Artificial Intelligence & Brain Implants: What Could Phenomenologically Go Wrong?

Frederic Gilbert (University of Washington)

Many experimental trials are currently testing in human novel medical brain implants operated by Artificial Intelligence (AI) (Reardon 2017; Ezzyat et al 2018). While these new generations of implantable AI-controlled brain devices moves rapidly ahead, ethical concerns about their potential effects on patients' sense of self, autonomy and identity is growing (Klein 2015; Brown et al. 2016; Kellmeyer et al. 2016; Goering et al 2017). AI-operated implants offer great control at the level of neural circuits, but the extent to which this grasp on neuronal function affects the patient's sense of control at the psychological level is still uncharted territory (Glannon and Ineichen 2016). A pressing ethical concern to explore is: how artificially intelligent brain implants affect the phenomenology of agency. Do they induce unprecedented vulnerabilities?

The purpose of this presentation is to explore the phenomenology of AI-controlled brain devices; in particular, we will use some results we obtained from a first-in-human trial involving AI-brain implant to make our case (Gilbert et al. 2017).

Overall, our study show that, on the one hand, artificially intelligent brain implants can positively increase a sense of the self and control; on the other hand, they can induce radical distress, feelings of loss of control, and a rupture of patient identity. Surprisingly, both observed scenarios generate an ethical issue: AI-operated implants seem to introduce an unprecedented decision-making vulnerability. We examine this phenomenon and discuss its ethical implications.

References

- Brown, T., Thompson, M. C., Herron, J., Ko, A., Chizeck, H., & Goering, S. (2016). Controlling our brains—A case study on the implications of brain computer interface-triggered deep brain stimulation for essential tremor. Brain Computer Interfaces, 3(4), 165–170.
- Ezzyat et al. (2018) Closed-loop stimulation of temporal cortex rescues functional networks and improves memory. Nat Commun. 2018 Feb 6;9(1):365. doi: 10.1038/s41467-017-02753-0.
- Gilbert et al (2017) Embodiment and Estrangement: Results from a First-in-Human "Intelligent BCI" Trial. Science and Engineering Ethics. DOI 10.1007/s11948-017-0001-5.
- Glannon, W., & Ineichen, C. (2016). Philosophical aspects of closed-loop neuroscience. In A. El Hady (Ed.), Closed loop neuroscience (pp. 259–270). London: Elsevier.
- Goering, S., Klein, E., Dougherty, D., & Widge, A. (2017). Staying in the loop: Relational agency and identity in next-generation DBS for psychiatry. AJOB Neuroscience, 8(2), 59–70.
- Kellmeyer, P., Cochrane, T., Müller, O., Mitchell, C., Ball, T., Fins, J. J., et al. (2016). The effect of closed-loop medical devices on the autonomy and accountability of persons and systems. Cambridge Quarterly of Healthcare Ethics, 25(4), 623–633.
- Klein, E. (2015). Informed consent in implantable BCI research: Identifying risks and exploring meaning. Science and Engineering Ethics, 22(5), 1299–1317.
- Reardon, S., (2017) AI-controlled brain implants for mood disorders tested in people, Nature 551, 549–550