Breaking the Silence with Direct-Speech Brain Computer Interfaces: Centering Communicative Disability in Ethical Recommendations for Mitigating Algorithmic Bias

Alice Finkelstein
Stanford University[1]

Abstract

Despite growing scholarly attention to the ethical implications of direct-speech brain-computer interfaces (BCIs), there remains a lack of concrete guidance on how to address these concerns—particularly for users with communicative disabilities. As neuroengineering continues to advance alongside developments in artificial intelligence and machine learning, the need for clear ethical frameworks and standardized protocols becomes increasingly urgent. This paper investigates the core ethical challenges of direct-speech BCIs for individuals with communicative disabilities, identifying algorithmic bias as a central and underexamined issue. Through analysis of existing neuroethical standards, policy proposals, and international legislation related to AI and neurotechnology, the paper exposes critical gaps in current guidance. Drawing from disability studies and related fields, it argues that mitigating bias and ensuring equitable BCI development requires a broader, more inclusive understanding of language and a commitment to user-centered design.

Keywords: direct-speech brain-computer interface (BCI), communicative disability, algorithmic bias, neuroethics, user-centered design, Crip linguistics, privacy, agency, assistive technology, inclusive AI design

Introduction

"Speaking is everything human."

Pat Bennett, who wrote these words in an interview with NBC's Today Show (Today.com, 2023) has amyotrophic lateral sclerosis (ALS), a progressive neurodegenerative disease that makes her unable to speak intelligibly. However, a breakthrough made by brain-computer interface (BCI) researchers at Stanford (Willett et al., 2023) now allows her to use speech-decoding technology to communicate at a rate of 62 words per minute, more than three times as fast as previous assistive communication devices. Known as direct-speech BCI, this technology functions by capturing an individual's neural signals through an electroencephalography cap or implantable chip placed directly on the brain, and then translating these signals into synthetic speech sounds and printing them on a computer screen to facilitate communication. Bennett's words illustrate how powerful the applications of this technology are; for those with communication impairments caused by traumatic brain injury, stroke, paralysis, or neurological disorders such as cerebral palsy

[1]Correspondence concerning this article should be addressed to Alice Finkelstein, Stanford University. Email: alicef@stanford.edu

and locked-in syndrome, among many other conditions, facilitating a conventional speed of communication would be transformative.

Direct-speech BCIs are just one example of the latest advances at the intersection of neuroscience and technology. From mice whose behaviors are controlled by light (Montgomery et al., 2015), to monkeys with brain implants that enable them to play video games (Wakefield, 2021), to patients with Parkinson's Disease who use deep brain stimulation (DBS) for reducing tremors (Hariz & Blomstedt, 2022), innovations in neurotechnology have surged in the past decade. The field is not solely being developed through academic research: companies such as Elon Musk's NeuraLink are already working on commercializing brain implants that aim to integrate AI with the human brain to allow for humans to control devices with their thoughts (Guarino, 2024). However, these neuro-technological developments, and the rate at which they are emerging, have raised alarms; ethicists and scientists are insisting on the need for proper legislative oversight in the field before it crosses dangerous boundaries, such as enhancing memory and brain function (Jangwan et al., 2022) or permanently changing an individual's identity (Müller & Rotter, 2017).

While the ethical concerns presented by brain-computer interfaces have been discussed by scholars on a broad level, with mention of general concerns such as autonomy and security (Davidoff, 2020), there is little conversation about more specific concerns such as bias in the algorithm responsible for decoding and translating brain signals into text. Algorithmic bias refers to a systematic error in a computer algorithm causing a preference for one outcome over another due to design limitations. The role of algorithmic bias in the perpetuation of social injustices and discrimination has been well-discussed in other contexts, such as in race and gender (Broussard, 2023), but there little conversation on how it might affect BCI users with communicative disabilities - an important group to address, as it is most at risk for misrepresentation due to it being more challenging for communication-impaired individuals to express their thoughts and feelings. There is also little discussion about how ethical concerns can be mitigated and addressed, and a lack of guidance that developers of BCI technologies should follow.

In an era where human beings and technology are more closely interconnected than ever before and further innovations in neuroengineering inevitably spur onwards, it is therefore imperative to outline clear guidelines and establish accepted protocols for the responsible development of brain computer interfaces. In this paper, I explore privacy, accountability, and agency as key ethical concerns surrounding the usage of direct-speech BCIs by individuals with communicative disabilities, uncovering algorithmic bias to be a critical underlying issue. I then analyze existing scholarly proposals and legislation aiming to address algorithmic bias, identifying gaps in current neuroethical standards. Finally, by applying the perspective of a disability lens, I provide recommendations for guiding responsible BCI development on a global scale. Illuminating the needs of individuals with communicative disabilities, an underrepresented group in the conversation, my research ultimately leads me to conclude that algorithmic bias should be addressed through the adoption of a broader view of language and the utilization of user-centered design.

Ethical Implications of Direct-Speech BCI for Users With Communicative Disabilities

Though great progress has been made in BCIs in recent years, there is still much that is unknown about how the human brain functions and how speech is generated. Rather than speech production being a single function localized to one brain region, evidence shows that there are multiple systems of speech production operating at distinct periods across a network of regions in the brain (Cooney et al., 2018). Additionally, while imagined speech retains a wide range of the phonological characteristics, or “featural richness” (Cooney et al., 2018, p. 109), associated with overt speech, the neuroanatomy of imagined speech differs among individuals and remains only partially understood. Correspondingly, the software used by direct-speech BCIs to translate neural signals is far from perfect in its ability to convert brainwaves into language, offering only the closest approximation of an individual’s thoughts (Klein et al., 2022). These scientific uncertainties have ethical consequences: if speech decoding software only approximates meaning, users—especially those with communicative disabilities—may be misunderstood or misrepresented. Such risks raise profound concerns about privacy, autonomy, and control, particularly for users who may be unable to correct or clarify what the machine outputs on their behalf. As BCIs become more integrated into communication support, addressing these concerns becomes essential to ethical development. Several review articles have described BCI-associated ethical issues at length (Rainey et al., 2020; Kellmeyer, 2018; Wolkenstein et al., 2018), converging on their discussions of BCI effects on privacy, accountability, and agency. The following sections further elaborate on these issues to reveal the need for clear, disability-centered ethical recommendations for direct-speech BCIs.

A. Privacy: Dangers Posed by Brain Data Collection and Unwanted Externalization of Personal Thoughts

Invasion of privacy is one of the principal concerns voiced by many researchers, ethicists, and BCI users; the potential for the personal brain data collected by BCIs to be used in unintended ways, and the highly personalized nature of the data which can result in increased identifiability, poses risks to users. Advocating for the importance of limiting and securing access to a patient’s data emerges as a critically important action, but is seen as an abstract concept due to the great ethical complexity in deciding what data should and shouldn’t be preserved (Kellmeyer, 2018). Machine-learning software further complicates this decision, analyzing data through creating algorithms that are extremely difficult, or impossible, to understand by humans.

Another privacy concern is associated with the mind-reading nature of direct-speech BCIs. Since mentalized speech is not always intended to be externalized, analyzing the neural activity associated with imagined speech is viewed as problematic because the resulting output of the BCI may bypass the user's intentions (Rainey et al., 2020). This situation is complicated by the predictive nature of artificial intelligence (AI) based language models that are currently under development. Although incorporating a predictive approach may increase the speed of the BCI model's text generation, the likelihood of erroneously externalizing speech that a user doesn't intend to express will also increase. For the user, this would be as though they are speaking without any sort of filter - in hypothetical scenarios in day-to-day communication, where people purposely prefer to not voice what they're thinking, this could translate into a user being unable to withhold thoughts that they'd want to keep private.

B. Accountability: Who's Responsible for BCI Output, User or Algorithm?

Because the BCI effectively acts on behalf of the user, ethical concerns with assigning accountability emerge. The ambiguous process that the software performs to convert from a person's thoughts to the technology's output results in what's been termed an "accountability gap" (Kellmeyer, 2018) or "responsibility gap" (Wolkenstein et al., 2018). Because control over BCI output is shared between the user and the computer software, it becomes difficult to discern whether the algorithm's output is truly the result of the user's intentions, or instead, the result of the computer's calculations and approximations. Wolkenstein et al. (2018) liken BCI algorithms to "black boxes" (p. 637): with constant mismatches between what the users believe about their mental state and what the direct-speech BCI detects, users' ability to separate their true intentions becomes increasingly hindered.

Moreover, in the specific case where direct-speech BCI is used by individuals with communication disabilities, there's often no evidence for whether or not an algorithm is properly translating the user's words, because it is much more challenging, if not impossible–such as in the case of a patient with locked-in syndrome (Snoeys et al., 2013)–for the user to independently communicate the algorithm's success or failure. This lack of real-time feedback makes reliance on algorithms especially risky, as it makes it even easier for a user to be misrepresented by the BCI.

C. Agency: Sense of Control Reduced By Inaccurate Algorithmic Outputs

With language being so varied based on each individual due to aspects such as unique regional mannerisms, personal sayings, and patterns of speech (Klein et al., 2022), the personalization of BCI poses a paradox: while the benefits of individually-tailored BCI algorithms include enabling people to find their "'true' preferences, or deepest emotions," if users are unable to discern their "real selves" (Wolkenstein et al., 2018, p. 641) when BCI outputs are unclear or unexpected, their autonomy is effectively reduced. Algorithms create evidence for a conclusion that a certain state of mind has occurred in a user, but there's always a degree of uncertainty to this conclusion; in the case of direct-speech BCIs, it can be easy to forget this element of uncertainty, leading users to become conflicted about their agency.

Agency encapsulates the concepts of cognitive liberty, the idea that people ought to be free to think their own thoughts (Sententia, 2004), and self-conception, described by Rainey et al. (2020) as the image an individual has of themselves. Rainey et al. (2020) compare the effect of a BCI on thought generation to someone keeping a diary, emphasizing how if a person was to realize that someone was able to read their diary, they might start to write entries that are less revealing. This process of refraining from potentially exposing thoughts would cause a "deformation of normal ways of thinking" (Rainey et al., 2020, p. 2301), stifling cognitive liberty by making users feel as though they aren't allowed to think certain thoughts. A defining characteristic of being human is the ability to act on reasoned opinions and reflect on options before making decisions; because direct-speech BCI algorithms regard a user's mental contents as public and then express these contents as language on the user's behalf, improper expression could translate into a diminishing of users' sense of what it means to be human.

Existing approaches offering solutions to BCI-associated ethical concerns

Identifying the ethical issues associated with direct-speech BCI is only a first step. This paper's examination of the prominent concerns of privacy, accountability, and agency point to a clear need for unbiased BCI algorithms that recognize personal variations in users' communication styles. There is little existing guidance on mitigating algorithmic bias in the context of direct-speech BCI users with communicative disabilities; due to this gap, exploring existing guidance from the scientific community on related neurotechnologies, neurorights policies in different countries, and related examples of algorithmic bias mitigation can serve to inform a set of guiding principles for ethical BCI development.

A. Suggestions for Bias Mitigation in Current Literature: Software Improvement, Code Transparency, and Representation of Users With Communication Disabilities

Mitigating algorithmic bias must undoubtedly include technical solutions improving the brain signal decoding software; while it is true that such solutions will be implemented only after scientists gain a greater understanding of speech production mechanisms and of the complex structure of the brain, scholars have proposed several currently actionable approaches to software improvement, such as the establishment of evaluation metrics that measure sources of bias (Liang et al., 2021) and increasing the transparency of algorithms (Bernal et al., 2021).

Specifically, this refers to adapting algorithmic design to move away from its current “black box” structure, where an unidentifiable process occurs between an input and output. Designing BCI software that outputs a justification for every action or nonaction would contribute to increased explainability of the algorithm, in turn allowing developers to identify bias and improve the software to reduce existing bias. Creating strong explanation frameworks remains a challenge, but future research must be done in this area, because making BCI algorithm function interpretable will lay the groundwork for systems that can symbiotically interact with humans (Rajpura et al., 2023), and aid in diminishing the aforementioned “accountability gap” between BCI users and software. In the context of users with communicative disabilities, interpretable and explainable algorithms become even more important, as communication between a user and the algorithm designer is already impeded.

Beyond the development of technical solutions that improve algorithm structure, non-technical solutions, such as the establishment of guiding principles, accepted standards, and laws, will play an equally essential role in reducing algorithmic bias. A principle that is well-advocated for in the context of not only artificial intelligence and machine learning, but also in the scientific community as a whole, is making software and collected data open-source (Fitzgerald, 2006), meaning that the code is publicly distributed in order to facilitate collaboration. Bernal et al. (2021) argue for the importance of BCI algorithms to be open-source, discussing how code and data transparency is vital for others to perform downstream analysis, replicate results, and understand brain function. Echoing this view, Mei et al. (2024) have recently created MetaBCI, an open-source platform for BCIs as part of an effort to encourage collaborative development and accelerate scientific progress. Expanding the amount of training data available, while also ensuring that it is standardized, audited, and regulated (Caliskan, 2021), will allow BCI algorithms to become more accurate due to greater representation, diminishing existing biases due to current limits in BCI training data.

Open-source code is one proposition aiming to reduce algorithmic biases by increasing BCI training data availability, but the limited amount of training data is itself an issue that needs to be addressed. Given the recent breakthroughs in modeling language with AI and the creation of generative AI language models such as ChatGPT (Chang & Kidman, 2023), researchers are currently exploring the potential of supplementing BCI speech decoding software with language models in order to better predict a user’s speech and make the BCI output text at a faster rate (Berezutskaya et al., 2023; Klein et al., 2022). Algorithmic bias often occurs when an algorithm is applied in a new context of use than the one it was designed for; this is precisely what can occur when individuals with communicative disabilities use direct-speech BCIs that utilize AI-based language models, as despite the existence of natural language processing (NLP) models trained on large datasets, there are currently no large datasets of language used by individuals who rely on augmentative and alternative communication (AAC), the technologies that supplement speech for those with speech impairments (Peters et al., 2022). The lack of adequate training data hence poses a great risk for contributing to algorithmic bias, as it increases the probability of incorrectly translating the imagined speech of those with communicative disabilities.

Concerns about the use of language models that improperly represent user speech are voiced by those with communicative disabilities. In a study evaluating the perspectives of individuals with neurological diseases that impair communication ability, Klein et al. (2022) found that although respondents supported personalization of BCI language models, they expressed worries about the possibility of algorithmic bias in the language models challenging their ability to maintain different registers by blocking the output of "familiar phrases or idiosyncratic word choices" (p. 8). Due to the reliance of direct-speech BCI algorithms on training data in order to learn to interpret neural signals and translate them into speech, if algorithms are primarily trained on users without any specific linguistic or cognitive disabilities, they may fail to accurately capture the neural activity of users who do have such disabilities. Therefore, acquiring large language datasets from a representative group of individuals is vital, along with ensuring that patients representing a disability perspective are able to provide input during the BCI development process and are actively involved in the creation of the end goals of the neurotechnology.

In addition to accurately representing the needs of users with communicative disabilities, it is also critical to include a diverse set of researchers who are conscious of the ethical implications associated with direct-speech BCIs. In a qualitative interview case study of neural engineering researchers working on BCI development, Sullivan et al. (2017) report that many researchers in the field are trained in basic or applied science, rather than assistive technology design, so their expertise is largely in the determination of neural processes instead of building technology to address a need. Sullivan et al. (2017) further describe that interviews of researchers suggested a need for more in-person engagement with patients with disabilities, and alongside other scholars (Berger & Rossi, 2023; Kellmeyer, 2019), argue for the importance of making an ethics curriculum a standard for all researchers engaged in neurotechnology development. Centering researcher training on addressing needs, and particularly on the unique ethical challenges posed by algorithmic bias, will be fundamental in encouraging responsible direct-speech BCI development.

B. Chile, China, and the EU: Neurotechnology and AI guidelines around the world

Apart from scholarly discussions on general recommendations for ethical concerns with algorithmic bias in direct-speech BCIs, new legislation on neurotechnologies has brought the field to the forefront of recent policy-making discussions: in 2021, Chile's Senate unanimously passed the first-ever constitutional amendment creating a "neurorights" law. This law enshrines mental privacy protection as a right every Chilean possesses, giving personal brain data equal status to an organ - meaning that it cannot be bought, trafficked, or manipulated in any way (Gúzman, 2023). The law was created to recognize five new human rights to mental privacy, personal identity, free will and self-determination, equal access to neurotechnologies, and protection from algorithmic decision making biases. While it is true that the law emphasizes a commitment to safeguarding human rights in the face of rapid advancements in AI and neurotechnologies, a number of scholars have critiqued the law, reflecting that the ruling is vague and fails to address more specific issues such as the ability of devices to neurophysiologically identify an individual, or how exactly individuals' rights would be affected by neurotechnologies (Cornejo-Plaza et al., 2024). Other scholars bring up additional criticisms: Zúñiga-Fajuri et al. (2021) argue that the law is premature, addressing inexistent threats due to neurotechnologies being far from realization, and that existing human rights protecting physical integrity suffice; Borbón & Ramírez-Gómez (2024) advocate that rather than creating new rights, a better approach is to promote legislative reforms and establish international conventions. Regarding algorithmic bias, though the law mentions a right to be protected from biases, it lacks any specifications on how this right will be enforced or carried out. This discussion shows that though Chile's neurorights law has brought attention to the ethical complexities posed by neurotechnology, there is a need for more specific and precise legislation that regulates concrete problems, as well as a need for more integrated dialogue between policymakers, scientists, and philosophers in order to facilitate effective neurotechnology development.

Chile is not the only country to introduce legislation to protect against harmful consequences that neurotechnologies may pose. China's Ministry of Science and Technology has also recently rolled out ethical guidelines for BCI research, which outline basic principles that should be followed by researchers, such as safeguarding health, respecting the subjects involved, adhering to fairness and ensuring equity, information disclosure, and strict regulation (Caiyu, 2024; Center for Security and Emerging Technology, 2024). These guidelines were primarily created to control companies from doing BCI research without government approval, standardize BCI research, and serve as a reference for researchers and other institutions internationally. Stressing a need for transparency and accountability and bringing attention to the sensitive nature of neural data collected by BCIs, China's ethical guidelines are certainly a step forward towards responsible BCI research; however, besides advocating that BCI research should primarily focus on restorative technologies rather than non-medical enhancement of brain function, it has no direction on how researchers should approach developing fair, equitable technology that takes into consideration the needs of those with disabilities. Additionally, though the guideline was developed by the Artificial Intelligence Ethics Subcommittee of the National Science and Technology Ethics Commission, it does not contain any mention of the potential effects of algorithmic bias or how it can be mitigated. The necessity for greater emphasis on the needs of individuals with disabilities and on addressing the impacts of algorithmic bias is further demonstrated by the European Disability Forum, which has called for a resolution of the European Union's Artificial Intelligence Act (Felix, 2023). The European Disability Forum emphasizes the importance of a resolution that would include individuals with disabilities and promote AI development that involves input from accessibility professionals. While the EU AI act is an important example of a national legislation that categorizes AI into risk levels and aims to begin facilitating a human-centered approach to AI regulation, it simultaneously serves as yet another reminder that the needs of those with disabilities must be brought to attention in regulatory documents pertaining to AI and neurotechnology.

The path forward: recommendations for responsible direct-speech BCI development

Addressing ethical concerns stemming from algorithmic bias in neurotechnologies may be considerably more challenging as compared to other AI applications for which it is already difficult to correct bias, such as facial recognition software, search engines, hiring algorithms, and healthcare tools (Johnson, 2021). This ethical complexity is largely because the majority of the signals that direct-speech BCIs collect are unconscious or outside of the direct control of the user, making it technically challenging to pinpoint the data that is collected and analyzed by the algorithm (Berger & Rossi, 2023). Moreover, troubleshooting and improving a BCI algorithm will inevitably involve analyzing user brain data, but as Jwa & Poldrack (2022) remark, the personal nature of brain data raises significant concerns in regard to privacy and mental integrity, as researchers would potentially gain access to intimate, personal thoughts that a user didn't intend to disclose. This paper's analysis of current ethical guidelines has illustrated that while countries around the globe have begun making efforts to control negative consequences of neurotechnology usage, there remain significant issues that must be addressed with respect to direct-speech BCIs.

Based on my examination of the ethical issues associated with direct-speech BCI and existing solutions to these issues proposed by scholars and by legislation, I will now outline recommendations for combating the identified gap in ethical guidelines and creating a more comprehensive, equitable approach to correcting algorithmic bias in direct-speech BCI.

A. Considering a Crip Linguistics Perspective

A framework calling for a bigger and more flexible understanding of language to include the language used by those with disabilities, known as Crip Linguistics, argues that instead of disabled people needing correction of their language styles, it is existing social perceptions of disabilities that must change (Henner & Robinson, 2023). Crip Linguistics challenges traditional notions of language by placing emphasis on valuing nonverbal, alternative, and adapted speech patterns, and recognizing them not as deviations to be corrected, but as their own form of language. Extending the arguments of Crip Linguistics to direct-speech BCIs, it is crucial to recognize that training BCI algorithms solely on language models of conventionally-accepted speech perpetuates the narrow framing of language into a single accepted form. Correcting algorithmic bias by identifying and correcting existing ableist assumptions of disabled language with the input of individuals of communicative disabilities should be a primary consideration in the development of direct-speech BCIs. It is only through an increased emphasis on understanding and accepting the language styles of those with communicative disabilities–accomplished through more rigorous experimental designs and further research on different gesticulation and communication styles–that equitable direct-speech BCIs can be implemented in day-to-day life.

B. User-Centered Design: Designing with, Not Designing For

Existing literature and guidelines already suggest a need for end-user representation in neurotechnology development; I argue that the inputs of end-users should not simply be "represented", but rather, that end-users should be at the core of the direct-speech BCI design process, providing input for each aspect of the technology. Having a representative group of users providing input and interacting with researchers at each stage of the design process, a concept known as user-centered design (UCD) (Nijboer, 2015), would serve as a corrective for BCI developers whose training puts them at risk for a selection bias in making predictions about end user priorities and needs. One such example of UCD is in the BrainGate system, a BCI that enables individuals with tetraplegia to control a robotic arm. BrainGate was built off of integrating user feedback to refine movement control strategies, improve comfort, and address usability concerns that were not initially apparent, and has enabled improved communication rates (Rubin & Hochberg, 2023). The UCD process should also include an emphasis on end-user literacy, ensuring that end-users fully understand the neurotechnology's capabilities, what data it will be collecting from them, and what the potential ethical concerns with the technology are.

UCD should also draw on insights from other domains of assistive technology that enable a customization of communication parameters. Sankaran et al. (2023) describe existing customizations used in assistive technologies, such as controlling how adaptive a language model is to a user's long-term communication pattern, altering the amount of input allowed from a language model to increase speed as opposed to prioritizing accuracy, and adding common "high-utility" (p. 5) phrases for reliable access. Incorporating such capabilities for customization into direct-speech BCI will be essential to create more accurate, ethical technology centered on user needs.

Conclusion: Where do we go from here?

This paper's argument for increased attention on the needs of direct-speech BCI users with communicative disabilities points to the need for a foundational change in how researchers approach neurotechnology development. Starting with an analysis of existing perspectives on the ethical concerns with direct-speech BCI technology, and moving to an examination of existing suggestions to address these concerns, my research has demonstrated a gap in current ethical guidelines, and informed my development of recommendations to address this gap. At the forefront of advances in AI and neurotechnology, direct-speech BCIs pose a significant challenge for ethical governance; we must embrace this challenge and take it as an opportunity to set a precedent for responsible innovation. Improving inclusion for individuals with communicative disabilities is the only way to harness the full potential of BCIs to build a better society – one in which those with communicative disabilities can "stay connected to the bigger world, perhaps continue to work, [and] maintain friends and family relationships", as Pat Bennett writes (as cited in Goldman, 2023).

Integrating diverse linguistic patterns into BCI algorithms will require significant advances in both computational modeling and dataset diversity. Training language models to recognize nonstandard speech patterns will be challenging and time-intensive, but a necessary step in order to ensure ethical BCI development. Balancing customization with usability poses another challenge, as overly complex interfaces may unintentionally create barriers for users. Addressing these issues will require interdisciplinary collaboration between linguists, researchers, and end-users in order to refine BCI algorithms.

Discussions of neuroethical challenges must continue at all levels–political, academic, and societal–in order to continuously iterate on frameworks and promote equitable use of neurotechnologies. Former Chilean Senator and lead sponsor of the Chilean neurorights bill Guido Girardi cautions that the state of neuroscience today is "similar to what atomic energy was in the 1950s. It may be used to develop a better society, but also to create weapons against humanity" (as cited in Strickland, 2021). We should not take Girardi's words lightly. It's time to take action: engaging with the current challenges posed by direct-speech BCIs will benefit not only BCI users with communicative disabilities, but humanity as a whole, equipping our society to tackle future ethical issues that will inevitably emerge with technological advancement.

References

Berezutskaya, J., Freudenburg, Z. V., Vansteensel, M. J., Aarnoutse, E. J., Ramsey, N. F., & van Gerven, M. A. (2023). Direct speech reconstruction from sensorimotor brain activity with optimized deep learning models.

Journal of Neural Engineering, *20*(5), 056010.

Berger, S., & Rossi, F. (2023). AI and neurotechnology: learning from AI ethics to address an expanded ethics landscape. *Communications of the ACM*, *66*(3), 58-68.

Bernal, G., Montgomery, S. M., & Maes, P. (2021). Brain-computer interfaces, open-source, and democratizing the future of augmented consciousness. *Frontiers in Computer Science*, *3*, 661300.

Borbón, D., & Ramírez-Gómez, J. A. (2024). Between politics and scholarship: the (un) settled debate over neurorights. *Frontiers in Political Science*, *6*, 1335561.

Broussard, M. (2023). More than a glitch: Confronting race, gender, and ability bias in tech. *MIT Press.*

Caiyu, L. (2024, February 15). *China rolls out ethical guidelines for Brain-Computer Interfaces Research*. Global Times. https://www.globaltimes.cn/page/202402/1306687.shtml

Caliskan, A. (2021, May 10). *Detecting and mitigating bias in Natural Language Processing*. The Brookings Institution, Artificial Intelligence and Emerging Technology (AIET) Initiative. https://www.brookings.edu/articles/detecting-and-mitigating-bias-in-natural-language-processing/

Chang, C. H., & Kidman, G. (2023). The rise of generative artificial intelligence (AI) language models-challenges and opportunities for geographical and environmental education. *International Research in Geographical and Environmental Education*, *32*(2), 85-89.

Chinese Communist Party Artificial Intelligence Ethics Subcommittee of the National Science and Technology Ethics Commission. (2024, March 11). *Ethics guidelines for Brain-Computer Interface Research*.

Center for Security and Emerging Technology (CSET). https://cset.georgetown.edu/publication/china-bci-ethics/

Cooney, C., Folli, R., & Coyle, D. (2018). Neurolinguistics research advancing development of a direct-speech brain-computer interface. *IScience*, *8*, 103-125.

Cornejo-Plaza, M. I., Cippitani, R., & Pasquino, V. (2024). Chilean Supreme Court ruling on the protection of brain activity: neurorights, personal data protection, and neurodata. *Frontiers in Psychology*, *15*, 1330439.

Davidoff, E. J. (2020). Agency and accountability: ethical considerations for brain-computer interfaces. *The Rutgers journal of bioethics*, *11*, 9.

Fitzgerald, B. (2006). The transformation of open source software. *MIS quarterly*, 587-598.

Felix, A. (2023, January 4). *Resolution on the EU Artificial Intelligence Act for the inclusion of persons with disabilities*. European Disability Forum. https://www.edf-feph.org/publications/resolution-on-the-eu-artificial-intelligence-act-for-the-inclusion-of-persons-with-disabilities/

Goldman, B. (2023, August 23). *Brain implants, software guide speech-disabled person's intended words to computer screen*. Stanford Medicine News Center. https://med.stanford.edu/news/all-news/2023/08/brain-implant-speech-als.html

Guarino, B. (2024, January 30). Elon Musk's Neuralink Has Implanted Its First Chip in a Human Brain. What's Next? *Scientific American,* https://www.scientificamerican.com/article/elon-musks-neuralink-has-implanted-its-first-chip-in-a-human-brain-whats-next/

Guzmán, L. (2022). Chile: Pioneering the protection of neurorights. *The UNESCO Courier*, *2022*(1), 13-14.

Hariz, M., & Blomstedt, P. (2022). Deep brain stimulation for Parkinson's disease. *Journal of internal medicine*, *292*(5), 764–778. https://doi.org/10.1111/joim.13541

Henner, J., & Robinson, O. (2023). Unsettling languages, unruly bodyminds: A crip linguistics manifesto. *Journal of Critical Study of Communication & Disability*, *1*(1), 7-37.

Jangwan, N. S., Ashraf, G. M., Ram, V., Singh, V., Alghamdi, B. S., Abuzenadah, A. M., & Singh, M. F. (2022). Brain augmentation and neuroscience technologies: current applications, challenges, ethics and future prospects. *Frontiers in Systems Neuroscience*, *16*, 1000495.

Johnson, G. M. (2021). Algorithmic bias: on the implicit biases of social technology. *Synthese*, *198*(10), 9941-9961.

Jwa, A. S., & Poldrack, R. A. (2022). Addressing privacy risk in neuroscience data: from data protection to harm prevention. *Journal of Law and the Biosciences*, *9*(2), lsac025.

Kellmeyer, P. (2018). Big brain data: On the responsible use of brain data from clinical and consumer-directed neurotechnological devices. *Neuroethics*, *14*(1), 83-98.

Kellmeyer, P. (2019). Artificial Intelligence in Basic and Clinical Neuroscience: Opportunities and Ethical Challenges.

Neuroforum, *25*(4), 241-250. https://doi.org/10.1515/nf-2019-0018

Klein, E., Kinsella, M., Stevens, I., & Fried-Oken, M. (2022). Ethical issues raised by incorporating personalized language models into brain-computer interface communication technologies: a qualitative study of individuals with neurological disease. *Disability and Rehabilitation: Assistive Technology*, 1-11.

Liang, P. P., Wu, C., Morency, L. P., & Salakhutdinov, R. (2021, July). Towards understanding and mitigating social biases in language models. In *International Conference on Machine Learning* (pp. 6565-6576). PMLR.

Mei, J., Luo, R., Xu, L., Zhao, W., Wen, S., Wang, K., ... & Ming, D. (2024). MetaBCI: An open-source platform for brain–computer interfaces. *Computers in Biology and Medicine*, *168*, 107806.

Montgomery, K. L., Yeh, A. J., Ho, J. S., Tsao, V., Mohan Iyer, S., Grosenick, L., ... & Poon, A. S. (2015). Wirelessly powered, fully internal optogenetics for brain, spinal and peripheral circuits in mice. *Nature methods*, *12*(10), 969-974.

Müller, O., & Rotter, S. (2017). Neurotechnology: Current Developments and Ethical Issues. *Frontiers in systems neuroscience*, *11*, 93. https://doi.org/10.3389/fnsys.2017.00093

Nijboer, F. (2015). Technology transfer of brain-computer interfaces as assistive technology: barriers and opportunities. *Annals of physical and rehabilitation medicine*, *58*(1), 35-38.

Peters, B., Eddy, B., Galvin-McLaughlin, D., Betz, G., Oken, B., & Fried-Oken, M. (2022). A systematic review of research on augmentative and alternative communication brain-computer interface systems for individuals with disabilities. *Frontiers in human neuroscience*, *16*, 952380.

Rainey, S., Martin, S., Christen, A., Mégevand, P., & Fourneret, E. (2020). Brain recording, mind-reading, and neurotechnology: ethical issues from consumer devices to brain-based speech decoding. *Science and engineering ethics*, *26*, 2295-2311.

Rajpura, P., Cecotti, H., & Meena, Y. K. (2023). Explainable artificial intelligence approaches for brain-computer interfaces: a review and design space. *arXiv preprint arXiv:2312.13033*.

Rubin, D. B., & Hochberg, L. R. (2023, February). BrainGate: An Intracortical Brain-Computer Interface for the Restoration of Communication and Functional Independence for People with Paralysis. In 2023 *11th International Winter Conference on*

Brain-Computer Interface (BCI) (pp. 1-3). IEEE.

Sankaran, N., Moses, D., Chiong, W., & Chang, E. F. (2023). Recommendations for promoting user agency in the design of speech neuroprostheses. *Frontiers in Human Neuroscience*, *17*.

Sententia, W. (2004). Neuroethical considerations: cognitive liberty and converging technologies for improving human cognition. *Annals of the New York Academy of Sciences*, *1013*(1), 221-228.

Snoeys, L., Vanhoof, G., & Manders, E. (2013). Living with locked-in syndrome: an explorative study on health care situation, communication and quality of life. *Disability and rehabilitation*, *35*(9), 713-718.

Soboroff, J. (2023). *Brain implant uses AI to give speech-disabled patients a voice*. TODAY.com. https://www.today.com/video/groundbreaking-device-uses-ai-to-help-speech-disabled-patients-200944709540

Strickland, E. (2021, December 18). *Worldwide campaign for Neurorights notches its first win*. IEEE Spectrum. https://spectrum.ieee.org/neurotech-neurorights

Sullivan, L. S., Klein, E., Brown, T., Sample, M., Pham, M., Tubig, P., ... & Goering, S. (2017). Keeping disability in mind: a case study in implantable brain–computer Interface research. *Science and engineering ethics*, *24*, 479-504.

Wakefield, J. (2021, April 9). *Elon Musk's Neuralink "shows monkey playing pong with mind."* BBC News. https://www.bbc.com/news/technology-56688812

Welker, Y. (2023). *AI Act and disability-centred policy: how can we stop perpetuating social exclusion?* OECD.AI. https://oecd.ai/en/wonk/eu-ai-act-disabilities

Willett, F. R., Kunz, E. M., Fan, C., Avansino, D. T., Wilson, G. H., Choi, E. Y., ... & Henderson, J. M. (2023). A high-performance speech neuroprosthesis. *Nature*, *620*(7976), 1031-1036.

Wolkenstein, A., Jox, R. J., & Friedrich, O. (2018). Brain–computer interfaces: Lessons to be learned from the ethics of algorithms. *Cambridge Quarterly of Healthcare Ethics*, *27*(4), 635-646.

Zúñiga-Fajuri, A., Miranda, L. V., Miralles, D. Z., & Venegas, R. S. (2021). Neurorights in Chile: Between neuroscience and legal science. *Developments in neuroethics and bioethics*, *4*, 165-179.