

Ethical Challenges at the Intersection of Artificial Intelligence and Neurotechnology

A. V. Shenderyuk-Zhidkov
Baltic Center for Neurotechnology
and Artificial Intelligence,
Immanuel Kant Baltic Federal University,
Kaliningrad, Russia

Vladimir A. Maximenko
Autonomous Nonprofit
Organization “Neimark”,
Nizhny Novgorod, Russia
maximenkovl@gmail.com

Alexander E. Hramov
Research Institute of Applied Artificial
Intelligence and Digital Solutions,
Plekhanov Russian University of Economics,
Moscow, Russia
hramovae@gmail.com

Abstract—The rapid co-evolution of artificial intelligence (AI) and neurotechnology presents unprecedented opportunities and challenges. This synthesis addresses key ethical issues emerging at their intersection, including data privacy, autonomy, fairness, transparency, and identity. We emphasize the unique risks neurotechnologies pose beyond AI, such as direct intervention in neural processes and threats to mental integrity, alongside shared concerns like bias and accountability. Recommendations focus on proactive regulatory frameworks, interdisciplinary collaboration, and the prioritization of human dignity to ensure responsible and equitable technological development.

Index Terms—artificial intelligence, neurotechnology, legal regulatory standards, ethical dilemmas

I. INTRODUCTION

Artificial intelligence (AI) and neurotechnology are transformative forces reshaping contemporary society. AI systems, particularly those based on machine learning and deep neural networks, have advanced capabilities in perception, decision-making, and language generation. Concurrently, neurotechnologies — encompassing brain-computer interfaces (BCIs) [1], neuroimaging [2], and neuromodulation [3], [4] techniques — enable direct interaction with, and modulation of, the human nervous system. These technologies are increasingly intertwined, with AI driving new neurotechnological applications and neurodata fueling AI models. Currently, this convergence is most visible in biomedicine [5], [6], but it is increasingly evident that these technologies are actively entering the everyday world through neuromarketing [7], gaming applications [8], advertising [9], education [10] etc.

However, this convergence entails significant ethical challenges [11]–[13]. Whereas AI primarily operates through analyzing external data and indirect influence, neurotechnologies have the unprecedented potential to intervene directly in cognition, emotion, and identity. This raises novel questions regarding personal autonomy, mental privacy, and societal impacts that extend beyond traditional AI ethics concerns.

This synthesis offers a concise overview of the emerging ethical issues at the nexus of AI and neurotechnology, highlighting their distinct and overlapping risks, and proposing avenues for their governance.

II. ETHICAL CHALLENGES IN AI AND NEUROTECHNOLOGY CO-EVOLUTION

A. Data Privacy and Security

Both AI and neurotechnologies depend heavily on large datasets, including sensitive personal and neural information. Neurodata, such as EEG or fMRI signals, reveal intimate details of mental states and cognitive functions, amplifying privacy concerns beyond typical personal data handled by AI systems. Unauthorized access to such data risks discrimination, psychological harm, and manipulation.

Ensuring robust data protection requires advanced anonymization, secure storage, and clear consent frameworks that inform users about neurodata usage. Regulatory gaps particularly exist around neurodata, necessitating specific legal provisions distinct from conventional data privacy laws.

B. Autonomy, Agency, and Mental Integrity

While AI systems may influence behavior indirectly via tailored content or recommendations, neurotechnologies can modulate neural activity in real time, impacting mental processes and self-perception. Interventions like deep brain stimulation or transcranial magnetic stimulation introduce questions about the preservation of individual autonomy and freedom of thought.

There is a risk that neurotechnological manipulation, especially if combined with AI-driven profiling, could covertly alter beliefs, preferences, or decisions without full user awareness or consent. The ethical imperative is to safeguard mental integrity and ensure technologies augment rather than undermine self-determination.

C. Fairness, Accessibility, and Social Justice

Technological advances often exacerbate existing inequalities if access is limited to privileged groups. AI systems trained on biased datasets can perpetuate discrimination in employment, credit, and justice contexts. Neurotechnologies, with their potentially enhancing effects, risk creating new forms of cognitive stratification—a “neuro-elite” advantaged by access to brain enhancement tools.

Moreover, the dominance of Western, Educated, Industrialized, Rich, Democratic (WEIRD) populations in data samples challenges the generalizability and fairness of AI

and neurotechnological applications globally. Equitable access and inclusive data collection strategies are critical to prevent marginalization and social division.

D. Transparency and Explainability

Complex AI models and neurotechnological interventions are often “black boxes,” limiting users’ and regulators’ understanding of decision-making processes and intervention mechanisms. Lack of explainability undermines trust, impedes accountability, and hinders informed consent.

Developing explainable AI (XAI) methods and transparent neurotechnology design are fundamental to ethical deployment. Users must be empowered with knowledge about how systems function and their implications for autonomy and rights.

E. Accountability and Governance

With increasing AI autonomy and the direct neural impact of neurotechnologies, it becomes difficult to assign clear responsibility for harms or malfunctions. Legal frameworks must clarify liability across developers, providers, and users.

Flexible governance incorporating principle-based regulations, risk assessment, and adaptive “sandbox” environments can balance innovation with protection. Cross-sector and international cooperation are vital for harmonized standards given the global reach of these technologies.

F. Identity and Human Dignity

Neurotechnologies have the extraordinary capability to shape cognition and personality, raising profound questions about what it means to be human. The alteration of memory, emotions, or behavior risks undermining personal identity and dignity.

Ethical approaches must prioritize respect for the intrinsic value of human life and promote technologies that enhance human flourishing without compromising core aspects of identity and humanity.

III. RECOMMENDATIONS FOR ETHICAL REGULATION

Addressing these multifaceted challenges demands an integrative approach:

- *Principle-based frameworks* that affirm human dignity, autonomy, justice, and transparency as foundational.
- *Specialized neurodata protections* recognizing the sensitivity and uniqueness of neural information.
- *Mandatory labeling* of AI-generated content and clear disclosure of neurotechnological interventions.
- *International standards* promoting cooperation across jurisdictions and cultures, including religious and ethical perspectives.
- *Inclusive participation* of diverse stakeholders, especially marginalized groups, in development and governance processes.
- *Promotion of explainable and auditable technologies* to foster trust and accountability.

- *Prohibition or strict regulation* of technologies capable of covert manipulation of consciousness or violating mental privacy.
- *Ongoing interdisciplinary research* to monitor evolving risks and societal impacts.

IV. CONCLUSION

The co-evolution of AI and neurotechnology opens unprecedented horizons for human advancement but also challenges critical ethical principles and social structures. Recognizing the unique risks posed by direct neural interventions alongside the well-documented concerns in AI ethics highlights the urgency of developing nuanced governance frameworks.

Responsible innovation must place the human being at the center, ensuring technologies respect autonomy, privacy, fairness, and identity. Only through proactive, collaborative, and interdisciplinary efforts can society harness the benefits of AI and neurotechnology while mitigating their potential harms.

REFERENCES

- [1] A. E. Hramov, V. A. Maksimenko, and A. N. Pisarchik, “Physical principles of brain-computer interfaces and their applications for rehabilitation, robotics and control of human brain states,” *Physics Reports*, vol. 918, pp. 1–133, 2021.
- [2] P. A. Bandettini, “What’s new in neuroimaging methods?” *Annals of the New York Academy of Sciences*, vol. 1156, no. 1, pp. 260–293, 2009.
- [3] M. Hallett, “Transcranial magnetic stimulation: a primer,” *Neuron*, vol. 55, no. 2, pp. 187–199, 2007.
- [4] A. L. Benabid and et al., “Deep brain stimulation of the subthalamic nucleus for the treatment of parkinson’s disease,” *The Lancet Neurology*, vol. 8, no. 1, pp. 67–81, 2009.
- [5] O. E. Karpov, E. N. Pitsik, S. A. Kurkin, V. A. Maksimenko, A. V. Gusev, N. N. Shusharina, and A. E. Hramov, “Analysis of publication activity and research trends in the field of ai medical applications: Network approach,” *International Journal of Environmental Research and Public Health*, vol. 20, p. 5335, 2023.
- [6] V. Khorev, S. Kurkin, A. Badarin, V. Antipov, E. Pitsik, A. Andreev, V. Grubov, O. Drapkina, A. Kiselev, and A. Hramov, “Review on the use of brain computer interface rehabilitation methods for treating mental and neurological conditions,” *Journal of Integrative Neuroscience*, vol. 23, no. 7, p. 125, 2024.
- [7] S. J. Stanton, W. Sinnott-Armstrong, and S. A. Huettel, “Neuromarketing: Ethical implications of its use and potential misuse,” *Journal of Business Ethics*, vol. 144, pp. 799–811, 2017.
- [8] D. Marshall and et al., “Brain-computer interface-based gaming: A step towards real-world applications,” *IEEE Transactions on Computational Intelligence and AI in Games*, vol. 5, no. 2, pp. 164–176, 2013.
- [9] C. Luna-Nevarez, “Neuromarketing, ethics, and regulation: An exploratory analysis of consumer opinions and sentiment on blogs and social media,” *Journal of Consumer Policy*, vol. 44, no. 4, pp. 559–583, 2021.
- [10] V. V. Grubov, M. V. Khramova, S. Goman, A. A. Badarin, S. A. Kurkin, D. A. Andrikov, E. Pitsik, V. Antipov, E. Petushok, N. Brusinskii, T. Bukina, A. A. Fedorov, and A. E. Hramov, “Open-loop neuroadaptive system for enhancing student’s cognitive abilities in learning,” *IEEE Access*, vol. 12, pp. 49 034–49 049, 2024.
- [11] S. Ruiz, L. Valera, P. Ramos, and R. Sitaram, “Neurorights in the constitution: from neurotechnology to ethics and politics,” *Philosophical Transactions B*, vol. 379, no. 1915, p. 20230098, 2024.
- [12] M. Sample and E. Racine, “Pragmatism for a digital society: The (in)significance of artificial intelligence and neural technology,” in *Advances in Neuroethics: Clinical Neurotechnology Meets Artificial Intelligence*. Springer, 2021.
- [13] N. Savage, “How ai and neuroscience drive each other forwards,” *Nature*, vol. 571, p. S15, 2019.