

# Internet of Thinks

Securing the Brain Computer Interface (BCI)





# Table of Contents

- **Executive**
- **Overview**
- **BCI Technology**
- **Related Te**
- **BCI** Applic
- Security Threat
- Performane
- Regulation
- $\overline{1}$  Conclusion
- About the
- Acknowled
- Further Rea



This whitepaper attempts to provide a high-level overview of technological development in Brain Computer Interfaces (BCIs), and to anticipate the security and privacy threats that these technologies introduce. With the pervasiveness of the digital world and its routine profiling and tracking of our behaviours, arguably one of the few remaining freedoms we have is our minds and thoughts – BCIs could pose a severe threat to these sacrosanct freedoms.

The trajectory of computing research has been to bring computers ever closer, more granular, and more tightly coupled to the lives, environments, and bodies of individuals. While computing was once large mainframes in centralized locations, it has progressively become home computers, then mobile devices, and now, the Internet of Things. As computers became smaller and more interconnected, so too did their tasks. Miniaturisation and smaller microchips now offer deployment of computing capabilities in places hitherto deemed to be the stuff of science fiction, such as within the human brain.

In this paper, we explore the accelerating development of BCIs (sometimes called Neural Interfaces), and what this means for security and privacy. We begin by outlining both the historical development and the current state-of-the-art in BCIs, exploring the different types and architectures of BCIs, and citing several commercially developed brainmachine interfaces and promising lab results.

We proceed to examine the potential social impact of BCIs, and explore the associated regulatory, policy, and ethical challenges. Finally, we examine the cybersecurity and privacy challenges of Brain Computer Interfaces, by threat modelling the end-to-end lifecycle and use of BCIs, and highlighting likely areas

of attack or compromise.

It is our hope that this paper will help direct engineers toward a more privacy and security-centric design ethos as they design neural interfaces, guide security researchers toward important BCI attack paths worthy of our study, and assist legislators as they create policy for these rapidly emerging technologies.

Brain-Computer Interfaces present profound ethical, legal, and existential questions, many of which cannot be considered independently from considerations of these systems' security and privacy properties. In this paper, we introduce a new term, **'Neural Security'**  and relatedly, **'Neural Privacy'** to describe the concepts of securing neural, cognitive, and psychological information, and ensuring its' confidentiality against

threat actors.

### Executive Summary

Indeed, it is the case that you cannot make assurances about systems that you do not control – and the only way to ensure that a system is protected from adversarial compromise and control is to secure it, by design, from the very beginning.

### **Overview**

6

*[Transhumanism](https://en.wikipedia.org/wiki/Transhumanism)* is a philosophical movement regarding the next evolutionary step in humankind, where that evolutionary step includes the integration of technology with the human organic matter of ourselves.

The idea of transhumanism is that the conjoining of body, mind and technology will help enhance cognitive and/or physical abilities, allowing humans to function far beyond their current abilities. The transhumanism domain is vast, and just one subset of it concerns the Brain Computer Interface (BCI) which is the focus of this whitepaper.

BCIs provide mechanisms for monitoring and decoding activity in the brain, and/or for sending signals directly to the brain by way of stimuli that bypasses typical human sensory substrates like vision and hearing.

The technology is gaining rapid attention and investment in R&D, building on decades of neuroscience research and leveraging developments in machine learning and artificial intelligence. For perspective, BCI funding in 2021 was circa \$300 million globally - triple the amount of funding from just two years [prior.](https://venturebeat.com/2021/08/14/brain-computer-interfaces-are-making-big-progress-this-year/)

- In 2017 Facebook announced their research efforts in using BCIs to decode speech directly from the brain, such that people would be able to simply 'think' their text-based inputs much more quickly than by typing on a [keyboard](https://fbthoughts.github.io/).
- **[Neuralink](https://neuralink.com/)** by Elon Musk has been making significant breakthroughs in BCIs embedded directly in the brain. To date the research and experiments have been performed on animals (pigs and monkeys), but the intention is for the technology to develop to support people with spinal cord injury, restore motor and sensory functions and help treat neurological disorders. Neuralink also has future goals of its BCI (The Link) in enabling users to communicate with their electronic devices simply through thought.

While perhaps just five years ago this topic might have existed almost entirely in the realm of science fiction, there are already a number of technology companies researching, developing and commercialising BCIs for real-world use.

For example:

We can only begin to imagine the potential application areas of BCIs and the impacts they will have on society – how we live, work and communicate. Where two or more people possess the appropriate BCIs, we can imagine a world where they are able to communicate with each other through conceptual [telepathy](https://twitter.com/elonmusk/status/1299512615001964544?lang=en). As R&D in this field will inevitably continue, so will the power and capabilities of BCIs.

BCIs bring with them security risks to confidentiality, integrity and availability, and moreover safety risks, where they may offer mechanisms to adversely affect the operation of a person's brain activity which could result in mental manipulation, long-term brain damage or worst case loss of life.

Despite the potential benefits of BCIs, the reality is that they involve integrating technology with our brains – technology can be insecure and vulnerable to attack, which may in turn put the privacy and integrity of an individuals' brain activity at risk. Thus the threat model of BCIs needs to be carefully understood, particularly within specific use-case contexts (e.g. thinking one's password to unlock a device). We explore how BCIs might change society and thus how we need to think about their data and clinical risks in the context of specific applications and usecase scenarios. These security considerations span the entire lifecycle of a BCI, from secure design, secure and safe surgery and implant (where BCIs are invasive), secure operation and secure decommission.

They also have the potential to impact individual privacy in ways that could dramatically alter our society and freedoms.

The aim of this paper is to raise awareness of the security, privacy and safety implications of BCIs and is written as part of NCC Group's commitment to studying the security and privacy implications of early-stage emerging technologies.

We also explore aspects of software resilience in this endeavour, data protection in the context of neuroprivacy and provide a high-level taxonomy of BCI-specific threats to support future threat modelling exercises on BCIs and their applications.

This paper is by no means exhaustive, but seeks to open the security discussion and raise awareness of the potential risks and threats to BCIs and their applications, with the intention of supporting principles of security by design, thus mitigating the potential risks which may be difficult if not impossible to easily remedy once a BCI has been manufactured, implanted or deployed.

### **Brief History of Brain Imaging and BCIs**

Berger analysed EEG traces to identify wave oscillations which correlated with brain activity. EEG is a method of recording electrical activity on the scalp, which is a representation of activity of the surface layer of the underlying brain.

German psychiatrist Hans Berger was the first person to record human brain activity by means of Electroencephalography (EEG) as [early as the 1920s](https://pubmed.ncbi.nlm.nih.gov/16334737/) .

Since Berger's work, neuroscience has continued to progress with ever-improving understanding of the human brain, through a combination of neuroimaging and cognitive-psychological techniques.

Research on BCIs gained traction in the 1970s by Jacques Vidal at the University of California, Los Angeles (UCLA) under grant from the National Science Foundation and subsequent contract from **[DARPA](https://www.darpa.mil/program/our-research/darpa-and-the-brain-initiative)**. Vidal's [papers](https://ieeexplore.ieee.org/document/1454811) mark the first appearance of the term "Brain Computer Interface" in scientific literature.

As neuroimaging (including but not limited to EEG monitoring) has improved over time, much focus has been made on using this knowledge and understanding in treating people with disabilities; such as those who have lost motor-neuron function as a result of stroke or brain injury. EEG monitoring, and brain stimulation methods have been successful in improving the quality of life of many. In addition to improving physical conditions, EEG technology and applications are also under continued research for diagnosis and management of **[neurological](https://dualdiagnosis.org/testing-assessments-comorbidity/electroencephalography/)** (autism, ADHD, dementia etc.) conditions or [mental](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5461831/) (depression, addiction etc.) disorders.

The application of EEG and BCI technology to diagnosis and therapy of physical and mental illnesses, disabilities and neurological disorders is a noble and a wondrous, positive application of technology. A natural leap in this technology however is to go beyond medical applications; whereby BCIs might enhance the cognitive abilities of able-bodied people.

Controversially, animals (pigs and monkeys in the case of Neuralink) are currently serving as the test subjects for BCI technology development; arguably so are those with disabilities, even though they will have consented to trial and the outcomes for them might be significant improvements to their quality of life. At the time of writing (early 2022), a number of technology companies already exist with different technology approaches to BCIs and with different intended use-cases, from transmitting speech to text via direct thought to controlling sprites and avatars in video games, to name but a few. We delve deeper into BCI use-cases later in this whitepaper.

### **Brief Overview of the Brain**

Before diving deeper into BCI technology we provide a very brief overview of the brain to ensure familiarity with common brain terminology and functions.

Chemical and electrical signals are sent and received by the brain, throughout a body. Different signals control different parts of a body (mental and physical); there are many parts (glands and organs) in and around the brain that provide many different functions for the [body](https://www.hopkinsmedicine.org/health/conditions-and-diseases/anatomy-of-the-brain), however there are three main parts to the brain:

1. Cerebrum – this is the largest part of the brain and its function is to coordinate movement and regulate temperature. Other functions within the Cerebrum include speech, judgment, reasoning, problemsolving, learning, emotions and processing of senses such as sight, sound and touch.

• Cerebral Cortex – this is the outer gray matter around the cerebrum and is comprised of two hemispheres; the right hemisphere controls the left side of a body

- 
- and vice-versa.
- 2. The brainstem (middle of brain) links the cerebrum with the spinal cord, allowing for messaging up and down the spine for motor functions for example.
- 3. Cerebellum this is at the back of the head and its function is to coordinate voluntary muscle movements and maintain balance. Neuroscience research continues to explore the cerebellum's role in aspects of thought, emotion and social behaviour.





## BCI Technology Overview

**1. Non-invasive BCIs** are usually sensors attached to the head or through use of a helmet or exoskeleton with an array of sensors (such as EEG) connected to a person's head. Such BCIs typically just read data from the brain, with limited to no direction of input or stimuli to the brain. Non-invasive BCIs are easy to wear and don't require surgery, however because they reside outside of the brain, they cannot effectively use higher-frequency signals because the skull presents resistivity, rendering reading of EEG activity less [effective](https://www.researchgate.net/publication/8459894_Effect_of_Skull_Resistivity_on_the_Spatial_Resolutions_of_EEG_and_MEG).

There are three main types of Brain Computer Interface and these can be categorised according to their physical invasiveness upon the human body and overall proximity to the affected user's brain:

Facebook reported that their BCI research efforts were focussed on non-invasive BCIs and that they were not interested in implants, but rather non-invasive techniques such as infrared [scanning](https://tech.fb.com/imagining-a-new-interface-hands-free-communication-without-saying-a-word/) from outside the skull. While non-invasive BCI R&D is typically focussed on use of sensing as close to the skull as possible (for EEG reading), conceivably more advanced technology could be capable of picking up EEG signals from further away, presenting a myriad of privacy concerns around brain activity [emanations](https://papers.ssrn.com/sol3/Delivery.cfm/SSRN_ID2427564_code2121264.pdf?abstractid=2427564&mirid=1).

**3. Invasive BCIs** require surgery to implant electrodes underneath the scalp for communicating brain signals directly into and out of the brain. Invasive BCIs present the most accurate brain readings; however disadvantages include intrusive surgery which carries greater risk than with less invasive BCIs.

**2. Partially invasive BCIs** are implanted inside the skull but rest just outside of the brain rather than within the brain's grey matter. Because partially invasive BCIs sit closer to the brain using Electrocorticography (ECoG) [techniques](https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.665.6280&rep=rep1&type=pdf), they produce better resolution signals than

Note that BCI users (or bearers) may not be subject to just one BCI implant. Because implants interface with specific neuroanatomic regions, conceivably there might be a need for multiple implants reading and stimulating different parts of the brain which correlate to different functions and aspects of brain activity. Neuralink has already successfully demonstrated two BCIs implanted in three of its pig test [subjects](https://www.reuters.com/article/us-tech-neuralink-musk-idUSKBN25O2EG)

non-invasive BCIs, and their position on the brain has lower risk of forming scar-tissue in the brain than fully invasive BCIs. Operationally they are less risky than implanting directly into the brain.

Back in 2006, researchers at Washington University in St. Louis successfully showed use of a partially invasive BCI in a young boy playing the Space Invaders video game, leveraging the signals from his brain to make the necessary movements on [screen](https://source.wustl.edu/2006/10/teenager-moves-video-icons-just-by-imagination/ ).

Invasive surgery could result in scar tissue forming on the brain which could lead to health-related issues such as [seizures](https://www.frontiersin.org/articles/10.3389/fnins.2016.00295/full).

Neuralink's BCI is invasive and termed 'The Link', which is the size of a coin and consists of very small probes containing more than 3,000 electrodes attached to flexible threads that are thinner than a human hair, which can then monitor the activity of 1,024 neurons.

Data is transmitted from and to The Link via Bluetooth. The Link retains power through periodic inductive charging. Brown University's BrainGate research group were the first to create wireless implanted and inductively charged BCIs back in

[2013](https://www.brown.edu/news/2021-03-31/braingate-wireless).

#### **Multiple BCI Implants**

#### **BCI Host/Near-Control Device (NCD)**

Most BCIs will communicate with an external host or Near-Control Device (NCD) – the NCD by virtue of typically being a more powerful computer will perform functions such as brain signal decoding, analysis, and potential relay of stimuli back to the BCI.

NCDs could be simple apps on a smartphone or more powerful computers such as laptops or desktops. The NCD is key in bridging inputs to and outputs from the brain with external computing and processing resources. In applications where multiple people with BCIs might interact in a BCI-based application, the NCD will serve as a form of proxy or router between all associated interacting BCIs.

and the contract of Thinks internet of Thinks 11 Internet of Thinks 11 Internet of Thinks internet of Thinks internet of Thinks

#### **BCI Communication Transport**

For BCIs that communicate, a transport mechanism is needed. Invasive BCIs will typically need to be of small form factor given their presence within the brain, thus commonly wireless transports and chipsets will be used for this purpose such as [Bluetooth Low Energy \(BLE](https://neuralink.com/approach/)).

Non-invasive BCIs will have greater options for communicating with NCDs, whether wired (connection from a helmet) or wireless but with higher bandwidth

protocols such as Wi-Fi.

#### **BCI Hardware**

#### **1. Unattended, autonomous operation** – this is where a BCI

Particularly for invasive and partially invasive BCIs, hardware plays a key role in their operation and pushes demand on miniaturisation to great lengths. BCI implants need to be small enough so as not to take up too much room inside a person's head and to be able to interface with a person's brain with great precision, while still presenting a number of functions and features for operation such as a CPU, bootloader, memory/storage, inductive charging and wireless interface such as Bluetooth. In essence these devices and their components closely resemble IoT devices, meaning that the conventional security issues with – and attacks against – IoT devices are equally applicable to BCI implant hardware.

 $12$  12  $\mu$  13  $\mu$  14  $\mu$  15  $\mu$  **only –** such a BCI would only transmit brain activity and data out of the brain, to a receiving Near-Computing Device (NCD). It wouldn't receive any incoming data or stimulus aimed at causing some level of change within the brain.

#### **Inbound, Outbound, & Bidirectional Brain-Computer Communications**

BCIs may communicate with the outside world in one of four main ways:

may be programmed with a specific task in terms of the brain activity it reads, and the electronic stimulus it produces based on pre-determined conditions. Apart from the need for routine charging of such a BCI, it wouldn't typically communicate outside of the brain, or receive inputs from outside of the brain – it would just operate uninterrupted based on its pre-programmed directive. An example might include a BCI detecting emerging patterns of seizure and blocking the seizure by generating counter electronic signals directly within the brain

#### **2. Outbound communication**

An example might be a gaming application which interprets signals from the visual cortex of a person and sends this to a computer which then controls elements on a visual display, such as moving a cursor or a video game character.

#### **3. Inbound communication only**

– this is where a BCI wouldn't transmit any data from the brain, but would receive incoming communication and stimuli. An example here could include a BCI application whereby a person can trigger a stimulus (e.g. via a connected app) based on their mood – perhaps when they feel depressed, then can trigger activity in the correct part of the brain that might alleviate anxiety or stress.

**4. Bidirectional** – such BCIs send data from the brain, and receive input to the brain. For example, Neuralink is very much focussed on bidirectional coupling whereby computers will receive people's brain activity and insert information into their neural circuitry by way of [response](https://pubmed.ncbi.nlm.nih.gov/32164851/) .

### Related **Technolgies** & Fields of Research

There are a number of complementary technologies and interfaces required to help realise and operate BCI applications, as well as other related fields which may inform or benefit from the development of neural interfaces which are commonly confused with BCIs.

A non-exhaustive list of complementary technologies, related disciplines, as well as common confounds, is presented in this section to provide a map of the territory in which BCIs are developed, as well as to disambiguate from similar terminology in other subfields.

#### **Artificial Intelligence & Machine Learning**

A key part of modern BCI technology and its development is the application of AI/ML to the field of neuroscience. On the surface, a lot of BCI applications can appear like magic – e.g. a monkey playing the video game pong through thought [alone](https://www.youtube.com/watch?v=rsCul1sp4hQ). However, the underlying technology isn't reading and understanding the exact thoughts of the monkey in these examples (e.g. "move the bat upwards"). Rather, AI/ML has been used to learn the brain activity associated with the Monkey's different movements of a joystick – over time the brain activity associated with 'move up' or 'move down' becomes learned and predictable, to the point that the joystick can eventually be removed and the monkey can simply think it's control of the game – in the background however, an AI/ML model is constantly classifying and predicting the monkey's neural activity to determine how and where to move the bat on the screen.

communication abilities in those with speech <u>[impairments](https://jneuroengrehab.biomedcentral.com/articles/10.1186/s12984-015-0071-z).</u> a Monero <del>mineral downloaded</del>downloadedd a brif a gweled a brif a gweled a brif a gweled a brif a gweled a brif a<br>Daeth a brif a gweled a brif • Eye-Tracking – some BCI applications relating to analysing user interaction with visual displays may utilise eye-tracking technology to support brain activity monitoring and decoding of where people are looking, to potentially increase

• Speech or Voice recognition - BCI applications may need to process speech or authenticate subjects from their [voice.](https://link.springer.com/content/pdf/10.1007/s13311-018-00692-2.pdf) to an IP in Russia. The file in the file in the file in question of the file in  $\mathcal{I}_1$ 

• Augmented Reality (AR) and Virtual Reality (VR) will likely be used to provide users with greater senses of immersion within (e.g. gaming virtual worlds or online forums). Coupled with BCIs, various new immersive experiences will likely arise through the combination of these [technologies](https://www.frontiersin.org/articles/10.3389/fnhum.2020.00144/full).

An important point to note here on the use of AI/ML is that because of the need for learning brain behaviour, it is not a given that training on one subject, will naturally translate to others. Brain activity may differ across different subjects; neurons might fire in different ways or with different speeds, or brain damage might mean certain brain functions aren't as optimal as they could be. While there will be an average commonality of brain activity across populations, ultimately our brains and their activities are unique (to the point of being a potential future biometric [identifier](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6011092/pdf/CIN2018-5483921.pdf)). This means that the performance of future COTS-based BCI products will likely differ per person, unless there is some element of bespoke training and customisation to the individuality and uniqueness of people's neural activity.

#### **Interface Types**

A number of technologies exist that help to provide an interface between neural activity and the outside world, and which may be application-specific.

#### Examples include:

- 
- 
- 

- [Wireless Technologies \(Bluetooth,](https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9390339)  [Wi-Fi, Zigbee etc.\)](https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9390339) – Most BCIs (whether invasive or non-invasive) will need to communicate with a NCD; the most convenient method for this will be wireless, particularly for invasive BCIs.
- Muscle-Sensing Pads, Exoskeletons & Neuroprosthetics – these components may attach to various parts of the body, and may also utilise wireless technologies for transmitting data to, and receiving data from, a [BCI](https://www.nature.com/articles/s41598-021-85134-4.pdf).



[Figure 2 - Key components in a BCI system architecture](https://cliparting.com/free-brain-clipart-4914/ ) 

#### **Neuroprosthetic Devices**

In neuroscience, neuroprosthetics is the use of artificial devices to replace the function of impaired nervous systems, brain-related problems, organs etc. Cochlear implants are an example of a neuroprosthetic, which electronically stimulate the cochlear nerve to improve hearing for those with hearing loss or impairment. Pacemakers are also a commonly-known type of neuroprosthetic device. We briefly mention neuroprosthetic devices here since they share common goals with BCIs and the terms can sometimes be used interchangeably, even though they not always or typically installed or located in or near the brain.

#### **Neurofeedback**

[Neurofeedback](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4892319/) (sometimes also referred to as neurotherapy) is a method used to assist subjects in exerting control over their own brain waves. During neurofeedback sessions, brain activity (often in the form of EEG) is recorded, and then extracted, decoded and a stimulus is fed back to a subject using real-time feedback in the form of video and/or audio. Neurofeedback typically involves non-invasive techniques, and whilst the effectiveness of the therapy has remained controversial in the medical [mainstream](https://digitalcommons.chapman.edu/cgi/viewcontent.cgi?article=1158&context=psychology_articles), permission for marketing a neurofeedback device for treatment of [Attention Deficit Hyperactivity Disorder](https://www.fda.gov/news-events/press-announcements/fda-permits-marketing-first-medical-device-treatment-adhd)  [\(ADHD\) was granted in 2019](https://www.fda.gov/news-events/press-announcements/fda-permits-marketing-first-medical-device-treatment-adhd), and research papers have shown some efficacy in the use of neurofeedback to treat anxiety [disorders](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8581142/) and post-traumatic stress [disorder \(PTSD\)](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8544423/).

neuromorphic computing is to develop new computing architectures that are performance-optimized for specific computing tasks over existing CPUs,  $GPUs,$  and other chips.  $\Box$ Neuromorphic computing involves the use of Very-Large-Scale Integration (VLSI) systems comprised of analogue circuits that mimic neuro-biological systems that are present in the human nervous [system](https://en.wikipedia.org/wiki/Neuromorphic_engineering). A neuromorphic computer or chip uses its artificial neuron composition to perform computations. The intention in

While these areas of research may overlap in terms of associated subfields, they are separate and generally non-converging lines of inquiry.

The technologies used for neurofeedback could be considered a type of non-invasive BCI, and it is possible that innovations in neural interfaces could improve the efficacy and precision of neurofeedback techniques in the future.

#### **Neuromorphic Computing**

triggered on a suspicious file While Neural Interfaces (BCIs) attempt to augment the human brain – or interface between it and external hardware – through the introduction of additional sensing and compute proximal to the brain, Neuromorphic Computing attempts to build brain-inspired computer architectures that are generally not intended to interface directly with the

human brain itself.

However, some researchers posit that developments in neuromorphic computing may allow for easier integration of invasive BCIs into existing neural pathways. In 2016, researchers showed successful development of a modular bi-directional BCI that used a compact neuromorphic processor as a [decoder](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5145890/pdf/fnins-10-00563.pdf) – this can help remove the need for external NCDs in the decoding of neural activity and thus create a more concise, closed-loop system attached directly to the brain.

#### **Computational Neuroscience**

Computational Neuroscience involves use of mathematical models, theoretical analysis and abstractions (simulations) of the brain to understand properties of brain development, structure, physiology and overall cognitive [abilities](https://en.wikipedia.org/wiki/Computational_neuroscience) .

We briefly mention computational neuroscience here due to the key role that the discipline will have on continued development of BCI technologies and applications, as a result of the continued research insights and developments in this field.

### BCI Applications, Industry & Societal Impact

Alleviating Physical **Disabilities** 

Alleviating Men Illness

| Health Monitor

Interfaces with Metaverse & \ **Worlds** 



BCIs are set to have significant impact on how we live and work – while some of these changes could improve quality-of-life and potentially assist individuals to help improve their lived experience with certain health issues and disabilities, there are also a number of dystopian visions that could be realized. In the near-term, it's clearer what types of application will become available and what they will offer.

Longer-term, there are many potential radical applications that might arise but are less clear at the time of writing, yet their developments will undoubtedly occur through advances in neuroscience and AI/ML and with unfettered imagination.

In this table we present just some of the potential BCI applications and their respective impacts on industries and society – however, we make no estimate on the likelihood of those applications actually being realised, nor do we necessarily endorse their creation or use.



### **Productivity & Cognitive Enhancement**









### Security Threat Lifecycle for Brain-Computer Interfaces

The volume of security threats to a BCI is not limited to attacks on actively operational implanted devices – as must be appreciated for all connected devices, security risks and mitigations spans the entire lifecycle of the BCI from design, implementation, surgical implantation, operation and potential future removal or decommissioning.

While throughout the BCI lifecycle the usual technology and infrastructurerelated security threats apply (computer, smartphone, cloud platform security etc.), in this section we will focus primarily on threats that are specific to neural interfaces and which may be more likely overlooked through traditional [threat](https://arxiv.org/pdf/1908.03536.pdf)  [modelling](https://arxiv.org/pdf/1908.03536.pdf).

In this section, we explore the endto-end threat and attack lifecycle for neural interfaces, broken down into 3 phrases: (1) Pre-Implantation Security, (2) In Vivo, Architectural, or Operational Security, and (3) Post-Operational or Post-Decommissioning Security.

Secure Software Development Lifecycles (SDLCs) are therefore very important for BCI device and BCI application development.

#### **1. Pre-Implantation Security**

Before a BCI device comes near a consumer, there are a number of threats and risk mitigations that should occur to preserve the safety and operational integrity of the device. The safety and regulatory issues unrelated to cybersecurity are beyond the scope of this paper, however the security-relevant concerns include designing the BCI with security in mind, ensuring the hardware and the development supply chain for the device and its firmware/software are resistant to compromise, and ensuring secure implantation and deployment of the device.

#### **1.1. Security by Design & the SDLC**

There is much software involved in a BCI system – it exists in the BCI itself, NCD apps and applications, surgical equipment and robots, AI/ML models, web and cloud services to name but a few.

It is important that users, regulators, device manufacturers and others understand the breadth of scope wherein software errors could present security vulnerabilities.

Guidance does exist for secure medical device development, such as the European Commission's Medical Device Coordination Group (MDCG) [guidance](https://ec.europa.eu/docsroom/documents/41863) and Principles and Practices for [Medical Device Cybersecurity](http://www.imdrf.org/docs/imdrf/final/consultations/imdrf-cons-ppmdc.pdf)  [by the International Medical Device](http://www.imdrf.org/docs/imdrf/final/consultations/imdrf-cons-ppmdc.pdf)  [Regulators Forum \(IMDRF](http://www.imdrf.org/docs/imdrf/final/consultations/imdrf-cons-ppmdc.pdf)) – however, the perspective of much of this guidance is that of external medical devices; some guidance exists for neuroprosthetics, but at time of writing, there is nothing specific to the secure development of BCIs.

#### **1.2. Supply Chain Security**

The production of hardware devices involves multiple suppliers at various stages of the production and support lifecycle. There is rarely an electronics manufacturer who manufactures every single component of a device in their own factory. As such, and has been demonstrated, these hardware and manufacturing supply chains introduce risk that threat actors could gain an opportunity to defraud, steal, or otherwise undermine the security and safety of the electronic devices

[produced](https://research.nccgroup.com/wp-content/uploads/2020/07/secure-device-manufacturing-supply-chain-security-resilience-whitepaper.pdf).

Further down the supply chain, those writing firmware and software for a device such as a neural implant often rely upon existing proprietary and/ or open source operating systems, libraries, and other components, all of which have their own security risks that need to be understood and

mitigated.

25

BCI manufacturers need to be diligent in their supply chain security, particularly due to the potential healthrelated safety risks intrinsic to BCIs. Due diligence therefore includes both an understanding of your own (and your suppliers') third-party risks and component procurement practices, as well as active security evaluation of upstream components which may include code review, code scanning, penetration testing, and protocol security and security architecture analysis. The transit of devices is also an integral component of securing this supply chain as components move between manufacturers, for even if a supplier satisfied all due diligence checks, it is conceivable that an adversary may seek to manipulate or modify components while in transit between places of manufacture. Use of trusted couriers (possibly with live goods tracking) is also important in this endeavour.

Supply chain threats concerning BCIs relate to the BCI devices themselves (whether invasive or non-invasive) and any accompanying host or host applications, so understanding what those components are, what their security risks may be, and having an inventory prepared so that you may more quickly identify when you are affected by a vulnerability in an upstream component is essential on the behalf of a BCI device manufacturer.

Overall, the main threat relating to BCI supply chain is the ability for adversaries to backdoor surgical equipment or the BCI devices themselves. This therefore calls for all BCI manufacturers to follow strict process and governance around supply chains, and to ensure the integrity of hardware and software throughout the development lifecycle.

#### **2.2. Novel Attack Classes for Neural Interfaces**

- Once a BCI is implanted or deployed and mutually authenticated, it is ready for operational use.
- Novel attack classes for neural interfaces could include:

**Mind Reading** – this is where attackers may attempt to gain unauthorised access to observe someone's brain activity, with the intent to extract information about things like their affective state, tacit knowledge (e.g. movement patterns or embodied skills), or explicit knowledge (e.g. passwords). The part of the target infrastructure that could be attacked to obtain this information varies widely, but could include attacks upon specific implanted sensors, wireless broadcasts, and interpreted data on host devices. Attack techniques could also vary widely, but would ultimately seek to influence either of the following:

- Tacit Knowledge attacks Seeking to extract, manipulate, or use information a person knows which is hard to codify, such as physical skills or behavioural patterns
- **Explicit Knowledge attacks** – Seeking to extract, manipulate, or use information a person explicitly knows and remembers, such as a password, personallyidentifying, or biographical information
- **Brain Control** (also termed [Brainjacking](https://www.sciencedirect.com/science/article/abs/pii/S1878875016302728)) – this is where adversaries would seek to make someone think, feel, and/or do something beyond their free will, or even the more speculative scenario in which an attacker seeks to – contingent upon a given BCI's capabilities - make use of the brain as a computational substrate for computational tasks of the attacker's choosing (e.g. a botnet composed of multiple compromised BCIs). The level of control could be broken down into the following types (which are themselves highly contingent upon the neurological substrates with which a given BCI interfaces, as well as the actual method of action of the BCI itself):
	- Movement Control – Interacting with an individual's motor cortex to induce a specific physical action by a person (e.g. move their limbs) beyond their free will

These attacks could seek to overcome traditional user authentication methods to make use of a human user as a mere pivot point in attacking larger systems to which they can authenticate; else they may be performed by adversaries for non-security reasons to do things like obtain private data about an individual, understand beliefs or attitudes of a group to enhance misinformation or political influencing, or as a dark new avenue for advertising technology, as just a few examples.

Users of BCIs may in many cases become reliant and dependent upon them, whether they are used to mitigate the negative effects of health issues like anxiety or epilepsy, or whether they are used for overall improved cognitive function or experience above an individual's normal baseline. This could be problematic in cases where a BCI manufacturer and system maintainer stops operating.

Such scenarios have sadly already occurred affecting real patients – in one example, a patient who had received an implanted BCI to relieve her from seizures had to have it removed because the company which had implanted it became [bankrupt](https://www.nature.com/articles/d41586-019-02214-2).

Once a device is deployed to (often surgically implanted into) a user, a range of new security risks arise which must be proactively mitigated.

**Mutual Authentication between** a BCI device and host or other platform to which it communicates is essential to prevent unauthorised parties from observing, modifying, or stealing sensitive information. For example, we might ask, how does a BCI and an authorised host perform mutual authentication and establish trust? What's stopping an attacker with the same app, attempting to pair with a BCI? Similarly, if a person loses their BCI host smartphone or other external device receiving brain telemetry, how do they reconnect, re-key any cryptography and authenticate to their BCI with a new device in a trustworthy way? Do BCI users receive some sort of notification (e.g. out of band) when an unknown or unauthorised device attempts connection to their BCI?

#### **1.3. Ensuring Software Resilience for Critical Healthcare Devices**

There are therefore many threats to continued operation and support in this domain, the effect of which could significantly regress the quality of people's lives.

This raises the importance of system and software escrow agreements with trusted third parties to ensure continued support in the event that a BCI manufacturer goes bankrupt or is for catastrophic reasons unable to continue their normal operations.

#### **2. In Vivo, Architectural, or Operational Threats to Neural Security**

These include ensuring robust mutual authentication, as well as a number of neural interface-specific attack types including the exfiltration of tacit or explicit knowledge, the control of movement, thoughts, or emotions, denial-of-service attacks on brain regions or BCI functions, or the impairment or erasure of segments of human memory. There are also a number of more traditional security attack avenues that must be mitigated, which we also discuss below.

#### **2.1. Mutual Authentication**

- **Emotion Control Interacting** with an individual's limbic system (including the hypothalamus and amygdala) to either artificially and directly invoke a specific emotion that is not their actual and natural current emotional state (e.g. invoke fear or paranoia in a victim), else to interact with an individual's perception pathways to induce a hallucination or synthetic experience which could trigger a specific emotional state
- Thought Control Interacting with an individual's higher cortical regions to induce specific thoughts or beliefs which deviate from an individual's natural thoughts and beliefs
- **Denial of Neurological** Function– Interacting with a specific neuroanatomical sub region to block specific functions of the brain (e.g. a Denial of Service) or temporarily denying

BCI operation such as in a ransomware scenario (Brainsomware), where a compromised BCI, BCI Near-Control Device host, or application is held to ransom

**• Memory Impairment** – Interacting with the neurological substrates of memory formation (particularly the hippocampus) to interrupt the creation of long term memories, or to attempt to compromise the integrity of memories (i.e. to create false memories) or to delete existing memories While these scenarios may at first glance seem exaggerated or unreasonable, history readily demonstrates both that attackers will opportunistically attack any sort of connected device with exploitable flaws, and furthermore that threat actors are keen to strategically target critical infrastructure and other targets in which denial or service, data exfiltration, or adversarial control would be ideologically impactful. As a consequence, we argue that it would be naïve to assume that these types of devices would not be a desirable target for attackers.

• What type of encryption is used to protect data in transit, and how are cryptographic keys and certificates managed (symmetric or asymmetric cryptography)?

#### **2.3. Securing Neural Interface I/O**

**• Broadcast Range** – what is the range of wireless broadcast, and can that be minimised/localised to as small a range as is strictly necessary for normal operation, so as to reduce the scope for

- On a related note, what broadcasts does a BCI perform during pairing attempts, and what information is in those broadcasts that might leak personal or sensitive information (e.g. BCI serial number, person's name etc.)?
- **• Vulnerabilities in the wireless stack (e.g. Bluetooth)** – what are the implications of a 0-day vulnerability being discovered in the Bluetooth stack and actively exploited? On the BCI-side, could such an exploit lead to remote code execution directly within the BCI? Recent Denial of Service vulnerabilities affecting several Bluetooth-enabled devices **(BRAKTOOTH)** show the potential negative impact such a vulnerability could have against vulnerable BCIs, allowing attackers the ability to remotely disrupt BCI [operation](https://asset-group.github.io/disclosures/braktooth/).

Most BCIs will need to communicate data into and/or out from the brain; as such this requires interfaces and transport mechanisms. Non-invasive BCIs may use wired connections, since these sit on top of the head and thus can easily connect to a host machine via a mechanism such as USB. Invasive BCIs will need to use wireless transports, since they sit inside of the skull. Most commercial non-invasive BCIs are equally likely to use wireless communication for ease and comfort  $-$  e.g. in gaming applications which might require head movements with VR headsets, a physical connection to a host machine would be less practical. Protocols such as Bluetooth Low Energy (BLE) are likely to be common amongst BCIs, owing to their low power, near-field operation and fairly decent data rates at around 2Mb/s. Neuralink's current experiments with animals show effective operation of invasive BCIs, paired with an app on a smartphone via Bluetooth. There is much to consider around the BCI interface and communications protocol security. Regarding potential threats, things to consider include:

#### **• Transport Layer Security** –

are communications between BCI and other devices properly secured at the transport layer to ensure confidentiality and integrity? Researchers have already shown BLE man-in-the-middle attacks against the Emotiv Insight commercial BCI. Their attacks allowed them to intercept and modify information, force the BCI to perform unwanted tasks and conduct replay attacks affecting the overall security of the [BCI](http://orapp.aut.ac.nz/bitstream/handle/10292/11449/SundararajanK.pdf).

• What protections are in place to mitigate replay attacks?

- 
- 
- interception?

#### **Radio Jamming (Brainjamming) & Interference** – attackers could disrupt BCI operation through wireless jamming or interference attacks. Jamming can be very difficult to block, and the implications for any health-related BCI applications could be severe. The small form factor and power capabilities of embedded BCIs means anti-jamming detection mechanism are not likely feasible within the BCI itself; however connecting hosts with more power and capabilities may be able to implement some form of jamming detection, alerting and/or [anti](https://scholarworks.montana.edu/xmlui/bitstream/handle/1/9021/jamming.pdf)[jamming](https://scholarworks.montana.edu/xmlui/bitstream/handle/1/9021/jamming.pdf). As in other fields, there may also be regulatory solutions to prevent at least some of these attacks.

**• Brain-to-Brain (BtB) Communication** – future BCI applications may allow for conceptual telepathy, allowing two or more users in a network of BCIs to communicate with each other through thought. Different network topologies for realising such functionality will demand different types of security control – i.e. such a capability could be realised through mesh computing and P2P networks connecting BCIs directly with each other, or a hub and spoke model whereby a central host manages the routing and relay of messages between different BCIs – this latter topology would of course present a single point of failure and thus might present demands for redundancy in such applications in order to maintain availability. Depending upon the design of BtB communication protocols, input validation will be important to prevent adversarial attack from untrusted peers.

#### **2.4. Securing Brain APIs: Know that Brain Data Will Become More Sensitive With Time**

Brain Computer Interface operation essentially involves obtaining readings from the brain (output – which could include patterns of brain wave activity, or other types of neuroimaging), and in some applications, writing information to the brain (input – which could include various ways of introducing stimulus to the target brain regions). APIs to support such tasks will provide for consistent methods of communication between BCI and NCD, and also allow for easier ways of extending those APIs with new features and functions. It is anticipated that Brain APIs will develop and hopefully standardise over time, both from a commercial and open-source perspective. BciPy is just one example of recent research and attempts at development of a generic BCI interface in [Python](https://github.com/CAMBI-tech/BciPy).

Vulnerabilities in Brain APIs could allow for a number or security threats such as unauthorised access to brain activity, or access to sensitive brain data.

As such, Brain API developers should ensure secure API implementation, by following guidance such as OWASP's API Security Project [Guidance](https://owasp.org/www-project-api-security/), and keep in mind that observable telemetry that may seem innocuous at the time of development may be demonstrated through further neuroscientific research to in fact reveal something specific and sensitive about the individual in which the activity is observed. Therefore, on an ongoing basis, it will be important for an interdisciplinary team of neuroscientists and computer scientists to translate academic brain research into an understanding of what kind of data can be produced, measured, modified, or extracted from the brain, not only to push forward technological advancement in BCI development itself, but also to mitigate the risks of exposing data about an individual which may become more sensitive as neuroscience advances.

We also mentioned earlier the importance of needing to mutually authenticate and authorise NCDs with their corresponding BCIs – being able to establish trust between a BCI and an NCD, and being able to detect and block unauthorised pairing attempts is crucial. It's likely that an NCD will be Internet-connected, whether through an app on a smartphone or a running application on a desktop or laptop computer.

This is a critical point which should not be underestimated: simply because a certain piece of brain telemetry is not meaningful or sensitive now does not mean that it will be this way forever.

Indeed, humanity still knows relatively little about how the brain works, but advances in neuroimaging and cognitive psychology are rapidly advancing the state of the art. This means that we should assume that on average, a given sample of brain telemetry will become more meaningful and sensitive as time progresses. Furthermore, we may also find that some of these patterns – particularly those which are unique to individuals in ways not currently appreciated – could be like other irreversible biometrics, such as fingerprints, in that they are both uniquely identifying, and also impossible to update. This type of data must be protected with utmost care.

### **2.5. BCI Host / Near-Control Device (NCD) Security**

Most BCIs will communicate with a host or Near-Control Device (NCD) – the NCD, by virtue of typically being a more powerful computer, will perform functions such as brain signal decoding, analysis, and potential relay of stimuli back to the BCI.

The NCD security is therefore paramount to the overall security of the BCI and its operation, since any unauthorised access to the NCD could provide for a number of different attacks against the overall system, including:

- Adversarial AI/ML attacks as mentioned earlier
- Manipulation of Bluetooth for interception/replay etc.
- Denial of Service or ransomware
- Brain control by sending specific stimuli to the BCD

This opens up the NCD's susceptibility to all manner of attacks and attempted compromise possible by remote attackers, via any flaw at any level of the technology stack or any phase of the supply chain.

The sobering realisation concerning Internet-connected NCDs is the fact that BCIs (and thus brains) essentially become part of the Internet, potentially accessible and routable from any other host on the Internet (or within our playful nomenclature here, the *Internet of Thinks*).

Other security considerations relating to NCDs include app and associated app store security  $-$  the consequences of a user being tricked into downloading and installing a fake or malicious BCI app could be severe and provide methods for various attacks against and intrusions into a person's BCI.

#### **2.6. Preventing Adversarial Input to Brain-Computer Interfaces**

Like most areas of computing, the application of machine learning to activities mediated by BCIs is a likely force-multiplier of their usefulness and impact. As such, the BCI ecosystem will comprise multiple AI/ML models; both generic to COTS BCI offerings and unique in the sense that a BCI user will have needed to train a model specific to their own brain's patterns of activity, in order to maximise the accuracy and performance of the BCI and its applications.

Research has already been performed in the adversarial AI space, showing actual attacks against existing BCI applications. One study has shown the use of adversarial perturbations against a BCI Speller (a BCI application that allows use of brain activity to output letters and words – specifically for users who have lost speech function), which are small enough not to be noticed or detected but that can mislead BCI spellers to spell arbitrary text as desired by an [attacker](https://arxiv.org/pdf/2001.11569.pdf).

A plethora of AI/ML-based adversarial attacks against BCIs are likely to surface in the near term, warranting an immediate need for research on how such attacks might be mitigated. Example attacks we might expect to see include model inference, model extraction and model poisoning to name but a few. However, depending upon the communications methods of the devices and the attacker privileges required, much of the impact of adversarial attacks can be potentially mitigated through thoughtful security-bydesign in the creation of BCIs – indeed, this is one of our primary motivations in writing this whitepaper.

#### **2.7. Secure Software/Firmware Updates**

As with most applications and hardware appliances, there may be occasional need for software or firmware updates to address any flaws or security vulnerabilities, or to enhance features and functionality. For non-invasive BCIs such activity is straightforward however complexities arise for the update of invasive BCIs, given their implantation in the body.

In the context of BCIs there is also the need to consider updates to the security of software and firmware of any surgical or robot implant equipment. Examples of things to consider include:

- What security is in place for the propagation of software and firmware updates (transport security, integrity checking, failsafe returns to known good state, etc.)?
- Where are software and firmware images stored, and how secure is that storage? (e.g. an attacker compromising an update server could potentially backdoor software versions and firmware images, which would affect all devices installing those updates)
- How and when are updates applied? For invasive BCIs, particularly those performing a health-related function, any downtime or error in the update process could be catastrophic, so the timing and initiation of patch application must have safeguards to mitigate operational risks

• For how long has the BCI manufacturer committed to providing security and other updates to the device, and what will happen once the device reaches end-of-life (EOL) and is no longer patched by the vendor?

### **2.8. Resisting Attacks on Inductive Device-Charging**

Invasive and partially-invasive BCIs will need power for operation, but will lack an ability to physically connect to a power source. As such, the only way to charge their batteries will be to use wireless or inductive charging. For example, Neuralink postulates that future BCI users may charge their BCIs overnight while they sleep, through use of inductive charging. A number of potential threats exist to BCI users in the domain of inductive charging. Firstly, the general dependence on recharging could be an issue for those who experience low to no BCI power, but lack immediate access to a charging source – by default, this would halt the operation of the BCI until such time it could be recharged. Proactively, attackers might seek to deliberately drain the power source of a BCI through repeated communication with it, reducing its performance or rendering it

inaccessible.

Because of the proximity to the brain, there are serious safety implications concerning power transmission to invasive and partially-invasive BCIs. With the design of an inductive charging setup there are a few options.

Firstly, there is usually a coil of wire and passive components to turn AC current into something for DC charging; this presents inherent dangers. The output voltage of a transformer in simple terms is A PTC is like a resettable fuse - as it the ratio of turns between the primary and the secondary. If the secondary (the side inside the BCI) has more winds of coil than the primary it will step the voltage. Likewise the inverse is true. However if the primary is say a coil resonating at the correct frequency and at 1,000 volts, the secondary components need to be able to withstand that voltage. In such a scenario there would be the need for clamping components. These components (like a TVS or Zener diode) are used to stop the voltage going above a certain point by clamping. The energy voltage currently clamped is turned into heat, which presents a problem. If the voltage is clamped (which is the easiest option), heat is generated.

The clamping components can dissipate *X* watts of heat for a period of time before they burn out. Ideally a clamping component will outlast the coil so in extreme cases the coil might burn out, yet there might still be enough heat to melt the [coil](https://techtellectual.com/fix-wireless-charger-overheating/).

Another design option would be to use a Positive Temperature Co-efficient resistor (PTC) and a clamping component. heats up the resistance goes higher and the current drops. As it cools off it can recover. This is all fine providing the output voltage of the coil used for charging is within the maximum voltage of the components. If the voltage goes too high the coil will arch and likely short circuit in which case the clamping components stop being effective. Eddy currents can still be generated in a shorted loop, the energy of which will be entirely converted to heat.

Yet another option is to isolate both ends of the coil, however again the voltage issue arises. All components have a maximum voltage and exceeding this could cause components to explode.

To implement these design options within a BCI means that components are tiny, arching distances are very short and charging is likely to be a few volts. The safety aspect of BCIs is likely to be centred around standard operation, and not considerate of handling deliberate misuse cases. Adversaries seeking to attack the inductive charging process might therefore aim to trigger [thermal](https://www.batterypoweronline.com/news/thermal-runaway-understanding-the-fundamentals-to-ensure-safer-batteries/)  [runaway](https://www.batterypoweronline.com/news/thermal-runaway-understanding-the-fundamentals-to-ensure-safer-batteries/), the consequence of which could severely damage the brain or result in loss of life.

Another potential threat to inductive charging is the use of communication protocols which can operate over the same medium. These protocols are used for the BCI to communicate things like how much power it needs or its current charge state. Software stack vulnerabilities relating to such communication could result in unauthorised access to, or corruption of, the [BCI](https://www.avnet.com/wps/portal/us/resources/article/watch-for-these-2-problems-in-your-qi-wireless-charging-project/).

### **2.9. BCI Forensics**

Regarding the operational lifecycle of BCIs, the topic of BCI forensics is interesting in a security context.

We can imagine a few scenarios whereby a BCI (invasive or otherwise) might become physically-accessible to adversaries or criminal investigators, and how there may be value in performing forensics on that BCI in attempts at gaining potentially useful information:

• Non-invasive BCIs may equally contain interesting and useful forensic information – while not needing surgery for access, conceivably they might contain data caches and identifiers that could be useful to attackers and/or law enforcement in a forensics capacity

Broadcasts from BCIs of missing persons (deceased or alive) also fall into the forensics domain for law enforcement.

- Surgical removal (authorised) an invasive BCI might be surgically removed for a number of potential reasons, such as malfunction, user consent withdrawal etc. If the BCI is not securely destroyed, conceivably there may be useful information on it that could be forensically recovered
- Surgical removal (unauthorised) in sinister scenarios, adversaries may seek to remove a person's invasive BCI without their consent (e.g. under duress/intoxication etc.) and extract potentially useful data from it
- Surgical removal (post mortem) a deceased person with a BCI may have it removed officially as part of a post mortem, e.g. to support a criminal investigation concerning the person's cause of death, as it may contain data or clues relating to the cause of death

In addition to BCIs, NCD forensics will also be of interest to adversaries, law enforcement and cybersecurity incident responders, and will likely contain even greater data caches and information concerning the affiliated BCI's operation. This also extends to forensics at the app layer (e.g. connecting BCI smartphone app), and any Internet-connected cloudbased services that perhaps consume and/or process data received from BCIs.

Invasive BCIs are unlikely to contain much storage and memory owing to their small form factors, yet they might still offer interesting information from forensic data recovery, such as serial numbers and personal identifiers. Likely there will exist registries of BCI serial numbers linked to specific identities, thus showing the value of being able to match any such identifiers with real identities (for attackers or law enforcement). Other potentially interesting data in the cache of a BCI could include logs of connection attempts with other NCDs, or caches of neural activity readings that could attest to a person's last mental state before BCI removal or power drain.

We can imagine scenarios where people with BCIs go missing and where law enforcement are actively searching for those people. If the BCI still has power and is broadcasting, then capture and enumeration of such broadcasts when in range could help find missing persons more easily.

#### **2.10. Responding to Vulnerabilities and Security Incidents**

A critical tenet of any sort of connected device security – particularly in safetycritical domains – is that it must be possible to remotely apply security updates to these devices, to ensure that as vulnerabilities are found, the devices can be updated to remain safe for their users.

As part of BCI operational security, BCI manufacturers will need to ensure they have robust security vulnerability mitigation and incident response plans, and the ability to address vulnerability disclosures from security researchers in a timely manner. Critically, this requires that manufacturers have clear (and long-lasting) support agreements for their devices, a well-defined mechanism for security researchers to submit vulnerability reports, reasonable patching and user notification policies, and secure software/firmware device update mechanisms to enable security patching post-implantation. We encourage medical device regulators to consider all of these factors in their development of regulation around implantable brain interface medical device security.

If a critical, remotely exploitable vulnerability were to be discovered in a common invasive or partially-invasive BCI, the manufacturer would need to action swift response in order to protect the security and safety of affected users. This presents a challenge for a device embedded in people's brains; recalling all customers and performing surgery on them to remove a vulnerable BCI

would be impractical. In these cases, BCI manufacturers and regulators can potentially borrow from best-practices around the recall of implanted, connected medical devices, provided that this information is robust and evidence-based with a comprehensive threat model of likely attacks, and provided that it takes into account the unique risks intrinsic to brain-specific connected devices.

Vulnerability remediation and incident response plans would need to understand the feasibility, safety, and user consent model of performing any necessary firmware updates to the affected BCIs, were that an option. Being able to track and confirm the number of people who had updated their BCIs (and those who hadn't) would be a critical dashboard function to mitigate safety risks, but also presents a privacy trade-off for affected users.

It may also be worth considering whether, for safety-critical medical devices, some form of logging and monitoring of the security of these devices by the manufacturer or a reliable third-party SOC may be appropriate.

How to notify and inform affected customers globally in a timely manner, and with actionable risk mitigation advice, would also need to be planned and documented as a process.

Given the safety-critical nature of BCIs, scenarios such as the one above raise the important question of what the appropriate fail-safes are for partially- or fully-invasive BCIs. For example: what happens when the device completely loses charge, encounters software errors or malformed input, performs an incomplete firmware update, is tamperedwith or damaged, or is End-of-Lifed by the manufacturer? Furthermore, should all BCIs offer a fail-safe on/off capability that only the bearer can initiate, and how can the integrity of this be ensured? These questions are of particular importance if the BCI is performing a critical health function; it may be less critical if BCI failure merely temporarily reduces a person's enhanced cognitive abilities achieved via the device.

#### **3. Post-Implantation or Post-Decommissioning Security Threats**

There will be an end to the operational life of a BCI, be that due to the bearer becoming deceased, or the bearer withdrawing their consent for its use or wanting it removed for some other reason. For invasive BCIs this demands manufacturers needing to offer the service of reversibility (removal of a BCI implant).

The concept of an on/off switch rears its head yet again in this domain. For a user wanting immediate removal of a BCI, they may be placated through the availability of an off switch that only they can control, until such point that they can undergo the necessary reversibility surgical procedure.

#### **3.1. Secure Decommission & Consent Withdrawal**

As mentioned in the earlier section on forensics, BCI removal will require secure disposal and destruction processes. This could involve secure wipe (in the event that the BCI may be reusable within other bearers), or secure physical destruction. Timely response to users wanting removal of a BCI or from withdrawal of consent will need to be considered. E.g. a six month waiting list for a reversibility procedure would not be appropriate for a user wanting BCI removal as soon as possible.

Similar secure decommission and wiping considerations are required on related BCI technology, such as NCDs and any cloud-related data storage concerning a BCI user. This will require BCI manufacturers to have strong data governance and understanding of all locations where customer data is stored, and thus from where it will need to be deleted.

Another potential negative impact on BCI performance could be from localised wireless and electrical interference. Mitigating such interference might be beyond the control of the BCI bearer, but the potential for such interference might at least be worth raising through awareness, in case there are options for minimising interference such as relocating or powering-down conflicting equipment.

## Performance & Environmental Factors

BCIs will have performance limitations that might be a result of a number of factors and potential environmental impacts, some of which may be beyond the control of the BCI bearer.

While not necessarily always a direct security concern, performance and environmental factors could ultimately affect or degrade system performance, which could present broader availability issues.

### **Model Training & Effectiveness**

BCIs won't necessarily work well as COTS products, out of the box across all individuals. Brain physiology will differ between people, thus requiring BCIs to learn and train on their bearers in order to maximise performance of specific tasks.

For example, Dr. Christian Herff of [Maastricht University](https://www.maastrichtuniversity.nl/c.herff) notes that it is unclear what happens to the speech areas of the brain in those who haven't spoken for years; their brains may have repurposed the speech part of the brain for some other tasks. Herff's research is yet to show application of data from one person to another in a BCI context.

He reports that a lot more data from a lot more brains is required before we might generalise neural models with reliable, predicable results across arbitrary individuals.

#### **Wireless & Electrical Interference**

Examples of interference might include wireless channel conflicts, mains electricity and poorly-shielded cables, or just electromagnetic emanations from NCDs such as laptops and [tablets](https://www.youtube.com/watch?v=TMeJyrPmwwM).

A curious thought on this topic is if/how in the future, performance might be affected for multiple BCI users gathered in the same place, such as at a large concert or in a busy airport.

Conceivably there could be many wireless conflicts and collective electromagnetic interference which could disrupt BCI operation for everyone in proximity.

#### **Mental & Physical States**

The mental state of BCI bearers could have a negative effect on an overall BCI operation. Neuro patterns may be more reliable and decodable when users are in specific moods or states of mental health.

Physically, changes in the body might affect performance of BCIs. As and when a BCI user might be multi-tasking or performing physical tasks, there may be deviations from normal brain patterns. Physical states of the body will also have an effect on BCI operation. For example, non-invasive EEG BCIs require electrodes to attach to or sit on the scalp of the bearer in order to read emanating neuro activity. If a person's skin is wellhydrated, there will be better conductivity and thus better readings from the brain – the corollary here is that a dehydrated non-invasive BCI user may experience poor or degraded system performance.

*Potential Health Impacts of BCIs (Safety)*

The health and safety aspects of BCIs should be governed by strict regulation, legislation and ethical frameworks. Invasive BCIs present the highest risk to health and safety due to the need for surgery and integration within the brain. Though the perceived risk of non-invasive BCIs is lower, there is still potential for long-term ill-effects from prolonged use. At this early stage of the technology we currently lack the data to be able to fully understand what the long-term effects and health implications might be.

Brain scarring  $-$  invasive and partially invasive BCI operations could result in scarring on the brain, which could lead to further complications such as memory loss, confusion, seizures etc.

A non-exhaustive list of potential health impacts of BCIs includes:

• Complications at surgery for invasive and partially invasive BCIs – this relates to implants, and any necessary maintenance or subsequent BCI

- removal
- 
- Burning prolonged use of noninvasive BCIs could result in excessive heat on specific parts of the scalp, causing a level of burning, rash and/or development of headaches
- Potential impact on free will some [research](https://www.slow-journalism.com/from-the-archive/the-billionaire-the-pig-and-the-future-of-neuroscience) on non-invasive techniques for brain stimulation has found that stimulating parts of the brain can trigger strange behaviour such as people feeling a sudden urge to move, or making involuntary movements. Unexpected urges and physical movements could put BCI users at risk
- Inductive charging invasive BCIs will require electrical charge for operation and because of their embedding within the brain, the only practical method of re-charging them is to use inductive (wireless) charging. For example, users may inductively charge their BCIs overnight while they sleep. Flaws in the design and implementation of BCI electronics could potentially result in increased heat and burning which could have severe impact on the brain



#### **BCI Health Regulation**

Deep Brain Stimulation (DBS) was approved in 1997 by the US Food and Drug Administration (FDA) for use in people with Parkinson's disease. Since then the technology has evolved to help treat many other conditions including Obsessive Compulsive Disorder (OCD) and epilepsy, and is being explored in its capacity to treat mental health conditions such as depression. More recently, the FDA has produced [guidance](https://www.fda.gov/media/120362/download) for device developers with both clinical and non-clinical testing considerations for invasive BCIs to be used by patients with paralysis or amputation. The FDA has also approved New York-based [Synchron](https://synchron.com/) to begin studies with six human subjects with severe paralysis. Synchron's BCI implantation procedure is not as invasive as other offerings such as The Link from Neuralink.

In the UK there is currently no regulation specific to BCI technology, yet BCIs could fall within scope of the EU Medical Devices [Regulation \(MDR\)](https://royalsociety.org/-/media/policy/projects/ihuman/report-neural-interfaces.pdf) if they are intended for medical use. The Royal Society has previously written a report with a number of recommendations on [BCI-based regulation](https://researchbriefings.files.parliament.uk/documents/POST-PN-0614/POST-PN-0614.pdf).

From May 2020, the MDR scope was extended to include non-invasive BCIs which use electrical currents, magnetic or electromagnetic fields to penetrate the skull and modify brain activity – it appears however that (EEG) headsets used for non-medical purposes such as gaming fall outside of the MDR's scope, meaning it is currently legal and permissible for anyone to manufacture and sell non-invasive, nonmedical use [BCIs](https://researchbriefings.files.parliament.uk/documents/POST-PN-0614/POST-PN-0614.pdf).

At the time of writing, a cursory review of global regulation concerning BCIs has revealed that there currently isn't much, and that which does exist is focussed on medical or clinical applications only, with most focus on invasive BCIs.

## Regulation, Legislation & Ethics

There doesn't appear to be any existing regulation concerning BCIs in a general consumer application context, and across the board, there is little to no security or privacy consideration in existing regulations.

#### **BCI Legislation**

Particularly in non-medical applications, the impact of BCI technologies on the brain and potentially people's thoughts and semblance of free-will is not currently well-understood so as to be able to understand the full legal implications and ramifications of BCI use. Coupled with security issues concerning potential unauthorised access to BCIs and their associated IT infrastructures, and what level of control that access could yield to an external attacker, leads to significant challenges for prosecutors, judges and policy makers.

#### *BCIs within Criminal Law*

Proving criminal responsibility for most crimes requires the prosecution to prove both a defendant's criminal act (*actus reus*) [and intention \(mens rea\)](https://www.legalcheek.com/lc-journal-posts/why-elon-musks-pigs-are-a-legal-headache/). Key questions here include how this would work for a defendant who used a BCI to commit a crime, whether willingly, or unknowingly (e.g. through an attacker's unauthorised access to a BCI and manipulation to invoke involuntary thoughts and actions)?

BCIs may require legislation to expand into the mental sphere. Historically, laws have supported the premise that people aren't punished for their thoughts (*cogitationis poenam nemo patitur*).

The potential impact on freedom of thought raises questions on whether it is acceptable to regulate BCIs, and if not, what the impacts would be on the rule of law when BCIs are *[involved](https://www.kilinclaw.com.tr/en/what-is-legal-dimension-neuralink-technology/)*. .

In examples where an attacker may have directly influenced a person to commit a crime through BCI-related activity, what are the implications around non-repudiation? Does the technology offer sufficient logging and auditing to be able to determine if any external unauthorised influence had occurred? Conceivably BCI technology might just malfunction due to implementation errors or electrical interference. As such, should BCIs be mandated to offer a level of logging and auditing of all actions, functions, inputs and outputs?

We can expect development of abilities to coerce information (potentially private) from BCIs and the data they generate. Research has already shown how subliminal visual cues, such as showing images of 4-digit PIN codes, dates of birth etc. has resulted in methods of inferring actual personal information from the brain patterns generated.

*["a subliminal attack in which, given that](https://arxiv.org/pdf/1312.6052.pdf)  [the visual probing lasts for less than 13.3](https://arxiv.org/pdf/1312.6052.pdf)  [milliseconds, the existence of any stimulus is](https://arxiv.org/pdf/1312.6052.pdf)  [below ones cognitive perception. We show](https://arxiv.org/pdf/1312.6052.pdf)  [that, even under such strong limitations, the](https://arxiv.org/pdf/1312.6052.pdf)  [attackers can still analyze subliminal brain](https://arxiv.org/pdf/1312.6052.pdf)  [activity in response to the rapid visual stimuli](https://arxiv.org/pdf/1312.6052.pdf)  [and consequently infer private information](https://arxiv.org/pdf/1312.6052.pdf)  [about the use](https://arxiv.org/pdf/1312.6052.pdf)r."* 

Use of such attacks (or methods) could conceivably become part of the toolkit of advanced interrogation techniques by government and law enforcement

authorities.

#### *People Tracking & Mind Warrants*

BCIs could present mechanisms for people tracking. Should BCIs use wireless broadcasts for communication with BCI hosts, then those broadcasts may contain identifiers or unique patterns that could be used to tag and track a person using various wireless interception techniques. Even without BCIs, conceivably technology may develop to be able to remotely pick up normal neural activity emanations from people's heads and thus be used as methods of identifying and/or tracking people. This will raise concerns for privacy campaigners, and could be

abused by criminals (or governments). For governments, investigatory powers legislation may need to advance to allow for tracking of criminals or suspects in such ways.

On a more invasive level, where remotely exploitable vulnerabilities might exist in BCIs and their related technologies, it's not too unimaginable to think that governments or law enforcement may wish to exploit those vulnerabilities to gain access to the mind and mind data of criminals or suspects. Such intrusive acts would surely require solid justification and (brain) warrants; this highlights where national security legislation and guidance may require significant overhaul in relation to BCIs.

#### *Equitable Access to BCIs*

Beyond safety and security, in the UK the National Institute for Health and Care Excellence (NICE) assesses medical technologies for efficacy and cost-effectiveness, and recommends whether devices should be available on the public health service (NHS) and if so whether their use should be mandatory to ensure equitable access.

Similar concepts might apply beyond just health applications – where BCIs and their applications may provide users with significant cognitive advantage, should there be legislation to ensure that exclusivity for the technology doesn't arise, rendering it accessible only to those with wealth?

#### *Brain Data Protection (Neural Privacy)*

There aren't likely current provisions specifically for BCI data protection across global data protection regulations and legislation; however, the nature of BCI data (both medical and potentially relating to private/personal aspects of thought) will surely fall under 'special category' personal data in frameworks such as GDPR.

As consumer markets for BCI technology will grow over the coming years, so will the volume of BCI data captured and stored by vendors and manufacturers. There will be significant value in such data, be that for the individual themselves (when used to train and fine-tune operation of their BCI), being used by broader neuroscience research for gaining new insights, insurance industry applications relating to health [insurance](https://www.bitbrain.com/blog/cybersecurity-brain-computer-interface) and even market researchers seeking to un-tap the thoughts, preferences and dislikes of consumers. How and where all of this data is stored will need to be understood as part of data mapping and governance activities. There will surely be huge elements of cloud storage of brain data at play, bringing with it the usual data governance issues around geographic regions of storage, who has access rights  $etc...$  43

Medical aspects and decoding methods of brain activity (brain data) are currently fairly well-understood. However, current understanding of deeper brain function is more limited, meaning that brain activity data cannot currently be used to 'read' a person's actual thoughts and memory. In theory, it may be possible to infer personal information from BCI data such as credit card numbers, secret sexual urges etc. However, current BCI technology doesn't offer this granularity of data.

Future technological advances however may lead to such a capability; that or, current methods of data capture of brain activity may mean that there are vast amounts of untapped, as of yet nondecoded personal information within those data sets, and conceivably as neuroscience develops, we may learn ways of extracting such information in the future, from data that is captured and stored today.

Relevant here will be legal aspects of investigatory powers – governments and law enforcement will surely want to obtain access to brain activity data from BCI companies, where that data relates to suspects and criminals.

The nuances of BCI data protection naturally segues to the aspect of consent. BCI companies will need to ensure they are clear in how their BCIs and associated applications work, and what data is captured, where it is stored, what it is used for etc. I.e. BCI users will need to consent to both use (or implant) of the BCI, and the associated data captures. The consent aspect is equally important for consent withdrawal – BCI manufacturers will need to be able to honour requests for data deletion and BCI removal (e.g. surgery) in a timely manner. Concepts of explicit on/ off switches for BCIs are therefore important in relation to consent. While not always appropriate for any realtime medical applications, certainly for non-medical applications, there should ideally be a mechanism for users to be able to disable their BCI access to the .

outside [world](https://fpf.org/blog/five-top-of-mind-data-protection-recommendations-for-brain-computer-interfaces)

#### **Neuroethics**

Neuroethics is a subfield of bioethics and is concerned with ensuring that technologies that directly affect the brain are ethically conceived and developed. Overall concerns within neuroethics relate to the potential for BCIs (or similar) to affect people's agency, autonomy and free will.

The longer-term effect on the brain from external influence and stimuli can also be unknown and fraught with health risk. Some research involving Deep Brain Stimulation (DBS) for treatment of Parkinson's disease has already seen examples where a minority of participants have become hypersexual, or become deeply [apathetic](https://plato.stanford.edu/entries/neuroethics/), even though other research suggests insufficient empirical studies to corroborate such [claims](https://link.springer.com/content/pdf/10.1007/s12152-018-9373-8.pdf). There can also be broader existential issues whereby subjects no longer perceive themselves in the same way, wondering how much of them and what they do is actually them, versus artificial control initiated by technology or external influence.

Neuroethics is perhaps well-understood in the context of medical applications; but it is less clear what the neuroethic priorities should be in relation to consumer applications for able-bodied people, the potential applications and use-cases of which are vast and not subjected to the same levels of regulation and oversight.

[Dr. Hannah Maslen](https://www.practicalethics.ox.ac.uk/people/dr-hannah-maslen#/) at Oxford University is contributing to regulation (in discussion with the [European](http://www.braincom-project.eu/)  [Commission](http://www.braincom-project.eu/)) concerning non-invasive consumer BCIs. Despite being noninvasive, such devices still send electrical current through people's scalps to stimulate brain activity. Maslen's research has found reports of such devices causing burns, headaches and visual disturbances.

### **Conclusions**

BCI R&D and technology is accelerating at [pace](https://www.cbinsights.com/research/neurotech-startups-to-watch/), yet the security and safety concepts of BCIs and their uses is not keeping pace, meaning that the security and safety of BCIs is at risk in the race to commercialisation.

This paper has merely scratched the surface of BCI security and safety, and how BCIs and their ecosystems should be secured.

We've seen that there are challenges across all BCI types, with significantly more risks and concerns with invasive BCIs.

There is little to no regulation on BCI security and safety requirements. A lack of regulation on consumer-grade BCIs means we are likely to see all manner of cheap and cheerful BCI offerings appear on marketplaces, developed to no safety or security standards yet potentially putting the safety and security of customers at significant risk – the same outcome that we've all experienced in the domestic IoT space in recent times.

From a legal perspective, we've touched on a number of areas where legislation may need significant overhaul in the medium to long term, commensurate with adoption rates of BCI technologies.

Our understanding and application of neuroprivacy concepts will be vital in tandem with BCI developments – this concerns our rights as individuals in relation to the imaging, extraction and analysis of our neural data. We might expect much future debate in this realm, and possibly unfortunate neuroprivacy breaches be they intentional through threat actor activities and/or unintentional through poor data security practices.

> As society's dependence on BCI technology will inevitably grow, so will the need for resilience in the software and services that underpin BCI operation.

and the contract of Thinks and th

We've touched on the various threats to BCIs throughout the BCI lifecycle, from design and manufacture (supply chain security), through to surgical implant, operational security and eventual secure decommission or reversibility. The security and safety threats throughout this lifecycle are vast.

Security and risk perspectives have a habit of projecting doom and gloom on a topic. Putting these perspectives aside, the convergence of mind, body and technology is fascinating and exciting, with potentially huge impact on humankind's evolution and enlightenment. BCI developments will naturally push the boundaries of computational and neuroscience, our growing understanding of the brain, how it works and how we can continually push its boundaries of capability.

BCI technology and AI/ML are mutually inclusive. The success of BCI technology will be predicated on AI/ML performance and the continued improvements in AI/ ML approaches – as such, all of the usual adversarial AI attacks become immediately relevant in the BCI domain.

For invasive BCIs that utilise inductive charging, we've highlighted how a number of different implementation choices might be exploited in ways that could result in overheating or explosion which could be detrimental to a person's health or mortality. We've briefly touched on environmental factors that might affect normal operation of BCIs – these need to be carefully considered by manufacturers and understood by BCI customers, in order to know how to achieve optimal operating conditions.

This demonstrates the need for software escrow and resilience services concerning BCI technology to ensure continued operation in the event of a BCI manufacturer's impacted business continuity such as bankruptcy.

We've seen throughout this paper a common requirement for an on/off switch capability for BCIs.

While this may not always be possible, for non-critical BCI applications, it provides the bearer with a level of control over who or what attempts access to their BCI (and thus brain), or allows them discretionary consent withdrawal capabilities.

How such an on/off switch might be realised in a secure manner is vitally important. We can imagine insecure implementations of such a feature which might provide adversaries with the ability to arbitrarily disable other people's BCIs. *I think therefore I am… part machine?*





## About the authors

**Jennifer Fernick** is a computer scientist and the SVP & Global Head of Research at NCC Group, and is a founding Governing Board and Technical Advisory Committee member of the Open Source Security Foundation. Before NCC Group she was Director, Information Security at a large global financial institution, after a tenure as their Senior Cryptographic Security Architect. She spent four years as a PhD researcher at the University of Waterloo, as a member of the Institute for Quantum Computing and the Centre for Applied Cryptographic Research, where her research focused on cryptography & quantum algorithms. Jennifer was a part of the 2018 cohort of the Berkman Assembly at Harvard University and MIT Media Lab, and was a 2019 Technologist Fellow at the National Security Institute at George Mason University. Her career has included designing and building satellite systems, working on bleeding edge cryptography research, building secure systems at massive scale, running incident response events for core pieces of critical infrastructure, and leading the development of global technology standards. She holds a Master of Engineering degree in Systems Design Engineering from the University of Waterloo, and an Honours Bachelor of Science in Cognitive Science & Artificial Intelligence from the University of Toronto. Jennifer spent multiple years as CFP Chair of Crypto & Privacy Village at DEF CON, and has served on the review boards of venues including USENIX CSET, USENIX Enigma, USENIX WOOT, multiple NeurIPS workshops, and IEICE Transactions Japan, and regularly speaks at major technology conferences including European Conference on Machine Learning, RSA, CFI-CIRT, DEF CON, O'Reilly Artificial Intelligence, the Linux Foundation Member Summit, and Black Hat USA. In 2021, she was named one of Canada's Top 20 Women in Cybersecurity by IT World Canada.

Binnendijk, A., Marler, T., & Bartels, E. M. (2020). Brain-Computer Interfaces: U.S. Military Applications and Implications, An Initial Assessment. RAND Corporation Research Report. https://www.rand.org/pubs/research\_reports/RR2996.html

## Acknowledgements

Sincere thanks to Stuart Kurutac (Connected Health Lead), Ivan Reedman and Verona Hulse (Senior Public Affairs Manager) for their invaluable contributions to this paper.

## Further Reading

Belkacem, A. N., Jamil, N., Palmer, J. A., Ouhbi, S., & Chen, C. (2020). Brain computer interfaces for improving the quality of life of older adults and elderly patients. Frontiers in Neuroscience. [https://www.frontiersin.org/articles/10.3389/fnins.2020.00692/full](https://www.frontiersin.org/articles/10.3389/fnins.2020.00692/full ) 

Drew, L. (2019). The ethics of brain-computer interfaces. Nature. <https://www.nature.com/articles/d41586-019-02214-2>

**Matt Lewis** is a computer scientist and experienced Technical Research Director. His specialisms include general security consultancy, scenario-based penetration testing, vulnerability research and development of security testing tools and methodologies. He is a Swansea University (BSc Hons) graduate and Oxford University MSc postgraduate (programming research group) with over twenty year's cybersecurity experience spanning national security and professional services (KPMG). His penetration testing, research and consultancy experience spans multiple technologies across all sectors and many FTSE 100 and Forbes 2000 companies. He is a public speaker with global recognition of knowledge and expertise in biometric security. He has presented research and insights at many international conferences and seminars on all manner of cybersecurity-related topics.