable to the components of ethical frameworks suggested for regulation of neurotechnologies more generally:

i. Avoidance of centralised control – “the principle according to which it is morally preferable to avoid centralised control on CT to prevent risks associated with unrestricted accumulation of capital, power, and control over the technology among organised groups such as large corporations or governments… Normative interventions aimed at limiting this risk of centralisation may be conceptualised as cyberethical counterparts of anti-trust laws.”

ii. Openness – “the principle of promoting universal access to (components of) the design or blueprint of cognitive technologies, and the universal redistribution of that design or blueprint, through an open and collaborative process of peer production.” Avoidance of centralised control (Principle I above) and openness “are critical requirements to make these same capabilities that will be recorded through or

---

[23] Greely et al, Neuroethics Guiding Principles, J.Neurosci., December 12, 2018, 38(50):10586 – 10588.

[24] “Democratizing cognitive technology: a proactive approach”, Marcello Ienca, 19 June 2018.

infused in cognitive technologies … available to everyone.” “In a more abstract sense, openness in CT involves the principle of infusing every application that we interact with, on any device, at any point in time, with (components of) cognitive technology.”

iii. Transparency – “the principle of enabling a general public understanding of the internal processes of cognitive technologies.”

iv. Inclusiveness – “the principle of ensuring that no group of individuals or minority is marginalised or left behind during the process of permeation of cognitive technology in our society … The principle of inclusiveness [applies to any] ethically relevant social bias that may emerge intendedly or unintendedly during CT development. These include cultural, political and language bias etc.”

v. User-centredness – “emerging cognitive technologies should be designed, developed and implemented according to users’ needs and personal choices … end-users (as widely as possible characterised, in accordance with the principles of openness and inclusiveness) [should be] involved in the design, development and implementation of cognitive technologies on an equal footage.”

vi. Convergence. “In the narrow sense, convergence is the principle of interoperability, intercommunication and ease of integration among all components of cognitive technology … [although] excessive interoperability might result in increased data insecurity … In a broader and more abstract sense, it is also the principle of converging different types of cognitive technology, especially neurotechnology, on the one hand, and artificial intelligence systems on the other hand.”

41. As regards neurotechnology designed specifically for military applications, it has been argued that “Although a global ban or moratorium on military neurotechnology appears ethically unjustified at present, softer and more calibrated regulatory interventions might be necessary to mitigate the risks of a disproportionate weaponization of neuroscience.” Again, this would imply an “urgent need for monitoring and careful risk assessment in the context of dual-use technology. Even though, at the moment, benefits seem to outweigh the risks, preventive mechanisms should be in place to promptly detect future variations in the risk-benefit ratio.”[25] It would nevertheless be necessary not to deprive oneself of the ability to defend oneself in the event of an adversary having use of these technologies.

42. There is a great deal of common ground between the field of ‘neuroethics’ and that of bioethics. In the field of bioethics, in 1997 the Council of Europe adopted the Convention for the protection of Human Rights and Dignity of the Human Being with regard to the Application of Biology and Medicine: Convention on Human Rights and Biomedicine (ETS No. 164, the ‘Oviedo Convention’), whose purpose is to “protect the dignity and identity of all human beings and guarantee everyone, without discrimination, respect for their integrity and other rights and fundamental freedoms with regard to the application of biology and medicine.” Amongst other things, the Oviedo Convention states that “the interests and welfare of the human being shall prevail over the sole interest of society or science”; obliges the Parties to “take appropriate measures with a view to providing, within their jurisdiction, equitable access to health care of appropriate quality”; and establishes that “any intervention in the health field, including research, must be carried out in accordance with relevant professional obligations and standards”. It also contains detailed provisions on consent, the permissible uses of predictive tests, and the permissible purposes of intervention on the human genome.

43. In November 2019, shortly after I presented my introductory memorandum to the committee, the Council of Europe’s inter-governmental ‘DH-BIO’ bioethics committee adopted a Strategic Action Plan on Human Rights and Technologies in Biomedicine (2020-2025). This document notes that “The application of emerging and converging technologies in biomedicine results in a blurring of boundaries, between the physical and the biological sciences, between treatment and research, and between medical and non-medical purposes. Although they offer significant opportunities within and beyond the field of biomedicine, they also raise new ethical challenges related to *inter alia* identity, autonomy, privacy, and non-discrimination.” Building on a foundation of co-operation (amongst Council of Europe bodies and with other relevant inter-governmental bodies) and communication (with external stakeholders), the Action Plan is structured around three pillars: governance (based on human rights, public dialogue, democratic governance and transparency), equity (in access to innovative treatments and technologies, to combat health disparities due to social and demographic change) and integrity (including strengthening children’s participation, safeguarding children’s rights and safeguarding the rights of persons with mental health difficulties).

44. In relation to BCI and related technology, the Action Plan notes that “Developments in neurotechnologies, such as deep brain stimulation, brain-computer interfaces, and artificial neural networks, raise the prospect of increased understanding, monitoring, but also control of the human brain, raising issues

---

[25] Ienca, Jotterand and Elger, op. cit.

of privacy, personhood, and discrimination … It therefore needs to be assessed whether these issues can be sufficiently addressed by the existing human rights framework or whether new human rights pertaining to cognitive liberty, mental privacy, and mental integrity and psychological continuity, need to be entertained in order to govern neurotechnologies. Alternatively, other flexible forms of good governance may be better suited for regulating neurotechnologies.” Unsurprisingly, DH-BIO shares my concerns and the broad outlines of my perception of the opportunities and risks that BCI technology represents. I welcome and support its future work on the key responses, namely targeted reinforcement of the human rights framework (see further below) and elaboration of a flexible regulatory regime that can support and channel research and development towards positive, constructive ends.

45. Other international organisations are also attentive to emerging neurotechnologies. In December 2019, the OECD Council adopted a Recommendation on Responsible Innovation in Neurotechnology. This recommendation “articulates the importance of (1) high level values such as stewardship, trust, safety, and privacy in this technological context, (2) building the capacity of key institutions like foresight, oversight and advice bodies, and (3) processes of societal deliberation, inclusive innovation, and collaboration.” It then calls on member and non-member States and “all actors” to promote and implement a series of nine “principles for responsible innovation in neurotechnology”, which are elaborated with details of specific actions. The principles are:

i. Promoting responsible innovation;

ii. Prioritising safety assessment;

iii. Promoting inclusivity;

iv. Fostering scientific collaboration;

v. Enabling societal deliberation;

vi. Enabling capacity of oversight and advisory bodies;

vii. Safeguarding personal brain data and other information;

viii. Promoting cultures of stewardship, trust across the public and private sector;

ix. Anticipating and monitoring potential unintended use and or misuse.

46. Neurotechnology in general, and BCIs in particular, have the potential to change fundamentally the relationship between the internal self and the outside world. Researchers (including Dr Ienca) have therefore called for innovative legal responses, including the creation (or specification) of four ‘new’ enforceable human rights: the right to cognitive liberty, the right to mental privacy, the right to mental integrity and the right to psychological continuity. ‘Cognitive liberty’ can be considered as “the right and freedom to control one’s own consciousness and electrochemical thought processes [and as such] is the necessary substrate for just about every other freedom”. In this respect, it is comparable to freedom of thought, which can be seen as a necessary predicate to the freedoms of religion, expression and association. Comparable, but fundamentally different and distinct: if freedom of thought is the right to think whatever one wants, cognitive liberty is its precondition – the right for one’s brain to generate thoughts without technological (or other) interference in this process. The right to mental privacy would protect individuals against non-consensual observation of their sub-conscious mental processes. The right to mental integrity would protect against harm in the form of ‘malicious brain-hacking’, giving control over the individual’s thoughts and actions. The right to psychological continuity would protect against actions that could affect “people’s perception of their own identity … [as] consisting in experiencing oneself as persisting through time as the same person”[26] – to remain psychologically oneself. The article in which these proposals appear is wide-ranging, detailed and thought-provoking. It addresses questions including whether or not these ‘new’ rights are already implicit in existing rights, and whether their ‘creation’ would amount to ‘rights inflation’.

47. One country, Chile, is already working on legal protection of ‘neurorights’, in collaboration with Professor Yuste and the NeuroRights Initiative of Colombia University (see above). A proposed amendment to article 19 of the Chilean constitution would define mental identity as a basic right that can only be altered in accordance with future laws. An accompanying ‘NeuroProtection’ bill would establish legal definitions of neurotechnology, brain computer interfaces and neurorights. All data obtained from the brain would be defined as neurodata and brought within the scope of existing legislation on organ donations, thereby prohibiting commerce in neuroData. All future use and development of neurotechnology would be subject to

---

[26] Ienca and Andorno, “Towards new human rights in the age of neuroscience and neurotechnology”, Life Sciences, Society and Policy, (2017) 13:5.

medical legislation. Alongside these initiatives, the Catholic University in Chile is working on ethical guidelines for the computer, AI and neuroengineering industries. All of these activities are combined with a public outreach campaign supported by the president, government ministers and parliamentarians.[27] This would make Chile the first country in the world to regulate and protect data that could be extracted from the human brain, so that the data can be used for altruistic purposes only.[28] It is said that the new Chilean legal framework would make technology such as Facebook's 'thought-to-type' project illegal.[29]

48. There is some scepticism towards neuro-ethicists' calls for legal protection of neurorights. Alan Mardinly of Neuralink, for example, hypothesised that a right to cognitive liberty might be breached by ordinary advertising, which was often deliberately designed and targeted to exploit subconscious predilections; likewise, treatment for addiction could also be seen as an external interference with an individual's freedom of choice to consume.

49. As regards the AI aspect of BCI technology, in September 2019 the Committee of Ministers established Ad Hoc Committee on Artificial Intelligence (CAHAI). The CAHAI has been instructed to examine the feasibility and potential elements of a legal framework for the design, development and application of artificial intelligence. Its work is based on Council of Europe standards of democracy, human rights and the rule of law, as well as other relevant international legal instruments and ongoing work in other international and regional organisations. Along with the usual participants representing Council of Europe member and observer States and other Council of Europe bodies (including the Assembly), the CAHAI has an exceptionally high level of involvement of representatives of private sector bodies, civil society, and research and academic institutions.

50. The CAHAI held its first meeting on 18-20 November 2019. Amongst other things, it decided that a key element of the future feasibility study would be a "mapping of risks and opportunities arising from the development, design and application of artificial intelligence, including the impact of the latter on human rights, rule of law and democracy". The CAHAI currently expects to adopt the feasibility study at its third meeting, scheduled for December 2020.

51. This is the institutional context within which the Assembly will debate the present and the various other AI-related reports currently under preparation in different committees. The Assembly has chosen to approach the topic on a contextual basis, examining the impact of AI in different areas. Within the Committee on legal affairs and human rights, for example, there are also reports on the impact of AI on "Justice by algorithm - the role of artificial intelligence in policing and criminal justice systems", on "Legal aspects of "autonomous" vehicles" and (in the early stages of preparation) on lethal autonomous weapons systems. The recommendations that the Assembly may adopt on the basis of these reports will thus provide important guidance for the CAHAI when mapping the risks and opportunities of AI and its impact on human rights, rule of law and democracy, and subsequently determining the need for a binding international legal framework.

6. Conclusions and recommendations

52. As with many technologies, the development of BCI technology creates both opportunities and risks. BCIs could be used to restore people's ability to move and communicate, to co-operate in the completion of tasks or to perform with greater efficiency and effectiveness; to enhance cognitive abilities by directly accessing data and harnessing supplementary computational power; or to experience novel sensory or even emotional situations. On the other hand, they could be used to bypass the rights to privacy, integrity and protection against self-incrimination, and the freedom of expression; to influence choice, behaviour and even long-term personality; or to undermine the fundamental characteristics of human equality and dignity.

53. Whilst neither scientific understanding nor technology are sufficiently advanced to produce all of these dangers, some of them are already conceivable. The examples above are all realistic, foreseeable consequences of progress that is rapidly under way. As many commentators from various perspectives have concluded, there is an immediate, urgent need to anticipate the potential risks and take regulatory action to mitigate or avoid them. As in the related field of AI, this may take the form of ethical charters, mandatory regulations, or even new rights; or, most likely, a combination of all three. BCI technology may increasingly rely on AI but it raises a separate set of concerns all of its own. The ethical principles and regulatory response required are therefore in some ways more complex, reflecting the significance of what it means for this technology to intrude into the very centre of our human being.

---

27. See further at https://nri.ntc.columbia.edu/projects.

28. "Computers Accessing Human Brain: Colombia Neuroscientist Calls for Regulation", Colombia Global Centers, 11 October 2019.

29. "Neurorights and Why We Need Them", Elena Blanco-Suarez, Psychology Today, 25 June 2020.

54. My practical and policy proposals are set out in the attached draft resolution and recommendation.

Appendix

Artificial Intelligence – description and ethical principles

There have been many attempts to define the term "artificial intelligence" since it was first used in 1955. These efforts are intensifying as standard-setting bodies, including the Council of Europe, respond to the increasing power and ubiquity of artificial intelligence by working towards its legal regulation. Nevertheless, there is still no single, universally accepted 'technical' or 'legal' definition.[30] *For the purposes of this report, however, it will be necessary to describe the concept.*

The term "artificial intelligence" (AI) is generally used nowadays to describe computer-based systems that can perceive and derive data from their environment, and then use statistical algorithms to process that data in order to produce results intended to achieve pre-determined goals. The algorithms consist of rules that may be established by human input, or set by the computer itself, which "trains" the algorithm by analysing massive datasets and continues to refine the rules as new data is received. The latter approach is known as "machine learning" (or "statistical learning") and is currently the technique most widely used for complex applications, having only become possible in recent years thanks to increases in computer processing power and the availability of sufficient data. "Deep learning" is a particularly advanced form of machine learning, using multiple layers of "artificial neural networks" to process data. The algorithms developed by these systems may not be entirely susceptible to human analysis or comprehension, which is why they are sometimes described as "black boxes" (a term that is also, but for a different reason, sometimes used to describe proprietary AI systems protected by intellectual property rights).

All current forms of AI are "narrow", meaning they are dedicated to a single, defined task. "Narrow" AI is also sometimes described as "weak", even if modern facial recognition, natural language processing, autonomous driving and medical diagnostic systems, for example, are incredibly sophisticated and perform certain complex tasks with astonishing speed and accuracy. "Artificial general intelligence", sometimes known as "strong" AI, able to perform all functions of the human brain, still lies in the future. "Artificial super-intelligence" refers to a system whose capabilities exceed those of the human brain.

As the number of areas in which artificial intelligence systems are being applied grows, spreading into fields with significant potential impact on individual rights and freedoms and on systems of democracy and the rule of law, increasing and increasingly urgent attention has been paid to the ethical dimension.

Numerous proposals have been made by a wide range of actors for sets of ethical principles that should be applied to AI systems. These proposals are rarely identical, differing both in the principles that they include and the ways in which those principles are defined. Research has shown that there is nevertheless extensive agreement on the core content of ethical principles that should be applied to AI systems, notably the following:[31]

- *Transparency.* The principle of transparency can be interpreted widely to include accessibility, explainability and explicability of an AI system, in other words the possibilities for an individual to understand how the system works and how it produces its results.
- *Justice and fairness.* This principle includes non-discrimination, impartiality, consistency and respect for diversity and plurality. It further implies the possibility for the subject of an AI system's operation to challenge the results, with the possibility of remedy and redress.
- *Responsibility.* This principle encompasses the requirement that a human being should be responsible for any decision affecting individual rights and freedoms, with defined accountability and legal liability for those decisions. This principle is thus closely related to that of justice and fairness.
- *Safety and security.* This implies that AI systems should be robust, secure against outside interference and safe against performing unintended actions, in accordance with the precautionary principle.
- *Privacy.* Whilst respect for human rights generally might be considered inherent in the principles of justice and fairness and of safety and security, the right to privacy is particularly important wherever an AI system is processing personal or private data. AI systems must therefore respect the binding

---

30. For a wide-ranging overview of attempts to define 'artificial intelligence', see *AI Watch: Defining Artificial Intelligence – Towards an operational definition and taxonomy of artificial intelligence,* Samoili S., López Cobo M., Gómez E., De Prato G., Martínez-Plumed F., and Delipetrev B., European Commission Joint Research Centre, 2020.
31. See *AI Ethics Guidelines: European and Global Perspectives,* Draft Report commissioned by the Council of Europe Ad Hoc Committee on Artificial Intelligence (CAHAI), Ienca M. and Vayena E., March 2020.

standards of the EU General Data Protection Regulation (GDPR) and the Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data of the Council of Europe (ETS No. 108 and the 'modernised' convention 108+, CETS No. 223), as applicable.

The effective implementation of ethical principles in relation to AI systems requires an 'ethics by design' approach, including a human rights impact assessment so as to ensure compliance with established standards. It is not sufficient for systems to be designed on the basis of technical standards only and for elements to be added at later stages in an attempt to evince respect for ethical principles.

The extent to which respect for these principles should be built into particular AI systems depends on the intended and foreseeable uses to which those systems may be put: the greater the potential impact on public interests and individual rights and freedoms, the more stringent the safeguards that are needed. Ethical regulation can thus be implemented in various ways, from voluntary internal charters for the least sensitive areas to binding legal standards for the most sensitive. In all cases, it should include independent oversight mechanisms, as appropriate to the level of regulation.

These core principles focus on the AI system and its immediate context. They are not intended to be exhaustive or to exclude wider ethical concerns, such as democracy (pluralistic public involvement in the preparation of ethical and regulatory standards), solidarity (recognising the differing perspectives of diverse groups) or sustainability (preserving the planetary environment).