Ethical Implications
of Brain Computer Interfaces:

EXECUTIVE SUMMARY

We see that neural interfaces today are one of the hottest pieces of technology. Elon Musk, with his company Neuralink, Sid Kouider with NextMind, and many other companies such as Neurable, Emotive, and BitBrain, strive to create brain interfaces that will eliminate the last intermediary between humans and computers.

However, alongside innovations, more and more experts start talking about the ethical implications of such interfaces. So what are the main ethical issues surrounding neural interfaces today?

> **Obtaining a fully sound patient consent, unexpected personality and behavior mutilation, and illicit exploitation are serious problems brain interfaces of today face.**

Contents

Author bios

Ivan Markov

Ivan is the Head of the Data Science department at Serokell. Previously he co-founded several AI-focused startups in the areas of customer behavior prediction and medical computer vision in both B2B and B2C.

Yulia Gavrilova

Yulia is an AI and tech ethics researcher at Serokell. She has a Master's degree in Gender Studies from Charles University where she wrote a thesis about ethical chatbot design.

Alexey Khachiyants

Graduated the machine learning faculty in HSE. Since then he worked as a data scientist in customer analytics department in Yandex and now is a ML expert with focus on computer vision in Serokell.

Dmitri Puzyrev

Has Skolkovo's masters degree in computer science, specifically NLP. He worked in Computational Pragmatics Lab and MTS AI as a researcher where he explored modern NLP problems and wrote several articles. Not so long ago he joined Serokell as the main NLP expert

Brief summary

What are the main ethical concerns about neural interfaces?

1. Obtaining consent from a BCI patient.
2. Potential change in attitude and reactions of a patient.
3. Illegal use of community-driven BCIs.
4. Enhanced inequalities caused by BCIs.
5. BCI animal testing.

Ivan Markov
Head of Data Science department at Serokell, points out several potential problems and respective ethical concerns:

1. Obtaining a fully sound and adequate consent from a BCI patient.

Most of the current experimental BCI trials are aimed at people with disabilities. In some cases, they are not fully aware of the implications of BCI usage but still agree because it's simply their only potential way out of their limitations. Jens Clausen in his paper provides an example of people with locked-in syndrome (LIS).

2. Invasive BCI procedure may cause different changes in attitude and reactions of a patient.

An example of such a case was described in the paper of Gilbert F., Embodiment and Estrangement: Results from a First-in-Human "Intelligent BCI" Trial. There he describes a patient that started gambling compulsively after this treatment. It is worth noting that this addiction was active only during the treatment. In the body of the research, Alexey Khachiyants mentions a great imaginary case of a person with BCI who commits a crime – how can we resolve whether it was committed due to side effects of the BCI impact or not?

3. Illicit exploitation of community-driven BCIs.

Importance of establishing safety protocols for BCI usage. Case studies: password memorability from EEG (link), malicious brain-hacking (link), personal information retrieval based on EEG (link).

Another potential breach is not information extraction but rather intervention into BCI functioning which may lead to incorrect signal reads and BCI device malfunctioning (measurement accuracy, processing speed, and distribution). Above mentioned exploits lead to the “dual-use dilemma of brain-hacking”. It states that BCI devices can be used both “for good” and “for bad” purposes. Dmitri Puzyrev mentions a great example of one of the buzzing examples of BCIs which is Neuralink by Elon Musk. They intend to develop a “Fitbit (fitness tracker) inside your skull” with the promise to be able to access the brain via your own smartphone. Suffice to say that if any third party can get access to the data from your Neuralink (or similar technology), it can potentially get a significant control over life: from harmless simple perception analysis and ad targeting to analysing your hazardness or even evaluating your social usefulness.

4. Human inequality due to the usage of BCIs & human definition.

A great question mentioned by Alexey Khachiyants is a potential society stratification based on the usage/availability of BCI interfaces (especially at the first stages of BCI expansion). Let's imagine that there's a BCI device that enhances brain activity and efficiency. First, it's going to be a very expensive device and only the richest of people will be able to install it. Which may cause a difference in welfare to get even worse. Another concern is whether a human remains human if he uses BCIs? Or what's the stage when BCI impact is not considered crucial in terms of "humanisation"? Here are the questions mostly from the sci-fi domain now and are still to be answered.

5. BCI animal testing.

Obviously, BCIs (esp. Invasive ones), as most of the medicine related tech, has to undergo a lot of trials and tests before being tested on living human beings. This raises a question of the ethicality of using animals as test subjects. Going back to the NeuraLink example, as Dmitri Puzyrev mentions, it was tested on a chimpanzee which led the general audience to realise how complex and multifaceted the problem is. First of all, installation of a chip is a complex invasive surgical procedure which can unlikely be reversed. Secondly, this trial was used as a large ad campaign for Neuralink technology which is kind of controversial.

The ownership of identity and thoughts is in peril

Yulia Gavrilova
Head of Content department at Serokell

Does your brain belong to you?

It's not hot news that global corporations use every bit of accessible information about their clients to drive profit. A neural interface that can read and write to your brain is a perfect source of information for marketing campaigns and a means for distributing advertisements.

In fact, neural marketing of the future can become as good at manipulation as ever before: you won't even know why you suddenly want a juicy burger from a local fast food restaurant at only half a price! Companies such as Neuralink predict that neural interfaces of the future will be able to respond to a person's commands and insert information into the brain.

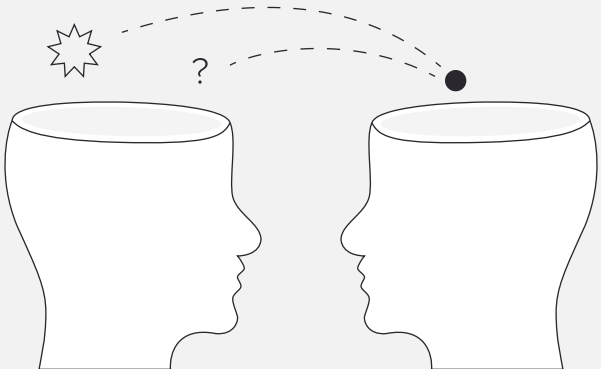
> **When you sign a contract with a brain interface company, you delegate part of control over your brain to a third party in exchange for a valuable service.**

The company decides what to do with this information and to whom they might sell it. If you want to keep using the services, you simply can't afford to break the contract.

Does your identity belong to you?

A neural interface has an effect on the way you think and behave. Today brain devices are already actively used for medical purposes to treat obsessive-compulsive disorder and nervous anorexia. However, is it ethical to change traits or behaviors that contribute to the person's identity? Does it mean that we will be able to treat people with 'unwanted' personality traits in the future? For example, it is proven that psychopaths have reduced connections between the ventromedial prefrontal cortex (vmPFC), the part of the brain responsible for sentiments such as empathy and guilt, and the amygdala, which mediates fear and anxiety. By stimulating these areas, we could, in theory, 'correct' the immoral people.

There are still some concerns about what a brain interface can do for us outside medical discourse. For example, people can already use neural interfaces to boost memory and alternate their mental state. Brain interface games teach people to activate specific areas of their brain through a feedback loop, and that game helps them relax and reduce levels of stress.



However, while neural interfaces become more researched, it might become easier to switch emotions in as much as one click. It goes without saying that feeling these emotions is as normal as feeling happy and energized, depending on the context. However, the device can step up and interfere with your decision-making.

Alexey Khachiynts
Senior Data Engineer at Serokell

Deep brain stimulation is a rather old method that was created to help patients with Parkinson's disease. It works via passing currents through specific regions of the brain. This method is extremely effective, but it is said that it might cause severe side effects. In paper Embodiment and Estrangement: Results from a First-in-Human "Intelligent BCI" Trial Gilbert F. describes a patient that started gambling compulsively after this treatment. It is worth noting that this addiction was active only during the treatment. This leads us to the question: can somebody with BCI really call themselves autonomous? Can they be sure that their actions were strictly out of their volition and not because the interface told them to? Though, this problem's main source is our lack of proper understanding of brain mechanics - when we stimulate a region, we currently can't say what exactly we are doing.

Are you still the agent of your life?

How would you feel if artificial intelligence, your boss, or your parents decided what mood you are in today? With neural interfaces, this scenario is unfortunately quite possible.

Our society doesn't welcome all moods equally. At large, we prefer to deal with people who are happy and satisfied or, at least, calm, relaxed, and can be reasoned with. When we express anger, disappointment, or sadness, other people often feel uncomfortable. You might feel obliged to give yourself a constant happiness enhancement just not to be a burden. Or, even worse, somebody else can do it for you. For example, your boss who decided to fire you and doesn't want to deal with your frustration can give you a mood boost just to smooth things out a bit.

How will the world look like where feeling happy and positive all the time becomes a new norm? We already live in this world in a way: look at your Instagram or Facebook feed where everyone seems to live a perfect life. However, internally, it's still up to you to experience and express different emotions. In a society where anyone can control their emotions at will, the medicalization of negative emotions can happen. You will be obliged to treat anger, sadness, or frustration just because other people feel uncomfortable – does it remind you of something (read, heteronormativity)?

Let's say sometimes some switching is necessary, for example, in treating depression. But does the patient understand that their personality was altered? Usually, coming out from depression happens gradually step after step through cognitive therapy and, in some cases, taking drugs. However, brain interfaces can just streamline this process. And it seems like this has consequences for our brain. There are examples of people who used brain interfaces to treat a mental disease but obtained other unhealthy patterns such as compulsive gambling that stopped only after the interface was turned off. Emotions are a tool 'given' to us by evolution and necessary for survival. What kind of mechanisms do we discard by switching off emotions in a click?

Dmitry Puzyrev
Data Science Engineer at Serokell

Brain activity sculpting can trigger differing emotional responses. ‘Does that change the personality of the client? Should the client be considered the same person by peers using that “personality change”? And, most importantly, can these changes be averted and who will be responsible for those changes, NeuraLink or the client?’

Who is accountable for your decisions?

Machine learning algorithms that learn from your historical data and guide users to the most logical choices are widely used for brain interfaces. For example, brain interfaces that help produce speech work a lot like an auto-suggest mode in text messages. However, if the algorithm constantly suggests what to say and do to the user, the user can simply approve that option all the time because the algorithm usually proposes the most optimal choice. In that case, it's unclear who becomes the author or the message and bears responsibility for it. A shared agency is born where only part of the decision-making comes from the user.

Alexey Khachiynts
Senior Data Engineer at Serokell

Suppose that a person with BCI did something illegal - for example, stole something from a shop. Who is at fault: the user or the interface?
We can't really say whether this action was caused by continuous stimulation that might have caused this behavior or it was done in a clear mind. Though, this question is more of a legal one.

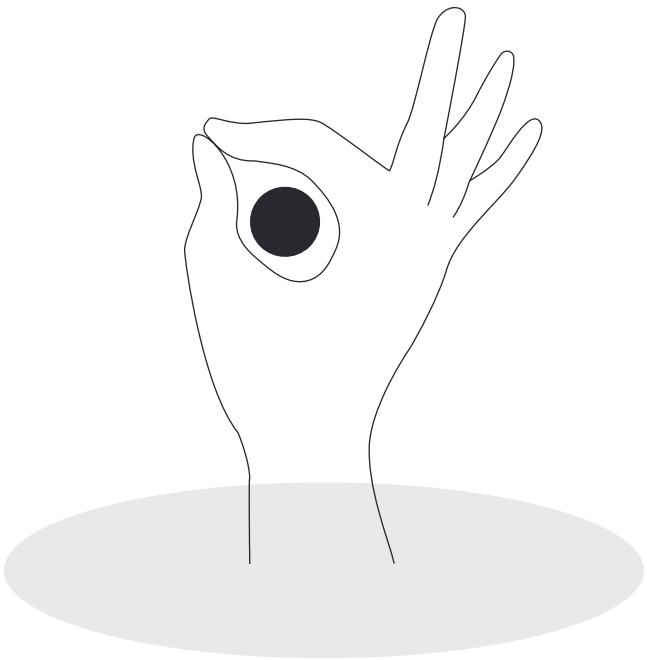
Neural interfaces are an attractive target for hackers

Today smartphones are our closest and most vulnerable devices. They contain private photos, messages, and other confidential information that is the desired target for hackers. However, the impact of hacking a phone can't even closely be compared to what would happen if hacking a brain interface were possible.

Neural interfaces pose serious privacy concerns. Once connected to a brain device, you find yourself in a symbiotic relationship with a computer that potentially anyone can get access to. Malware can mingle with your behavior, mood, thoughts, and body movements, and the effects of such an invasion on a person's health can't be told until we have precedents.

Neural interfaces raise questions about accessibility and the distribution of resources

There is no doubt that the development of neural interfaces can bring tremendous benefits to humanity. However, who gets to have an AI-enhanced brain, and who doesn't? In most cases, only the rich and the privileged get to experience technological innovations at their emergence (think, for example, of space tourism). Will they get even more rich and privileged?



If some people get to use neural interfaces and others do not, the problem is apparent. However, what happens when everybody becomes a super-genius with a fantastic memory and improved cognitive abilities? Today's compensation on the market is directly connected to what you're worth as a professional, specifically because not everybody has the same skills. However, with neural interfaces, there is a possibility to become genuinely equal in a certain way; how would that affect wealth distribution and the economy as a whole?

Dmitry Puzyrev
Data Science Engineer at Serokell

Adds that now focusing on medical assistance, NeuraLink states ambitious goals in constructing ‘brain overlay connected with artificial devices’. It’s easy to imagine people wanting to use the technology for mental enhancement. Will the cognitive boost created by the device count as unfair competition practice? In the same fashion, as steroids are exploited in sports, neural assistance can be abused.

Finally, neural interfaces are actively tested on animals

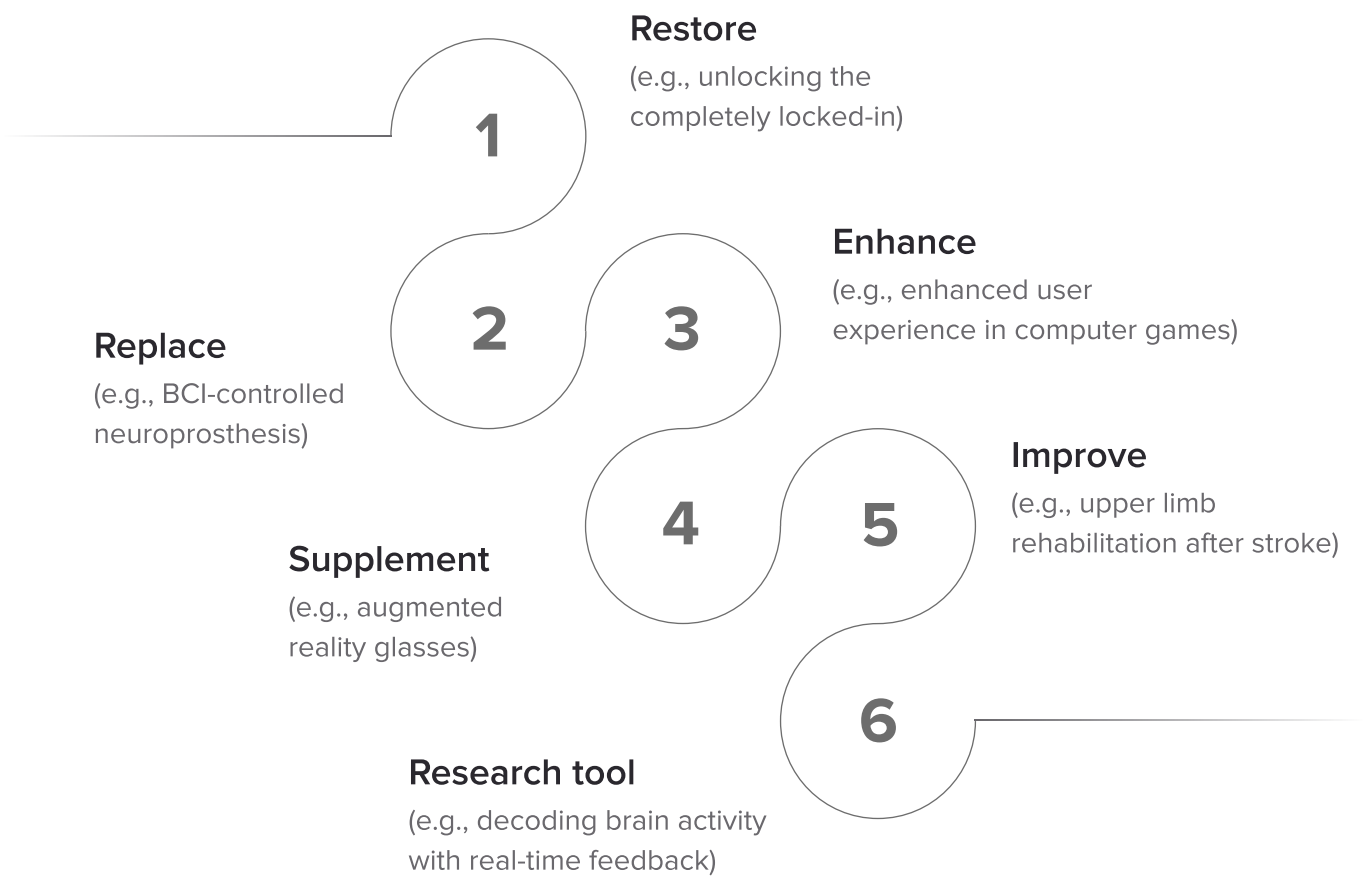
Invasive brain interfaces, such as the one proposed by Neuralink, are actively tested on animals. Dmitry Puzyrev, Data Science engineer and AI ethics enthusiast at Serokell, notices that while this is quite standard medical practice, this spawned a wave of criticism from animal rights organizations. It should be noted that the system is transmitted via surgical invasion and can't be easily removed. Not stopping with closed experiments, a video of a monkey playing pong was released. This raises accusations of turning animal testing into a "show" for media advertisement.

It is not our intention as ethicists to police innovations. But, it goes without saying, neural interfaces are the technology that should be developed responsibly.

BCI progress

Ivan Markov
Head of Data Science department at Serokell

Six promising BCI themes (Brunner et al., 2015):



Example of different BCI applications with respective categorisation:

Feature	EEG	MEG	ECoG	Intracortical Recording	fMRI	fNIRS	PET
Activity type	Electrical	Magnetic	Electrical	Electrical	Metabolic	Metabolic	Metabolic
Measurement type	Direct	Direct	Direct	Direct	Indirect	Indirect	Indirect
Invasiveness	Non-invasive	Non-invasive	Invasive	Invasive	Non-invasive	Non-invasive	Non-invasive
Portability	Yes	No	Yes	Yes	No	Yes	No
Temporal resolution	~0.05 s	~0.05 s	~0.003 s	~0.003 s	~1 s	~1 s	1-2 min
Spatial resolution	~10 mm	~5 mm	~1 mm	~0.5 mm (LFP) ~0.1 mm (MUA) 0.05 mm (SUA)	~1 mm	~5 mm	~4 mm
BCI applicability	Acceptable spatio-temporal resolution with high-density electrodes	Mobility constraint	Unfavorable for healthy BCI users	Unfavorable for healthy BCI users	Slow and mobility constraint	Slow, but mobile and a potentional alternative to fMRI	Limited potentiality

EEG, electroencephalography **MEG**, magnetoencephalography **ECoG**, electrocorticography **PET**, positron emission tomography

fNIRS, functional near infared spectroscopy **fMRI**, functional magnetic resonance imaging

Around 15–30% of individuals are inherently not able to produce brain signals robust enough to operate a BCI (Cecotti, 2020). Also, there's a large discrepancy in terms of response to BCI techniques meaning specific cases require specific BCI approaches (e.g. rehabilitation of stroke survivors, Park et al., 2016). Even different human traits such as emotionality, fatigue, stress, lifestyle, age, gender, and motivation affect results in BCI testing (look for P300-BCI paradigm).

Our brain-computer interface market estimations

Yulia Gavrilova
Head of Content department at Serokell

Every day, more and more brain interface companies continue to appear. The majority of promising brain interface startups are located in North America. In 2020, the brain interface market in the USA was valued at 1.4 billion. Based on historical data about the market's growth rate (2016-2018), we expect the market to reach $6 billion by 2030.

Even though invasive neural interfaces (aka those requiring brain surgery to be installed) prove to be more effective today, we believe that the future is after wearable technology. Companies that produce non-invasive brain interfaces today are NextMind, Kernel, and BrainScope.

The major industries where brain interfaces will be in demand are MedTech, military, education, and entertainment. Kernel that develops interfaces to help people with memory problems, BrainScope's solution for non-invasive medical brain assessment, and Halo Neuroscience with their wearable device for improving cognitive performance in the healthy and impaired have the largest capitalization right now ($100 million, $62 million, and $42 million respectively). However, companies like Neurable and NextMind that concentrate on gaming and entertainment are also quite well off. We expect them to triple in size by 2030 and attract more investments to realize new projects.

Conclusion

A broad consensus on ethical issues and beneficial socioeconomic application of BCI technology is urgently needed (somewhat similar to the ethical usage of AI in medicine, self-driving cars or in other domains where lives of people are dependent on AI decisions).

Established transparent safety protocols for BCI devices to prevent any hacking and/or redirecting BCI data to other parties. Also ensure that those protocols are followed precisely for any publicly available BCI tech.

Important to develop more rigorous procedures for informed consent should be implemented to increase the user's understanding of the risk–benefit ratio.

Brain-hardware interfaces need continuing ethical discussion to realize their full beneficial potential and avoid the pitfalls of hasty application
Ideally, the above mentioned points should involve the collaboration of ethicists, neuroscientists, engineers, computer scientists, cybersecurity experts, lawyers and other significant stakeholders and inform regulators and policymakers.