

Artificial Intelligence, Cybernetics and Philosophy: Ramifications of Human - Machine Integration

Dr Nidhi Nema¹ , Dr Swati A. Sharma²

¹Assistant Professor (HoD, Dept. of English) Govt. Kamla Nehru Girls College, Balaghat (MP).

²Professor/Dean Sanjeev Agrawal Global Educational (SAGE) UNIVERSITY, BHOPAL

ABSTRACT:

Recent decades have witnessed remarkable advancements in Artificial Intelligence (AI) and Cybernetics, leading to a transformative convergence of human and machine capabilities. This abstract explores the intricate interplay between these domains, with a keen focus on the philosophical dimensions of this evolution.

AI, powered by deep learning, neural networks, and data analytics, is rapidly encroaching on tasks once deemed exclusive to human intelligence. It has profound implications for our understanding of consciousness, decision-making, and even the definition of creativity. Simultaneously, Cybernetics, which delves into the dynamics of control and communication within living organisms and machines, further deepens the man-machine relationship. Philosophy, as the reflective cornerstone of human thought, emerges as the arena for grappling with these transformations. Questions about the essence of personhood, ethics in AI-driven decision-making, and the very nature of intelligence become paramount. This study provides a brief overview of this interdisciplinary exploration, covering the historical development of AI and Cybernetics, ethical dilemmas, and the implications for philosophical discourse. Understanding the convergence of human and machine is not merely an academic endeavour but a profound reflection on the changing landscape of human existence. This research aspires to contribute to the ongoing dialogue about the future of AI, Cybernetics, and the philosophical inquiries that underpin this fascinating convergence.

KEYWORDS: Cybernetics, Posthumanism, Philosophy, Ethics.

Introduction

We have embarked on the next phase in cyborg development with the initiation of clinical trials of *Neuralink*¹, Elon Musk's venture to merge human brain's neurons with electronics to develop an enhanced brain – computer interface that can help biological brains to interact with external technology. Before analysing the biological, social and moral ramifications, let's grasp the trajectory it is taking and the overarching mission we are striving to accomplish. All the information would either be digitized into our consciousness or all of our consciousness would be digitized and could be broken into readable datasets. By taking measures to minimize disparities and harmonize the attributes of both humans and machines in the process, we aspire to create computers endowed with nervous systems and addressing issues of consciousness and identity, while concurrently enhancing human responses with mechanical precision and bestowing them with superhuman capabilities.

Founded by Elon Musk, Neuralink Corp. is an American neurotechnology company that is developing ¹ implantable brain-computer interfaces. (<https://neuralink.com>)



In the realm of modern science and technology, the convergence of neural implants and AI-enabled prosthetics represents a profound leap forward in human capability. This amalgamation of cutting-edge innovation is propelling us beyond the realms of Lewis Carol's whimsical imagination, William Gibson's fantastical exploration of cybernetic enhancements in his cyberpunks², or the cautionary tale of unchecked scientific experimentation in Frankenstein³. Instead, we are witnessing the dawn of a new era characterized by AI-human hybridization. Today, prosthetic limbs equipped with microprocessors and AI algorithms not only restore lost functionality but also enhance human abilities to unforeseen levels. This synergy between man and machine is, on one hand, a testament to human ingenuity and the potential for a harmonious coexistence of technology and biology, ushering in a future where limitations are redefined and human potential knows no bounds, and on the other hand, brings all our constructs – knowledge, business, society and self - into immediate reconsideration. AI and robotics have already sparked profound debates concerning their ethical use, their inherent behavior, and their enduring risks. They also disrupt the conventional human perception of ourselves as the superior and intelligent species on earth. While AI-human convergence is bound to be wrought with issues relating to legislation, technicalities, medical, political, and religious considerations, cybersecurity and privacy, that would undoubtedly demand reevaluation of all our existing regulatory systems and frameworks, we propose to focus on philosophical inquiries. This new development necessitates a philosophical examination that can inform our responses, offering insights not only for new challenges but also for traditional philosophical dilemmas.

The roots of cyborg⁴ ontology can be traced back to early notions of merging human and machine in popular imagination - myths, fiction and movies. Ancient myths and legends often portrayed humans with enhanced abilities or prosthetic limbs, blurring the line between natural and artificial. According to Adrienne Mayor (2018), the earliest notions of automata or cyborg- like organisms can be located in Greek, Roman, Indian and Chinese mythology. Mayor records *ideas* about making artificial life, beings that were “made, not born” and replicating nature in ancient Greek myths. However, it was not until the 20th century that the concept of cyborgs gained traction. Science fiction and films, such as *Robocop*, *Blade Runner* and *The Six Million Dollar Man*, explored the possibilities and consequences of human-machine integration, shaping our understanding of cyborg entity.

NASA consultants Nathan Kline and Manfred Clynes are attributed with coining the term cyborg in 1960. They published a paper called 'Cyborgs and Space' (Clynes & Kline 1960) where they proposed that rather than trying to construct earth-like environments for humans to live in, we should try altering the humans themselves. They proposed a Cyborg or Cybernetic Organism, that was self-regulating and would take care of the bothersome issues of breathing, metabolism, sleep, circulation, and any other biological necessities, “leaving man free to explore, to create, to think, and to feel.” Since then, the concepts like cyborgs, short for cybernetic organisms, biomechanics, trans humanism, bio hacking have emerged as central theme of discussion or area of research in the circles of future science and technology. These subjects have also prompted investments in philosophy, among other fields of knowledge, as philosophers are curious to analyse the foundational and complex nature of existence of these newer entities and are delving deeper into an exploration of the nature of their existence and the construction of their identity. These concepts blur the traditional boundaries between humans and machines, raising profound questions about the nature of its ontology and identity. By examining historical perspectives, the mind-body problem, the influence of technology and epistemological implications and ethical considerations, it is imperative to shed light on the intricate aspects of entities born out of human – machine integration.

Gibson is a widely recognized pioneer of the science fiction subgenre known as cyberpunk. Gibson is an ² American-Canadian author of essays and speculative fiction. His early writings, which were noir, near-futuristic tales that examined how technology, cybernetics, and computer networks affected people. In the titular novel (1818) written by Mary Shelley, narrates the havoc created by the ‘sapient monster’ created ³ by the unchecked scientific experiment of young Victor Frankenstein. The term “cyborg” is formed with combination of the terms “cybernetic” and “organism” ⁴



One of the central challenges in cyborg ontology lies in addressing the mind-body problem. Traditionally, Cartesian dualism has posited a strict separation between the mind and the body. It initiated the notion of difference between mind and body, where body was treated as physical and mind/reason was regarded as disembodied hence transcendental. But human-machine integration disrupts the dichotomy of mind and body as it embodies the fusion of human consciousness and technological components, though such integration raises serious questions about the nature and location of consciousness, subjective experiences, and the boundaries of personal identity. Further, the notion of embodied cognition, emphasizes the role and importance of an agent's physical body in their cognitive capacities and, in various ways, rejects or reformulates the computational commitments of cognitive science. The premise that the body or the interactions of the body with the environment comprise or contribute to cognition in ways that necessitate a new framework for its investigation unites researchers of embodied cognition. Not all mental processes involve calculation or just computation. The seat of cognition or a computer are not the brain.

Building upon the idea of embodied cognition, the extended mind theory posits that cognitive processes can extend beyond the biological brain and incorporate external tools and technologies. We are, in fact, 'human-technology symbionts'. Andy Clark's thesis underscores the significance of this symbiotic relationship, highlighting that our distinct humanity is defined by our remarkable capacity to engage in thinking that is both influenced by and in collaboration with our technologies. Cybernetics embody this extended mind perspective, as it relies on technological devices and interfaces to enhance cognitive abilities. This raises profound questions about the nature of personal identity. If we augment our cognitive or physical abilities with technological enhancements, does it alter our sense of self? What does it mean to be an authentic human in a world where human-machine interactions are increasingly prevalent? Philosophical inquiries into personal identity and authenticity explore how the blurring of human and machine boundaries challenges our understanding of what it means to be human and the continuity of personal identity over time.

So now the next bigger question here is - can machines possess consciousness? As AI systems become more sophisticated, capable of learning, reasoning, and mimicking human behaviour, the philosophical debates explore whether consciousness is solely a product of biological processes or if it can emerge in artificial systems. The exploration of machine consciousness raises fundamental questions about the nature of subjective experience and the boundaries of personhood. If external technologies become integral to cognitive processes and shape the experiences and sense of self, they become integral to the shape and elements of what we understand as consciousness. As per Noschka (2014), one aspect that the post/human brings to our attention, for example, is the language of spiritualism present in post human discourse. This is especially the case with regard to strains of post humanism that point towards the annihilation of the human body through bio enhancement or trans humanist visions of downloading human consciousness into machinery.

In Cartesian dualism theory, it is the event of mind and is unconnected from body; Empirical approach of Locke makes it body (experience) dependent; Leibniz's 'apperception' notion⁵ denies the mechanization of consciousness into any process; For Nietzsche it is 'epiphenomenal,' a by-product of brain activity; Hegel makes it relational, i.e., self is aware of itself only through the eyes of another self. As per NCC (Neural Correlates of Consciousness)⁶, it is the relationship between mental states and neural states, and 'may be thought of as a state-dependent property of an undefined complex, adaptive, and highly interconnected biological system' (Squire, 2008). David Chalmers, an Australian

This famous definition is presented in section 4 of the *Principles of Nature and of Grace* (1714), where Leibniz ⁵ says that apperception is "*consciousness*, or the reflective knowledge of this internal state." (*The Stanford Encyclopaedia of Philosophy*)

The minimum brain mechanisms that combined are required and sufficient for experiencing any conscious ⁶ perception have been dubbed the neural correlates of consciousness.

cognitive scientist, working in the field of philosophy of mind has formulated the concept of the hard question of consciousness.⁷

...even when we have explained the performance of all the cognitive and behavioral functions in the vicinity of experience—perceptual discrimination, categorization, internal access, verbal report—there may still remain a further unanswered question: Why is the performance of these functions accompanied by experience? (Chalmers, 1995)

The emergence of cyborgs has also intersected with the philosophical framework of post human ontology and the notion of ‘self’ that transcends human limitations through technology. Post human ‘self’ embodies post humanist ideals and post humanist perspective that redefines human identity by embracing hybridity, fluidity, and the continuous negotiation of their relationship with technology. This is where it becomes metaphysical as well as political. By challenging the conventional power dynamics and hierarchies by blurring the boundaries between natural and artificial, human and machine, post human embodies political agency, as their existence disrupts established norms and invites reconsideration of societal structures, power relations, and cultural contexts.

Apart from this, the convergence of technology with nature raises significant ethical implications concerning the moral obligations associated with post human philosophical conditions. An exploration of ethical considerations, especially around cyber – human entity, surrounding this new impending ontology and identity, discussing issues such as autonomy, inequality, and human enhancement is inevitable. The integration in human bodies will ensue debates on primary questions of human agency, informed consent and autonomy for the nature and extent of technological enhancements and modification to the body, ethical guidelines and frameworks for protecting bodily integrity. The pursuit of human enhancement technologies raises ethical questions regarding the establishment of normative standards, nature of self and self-determination. The most pertinent questions to be answered are - should society dictate the boundaries of acceptable enhancements, or should individuals be free to pursue modifications that redefine their physical and cognitive capacities? Do modifications and technological enhancements compromise the core essence of being human? The integration of technology into human bodies, especially in the case of cyborgs, changes the very nature and fabric of humanness in humans, and blurs the notion of ‘selfhood’. So far, the practiced ethical considerations have dictated that an equitable society recognizes the diversity of identities and respects individuals' self-perception and chosen modifications. Several authors adopt a more relaxed interpretation of the term "artificial moral agent," influenced by its usage in software engineering. In this context, issues regarding responsibility and rights are not a concern, as noted by Allen, Varner, and Zinser in their paper entitled “Prolegomena to Any Future Artificial Moral Agent” (2000). James Moor in his research “The Nature, Importance and Difficulty of Machine Ethics” (2006) opines that –

The question of whether machine ethics exists or might exist in the future is difficult to answer if we can't agree on what counts as machine ethics. Some might argue that machine ethics obviously exists because humans are machines and humans have ethics. Others could argue that machine ethics obviously doesn't exist because ethics is simply emotional expression and machines can't have emotions.

Explaining the relationship between physical phenomena, like brain functions, and experience, or phenomenal⁷ consciousness—that is, mental states or experiences having phenomenal qualities or *qualia*—is known as the “hard problem of consciousness” (Chalmers, 1995).



Moor categorizes machine agents into four types: ethical impact agents, implicit ethical agents, explicit ethical agents, and full ethical agents.⁸

The rise of AI challenges our notions of free will and moral responsibility. Furthermore, if AI systems influence human decision-making processes through algorithmic recommendations or persuasive techniques, does it undermine human agency and responsibility? Philosophical discussions delve into the concept of moral agency in the context of AI, considering the interplay between human intentions, machine determinism, and the ethical implications of human-machine collaborations. Opacity and bias are central issues in what is now sometimes called “data ethics” or “big data ethics” (Floridi and Taddeo 2016; Mittelstadt and Floridi 2016). The black-box nature of some AI algorithms poses challenges in ensuring transparency and accountability. Ethical guidelines should be established to encourage the development of explainable AI models, enabling users and regulators to understand the reasoning behind AI-generated outcomes.

The advancement of AI, cybernetics, and cyborg technologies carries significant implications for society, including the redistribution of power and social structures. As per Young (2017), ‘AI technologies are already used to dynamically personalize an individual's choice environments, to paternalistically nudge, deceive, and even manipulate behaviour in unprecedented manners.’ As AI systems gain decision-making capabilities, they may influence economic, political, and social systems. The concentration of power in the hands of those who control AI technologies raises concerns about societal inequalities and the potential for exploitation. Philosophical discussions revolve around the ethical dimensions of these transformations, seeking to ensure that the integration of AI and machines into society promotes justice, equality, and human well-being. Privacy concerns arise when AI algorithms process and analyze sensitive information, potentially leading to unauthorized access or misuse of personal data. Protecting individual's privacy rights requires robust data protection measures, including secure data storage, informed consent, and transparent data handling practices. Ethical guidelines should prioritize privacy protection, empowering individuals with control over their data and promoting responsible data practices in AI development and deployment.

The future of AI, cybernetics, and cyborgs holds immense potential for transformative advancements that can reshape various aspects of human life. Technological progress in these fields is expected to continue at an accelerated pace, leading to speculative philosophical implications that challenge our understanding of reality, consciousness, and human nature. However, balancing ethical considerations and ensuring responsible development and deployment of these technologies will be crucial in shaping a future that aligns with human values and promotes societal well-being.

As AI and cybernetic technologies advance, there is speculation about the potential emergence of a technological singularity. The singularity refers to a hypothetical point where AI systems surpass human intelligence and initiate a self-sustaining feedback loop of technological progress. This notion was popularized by mathematician and computer scientist Vernor Vinge (1993)⁹, and has been a topic

⁸ethical impact agents (e.g., robot jockeys), implicit ethical agents (e.g., safe autopilot systems), explicit ethical agents (which employ formal methods to assess utility), and full ethical agents (capable of making explicit ethical judgments and generally possessing the ability to reasonably justify them, akin to the ethical capacity of an average adult human).

According to Vinge, this point in time will lead to a breakdown in humans' ability to model their own ⁹ future, and such Singularity is estimated to occur within 40 years.





of debate and speculation among futurists, scientists, and philosophers. One of the primary philosophical hazards of the AI singularity is the loss of control over super intelligent AI systems. If AI systems can self-improve without human intervention, they may quickly become autonomous which further poses a bigger 'alignment problem' that involves defining and embedding human values into AI systems to prevent them from pursuing objectives that are harmful or misaligned with human interests.

At the philosophical level, this concept raises questions regarding the implications of super intelligent machines and their impact on human existence. Philosophical discussions explore the possibilities of the singularity, examining its potential risks, benefits, and the ethical dimensions of a world dominated by super intelligent machines. Philosophical inquiries delve into the realms of metaphysics, epistemology, and ethics, contemplating questions such as the nature of machine consciousness, the blurring of human-machine boundaries, the limits of human cognition, and the possibilities of post-human existence. Speculative philosophical discussions stimulate our imagination and provoke critical reflection on the potential philosophical transformations brought about by advanced technologies.

As we navigate the future of AI, cybernetics, and cyborgs, striking a balance between ethical considerations and technological progress becomes crucial. While these technologies hold great promise, they also carry risks and challenges that demand careful ethical assessment and regulation. Ethical considerations encompass a wide range of issues, including transparency, fairness, privacy, accountability, social impact, and the preservation of human values and rights. Integrating ethical frameworks into the development and deployment of these technologies ensures that they are aligned with human well-being, avoid harm, and contribute to a just and inclusive society. AI systems can unintentionally perpetuate bias and discrimination, reflecting the biases present in the data they are trained on. If biased data is used to train AI algorithms, the resulting models may make decisions that disproportionately disadvantage certain groups based on race, gender, or other protected characteristics. Ensuring fairness in AI requires careful attention to data collection, algorithm design, and evaluation processes. Ethical frameworks should prioritize fairness and promote efforts to mitigate bias, ensuring that AI systems do not perpetuate or exacerbate existing societal inequalities.

Conclusion:

The intersections of AI, cybernetics, and cyborgs with philosophy touch on profound questions about the nature of human existence, consciousness, free will, personal identity, and societal transformations. These interdisciplinary inquiries challenge traditional philosophical frameworks and demand thoughtful engagement with the ethical, metaphysical, and epistemological implications of human-machine interactions. By critically examining the philosophical aspects of AI and related fields, we can navigate the ethical, social, and existential challenges they present, ultimately shaping a future that reflects our values and aspirations as a technologically augmented society.

The future of AI, cybernetics, and cyborgs presents exciting possibilities that challenge our philosophical understanding of reality, consciousness, and the nature of humanity. Speculative philosophical discussions allow us to explore the potential implications of these technologies, encouraging critical reflection and ethical analysis. Balancing ethical considerations with technological progress is essential to ensure that these advancements are guided by human values, promote societal well-being, and respect fundamental rights and principles. By actively engaging in ethical discourse and responsible decision-making, we can shape a future that harnesses the transformative potential of AI, cybernetics, and cyborgs while upholding our shared human values. While AI and cybernetic ethics will play a pivotal role in shaping the responsible development and deployment of artificial intelligence. Addressing issues such as transparency, fairness, accountability, privacy, and the impact on human employment is essential to ensure that AI systems benefit society as a whole. Ethical frameworks, informed by interdisciplinary collaboration, should guide the design, implementation, and regulation of AI technologies, promoting their alignment with human values and societal well-being. By proactively addressing AI ethics, we can harness the



potential of AI while minimizing potential harms and fostering a future that is both technologically advanced and ethically grounded.

References:

1. Allen, Colin, Gary Varner, and Jason Zinser (2000), 'Prolegomena to Any Future Artificial Moral Agent,' *Journal of Experimental & Theoretical Artificial Intelligence*, 12(3): 251–261. doi:10.1080/09528130050111428
2. Chalmers, David (1995), 'Facing up to the problem of consciousness,' *Journal of Consciousness Studies*. 2 (3): 200–219.
3. Clark, A. (2003), *Natural-Born Cyborgs: Minds, Technologies, and the Future of Human Intelligence*, Oxford: OUP.
4. Clynes, Manfred E. & Nathan S. Kline (1960), 'Cyborgs and Space,' *Astronautics* 5(9): 26-27, 74-76.
5. Floridi, Luciano and Mariarosaria Taddeo (2016), 'What Is Data Ethics?,' *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 374(2083): 20160360. doi:10.1098/rsta.2016.0360
6. Haraway, D. (1991), 'A Cyborg Manifesto: Science, Technology, and Socialist-Feminism in the Late Twentieth Century,' *Cyborgs and Women: The Reinvention of Nature*. New York: Routledge.
7. Kulstad, Mark and Laurence Carlin, 'Leibniz's Philosophy of Mind,' *The Stanford Encyclopedia of Philosophy* (Winter 2020 Edition), Edward N. Zalta (ed.), URL = <<https://plato.stanford.edu/archives/win2020/entries/leibniz-mind/>>.
8. Mayor, Adrienne (2018), *Gods and Robots: myths, machines and ancient dreams of technology*, PRINCETON: Princeton University Press.
9. Mittelstadt, B.D. & Floridi, L. (2016), 'The Ethics of Big Data: current and foreseeable issues in biomedical contexts,' *Science and Engineering Ethics*, 22(2).
10. Moor, James H. (2006), 'The Nature, Importance, and Difficulty of Machine Ethics,' *IEEE Intelligent Systems*, 21(4): 18–21. doi:10.1109/MIS.2006.80
11. Noschka, Michael (2014), 'Extended Cognition, Heidegger, and Pauline Post/Humanism,' *Literature and Theology* Vol. 28, No. 3.
12. Shapiro, Lawrence and Shannon Spaulding, 'Embodied Cognition,' *The Stanford Encyclopaedia of Philosophy* (Winter 2021 Edition), Edward N. Zalta (ed.), URL <<https://plato.stanford.edu/archives/win2021/entries/embodied-cognition/>>.
13. Stevenson, Melissa Colleen, 'Trying to Plug In: Posthuman Cyborgs and the Search for Connection,' *Science Fiction Studies*, Vol. 34, No. 1 (Mar., 2007), pp. 87-105
14. Squire, Larry R. (2008), *Fundamental neuroscience* (3rd ed.), Academic Press.
15. Susser, D., Roessler, B., and Nissenbaum, H. (2019), *Online Manipulation: Hidden Influences in a Digital World*. Georgetown L. Techn. Rev. 4 (1), 1–45. doi:10.2139/ssrn.3306006
16. Vinge, Vernor (1993), 'The coming technological singularity: How to survive in the post-human era,' *Whole Earth Review*.
17. Yeung, Karen (2017), '“Hypernudge”: Big Data as a mode of regulation by design,' *Information, Communication & Society*, 20:1, 118-136, DOI: [10.1080/1369118X.2016.1186713](https://doi.org/10.1080/1369118X.2016.1186713)