Enter the Matrix:

What the new brain-computer interfaces teach us about agency, privacy, and human subjectivity

Benjamin Dueck

Abstract:

Within the last decade, the global brain-computer interface (BCI) industry has experienced significant economic growth and captured the public's imagination through enthusiastic coverage in the popular media. Responding dialectically to this growing public discourse, this paper overviews the social, ethical, and philosophical aspects of BCIs that pose the greatest concern for information policymakers. By reading the industry's techno-economic trends against contemporary critical philosophy, I argue that these technologies enable new "posthuman" forms of surveillance and control that challenge democratic notions of freedom, privacy, and equality.

Keywords:

Brain-Computer Interfaces, Surveillance Capitalism, Philosophy, Neuroscience, Ethics.

DOI: 10.33137/ijournal.v6i2.36453

© 2021, Dueck, B. "Enter the Matrix: What the new brain-computer-interfaces teach us about agency, privacy, and the boundaries of human subjectivity. This is an Open Access article distributed under CC-BY.

Lana and Lilly Wachowski's classic science fiction film *The Matrix* (1999), depicts a dystopian future where human beings have become biological batteries for a race of intelligent machines. At some point during the early 21st century, these machinic overlords created the titular Matrix, a simulated reality in which humanity unknowingly participates while their biological energy is siphoned. In one of the film's most iconic scenes, the protagonist Neo awakens from virtual slumber in an alien pod filled with glowing amniotic fluid. As he regains consciousness, pulling glistening tendrils from his mouth, neck, and spine, the camera pans out to reveal thousands of synthetic wombs, organized like a cybernetic insect hive.

Evoking this scene in the recent *Hegel in a Wired Brain* (2020b), philosopher Slavoj Žižek poses an interesting question about the motivations of this machine civilization. If human beings function as mere energy sources for its operation, why go through the trouble of creating a simulated, intersubjective reality at all? Would it not be simpler to sustain the human body in a comatose state and harvest energy that way? The only satisfactory answer, Žižek argues, is that the Matrix feeds off human jouissance, deriving energy from our continuous silent enjoyment (p. 156).

In this way, the Matrix stands as an astute metaphor for the era of networked platform capitalism (Srnicsek, 2016), which functions by extracting surplus value from the surveillance and control of human desire. From meal delivery services and online dating apps to the growing Amazon empire, the present capitalist subject is incorporated into the circuits of production and consumption while they chase satisfaction and pleasure, paradoxically experiencing this unequal exchange as their "highest exercise of freedom" (Žižek, 2020a, p. 747).

A new kind of matrix...

In recent decades, the global brain-computer-interface (BCI)¹ industry has experienced significant economic growth and is projected to reach \$1.84 billion USD by the year 2023 (Knowledge Sourcing Intelligence LLP, 2017). A growing number of media sources have penned enthusiastic descriptions about the possibilities occasioned by these technologies: telepathic communication (Cuthbertson, 2019), the effortless control of devices and applications (Gonfalonieri, 2020), and the ability to share life experiences with loved ones as if they were digital photographs (Winkler and Austin, 2020). These idyllic promises, however intoxicating, raise a more interesting set of questions: what is concealed by the utopian vision of human-machine symbiosis that surrounds these technologies in the media? How will BCIs complexify the notions of agency, priva-

1 Sometimes referred to as neural-control interfaces (NCIs), mind-machine interfaces (MMIs), or direct neural interfaces (DNIs) (Žižek, 2020, p. 747).

cy, and equality foundational to democratic life?

This paper explores the emerging sector of brain-computer interface (BCI) technology, overviewing the social, ethical, and philosophical dilemmas brought to light by these devices. More specifically, I argue that the new forms of surveillance and control occasioned by these technologies call into question the nature of democratic individuality and make possible a new “post-human”² form of subjectivity that must be properly understood before information policymakers can regulate this industry effectively.

The paper is divided into three distinct sections. In the first, I provide an overview of the emerging BCI sector and its primary avenues of research and development. In part two, I take up ideas explored by philosopher Slavoj Žižek in his recent book *Hegel in a Wired Brain* (2020b) and examine how BCIs complexify our philosophical understanding of subjectivity, identity, and free will. In the final section, I describe some of the regulatory problems that information policymakers will have to consider as this technology becomes widespread in coming decades. While we are a long way away from a Matrix-style dystopia, the developmental vector of this industry and its impact on democratic life will be defined by the choices we make today.

Introducing the BCI industry

The term “brain-computer interface” originated in the neuroscientific literature of the 1970s and appeared in the English-speaking media for the first time in 1993 (Gilbert et al, 2019, p. 49). Wolpow (2012) defines the BCI as a technical system that measures electrical activity in the central nervous system (CNS) and converts this information into an artificial output. Neural engineering for BCIs—which can be invasive (requiring surgical implants) or non-invasive (such as wearable headsets)—can be utilized to “replace, restore, enhance, supplement...or improve” natural CNS output (p. 384).

While it will likely take decades before BCIs become part of everyday life, the industry is already showing steady signs of growth. A report published by Nature in 2017 tells us that

2 In this paper, I use the term “posthuman” to figuratively group the disruptions to the enlightenment-humanist notion of subjectivity likely to accompany BCI technology and its dissemination. In her pioneering overview of philosophical posthumanism, Francesca Ferrando (2019) charts how the concept of the posthuman has developed to “cope with the urgency for an integral redefinition of the notion of the human” occasioned by the technological, scientific, and social developments of the 21st century (p. 1). Though the various posthumanism(s) are often confused with the related transhumanist movement (which envisions how the human individual can be enhanced using science and technology), they differ in their critical treatment of the construct of rational, bounded individuality underlying the legal and regulatory structure of liberal democracies.

there are more than a dozen major companies working on neurotechnology that will mediate bidirectional information transfer between the human brain and a computerized device (Yuste et al., 2017, p. 160). The authors estimate that private-sector spending in this industry has already exceeded \$100 million USD per year, with major companies like Compumedics, Nihon Kohden Corporation, ANT Neuro, and Elon Musk's startup firm Neuralink accounting for most of this cash flow (p. 160). The direct to consumer (DTC) neurotechnology industry is even larger and is projected to surpass \$3 billion USD by 2020 (Wexler, 2019). A similarly significant flow of funds (about \$500 million) has been invested into the BRAIN initiative by the US Department of Health and Human Services to develop novel neurotechnology for biomedical purposes (US Department of Health and Human Services, 2020).

The function of BCIs

A report from Knowledge Sourcing Intelligence LLP (2017) explains that the primary focus of the BCI industry has been developing neuroprosthetics for people living with paralyzed limbs. For example, after participating in the US government's multi-institution BrainGate program, Dennis Degray—who has been paralyzed from the collarbones down for over a decade—had his motor cortex surgically implanted with Utah arrays, tiny squares of silicon joined with metal electrodes protruding from the skull (Corbyn, 2019). The Utah arrays function by recording electrical activity in areas of the brain responsible for limb movement. Using the device, Dennis is able to control a computerized joystick that allows him to send text messages, type with an on-screen keyboard, and purchase products using Amazon (Corbyn, 2019).

Aside from these therapeutic, biomedical applications, investors in the private sector have more ambitious plans for the development of BCIs. The most infamous and publicly discussed initiative has been Elon Musk's Neuralink, which is premised on the idea that BCIs can be used to achieve "symbiosis with artificial intelligence" and augment biological processes like memory, intelligence, and response time (Mull, 2019). Neuralink works via an invasive process called deep brain stimulation (DBS) and involves a surgical procedure "designed to be implanted in an outpatient setting, akin to LASIK in terms of speed and discomfort" (Dadia and Greenbaum, 2019, p. 188). According to Neuralink Corporation's whitepaper, this experimental procedure will involve a series of small, flexible electrode threads (made from a biocompatible polymer material) being "sewn in" to the brain by a neurosurgical robot (Musk and Neuralink, 2019). After the link is made, "full-bandwidth data streaming" can be carried out using a standard USB-C cable, which gives the technology an unprecedented degree of scalability (Musk and Neuralink, 2019).

While these claims might strike readers as farfetched, premature, or even delusional, there is no a priori reason that prevents this technology from being actualized in the future.³As public policymakers will eventually be tasked with regulating this rapidly evolving space, it is imperative that they take proactive steps today to imagine a world where the human brain is linked to a computational system. When it is possible to directly manipulate and decode the mental processes underlying decisions, emotions, intentions, and communication—not only for individuals, but for groups of diverse people—we may indeed become like gods; gaining the ability to heal the sick and pursue our highest human potentials. But left unchecked, these technologies pose the danger of exacerbating existing social inequalities, giving a host of public and private actors unprecedented access to the very contours of human subjectivity. In order to fully understand the dilemmas that policymakers will grapple with in the near future, it is important to reflect on the philosophical concerns brought to light by BCIs and envision how they destabilize not only our self-experience as free human individuals, but our very status as free human individuals.

Philosophical implications of the wired brain

In *Hegel In a Wired Brain* (2020b), philosopher Slavoj Žižek warns us of the threat that BCIs pose to human individuality by revisiting an experiment carried out at New York University in 2002. During the study—which was spearheaded by neuroscientist John K. Chapin to develop prosthetics for paralyzed people—scientists attached a computer chip directly to the brain of a rat, allowing them to operate the animal as if it were a remote-controlled toy (Žižek 2020b, p. 46). The computer chip functioned by stimulating the rat’s whisker receptors, sending signals to a mechanism used by the scientists to steer the animal across tunnels, ramps, and tree trunks (Guterman, 2002). For Žižek, the major philosophical questions raised here relate to the sense of agency experienced by a subject when its neurological processes are hijacked by an external entity. Did the rat experience its controller’s commands as something invasive or did the electrical signals manifest within its mind as an internal drive, or more disturbingly, an act of free will? (Žižek, 2020b, p. 46).

Žižek describes how when BCI’s are discussed in the mainstream media, they are often depicted as a bridge to what Ray Kurzweil calls “The Singularity”, a global space of shared, post-human awareness made possible by the exponential development of intelligent machines (Žižek, 2020a, pp. 745-746). A recent study by Gilbert et al. (2019) confirms this insight, reveal-

³ A Neuralink progress update, live streamed to YouTube last summer, shows that the technology has already been successfully embedded into the brains of pigs (Neuralink, 2020).

ing that 76.91% of articles covering BCIs from 1993-2017 portrayed the technology positively with 25.27% containing “overtly positive” utopian rhetoric. Although there has been much debate as to whether the utopian benefits of BCIs are exaggerated, Žižek (2020b) is most concerned with the political implications of the moment when a direct link is formed between human mental processes and a digital machine:

“When people reflect on the implications of BCI, they usually focus on how our immersion in Singularity will affect us, making us a homo deus: like a divine being...But we should rather take a step back and ask the question: who will control the chips in our brain which sustains the BCI?” (pp. 46-47).

Expanding on Zuboff’s (2019) well-known analysis of surveillance capitalism, Žižek describes how the prospect of a wired brain holds untapped potential for the monitoring and control of human behavior (Žižek, 2020a, p. 747). According to Žižek, when a society’s daily activities are permanently registered by a complex network of digital machines, it assumes the form of an inverted totalitarian police state where subjects experience their social control not as dominion from an overarching authoritarian agent, but as a state of uninhibited freedom (2020, p. 29).

In his public statements on the issue of neurological surveillance, Elon Musk insists that although BCIs like Neuralink will allow individuals to “register and/or share” their thoughts and feelings, they will have to provide meaningful consent for this to occur: “People won’t be able to read your thoughts—you would have to will it. If you don’t will it, it doesn’t happen. Just like if you don’t will your mouth to talk, it doesn’t talk” (Urban, 2017). This vague definition of willed consent echoes the logic utilized today by platform developers working in the data-collection industry. For example, a smartphone application asks its user for permission to access their device’s location, despite already possessing this access for all intents and purposes. Reflecting on the form consent may assume in the coming era of BCIs, Žižek (2020a) raises the disturbing possibility that the process of surveillance will be ambiguous, if not invisible:

“Is it not much more reasonable to surmise that I will not even be aware, when plugged into BCI, whether or not my inner life is transparent to others? In short, does BCI not offer itself as the ideal medium of (political) control of the inner life of individuals?” (p. 751)

Returning to the example of the remote-controlled rat, when an individual is connected to a BCI, to what degree will they be aware that their thoughts are being controlled and surveilled? In all

likelihood, it will be difficult to tell. In a 2016 study on deep-brain stimulation, a man using an electrode-based brain stimulator to treat his depression reported having a difficult time differentiating which of his actions originated from the device, his depression, and his own agency (Klein and Nam, 2016). When the limitations that define our status as individuals—the distance between thought and action, the ability to think and dream in the privacy of one’s mind, the necessity of communicating through symbolic representation—we effectively enter a new “posthuman” state of subjectivity. With these possibilities in mind, we can explore some of the steps that policymakers can take today to ensure that fundamental aspects of the human condition survive this technological transition.

Concerns for information policymakers

In a recent report for Nature, Yuste et al. (2017) describe a future where BCIs and wearable neurotechnologies become part of everyday life. While they praise the biomedical potential of these devices, they argue that it is crucial to remain aware of the ways they might exacerbate existing social inequalities. Policymakers working in the neurotechnology sector face a set of challenges that will only complexify as these devices become more sophisticated. For Dadia and Greenbaum (2019), the biggest roadblock is often the simple matter of determining an appropriate moment to act. They speak of the concept of “Neurohype” the tendency for the academic and popular press to exaggerate the feasibility of neurotechnology, leading to hastily made policy grounded in unsubstantiated claims (Dadia and Greenbaum, 2019). Another challenge is that these technologies are often intentionally classified as “recreational or wellness devices” to sidestep regulatory oversight, which contradicts their marketing as instruments of human enhancement (p. 187). Despite the regulatory complexity of this industry, there are three key areas that policymakers should consider when thinking about its future development.

1.) Protecting privacy and agency

Given the startling amount of information that can already be gleaned from the data trails of individuals, it is important for policymakers to ensure that individual privacy is protected when neural devices connected to the internet become commonplace. Yuste et al. (2017) describe a phenomenon called brainjacking, “the possibility of individuals or organizations...tracking or even manipulating an individual’s mental experience” (p. 161). This brainjacking need not be overt and could function in subtle ways, such as building “auto-complete” functions into neuro-user interfaces to shorten the gap between intention and action, gradually engineering an individual’s habit-

ual responses to the external world (p. 162).

Questions of agency will be further complicated if these devices are used for criminal activity. How for example, would a court determine agency in a situation where an individual commits a crime using their Neuralink device? Nakar et al. (2015) describe a hypothetical situation where a patient uses a BCI to operate a robotic arm and physically harms another person. In order to determine whether the action stemmed from individual intention or technological malfunctioning, police would need to have the right to access the suspect's neurological data, opening up a new domain for privacy violation (Nakar et al., 2015). In order to prevent situations like this one from occurring, Yuste et al. (2017) propose that citizens retain the right to opt-out of the collection of data as a default (p. 161). Additionally, there should be clear regulations controlling the way neural data is stored and processed, by enforcing open-source principles, or making use of block-chain-based "smart contracts" to make this data transparent and auditable (pp. 161-162).

2.) Augmentation and human enhancement

If the utopian rhetoric of Elon Musk is any indication, the coming neurotechnologies may deepen the chasm between those with and without access to technology. It is not hard to imagine a new digital divide forming around BCIs that allow individuals to augment their sensory and mental acuity for social, professional, and cognitive advantages. Even more troubling is the potential that these technologies will be utilized by military actors to engineer superhuman soldiers for use in combat settings. It is important that policymakers work imaginatively to separate truth from fiction and define the context in which forms of augmentation and enhancement can occur. Yuste et al. (2017) point towards current regulations surrounding gene-editing in human beings as a potential model. Indeed, notions of individualism differ between cultures, making it necessary to balance global guidelines with citizen participation at the local level (p. 162).

3.) Algorithmic bias

As scholars like Sofia Noble have already demonstrated, the algorithmic architecture underlying media ecosystems subject those living as minority groups to dehumanization and exploitation. In *Algorithms of Oppression*, Noble (2018) calls "the benign instrumentality of technologies" into question, examining how the racist and sexist nature of contemporary communication media has been fueled by decades of neoliberal technology policy (p. 29). In order to prevent these biases from being embedded into neurotechnology, it is essential that engineers incorporate ethical decision making into their research and development process, and do not simply "tack on" these considerations retroactively. Any BCI development must include the voices and

perspectives of user groups who are already marginalized. Indeed, these processes are time consuming and often go against the profit motives that fuel the technology industry, raising the need for active coordination and regulation at the international level.

Conclusion

In another famous scene from *The Matrix*, the character Cypher—one of the rebellious human beings who manages to escape his simulated reality—becomes disillusioned by the bleak nature of the world outside the Matrix. In a pivotal sequence, Cypher enters the Matrix to meet with Agent Smith—a program created by the machinic overlords to identify and eliminate rebellious humans—and cuts a deal with him, agreeing to sell out his liberated crew in exchange for a position of comfort within the virtual world. “Ignorance is bliss” he exclaims while marveling at the juicy tenderness of an unreal piece of steak.

When it comes to issues of privacy and surveillance, many of us feel more comfortable adopting the attitude of Cypher then facing the uncomfortable reality of our own subjection. If I feel as if I have nothing to hide, who cares if the private sector gets to survey my neurological data? In order to prevent a dystopic posthuman future of the kind Žižek describes from occurring, it is essential that the public begins to make their voices heard on the issue of neurotechnology. To facilitate this widespread discussion, researchers and technology analysts alike must work to provide the public with sober and digestible overviews of developments in the BCI sector. Scholars like Anna Wexler (2020) have already begun reflecting on the ethical, legal, and social issues surrounding neurotechnology, balancing academic rigor with issues relevant to the public interest. As Corbyn (2020) points out, although there is still time for society to ponder questions of neuro-technical ethics, the clock is quickly running down. It is likely that basic BCIs will become available for people living with disabilities within the decade and be made accessible to the public in the next 20 years (Corbyn, 2020). Although the societal, biomedical, and creative potentials of neurotechnologies are vast, in order to align them with our highest human potentials, we must take a step back from our utopian dreams and reflect on whether the desire to become posthuman benefits the common good.

References

Corbyn, Z. (2020, September 22). Are brain implants the future of thinking? *The Guardian*. <https://www.theguardian.com/science/2019/sep/22/brain-computer-interface-implants-neuralink-braingate-elon-musk>

Cuthbertson, A. (2019, September 10). Brain-computer interface will make people telepathic, *iJournal*, Vol 6, No. 2,

scientists say. *The Independent*. <https://www.independent.co.uk/life-style/gadgets-and-tech/news/brain-computer-interface-neuralink-elon-musk-telepathy-a9097821.html>

Dadia, T., & Greenbaum, D. (2019). Neuralink: The ethical ‘rithmetic of reading and writing to the brain. *AJOB Neuroscience*, 10(4), 187-189. <https://doi.org/10.1080/21507740.2019.1665129>

Ferrando, F. (2019). Philosophical posthumansim. *Bloomsbury Academic*.

Gilbert, F., Pham, C., Viaña L., & Gillam, W. (2019). Increasing brain-computer interface media depictions: pressing ethical concerns, *Brain-Computer Interfaces*, 6(3), 49-70. <https://doi.org/10.1080/2326263X.2019.1655837>

Gonfalonieri, A. (2020, October 6). What brain-computer interfaces could mean for the future of work. *Harvard Business Review*. <https://hbr.org/2020/10/what-brain-computer-interfaces-could-mean-for-the-future-of-work>

Guterman, L. (2002). Robots with whiskers: Remote-controlled rats give critics the creeps. *The Chronicle of Higher Education*, 48(37). <https://www.chronicle.com/article/robots-with-whiskers/>

Klein, E., & Nam, C. S. (2016). Neuroethics and brain-computer interfaces (BCIs). *Brain-Computer Interfaces*, 3(3), 123-125. <https://doi.org/10.1080/2326263X.2016.1210989>

Knowledge Sourcing Intelligence LLP. (2017). Brain computer interface market- Industry trends, opportunities and forecasts to 2023. Research and Markets. <https://www.researchandmarkets.com/reports/4451970/brain-computer-interface-market-industry#adaptive>

Mull, A. (2019, July 18). Elon Musk’s next wild promise. *The Atlantic*. <https://www.theatlantic.com/health/archive/2019/07/neuralink-elon-musk-silicon-valley-promises/594316/>

Musk, E., & Neuralink. (2019). An integrated brain-machine interface platform with thousands of channels. 1-12. <https://www.biorxiv.org/content/10.1101/703801v4>

Nakar, S., Weinberger, S. & Greenbaum, D. (2015). Legal and social implications of predictive brain machine interfaces: duty of care, negligence and criminal responsibility. *AJOB* 6(4), 40-42. doi: 10.1080/21507740.2015.1094558.

Neuralink. [Neuralink]. (2020, August 28). Neuralink progress update, Summer 2020. [video] YouTube. https://www.youtube.com/watch?v=DVvmgjBL74w&feature=emb_title

Noble, S. (2018). Algorithms of oppression: How search engines reinforce racism. New York University Press.

Srnicek, N. (2017). Platform capitalism. Polity Press.

Urban, T. (2017). Neuralink and the brain’s magical future. Wait But Why. <https://waitbutwhy.com/2017/04/neuralink.html>

US Department of Health and Human Services. (2020). The BRAIN Initiative. Retrieved November 26, 2020, from <https://braininitiative.nih.gov/about/overview>

Wexler, A. (2020). About Anna Wexler. Anna Wexler. <http://www.annawexler.com/about.html>

Wexler, A., & Reiner, P. B. (2019). Oversight of direct-to-consumer neurotechnologies. *Science*, 363(6424), 234–235. doi:10.1126/science.aav0223.

Winkler, R. & Austin, S. (2020, April 27). Five things Elon Musk said about linking brains to iJournal, Vol 6, No. 2,

computers. Wall Street Journal. <https://www.wsj.com/articles/five-things-elon-musk-said-about-linking-brains-to-computers-1492803134>

Wolpaw, J. (2012). The future of BCIs: Meeting the expectations. In: Wolpaw J., Wolpaw E.W., eds. Brain-computer interfaces: principles and practice. Oxford University Press; pp. 387-392.

Yuste, R., Goering, S., Arcas, B. A., Bi, G., Carmena, J., Carter, A., Fins, J., Friesen, P., Gallant, J., Huggins, J., Illes, J., Kellmeyer, P., Klein, E., Marblestone, A., Mitchell, C., Parens, E., Pham, M., Rubel, A., Sadato, N., ..., & Wolpaw, J. (2017). Four ethical priorities for neurotechnologies and AI. *Nature*, 551, 159-163. <https://www.nature.com/news/four-ethical-priorities-for-neurotechnologies-and-ai-1.22960>

Žižek, S. (2020a). The apocalypse of a wired brain. *Critical Inquiry*, 26(4), 745-763. <https://www.journals.uchicago.edu/doi/pdf/10.1086/709222>

Žižek, S. (2020b). *Hegel in A wired brain*. Bloomsbury Academic.

Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. Profile Books. <https://doi.org/10.1080/24701475.2019.1706138>