The Foundational Role of Autonomy for Brain-Computer Interfaces

Dr. Florian Richter

florian.richter@thi.de

Technische Hochschule Ingolstadt

AImotion Bavaria

1. Introduction – Problem Situation

- Invasive brain-computer interfaces (BCIs) present a challenge to describe human-system interactions because such systems have a depth of engagement that can hardly be disentangled, which is particularly problematic in cases where clear authorship of decisions is needed
- This presupposes that such devices might also decide autonomously in certain situations.

2. Objective

- Although the use and implementation of such devices are highly disputed in practical ethics, they seem at least to offer therapeutic benefits for persons with, e.g., Parkinson's disease.
- That is why it is indispensable to develop a conceptual framework that allows a description of the system's interactions with the human agent, particularly considering the concept of "autonomy."

3. Methodological Background

1. Approaches in General:

The effects of brain-computer interfaces on patients are assessed empirically in studies and hypothesized on a more speculative level via thought experiments.

2. Model from Schönau et al. (2021):

They proposed a model to operationalize dimensions of agency, such as responsibility and privacy.

Values "Dimensions"	Criteria "Agency Competency"	Indicators "Agency Inquiry Prompts"
"Responsibility"	"Exercising Control"	"Do you feel that the (device mediated) movements you perform are under your intentional control?" "Do you sometimes feel that you share control with the AI/device?"
"Privacy"	"Negotiating Access"	"Do you feel that you are active in determining who has access?"
"Trust"	"Fostering Self-Trust"	"Do you trust yourself when you are using the device?"

Table. Schönau et al. (2021, 182) call it "Qualitative Agentive Competency Tool (Q-ACT)". This is just a selection from their table.

All dimensions of agency stand in relation to aspects of autonomy:

- independence to negotiating access
- self-dependence to fostering self-trust
- self-determination to exercising control

4. Challenge

- Nevertheless, models of agency are inchoate if they are not related to autonomy. It is furthermore widely assumed that autonomy is not only a characteristic of the human agent but also of the technological system (at least to a certain degree).
- Consequently, a more complete model needs to be developed by classifying different levels of autonomy (operational, strategic, and moral) that cover different forms of performances by systems and humans. This tripartite classification can be used to develop methods for evaluating the ethical side of BCIs.
- Technological systems can only operate on the first two levels because their behavior is based on dispositions, while human agency is based on deontic structures (expressed via normative attitudes).

- Disclosures: None; Conflict of Interest: None

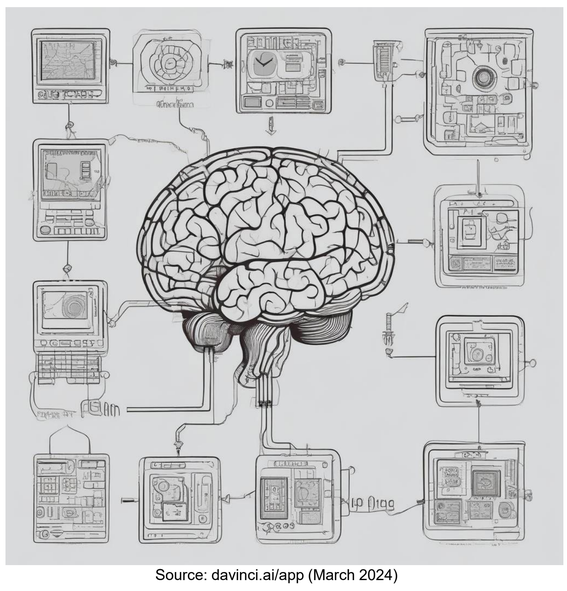
5. Types of Autonomy

Operational
- Systems have certain degrees of freedom to select the means for an objective to improve the effectiveness and efficiency of an action.
- Responsibility on the operational level can be almost completely delegated to systems within certain limits.

Strategic
- Systems have certain degrees of freedom in the selection of the optimal ends or objectives within a general framework of goals.
- Responsibility on the strategic level can be delegated partly to the systems; interventions by humans must be possible.

Moral
- Moral autonomy is based on freedom where one does not only know the goals but also recognizes the goals and principles.
- Responsibility of this type should not and cannot be delegated to the systems.

See Hubig (2019), 282/283



Source: davinci.ai/app (March 2024)

6. The Foundational Role of Autonomy

- Humans react or act autonomously on an operational and strategic level and not just on a moral level.
- Autonomy: set yourself in relation to performances, attitudes, claims, beliefs, etc.
- Transcendental claim: "The, I think, must accompany all my representations," as Immanuel Kant stated. (*Critique of Pure Reason*, B 131/132)
- Relations to attitudes and dissociations, such as "I feel like a robot," can only be experienced if a self is presupposed.

7. Use Cases

1. Epileptic Seizures: Safety and Responsibility

- Warning when an epileptic seizure might occur: gain in control and improvement of quality of life.
- The patient can either cease certain activities to be safer or decide to use the medicine.
- While it is an advantage to enhance the strategic autonomy of the patient "with respect to her seizures" by giving "the patient as much knowledge about the state of her system as possible," it also has "the disadvantage of requiring the patient to act in a way that maximizes safety." (Kellermeyer et al., p. 627)
- The person who is in the loop has more *strategic* autonomy but thus also more responsibility. In the case that the person is out of the loop, it decreases her autonomy to an *operational* level, and a certain form of "informed consent" is needed in the beginning.

2. ALS: Control and Privacy

- Certain devices can preserve or regain personal autonomy by enabling control for patients with neurodegenerative diseases such as amyotrophic lateral sclerosis (ALS).
- The system only reacts to the input of the patient, and this input is trained to improve its performance. This is, therefore, more of an assist system that functions on an *operational* level of autonomy, yielding an increase in the personal autonomy or control of the patient.
- Nonetheless, it seems to pose a threat to the privacy of the patient. One could imagine a case in which the system detects and outputs thoughts of the person that should not be heard by others. This is a hypothetical situation that goes beyond the operational level of such systems.
- However, on the operational level, the collection of the data does not pose so much of a threat directly to the patient as long as sufficient security measures are taken that there might be no leakage of the data through, for instance, hacking.

References

- O. Friedrich, E. Racine, S. Steinert, J. Pömsl and R. J. Jox, "An Analysis of the Impact of Brain-Computer Interfaces on Autonomy," *Neuroethics*, vol. 14, pp. 17-29, 2021.
- F. Gilbert, "A Threat to Autonomy? The Intrusion of Predictive Brain Implants," AJOB Neuroscience, vol. 6, no. 4, pp. 4-11, 2015.
- C. Hubig, *Die Kunst des Möglichen II: Ethik der Technik als provisorische Moral*, Bielefeld, 2007
- C. Hubig, *Die Kunst des Möglichen III: Macht der Technik*, Bielefeld, 2015.
- C. Hubig, "Haben autonome Maschinen Verantwortung?," in *Autonome Systeme und Arbeit: Perspektiven, Herausforderungen und Grenzen der Künstlichen Intelligenz in der Arbeitswelt*, H. Hirsch-Kreinsen and A. Karacic, Eds., Bielefeld, 2019, pp. 275-298.
- I. Kant, *Kritik der reinen Vernunft*, Hamburg, 1998.
- P. Kellmeyer, T. Cochrane, O. Müller, C. Mitchell, T. Ball, J. J. Fins and N. Biller-Andorno, "The Effects of Closed-Loop Medical Devices on the Autonomy and Accountability of Persons and Systems," *Cambridge Quarterly of Health Care Ethics*, vol. 25, no. 4, pp. 623-633, 2016.
- F. Kraemer, "Authenticity or autonomy? When deep brain stimulation causes a dilemma," *Journal of Medical Ethics*, vol. 39, pp. 757-760, 2013.
- A. Schönau, I. Dasgupta, T. Brown, E. Versalovic, E. Klein and S. Goering, "Mapping the Dimensions of Agency," *AJOB neuroscience*, vol. 12, no. 2-3, p. 172–186, 2021.