European Journal of AI, Computing & Informatics

Vol. 1 No.1 2025

Review **Open Access**

Transhumanism and the Future of Consciousness: The Integration of Humans and Machines

Bat-Erdene Togooch [1,*] and Gansukh Norovsambuu [1]

[1] Graduate University of Mongolia, Sukhbaatar Square 20/6, Academic Palace, Ulaanbaatar, Mongolia

* Correspondence: Bat-Erdene Togooch, Graduate University of Mongolia, Sukhbaatar Square 20/6, Academic Palace, Ulaanbaatar, Mongolia

Received: 09 March 2025
Revised: 15 March 2025
Accepted: 25 March 2025
Published: 31 March 2025

Abstract: Transhumanism envisions a future where technology surpasses biological limitations, integrating artificial intelligence (AI), brain-computer interfaces (BCIs), and advanced prosthetics into human life. This paper explores how these advancements impact consciousness, society, and ethics. While AI exhibits remarkable computational capabilities, it remains fundamentally distinct from human cognition and self-awareness. The application of transhumanist technologies in healthcare, education, and industry has the potential to transform human capabilities, yet it also introduces profound ethical dilemmas, including disparities in access, privacy concerns, and existential risks. This paper argues that a multidisciplinary approach — incorporating neuroscience, philosophy, and technology policy — is essential to ensure that the integration of humans and machines benefits society as a whole.

Keywords: transhumanism; artificial intelligence; consciousness; brain-computer interface; ethics; human-machine integration

1. Introduction

Transhumanism is no longer confined to speculative fiction — it is a tangible movement seeking to enhance human capabilities through technological means. Brain-computer interfaces (BCIs), AI-driven prosthetics, and neural augmentation are reshaping the boundaries between biological and artificial intelligence. However, the question remains: Will these advancements simply amplify human abilities, or will they redefine human existence itself?

Consciousness — the deeply subjective experience of awareness and self-perception — remains an unsolved mystery in neuroscience and philosophy. AI has demonstrated its ability to mimic certain cognitive functions, yet it lacks self-awareness, emotions, and the intrinsic sense of existence that defines human consciousness. If transhumanist technologies advance far enough, could we eventually create machines that "think" or "feel" as humans do? And if human cognition becomes augmented by AI, will the line between organic and artificial intelligence blur?

This paper aims to examine these critical questions by drawing from neuroscience, cognitive science, and ethics. It will explore how AI and brain augmentation technologies impact human cognition, analyze the potential benefits and dangers of transhumanist advancements, and propose strategies for responsible technological integration.

2. The Brain, Consciousness, and AI

2.1. Understanding the Brain and the Nature of Consciousness

The human brain, with its approximately 86 billion neurons, operates as a highly complex network, coordinating thoughts, emotions, and bodily functions [1]. While neuroscience has provided detailed insights into neural mechanisms, the nature of consciousness remains one of the most debated topics in science and philosophy.

Philosophically, dualist theories, such as those proposed by Descartes (1641), suggest that the mind and body are separate entities. In contrast, materialist perspectives, such as those advanced by Searle [2], argue that consciousness arises purely from physical processes within the brain. More radical theories like panpsychism propose that consciousness is a fundamental property of the universe, potentially existing even in non-biological entities [3].

Neuroscientific research indicates that consciousness is closely associated with the cerebral cortex and thalamus [4]. Studies on coma patients reveal that disruptions in these areas significantly impact awareness. However, while scientists can correlate brain activity with conscious experience, they have yet to explain why neural processes generate subjective experience — a challenge known as the "hard problem of consciousness" [5].

If consciousness is purely a product of neural computation, could an artificial system achieve self-awareness by replicating these processes? Or is there an intrinsic biological component that machines can never replicate?

2.2. AI: Mimicking Cognition Without Awareness

AI has demonstrated remarkable capabilities in specific cognitive tasks, surpassing human performance in fields such as medical diagnosis and strategic gaming. For example, DeepMind's AlphaGo defeated a world champion by recognizing complex patterns beyond human intuition [6], and AI-driven medical systems can detect cancer with 94% accuracy.

However, these systems operate within a predefined scope — known as "narrow AI" — and lack the general intelligence and self-awareness associated with human cognition. Searle's "Chinese Room" argument illustrates this distinction: an AI might convincingly process language, but it does not understand meaning the way a human does [2].

Recent advancements in neuromorphic computing aim to replicate the architecture of the human brain using artificial neurons [7]. Quantum computing offers another potential breakthrough, with the ability to process vast amounts of information simultaneously [8]. Additionally, bioengineered neural networks — lab-grown "mini-brains" — are being explored as a hybrid approach to artificial cognition [9].

Yet, despite these advancements, AI remains fundamentally different from human cognition. While AI excels in pattern recognition, data processing, and predictive analytics, it lacks emotions, self-reflection, and intrinsic motivation. True artificial general intelligence (AGI), capable of independent thought and learning across diverse domains, remains theoretical.

3. Transhumanism: Potential and Challenges

3.1. Enhancing Human Abilities Through Technology

Transhumanist technologies are already improving human lives in profound ways. Brain-computer interfaces (BCIs) are enabling paralyzed individuals to regain motor control. For example, Neuralink has developed brain implants that restore movement by directly interfacing with neurons. Similarly, cochlear implants have restored hearing to over 700,000 individuals worldwide (NIDCD, 2023). Advanced prosthetics, such as AI-driven robotic limbs, now integrate with neural signals, allowing users to control them as if they were natural limbs.

Beyond medical applications, BCIs could enhance cognitive functions, improving memory, decision-making, and problem-solving skills. Exoskeletons, initially designed for rehabilitation, are now used in industries and the military to enhance human strength and endurance. Furthermore, gene-editing technologies like CRISPR hold promise for eliminating genetic disorders, potentially extending the human lifespan.

3.2. Ethical and Socioeconomic Concerns

Despite its transformative potential, transhumanism raises significant ethical and social concerns. One of the primary issues is inequality — advanced augmentations remain prohibitively expensive, leading to concerns about a growing divide between enhanced and non-enhanced individuals. For instance, in low-income countries, only 10% of those in need of prosthetic limbs have access, compared to 80% in developed nations (American Prosthetics Association, 2021). This disparity raises concerns about the emergence of a biotechnological elite.

Additionally, automation and AI-driven augmentation could disrupt labor markets. Studies predict that up to 800 million jobs could be eliminated by 2030 due to automation (McKinsey Global Institute, 2017). While new jobs may be created, economic disparities could widen before society adapts [10].

Another major concern is privacy and security risks. As brain-connected technologies become more prevalent, hacking BCIs could lead to manipulation of thoughts and memories [11]. Governments and regulatory bodies must establish strict cybersecurity measures to prevent these potential threats.

4. Risks and Safeguards

4.1. The Scary Bits: Ethical, Security, and Existential Risks

While transhumanist technologies promise to extend human capabilities, they also introduce significant risks, ranging from ethical concerns to security threats and existential dangers.

4.1.1. The Ethical Dilemma of Human Augmentation

One of the biggest ethical concerns in transhumanism is the divide between enhanced and unenhanced individuals. As technologies like brain-computer interfaces (BCIs), AI-driven cognitive enhancement, and genetic engineering become more advanced, those with access to these technologies may gain significant intellectual, physical, and economic advantages over others. This could lead to the emergence of a new social class hierarchy, where the enhanced enjoy privileges inaccessible to the unenhanced.

For example, CRISPR gene-editing technology has the potential to eliminate genetic disorders, but it could also be misused to create "designer babies" with superior intelligence, strength, or beauty. If only the wealthy can afford such enhancements, societal inequality could reach unprecedented levels, reinforcing existing power imbalances.

Additionally, the modification of human cognition through AI augmentation raises concerns about personal identity and autonomy. If an individual’s memory, thoughts, or cognitive processes are altered by an external AI system, to what extent do they retain their original sense of self? Scholars have debated whether neural implants could fundamentally alter human consciousness, effectively creating a new form of post-human intelligence that is no longer truly "human" [12].

4.1.2. Security Threats: The Risk of Hacked Minds

The integration of AI with neural interfaces introduces cybersecurity threats unlike anything humanity has faced before. While traditional cyberattacks target devices and networks, brain-connected implants could make human thoughts and emotions vulnerable to hacking [11]. Potential threats include:

Mind manipulation: If a brain-computer interface (BCI) can read and write neural signals, a hacker could implant false memories, alter emotions, or influence decision-making. Research has already shown that electrical stimulation of certain brain regions can alter moods, raising concerns about potential misuse in military, political, or commercial contexts.

Cognitive espionage: Just as hackers steal sensitive corporate data today, future attacks could involve extracting confidential information directly from a person's mind. Imagine a world where corporate secrets, military intelligence, or personal thoughts can be accessed by cybercriminals.

AI-powered surveillance and thought policing: Governments and corporations could use BCIs to monitor citizens' thoughts, potentially criminalizing certain ideas or dissenting opinions. While dystopian, this concern is not far-fetched — China has already tested AI-based emotional monitoring on workers to detect stress levels and improve productivity (Xinhua News, 2018).

Without strong cybersecurity measures, BCIs and AI-augmented cognition could become tools of oppression rather than liberation. Governments, tech companies, and regulatory bodies must develop international security protocols to prevent these threats.

4.1.3. The AI Control Problem and Existential Risk

A more long-term concern is the existential risk posed by superintelligent AI. While current AI systems are specialized (narrow AI), scholars such as Stuart Russell warn that if AI reaches general intelligence — surpassing human cognitive abilities — it could become uncontrollable [13].

Potential risks include:

AI misalignment: If an AI system is designed to optimize a specific goal without considering human values, it might take extreme actions to fulfill its objective. A well-known thought experiment is the "paperclip maximizer" scenario, where an AI tasked with producing paperclips optimizes so aggressively that it consumes all resources — ultimately leading to human extinction [12].

Loss of human control: Autonomous AI systems used in finance, healthcare, or military operations may begin making decisions beyond human comprehension. Automated weapons systems, for example, could escalate conflicts without human intervention, leading to unintended wars [13].

Economic disruption and mass unemployment: As AI continues to replace human labor, mass job displacement could destabilize economies and societies. The World Economic Forum predicts that AI-driven automation could eliminate 85 million jobs globally by 2025 while creating 97 million new roles, but the transition could leave millions unemployed in the short term (WEF, 2020).

Given these risks, researchers argue that strict AI alignment protocols, kill-switch mechanisms, and ethical AI development frameworks are necessary to prevent catastrophic outcomes.

4.2. Keeping It in Check: Policies and Ethical Safeguards

4.2.1. Global AI and Transhumanism Regulations

To ensure the safe and equitable integration of transhumanist technologies, governments must establish clear international regulations. Several initiatives are already in progress:

The EU AI Act (2021): This framework classifies AI applications into different risk levels, restricting high-risk technologies such as facial recognition and autonomous weapons (European Commission, 2021).

The UNESCO AI Ethics Guidelines (2022): These guidelines emphasize human dignity, privacy, and transparency, advocating for the responsible development of AI-driven augmentation technologies (UNESCO, 2022).

The UN's Call for a Ban on Autonomous Lethal Weapons: The United Nations has been pushing for a ban on autonomous weapons that can kill without human intervention, recognizing the dangers of AI-controlled warfare (UN Report, 2022).

However, enforcement remains a challenge, as many governments and corporations prioritize economic and military advantages over ethical concerns.

4.2.2. Economic and Social Policies for an AI-Integrated Future

To prevent extreme wealth inequality in a transhumanist future, economic reforms must ensure access to enhancements is not limited to the elite. Possible solutions include:

AI and automation taxation: Countries like Sweden have experimented with taxing AI-driven companies to fund worker retraining programs. A global AI tax could help offset the economic disruptions caused by automation.

Universal basic income (UBI): Some economists propose UBI as a safety net for workers displaced by AI. While controversial, pilot programs in Finland and Canada have shown improved well-being and economic stability among participants (OECD, 2020).

Subsidized transhumanist technologies: Governments could fund public access to BCIs, prosthetics, and AI-driven education tools, ensuring that enhancements are available to a broader population rather than an elite minority.

4.2.3. Ethical Frameworks for AI and Human Augmentation

The ethical implications of merging AI with human cognition demand strong interdisciplinary oversight. Bioethicists, philosophers, and neuroscientists must work alongside technologists to develop principles guiding the ethical use of augmentation technologies. Some proposals include:

Human autonomy safeguards: All brain implants and AI augmentation tools should have manual override mechanisms, ensuring individuals retain control over their cognitive functions.

Transparency in AI decision-making: AI systems affecting human cognition should be required to explain their decisions, preventing black-box scenarios where users don’t understand AI-driven alterations to their thoughts or emotions.

Strict data privacy laws: Governments should enforce stringent BCI data protection laws, ensuring that brain data is not exploited by corporations or governments.

4.2.4. Promoting Open-Source AI Development

One way to democratize AI and transhumanist advancements is to promote open-source AI models. By making AI research publicly accessible, independent researchers and smaller nations can compete with tech giants, reducing monopolistic control over AI-driven human augmentation. Organizations like OpenAI have already taken steps toward shared AI research, but further efforts are needed to balance innovation with responsible governance.

5. Conclusion

The integration of transhumanism and artificial intelligence presents a paradigm shift in human evolution, offering unprecedented opportunities alongside substantial risks. AI has demonstrated remarkable capabilities in cognitive tasks, yet it remains fundamentally distinct from human consciousness. While technologies such as Neuralink and AI-driven prosthetics have the potential to redefine healthcare, education, and human augmentation, they also raise ethical, economic, and existential concerns that cannot be ignored.

One of the most pressing challenges is ensuring equitable access to these advancements. As discussed, the financial burden of cutting-edge prosthetics, brain-computer interfaces, and AI-driven cognitive enhancement could exacerbate existing socio-economic inequalities, leading to a divide between the technologically enhanced elite and those left behind. Addressing this requires proactive policymaking, such as government subsidies for essential transhumanist technologies, ethical AI governance, and equitable education reforms to prepare the workforce for an AI-integrated future.

Moreover, the question of consciousness remains unresolved. While AI can simulate cognitive processes, it lacks the subjective experience that defines human awareness. Future research should focus on collaborations between neuroscientists, AI ethicists, and philosophers to deepen our understanding of what it means to be conscious and how human identity might evolve alongside machine intelligence. Interdisciplinary frameworks are essential to prevent a future where human augmentation is dictated solely by corporate and military interests, rather than collective human well-being.

Ultimately, transhumanism is not merely about technological advancement — it is about reshaping the essence of what it means to be human. The choices we make today will determine whether AI and augmentation technologies lead to a more inclusive and empowered society or a divisive and uncontrolled future. While uncertainties remain, one thing is clear: the future of human-machine integration is ours to shape, and the responsibility lies with all of us to ensure it unfolds in a manner that upholds human dignity, equity, and ethical integrity.

References

1. F. A. Azevedo, et al., "Equal numbers of neuronal and nonneuronal cells make the human brain an isometrically scaled-up primate brain," *J. Comp. Neurol.*, vol. 513, no. 5, pp. 532-541, 2009, doi: 10.1002/cne.21974.
2. J. R. Searle, "Minds, brains, and programs," *Behav. Brain Sci.*, vol. 3, no. 3, pp. 417-424, 1980, doi: 10.1017/S0140525X00005756.
3. P. Goff, *Galileo's Error: Foundations for a New Science of Consciousness*, Pantheon, 2019. ISBN: 9781524747961.
4. G. Tononi, M. Boly, M. Massimini, and C. Koch, "Integrated information theory: from consciousness to its physical substrate," *Nat. Rev. Neurosci.*, vol. 17, no. 7, pp. 450-461, 2016, doi: 10.1038/nrn.2016.44.
5. D. Chalmers, "The hard problem of consciousness," in *The Blackwell Companion to Consciousness*, pp. 32-42, Oxford, UK: Wiley-Blackwell, 2017. ISBN: 9780470674062.
6. D. Silver, et al., "Mastering the game of Go with deep neural networks and tree search," *Nature*, vol. 529, no. 7587, pp. 484-489, 2016, doi: 10.1038/nature16961.
7. C. D. Schuman, T. E. Potok, R. M. Patton, J. D. Birdwell, M. E. Dean, G. S. Rose, and J. S. Plank, "A survey of neuromorphic computing and neural networks in hardware," *arXiv preprint arXiv:1705.06963*, 2017, doi: 10.48550/arXiv.1705.06963.
8. J. Preskill, "Quantum computing in the NISQ era and beyond," *Quantum*, vol. 2, p. 79, 2018, doi: 10.22331/q-2018-08-06-79.
9. D. L. Yamins and J. J. DiCarlo, "Using goal-driven deep learning models to understand sensory cortex," *Nat. Neurosci.*, vol. 19, no. 3, pp. 356-365, 2016, doi: 10.1038/nn.4244.
10. N. Selwyn, *Should Robots Replace Teachers?: AI and the Future of Education*, John Wiley & Sons, 2019. ISBN: 9781509528950.
11. M. Ienca and P. Haselager, "Hacking the brain: brain–computer interfacing technology and the ethics of neurosecurity," *Ethics Inf. Technol.*, vol. 18, pp. 117-129, 2016, doi: 10.1007/s10676-016-9398-9.
12. B. Yvert and E. Fourneret, "Neuromorphic brain interfacing and the challenge of human subjectivation," *Nat. Rev. Bioengineering*, vol. 1, no. 6, pp. 380-381, 2023, doi: 10.1038/s44222-023-00041-9.
13. S. Russell, "Artificial Intelligence and the Problem of Control," in *Perspectives on Digital Humanism*, pp. 19-24, Springer, 2022. ISBN: 9783030861438.