Cognitive Computation (2026) 18:20
https://doi.org/10.1007/s12559-026-10569-8

RESEARCH

Artificial Intelligence as the Fourth Decentering Revolution: From Cosmic, Biological, and Psychological Displacement to Cognitive Decentering

Erik Cambria[1] · Rui Mao[1] · Nicola Bianchi[2] · Amir Hussain[3] · Keith Oatley[4] · Geoffrey Hinton[4]

Received: 2 February 2026 / Accepted: 16 February 2026

Abstract
Artificial intelligence (AI) is commonly framed as a technological force transforming economies, labor markets, scientific discovery, and social interaction. While these impacts are profound, they represent only the surface of a deeper historical and philosophical shift. We argue that AI constitutes a fourth major decentering revolution in human self-understanding. The Copernican revolution displaced humans from the center of the universe, the Darwinian revolution displaced humans from the pinnacle of biological life, the Freudian revolution displaced the conscious ego from the throne of the mind, and today AI initiates a cognitive decentering revolution, challenging the long-held assumption that humans occupy a unique and unassailable position at the apex of intelligence. By situating AI within historical, philosophical, and technological contexts, incorporating detailed case studies, illustrative examples, and policy analysis, this paper provides a comprehensive account of cognitive decentering and its societal implications.

Keywords Artificial intelligence · Decentering revolutions · Cognitive decentering · Philosophy of AI · Human intelligence · Ethics · Policy · History of science

Introduction

AI has undergone a rapid transformation from a niche research endeavor within computer science and cognitive modeling to a pervasive societal and cultural force with profound implications for virtually every domain of human activity. Modern AI systems are no longer confined to laboratories or specialized technical applications; they now influence economic production through automation, optimization, and predictive analytics, accelerate scientific discovery by identifying patterns and generating hypotheses at scales impossible for human researchers, and reshape healthcare by supporting diagnostics, treatment planning, and personalized medicine. In education, AI facilitates adaptive learning, automated assessment, and individualized tutoring, while in governance, AI-assisted decision-making informs policy design, resource allocation, and public administration. Beyond these functional domains, AI also exerts a transformative influence on creative expression, producing music, literature, visual art, and even game design, and it permeates public discourse by shaping

✉ Erik Cambria
cambria@ntu.edu.sg

Rui Mao
rui.mao@ntu.edu.sg

Nicola Bianchi
bianchi@lnf.infn.it

Amir Hussain
amir.hussain@phc.ox.ac.uk

Keith Oatley
keith.oatley@utoronto.ca

Geoffrey Hinton
geoffrey.hinton@gmail.com

[1] Nanyang Technological University, Singapore, Singapore

[2] National Institute for Nuclear Physics, Rome, Italy

[3] University of Oxford, Oxford, UK

[4] University of Toronto, Toronto, Canada

Published online: 26 February 2026

information consumption, recommendation systems, and the dynamics of social media engagement.

While popular and academic discussions often foreground issues such as job displacement, economic inequality, ethical dilemmas, and regulatory challenges, these practical concerns, although pressing, obscure a deeper and more profound phenomenon. AI is not merely a tool or technology that extends human capabilities; it represents a transformative force in the human understanding of intelligence itself. Specifically, AI constitutes what can be called a fourth major decentering revolution in the history of human self-conception. Just as the Copernican revolution displaced humans from the center of the universe, the Darwinian revolution displaced humans from the pinnacle of biological life, and the Freudian revolution undermined the sovereignty of the conscious self by exposing the unconscious determinants of the mind, AI now initiates a cognitive decentering, challenging long-held assumptions that human intelligence is uniquely superior, inherently central to decision-making, and qualitatively distinct from computational or artificial forms of cognition.

We call this the "Intuitive AI" revolution to recognize the fact that the current form of AI is based on biologically inspired neural networks that excel at modeling human intuition. Earlier attempts at AI were mostly inspired by logic. They viewed reasoning as the highest form of intelligence and they assumed that reasoning must be done by using rules to manipulate symbolic expressions in some unambiguous internal language. The difference between these two approaches is exemplified by the following problem which cannot be solved by logic. Suppose you have to choose between two biologically absurd possibilities. Either all cats are male and all dogs are female or all cats are female and all dogs are male. It is just intuitively obvious to most men in western culture that the second possibility is more plausible. Dogs are big and noisy and chase cats. The pattern of neural activity in our brains (or in an artificial neural network) that represents dog is more similar to the pattern for man than for woman, whereas the pattern for cat is more similar to the pattern for woman than for man. The similarities between the patterns of neural activity that we use to represent the meanings of symbols make it easy for us to see analogies which form the basis of intuitive reasoning (Fig. 1).

What is striking about this revolution is that it represents far more than a technological advancement: it constitutes a profound philosophical disruption. AI challenges the traditional, human-centered understanding of intelligence and the mind, forcing us to reconsider long-held assumptions about cognition, intuition, agency, and the uniqueness of human thought. Unlike earlier revolutions, in fact, the

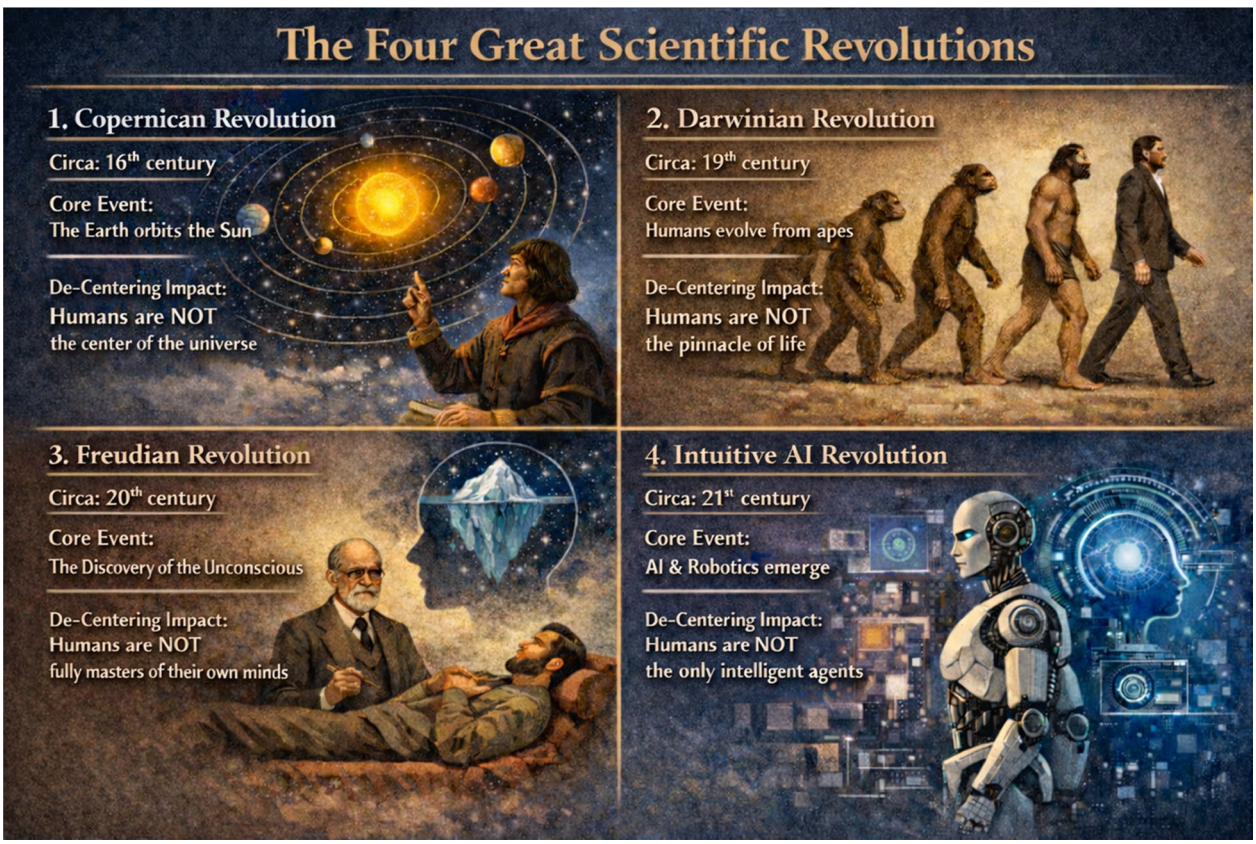

Fig. 1 Decentering revolutions in human history

decentering introduced by AI is uniquely intimate, immediate, and pervasive. Whereas Copernican, Darwinian, and Freudian decenterings required extensive study and reflection before their implications were widely absorbed, cognitive decentering unfolds in daily human experience. People encounter AI systems through mundane interactions: predictive text, personalized recommendations, automated translation, and problem-solving tools. These everyday systems now touch multiple cognitive functions that were once treated as distinctly human, e.g., perception, memory, reasoning, planning, language [1].

These interactions implicitly challenge the notion that humans hold exclusive mastery over reasoning, pattern recognition, or creativity. Moreover, AI's encroachment into domains previously considered the exclusive province of human expertise (e.g., scientific research, strategic planning, artistic creation, and complex decision-making) forces a re-evaluation of human identity, professional roles, and social authority in real time. The transformative power of AI extends beyond individual cognition into collective and institutional domains. Scientific communities, for instance, now collaborate with AI in data-driven hypothesis generation, model selection, and experimental design, reshaping epistemic practices and knowledge production. Similarly, governance and business institutions are increasingly structured around AI-assisted predictive analytics, algorithmic decision-making, and automated monitoring systems, raising ethical and legal questions about accountability, transparency, and societal responsibility. Cultural production is likewise affected: AI-generated literature, music, and visual art blur the lines between human and artificial creativity, prompting reconsideration of authorship, originality, and value. While this article takes an optimistic view of AI's cognitive potential and future progress, we avoid claims of broad "cognitive superiority" and instead characterize current advances as strong task performance under specified conditions.

This paper situates AI within the historical lineage of decentering revolutions, drawing parallels between the existential, societal, epistemic, and cognitive implications of Copernican, Darwinian, Freudian, and AI-driven shifts. By examining historical resistance and adaptation, analyzing philosophical and epistemic ramifications, and providing concrete illustrative examples, the paper highlights the uniquely disruptive nature of cognitive decentering. Furthermore, it offers detailed policy and societal recommendations for navigating this transformation responsibly, emphasizing the need for ethical frameworks, educational reform, and governance structures that acknowledge the decentered nature of intelligence while preserving human dignity and fostering collaboration between human and artificial cognitive agents [2, 3].

Historical Decentering Revolutions

The Copernican Revolution

The Copernican revolution, initiated by Nicolaus Copernicus with the publication of *De revolutionibus orbium coelestium* in 1543 [4], fundamentally displaced humanity from its perceived cosmic centrality. Prior to Copernicus, the Ptolemaic geocentric model dominated both scientific and philosophical thought, positioning the Earth (and by extension, humanity) at the center of the universe. This framework was deeply intertwined with theological and metaphysical assumptions, reinforcing the idea of human exceptionalism and divine favor. By proposing a heliocentric model, in which Earth is merely one planet among many orbiting the Sun, Copernicus challenged centuries of received wisdom and forced a radical re-evaluation of humanity's place in the cosmos.

The impact of this shift extended beyond astronomy. It signaled the beginning of a profound epistemic transformation: knowledge could no longer rely solely on sensory perception or inherited authority but required rigorous observation, mathematics, and predictive modeling. The Copernican model introduced a new framework in which the cosmos was governed by universal laws rather than human-centered design. This decentering initiated the intellectual process that would eventually culminate in modern scientific methodology.

Philosophical and Cultural Implications

The Copernican revolution had profound philosophical and cultural consequences. By undermining the anthropocentric worldview, it fostered epistemic humility, demonstrating that human intuition and sensory experience were fallible guides to understanding the universe. This shift destabilized traditional hierarchies of knowledge and power, as previously unquestioned authorities (both religious and philosophical) were now challenged by empirical observation and mathematical reasoning.

Culturally, the Copernican revolution provoked widespread intellectual and existential anxiety. The realization that Earth was not the fixed center of creation was experienced by many as a loss of cosmic significance, what later psychologists and historians have described as the first of humanity's "narcissistic wounds" [5]. Literature, art, and public discourse of the period reflect this tension: cosmological uncertainty became intertwined with theological and moral concerns, as people struggled to reconcile religious doctrine with empirical findings.

Philosophers such as Giordano Bruno further expanded on Copernicus' ideas, positing that the universe was infinite

and that the Sun was just one star among countless others. Such radical notions were met with fierce resistance, highlighting the psychological and societal difficulty of accepting decentering knowledge. The Copernican shift also laid the groundwork for subsequent philosophical developments, including Descartes' mechanistic universe and Newton's laws of motion, by establishing that natural phenomena could be studied independently of human-centric interpretations.

Historical Case Study: Galileo Galilei

Galileo Galilei's contributions provide a concrete illustration of the Copernican revolution in action. Utilizing telescopic observations, Galileo confirmed key predictions of the heliocentric model, including the phases of Venus and the moons of Jupiter [6]. These empirical observations provided irrefutable evidence that the Earth was not the center of the cosmos, challenging both Aristotelian physics and Ptolemaic cosmology.

Galileo's trial by the Roman Catholic Inquisition in 1633 exemplifies the cultural and institutional resistance to this decentering. The conflict was not solely about empirical facts but about the perceived threat to the social and religious order. Galileo's advocacy for Copernicanism exposed the tension between emerging scientific reasoning and established doctrinal authority, highlighting the broader societal implications of intellectual decentering.

The case of Galileo demonstrates how the Copernican revolution was both a scientific and a cultural event. It reshaped epistemology, challenging assumptions about observation, evidence, and authority. At the same time, it forced society to confront the unsettling idea that humans were not the center of creation, initiating a profound transformation in human self-understanding that resonates with later decentering revolutions in biology and cognition.

The Darwinian Revolution

Charles Darwin's seminal work, *On the Origin of Species* (1859) [7], started a second major decentering revolution in human history, this time within the biological realm. Prior to Darwin, prevailing religious and scientific thought tended to place humans at the pinnacle of creation, as having been made directly by God, with similarities to their creator. This idea was reinforced by theology, philosophy, and common sense. Darwin's theory of evolution by natural selection challenged this assumption by proposing that human beings are a natural species, generated by a random exploratory process with these three components.

Superabundance: when they reproduce members of each species do not just replace themselves with a single being, they generate many offspring. Variation: the offspring tend to have differences. Selection: some offspring fit well into the niche into which there are cast, and thereby are selected to survive and reproduce, while others do not fit well and die before reproducing. This proposal displaced humans from the apex of beings, situating them within a continuum that includes plants and animals. The implications of this shift were both profound and unsettling. If humans were no longer ontologically distinct, traits such as reasoning, moral judgment, and social behavior could no longer be assumed to be uniquely human. Cognitive, emotional, and social capacities were now interpreted as evolved adaptations, existing on a spectrum rather than as categorical differences. This understanding paved the way for modern comparative psychology, ethology, and evolutionary neuroscience, which investigate intelligence, emotion, and behavior across species using empirical and theoretical frameworks grounded in evolutionary principles.

Resistance and Integration

The Darwinian revolution provoked intense resistance, both from religious authorities and from segments of the philosophical community. Many perceived the theory as a direct threat to foundational beliefs about human exceptionalism, divine creation, and moral order. Critics argued that if humans were merely another species, traditional ethical frameworks and theological doctrines might be undermined, raising existential and cultural anxieties that extended far beyond the scientific domain. Newspapers, pamphlets, and public debates reflected the social tension of the period, with arguments ranging from literal interpretations of scripture to moral and metaphysical objections.

Despite early resistance, the principles of evolution came to reshape the biological sciences. The Darwinian framework provided a unifying theory for understanding species diversity, adaptation, and speciation. It also catalyzed the development of fields such as genetics, ecology, and evolutionary psychology. By reframing human cognition and behavior as products of natural processes, Darwinian thought introduced an empirical and comparative methodology for studying intelligence, morality, and social structures across species. Importantly, it also encouraged a recognition of the deep interconnectedness of living beings on Earth, fostering a perspective that emphasized continuity and adaptation rather than hierarchy and centrality.

Historical Case Study: Huxley vs. Wilberforce

A particularly illustrative example of societal confrontation with Darwinian ideas is the famous 1860 debate at Oxford between Thomas Henry Huxley, known as "Darwin's

Bulldog", and Samuel Wilberforce, the Bishop of Oxford. Huxley defended Darwinian evolution with empirical reasoning and scientific evidence, emphasizing the continuity of humans with other species, while Wilberforce attacked the theory on theological and moral grounds, ridiculing its implications for human dignity. The debate exemplified the cultural and intellectual friction inherent in accepting a decentered view of humanity. Although contemporary accounts vary on the rhetoric and outcome, the encounter has become emblematic of the broader societal struggle to integrate evolutionary science into existing belief systems.

This case also highlights the broader psychological and cultural processes triggered by decentering revolutions. Resistance often manifests as denial, ridicule, or selective acceptance, followed by gradual integration as evidence accumulates and societal paradigms shift. Modern parallels can be drawn with contemporary reactions to AI, where anxieties over human obsolescence, loss of authority, or cognitive displacement echo the patterns observed during Darwinian debates. The Huxley-Wilberforce confrontation thus serves not only as a historical episode but also as a conceptual lens for understanding human adaptation to successive decentering revolutions, from cosmic to biological to cognitive.

Philosophical and Cultural Implications

The Darwinian revolution also prompted significant philosophical reflection. By situating humans within a natural continuum, it challenged metaphysical claims about essence, purpose, and moral centrality. Philosophers and social theorists began to reconceptualize humanity not as a fixed endpoint of creation but as an evolving entity, subject to the same contingent processes that shape all life.

An important proposal by Michael Tomasello and his research group (see e.g. Tomasello 2019, Becoming Human: A Theory of Ontogeny. Harvard University Press) has been to show that a recent move in human evolution has been for humans to make co-operative arrangements with each other, and then fulfil them, with "we" becoming, for a shorter or longer period, more important than "I". This change has laid the way for human language, joint explorations, and shared use of technology an early example of which is likely to have been the bag, used to carry food that had been gathered in order to take it back others so that they could all eat together. Our closest relatives, chimpanzees do nothing similar: groups of them do search for food together but, when they find some, each grabs what she or he can (with the alpha male tending to get first pick), then going to eat it alone.

Science, philosophy, literature, art, and popular discourse, have reflected this evolutionary move. Writing this paper together is an example, with "we" hoping that "you" as readers will join us. Other themes of this kind are human fragility and adaptation. The intellectual legacy of Darwin's decentering endures today, forming the conceptual groundwork for contemporary discussions of cognition, intelligence, and the ethical status of non-human agents, including artificial intelligences.

The Freudian Revolution

If the Copernican and Darwinian revolutions displaced humanity in space and nature, the Freudian revolution displaced humanity within itself. Psychoanalysis is often caricatured as a doctrine of sexuality or a mythology of family romance. Yet its truly revolutionary claim is epistemological: the subject is not transparent to itself [8]. The human being is not, even in principle, the sovereign author of its own reasons. What we call "mind" is not a unified chamber of clear thoughts, but a stratified economy of forces, e.g., desires, prohibitions, memories, and fantasies, whose negotiations occur largely outside conscious report.

In this sense, Freud introduced a decentering more intimate than astronomy or biology: he relocated the center of agency away from the reflective "I" and into a conflicted, partially opaque psychic apparatus. The psychoanalytic thesis is not merely that there exists an unconscious, but that conscious life is continuously constructed as a compromise with it. Consciousness, for Freud, is not the executive summit of cognition. It is a fragile surface phenomenon, an interface that must present a coherent narrative to sustain action and social belonging. Psychoanalysis therefore reconceives rationality as a local achievement rather than a global condition. The mind is not primarily a truth-tracking instrument. It is, before all else, a survival system tasked with regulating anxiety, preserving attachment, and maintaining a tolerable image of the self. From this perspective, cognition is inseparable from emotion: belief, attention, and interpretation are not neutral operations but are recruited into the management of intrapsychic conflict.

Philosophical and Cultural Implications

The philosophical import of psychoanalysis is frequently misunderstood because it is judged by the wrong ideal: the ideal of a mind that ought to be transparent, self-identical, and finally accountable to its own conscious reasons. Freud's proposal was to replace this ideal with something more unsettling. The human subject is not an epistemic sovereign but an interpretive achievement, which is a narrative equilibrium maintained under pressure. To live as a person is to continuously negotiate between several demands, some of which are incompatible with others: pleasure and reality,

intimacy and autonomy, aggression and moral constraint. Psychoanalysis names this condition as the basic architecture of the psyche. The "I" is therefore not a simple point of origin. It is a diplomatic institution, improvising treaties among inner powers that do not always share the same assumptions.

This shift has immediate consequences for epistemology. If cognition is recruited into the service of psychic survival, then knowledge is never merely descriptive. It is also defensive. The Freudian insight is not just that people sometimes lie, but that they can lie to themselves, with a deeply felt sincerity. A rationalization is not a conscious deception. It is an explanation that functions as anesthesia. Psychoanalysis thus reveals a category of error that classical philosophy struggled to name: error as adaptation. There is a sense in which the psyche can be "right" to be wrong, when truth threatens the coherence required for action. This does not sanctify illusion. It diagnoses the costs of insight, and thereby explains why cultures, not only individuals, often cling to flattering fictions.

Culturally, this diagnosis reconfigures the modern understanding of responsibility. Before Freud, moral discourse could still fantasize that wrongdoing primarily reflects ignorance or vice in the transparent will. After Freud, the will becomes porous: motives can be hidden from the agent, and self-knowledge becomes a task rather than a premise. This does not absolve responsibility. It deepens responsibility. The ethical demand shifts from "be rational" to "be interpretable": cultivate practices and institutions that make motives legible, conflicts speakable, and defenses revisable.

Historical Case Study: Talking Cure and Self-Knowledge

The clinical scene is psychoanalysis's most radical laboratory, because it stages a new kind of evidence. In the talking cure, what counts is not only what is said, but what is not said; not only what is remembered, but what returns in disguise. The symptom becomes a text, and interpretation becomes an instrument for tracing the logic of compromise. This is an epistemic inversion: rather than treating irrationality as mere noise to be eliminated, psychoanalysis treats it as structured output. The slip, the repetition, the sudden affective surge are not just accidents but signatures of an underlying organization.

A principal revolutionary element here is transference. The projection of a mental model and attitude toward a therapist that derive from a significant person in the past, typically a parent. Psychoanalytic therapy is principally based on the therapist making interpretations of this when it occurs, to move the mode of relating from the unconscious into consciousness. It occurs not just in psychoanalysis, but every form of therapy, and in consultations with doctors who, perhaps like our mother, we hope will make everything better. It occurs between students and teachers. It occurs in adult sexual relationships. When we were infants, we might have been needy, affectionate, controlling, or irritable. As adults we can do the same with a loved one. Although we may think that we are acting in response to a current situation, unconsciously the action can derive from our childhood past.

In this sense, psychoanalysis offers a prototype for a broader theory of human rationality: rationality is not a static faculty but a negotiated stance within a conflicted system. The ego, in Freud's later vocabulary, is neither hero nor villain. It is a regulator with limited resources, compelled to manage internal drives and external demands under conditions of uncertainty. The analytic encounter makes visible how easily cognition becomes a servant of defense, and how difficult it is to transform defense into reflection. Case studies, therefore, are not merely about therapy. They are about models of the mind appropriate to the present.

From Freudian Decentering to AI's Cognitive Decentering

Seen through this lens, the contemporary shock of AI is not that machines can compute, but that they can occupy roles once reserved for the ego's privileged functions: composing narratives, proposing reasons, simulating empathy, and suggesting plans.

Freud warned that the ego's sovereignty is fragile; AI reveals that even the *form* of sovereignty (coherent articulation, strategic foresight, linguistic fluency) can be instantiated outside the human body. This is why cognitive decentering feels less like an external technological shift and more like an intrusion into the architecture of subjectivity.

Psychoanalysis helps diagnose two predictable cultural reactions. One is denial: treating AI as mere tool, insisting that "real" understanding remains untouched. The second is catastrophic projection: imagining AI as a rival superagent, externalizing our own fears of dependency and loss of control. Both reactions echo the logic of defense. Both reactions are defensive. They attempt to preserve the narcissistic image of the human as unchallenged center by either trivializing the challenger or demonizing it.

A more mature integration would treat AI not as an enemy of the self, but as a new mirror for the self's already divided condition. AI systems can amplify rationalization, automate avoidance, and industrialize confirmation bias; they can become prostheses for denial. Yet they can also externalize our interpretive habits, making visible the patterns by which we confabulate reasons and manage discomfort. If psychoanalysis taught that emancipation requires confronting the unconscious, then AI forces a further confrontation: the recognition that many "higher" cognitive performances are not

the private property of human interiority, but reproducible strategies in a broader space of intelligences.

The Freudian revolution thus remains indispensable to understanding AI's decentering. It equips us with a vocabulary of conflict, defense, and narration (concepts that explain not only individual resistance to AI but also institutional and cultural resistance). Most importantly, it reframes the question of human dignity. Dignity need not be grounded in cognitive supremacy. Freud already stripped supremacy from the self; what remained was a more demanding ideal: the capacity for self-interpretation, ethical responsiveness, and the courage to live without flattering illusions. In the age of AI, that ideal becomes not a consolation prize, but the core of what it could mean to be human in a decentered cognitive ecology.

The Fourth Revolution: Intuitive AI

Intuitive AI introduces a fourth form of decentering, directly challenging the long-standing assumption that humans occupy the apex of intelligence. Whereas previous decenterings (such as the Copernican decentering in astronomy, the Darwinian decentering in biology, and the Freudian decentering in psychology) displaced humans from the center of the universe and the pinnacle of an evolutionary hierarchy, cognitive decentering displaces humans from the privileged position of intellectual and problem-solving superiority. This decentering does not depend on whether AI possesses consciousness or self-awareness; rather, it challenges the functional and operational definition of intelligence as a domain where humans were once assumed to be uniquely dominant.

Defining Cognitive Decentering

Cognitive decentering refers to the displacement of humans from the perceived apex of intelligence. It emerges from the realization that cognitive abilities (ranging from pattern recognition and strategic reasoning to creative ideation) can be instantiated in non-biological substrates. In this context, intelligence is increasingly understood not as a property tied to human biology but as a functional capacity that can be realized in diverse systems, including machines and human-AI hybrids [9, 10].

This redefinition has profound epistemological and philosophical implications: it forces a reconsideration of human exceptionalism and questions the assumption that intelligence, creativity, and insight are inherently human traits. Cognitive decentering highlights that intelligence does not require biology or evolution. It can emerge in machines that have sufficient data, computational power, and architectural sophistication and it can already outperform humans in many specific cognitive domains.

Illustrative Examples and Case Studies

Scientific Discovery

One of the clearest demonstrations of cognitive decentering occurs in scientific discovery. AI systems such as AlphaFold have transformed protein structure prediction, accomplishing tasks in hours that would take human teams months or years [11]. AI researchers have contributed to the generation of novel scientific hypotheses across multiple fields, including the social sciences and chemistry [12]. The speed, scale, and accuracy of such AI-driven predictions reveal that knowledge production itself can be automated, not merely assisted. In fields like genomics, drug discovery, and materials science, AI increasingly generates hypotheses, analyzes experimental data, and proposes solutions that extend beyond conventional human expertise. This challenges the notion that advanced scientific reasoning is exclusively human, showing that intelligence is not species-bound.

Creative Domains

AI's impact extends to creative and expressive domains. Generative models, such as those based on diffusion or transformer architectures, produce visual art, music, and literature at levels that are beginning to rival professional human creators [13, 14]. Notably, AI-generated artworks have been exhibited and sold in high-profile galleries, and AI-assisted music compositions have gained mainstream attention. These developments provoke fundamental questions about creativity, authorship, and aesthetic judgment. Cognitive decentering in creative domains demonstrates that imaginative problem-solving is not an exclusive marker of human identity; rather, creativity can emerge from the interaction of data, computation, and generative algorithms.

Strategic Decision-Making

In strategic reasoning, AI systems have repeatedly outperformed human experts. Landmark examples include AlphaGo and its successors, which defeated professional Go players in matches once thought to require uniquely human intuition and experience [15]. Beyond games, AI systems are increasingly deployed in high-stakes environments such as financial trading, clinical decision-making, and legal reasoning, where AI analyses can identify patterns, optimize strategies, and predict outcomes more efficiently than human experts. These applications underscore

how cognitive decentering extends to domains historically associated with human judgment, experience, and authority.

Multimodal Cognitive Systems

Modern AI increasingly integrates multiple modalities (vision, language, and even physical manipulation) to solve complex tasks that exceed the capability of human specialists working in isolation. For instance, multimodal models can interpret medical images, analyze textual records, propose diagnostic decisions, effectively combining perception, reasoning, and memory in ways that human cognition cannot [16]. Such systems exemplify the distributed, substrate-independent nature of intelligence: cognition emerges not from a singular, human-centered mind but from the orchestration of data, learning algorithms, and computational architectures capable of solving problems across domains.

Implications of Cognitive Decentering

Cognitive decentering has implications that extend beyond performance metrics. It reshapes how society understands expertise, authority, and agency. As AI systems assume tasks once reserved for human cognition, humans are increasingly collaborators rather than sole originators of knowledge. This shift requires rethinking educational priorities, governance structures, and ethical frameworks for AI deployment. Furthermore, it challenges philosophical assumptions about what it means to be intelligent, creative, or rational, emphasizing that these traits are functional rather than ontological.

The Disruptive Nature of Cognitive Decentering

Cognitive decentering will require a profound shift in human self-understanding and societal structures. Unlike prior decenterings (such as the Copernican shift, which repositioned humanity within the cosmos, or the Darwinian revolution, which challenged assumptions of biological superiority) cognitive decentering is uniquely disruptive because it directly targets intelligence, a core attribute intimately tied to human identity, agency, and social authority. Intelligence is not merely a functional trait; it underpins notions of self-worth, ethical responsibility, and leadership. Consequently, as AI encroaches on domains previously considered the exclusive province of human cognition, individuals and societies confront existential, cultural, and practical challenges.

This disruption is multifaceted: it affects how humans perceive their own abilities, how authority is assigned and maintained, and how societal institutions are structured. It operates not in abstract terms but in everyday lived experiences, reshaping interactions, expectations, and aspirations. Cognitive decentering is therefore not a distant theoretical concern: it is an immediate, intimate, and pervasive force that challenges foundational assumptions about human primacy in cognitive domains.

Societal and Cultural Dimensions

The societal and cultural impacts of cognitive decentering are experienced in everyday life through interactions with AI systems. Writing assistants, recommendation algorithms, diagnostic tools, and autonomous decision-making systems provide continuous, tangible demonstrations of human cognitive displacement. These technologies produce both enhancement and anxiety: on the one hand, they augment human capabilities, improving efficiency, accuracy, and insight; on the other, they raise concerns about dependency, obsolescence, and the erosion of uniquely human expertise.

Media discourse plays a central role in shaping public perception of cognitive decentering. Some analyses highlight potential opportunities for collaboration and empowerment whilst others highlight the possibility of mass unemployment. Some highlight the enormous potential benefits in healthcare and education whilst others highlight the possibility that superintelligent AI will make humans irrelevant or extinct.

Popular culture (from films depicting superintelligent AI to literature exploring post-human futures) reflects and reinforces these tensions, influencing societal attitudes toward intelligence, creativity, and decision-making. In turn, these cultural dynamics affect policy debates, educational priorities, and the broader social understanding of AI's place in human life.

Furthermore, cognitive decentering intersects with identity and social cohesion. As AI demonstrates superior performance in certain cognitive domains, individuals may experience shifts in self-conception, re-evaluating what it means to be intelligent, capable, or creative. Professional hierarchies are disrupted, as roles once reserved for human experts (such as medical diagnosis, legal analysis, or scientific research) become augmented or replaced by AI systems. These changes necessitate a re-examination of cultural norms, societal values, and the symbolic significance of human intelligence.

Power, Governance, and Ethics

Cognitive decentering also raises critical questions about power, governance, and ethical responsibility. The development and deployment of advanced AI systems are

concentrated within a small number of corporations, research institutions, and governmental agencies, creating asymmetries in knowledge, capability, and influence. These asymmetries amplify the social consequences of cognitive decentering: those with access to AI's cognitive enhancements can consolidate economic, political, and cultural power, while others may face marginalization or dependency.

Ethically, cognitive decentering challenges conventional frameworks of responsibility and accountability. If AI systems make decisions or generate knowledge that surpasses human understanding, questions arise: who is accountable for errors or harms? How should liability be allocated between designers, deployers, and end-users? Additionally, cognitive decentering intersects with distributive justice, highlighting disparities in access to AI capabilities. Societies must grapple with the ethical imperative to ensure that cognitive augmentation is not a privilege of the few but a resource that benefits humanity broadly. Furthermore, cognitive decentering can be understood as a shift in epistemic control, e.g., who sets goals, verifies outputs, and ultimately defers to AI recommendations, making calibration and appropriate reliance central governance concerns. This shift can be tracked through measurable indicators such as reliance rates under uncertainty and trust calibration (e.g., over/under-deference patterns), which help identify when authority displacement increases the risk of error, harm, or unjust outcomes.

Policy frameworks must therefore address not only technical safety and robustness but also societal fairness, transparency, and inclusivity. Mechanisms such as algorithmic auditing, participatory governance, ethical review boards, and regulatory oversight are essential to mitigate risks while fostering responsible innovation. Furthermore, international coordination becomes critical, as cognitive decentering is not confined to local jurisdictions but is a global phenomenon, with AI systems and their impacts transcending national and cultural boundaries.

Psychological and Existential Implications

Beyond societal and institutional dimensions, cognitive decentering carries profound psychological and existential consequences. Individuals may confront challenges to their self-efficacy, professional identity, and sense of purpose. The recognition that machines can outperform humans in cognitive tasks can produce feelings of inadequacy or redundancy or it can open opportunities for new forms of collaboration and creativity.

Philosophically, cognitive decentering prompts re-evaluation of what constitutes meaningful human experience. If intelligence is decoupled from biological embodiment and realized in artificial substrates, the traditional markers of human uniqueness (reasoning, creativity, problem-solving) must be reconceptualized. This raises questions about value, agency, and the ethical treatment of intelligent systems, as well as the responsibilities humans bear in a world where cognitive primacy is no longer guaranteed.

At the individual level, these existential shifts are often mediated by epistemic reliance: as people defer more to AI judgments, the locus of cognitive control can move from internal deliberation to external validation, reshaping confidence, agency, and identity. Over time, this reliance may alter how people learn and maintain expertise, encouraging verification and tool orchestration in some settings, but risking deskilling and over-trust in others, making calibration under uncertainty a central psychological dimension of cognitive decentering.

Philosophical Context: Narcissistic Wounds and AI

The idea of this paper actually stems from Sigmund Freud's concept of "narcissistic wounds" to human self-importance, each progressively decentering humanity from its perceived centrality in the universe. The first, the "Copernican wound", revealed that the Earth is not the center of the cosmos, challenging the anthropocentric worldview and redefining humanity's spatial and cosmological significance. The second, the "Darwinian wound", displaced humans from the apex of the biological hierarchy, demonstrating that humans share common ancestry with other species and are subject to the same evolutionary forces as all life [17]. The third, the "Freudian wound", confronted humans with the unconscious mind, highlighting the limits of conscious self-knowledge and the hidden forces that shape thought, behavior, and desire.

AI introduces a "fourth narcissistic wound", or perhaps more accurately, a reframing of the concept of intelligence itself. Cognitive decentering challenges the assumption that humans are the default measure of reasoning, problem-solving, creativity, and decision-making. Unlike previous wounds, which displaced humans in relation to the universe, biology, or the mind, AI decentering directly confronts the *functional and operational domains of intelligence*, striking at the core of human identity, agency, and social authority. Philosophers such as Daniel Dennett, Luciano Floridi, and David Chalmers have emphasized that AI requires a fundamental rethinking of concepts such as intelligence, moral agency, and responsibility, as well as our ethical and epistemological

frameworks for evaluating action and decision-making in a world shared with non-human cognitive agents [9, 10, 18].

This philosophical context is essential for understanding the disruptive nature of AI. Where earlier narcissistic wounds were primarily epistemic (altering what humans believed about the universe, nature, and the mind) cognitive decentering has practical, ethical, and existential consequences. It demands reflection on questions of value, purpose, and the boundaries of human uniqueness.

Thought Experiment: AI and Human Identity

To illustrate the stakes of cognitive decentering, consider the following thought experiment: imagine a world in which AI systems can reliably outperform humans in all cognitive domains, from scientific reasoning and strategic planning to artistic creation and ethical deliberation. In such a world, which traits remain uniquely human, and what meaning does humanity derive from its traditional markers of intelligence?

This thought experiment surfaces several philosophical and psychological questions:

- **Identity and self-conception:** If cognitive performance is no longer a distinguishing feature, humans must reassess what constitutes self-worth, personal achievement, and collective identity.
- **Agency and moral responsibility:** With machines capable of making more accurate and informed decisions than humans, the attribution of responsibility becomes complex. Who is morally accountable for actions driven or mediated by AI systems?
- **Value and purpose:** If humans are no longer cognitively preeminent, traditional notions of purpose, vocation, and the hierarchy of labor may need re-evaluation. Societies may prioritize collaboration with AI, creativity in new domains, or ethical stewardship over computational supremacy.
- **Psychological effects:** Experiencing decentering can produce existential anxiety, imposter syndrome at a societal scale, or a sense of obsolescence. Conversely, it can also catalyze cognitive humility, openness to collaboration, and the exploration of capacities such as empathy, emotional intelligence, and moral imagination that many people still consider to be uniquely human.

This thought experiment underscores that cognitive decentering is not merely a technical or economic phenomenon; it is fundamentally philosophical and psychological. It challenges humans to redefine their self-conception, moral frameworks, and the very meaning of intelligence. Moreover, it illuminates the urgency of developing ethical, societal, and governance frameworks capable of managing the transformative potential of AI without undermining human dignity or agency.

AI as a Mirror of Human Cognition

In addition, AI serves as a mirror reflecting the nature and limitations of human cognition [19]. By externalizing pattern recognition, intuition, reasoning and creativity into artificial systems, AI forces humans to confront both the strengths and fragilities of their own cognitive processes. This mirroring effect provides a unique opportunity for self-reflection, allowing societies to critically examine biases, heuristics, and assumptions that were previously taken for granted. In this sense, AI's cognitive decentering is simultaneously a challenge and a pedagogical tool, prompting a reassessment of what it means to think, to create, and to act responsibly in an increasingly technologically mediated world.

Discussion

The phenomenon of cognitive decentering challenges deeply ingrained assumptions about human primacy in intelligence and agency. Initial reactions to AI's capabilities (ranging from fear and skepticism to awe) are natural psychological responses to the displacement of humans from domains traditionally considered uniquely human. Historical analogues, such as the Copernican, Darwinian, and psychoanalytic revolutions, suggest that while initial resistance may be intense, societies eventually integrate these insights into more mature and sophisticated worldviews [9, 17]. Cognitive decentering through AI, however, is qualitatively distinct: it operates directly on the substrate of cognition itself, affecting daily work, creativity, decision-making, and knowledge production.

Recognizing AI as a decentering revolution opens the possibility for humans to redefine value and purpose in ways that transcend raw cognitive performance. Instead of measuring worth by the ability to outcompute or outperform AI, humans can emphasize ethical responsibility and collective flourishing. This reframing encourages a more holistic conception of intelligence, where human-AI collaboration becomes a source of innovation rather than competition, and societal structures evolve to leverage complementary strengths. As Tomasello and Harari have emphasized, humans differ from the other great apes in having a highly advanced capacity to cooperate with one another. This capacity may now be crucial for cooperating with AI.

Extended Case Studies

AlphaFold in Biology

AlphaFold represents a paradigm shift in biological research. Its deep learning architecture predicts protein folding with unprecedented accuracy, surpassing the capabilities of expert human biochemists in speed, scale, and reliability [11]. The implications are profound: drug discovery pipelines can be accelerated, previously intractable biological problems become tractable, and human experts are liberated to focus on hypothesis generation, ethical evaluation, and interdisciplinary synthesis rather than exhaustive computation. AlphaFold exemplifies hybrid knowledge production, where human creativity and AI computational power synergize to advance scientific frontiers.

GPT Models in Creativity and Communication

Generative Pretrained Transformer (GPT) models illustrate cognitive decentering in the domains of language, creativity, and communication [13]. These models produce high-quality text, technical documentation, poetry, and narrative content, sometimes indistinguishable from human authorship. Their integration into creative workflows fosters hybrid authorship, where humans provide context, intent, and ethical direction, while AI offers speed, pattern recognition, and ideation. This partnership reshapes notions of originality, authorship, and intellectual labor, challenging traditional hierarchies in publishing, journalism, and creative arts. It also raises questions about responsibility, copyright, and cultural impact, emphasizing the need for ethical frameworks in hybrid cognitive systems.

Autonomous Decision Systems in Governance

AI-assisted decision-making is increasingly applied to governance, policy design, and resource allocation. Autonomous systems can analyze vast datasets, simulate policy outcomes, and optimize interventions in domains such as healthcare, urban planning, and disaster response [20]. These applications demonstrate the societal impact of cognitive decentering, as human decision-makers are no longer the sole arbiters of complex policy outcomes. AI augments human judgment, but also challenges traditional notions of accountability, authority, and transparency. Hybrid governance models, which combine AI-generated insights with human ethical oversight, exemplify how cognitive decentering can be harnessed for societal benefit while maintaining human agency and responsibility.

Implications for Human Identity and Societal Structures

Cognitive decentering prompts reflection on what remains uniquely human in a world of increasingly capable AI. Human identity may shift from performance-centered definitions to qualities such as wisdom, ethical reasoning, empathy, and collaborative problem-solving. Societal structures (educational, professional, and cultural) must adapt to prioritize these capacities. Education systems should foster meta-cognition, creativity, and critical reasoning, workplaces must integrate human-AI collaboration, and cultural narratives should normalize intelligence as a distributed, substrate-independent property.

Moreover, cognitive decentering underscores the importance of resilience and adaptability. Humans must learn to thrive not by competing with AI in raw cognitive tasks but by cultivating complementary skills, ethical foresight, and the ability to contextualize AI outputs within societal and humanistic frameworks. This approach aligns with historical patterns in which humanity has assimilated decentering insights (from the cosmos, biology, and the unconscious mind) to expand knowledge, self-understanding, and collective capability.

Ethical and Policy Considerations

The discussion of case studies and societal implications also emphasizes the need for robust ethical and policy frameworks. Ensuring transparency, fairness, accountability, and inclusivity in AI deployment is critical. Policies must address potential risks of misuse, bias, and concentration of decision-making power, while promoting AI applications that augment human judgment, creativity, and problem-solving capacity. The integration of cognitive decentering into governance and cultural discourse allows society to harness AI as a tool for human flourishing rather than as a source of alienation or ethical uncertainty.

Future Directions

Future research and practice should explore human-AI co-evolution: examining how humans and AI can jointly redefine the boundaries of knowledge, creativity, and decision-making. Interdisciplinary collaborations across computer science, philosophy, psychology, education, and governance will be essential to guide responsible AI deployment, maximize societal benefit, and cultivate a shared understanding of intelligence that transcends human-centered paradigms. Cognitive decentering, while initially disruptive, may ultimately catalyze a new era of human potential, ethical sophistication, and collaborative innovation.

Because cognitive decentering ultimately manifests as shifts in cognitive labor, epistemic authority, and control allocation, it demands evaluation frameworks that can quantify these changes. A concrete cognitive-computation agenda may operationalize this by 1) benchmarking human-only, AI-only, and joint human–AI performance to reveal emergent strengths and failure modes, 2) testing robustness under distribution shift to distinguish laboratory competence from real-world reliability, 3) measuring uncertainty calibration and reliance dynamics, including whether explanations improve verification or merely increase deference, and 4) conducting longitudinal designs that track whether sustained AI use produces deskilling, reskilling, or a reallocation of expertise toward evaluation and orchestration.

Conclusion

AI represents a profound decentering in human self-perception, constituting what can be understood as the fourth decentering revolution. Unlike previous revolutions (Copernican, Darwinian, and Freudian) which displaced humans from the perceived center of the cosmos, biology, and consciousness, cognitive decentering through AI directly challenges the notion of humans as the apex of intelligence. This shift is not merely technological; it has deep philosophical, ethical, and societal ramifications.

Cognitive decentering compels a re-evaluation of core assumptions about human uniqueness, agency, and responsibility. Humans are no longer the sole bearers of complex problem-solving, strategic decision-making, or creative production. AI systems demonstrate superhuman performance in domains ranging from scientific discovery to generative art and language synthesis. This transformation necessitates a reorientation of human roles toward a complementary relationship with AI rather than a competitive or adversarial one.

The implications of this decentering are far-reaching. In ethics and governance, humans must ensure AI is deployed responsibly, with transparency, fairness, and accountability, while mitigating risks of bias, centralization, and misuse. In education, curricula must evolve to prioritize AI literacy, critical thinking, creativity, social intelligence, and metacognition, equipping learners to thrive in hybrid human-AI environments. In work and knowledge production, humans must integrate AI as a collaborative partner, enabling faster discovery, enhanced creativity, and more effective decision-making. Culturally, cognitive decentering invites societies to reconsider notions of intelligence, value, and human identity, fostering humility and ethical stewardship.

Recognizing AI as a decentering revolution encourages humans to embrace strategic collaboration with non-biological intelligences. Rather than perceiving AI solely as a competitor, society can leverage its strengths to augment human capabilities, expand collective knowledge, and address complex global challenges. In the immediate future, this perspective aligns with an emerging paradigm of hybrid intelligence, in which human judgment, creativity, and ethical sensibilities operate in tandem with AI's computational power, pattern recognition, and generative capacities. In the much longer term, it is still unclear how or if human-AI collaboration will work.

Ultimately, cognitive decentering is both a challenge and an opportunity. It challenges humans to confront discomforting truths about their cognitive limits, requiring humility and adaptability. Simultaneously, it offers the opportunity to redefine value, purpose, and progress in ways that transcend individual cognition. By fostering ethical collaboration, promoting inclusive governance, and integrating AI into education and cultural imagination, humanity can harness this fourth decentering to create a more just, innovative, and resilient society: one in which AI serves as a partner in human flourishing rather than a source of alienation or displacement.

In conclusion, the fourth decentering revolution urges a fundamental rethinking of what it means to be human in an AI-augmented world. It is a call to embrace humility, cultivate capacities that are uniquely human for the time being, and strategically collaborate with artificial intelligences to expand the horizons of knowledge, creativity, and ethical responsibility. By acknowledging and integrating cognitive decentering into societal structures, education, governance, and culture, humanity can navigate this transformative era with foresight, wisdom, and collective resilience.

Author Contributions Erik Cambria: Conceptualization, Investigation, Visualization, Writing – Original Draft; Rui Mao: Investigation, Writing – Review & Editing; Nicola Bianchi: Investigation, Writing – Review & Editing; Amir Hussain: Investigation, Writing – Review & Editing; Keith Oatley: Investigation, Writing – Review & Editing; Geoffrey Hinton: Investigation, Writing – Review & Editing.

Funding This research received no external funding.

Data Availability No datasets were generated or analysed during the current study.

Declarations

Competing Interests The authors declare no competing interests.

References

1. Mao R, Liu Q, Li X, Cambria E, Hussain A. Bridging minds and machines: Toward an integration of AI and cognitive science. 2025. arXiv preprint arXiv:2508.20674
2. Zhang T, Zhou X, Wang Y, Cambria E, Traum D, Mao R. Individualized cognitive simulation in large language models: Evaluating different cognitive representation methods. 2025. arXiv preprint arXiv:2510.20252
3. Chen M, Cambria E, Laschi C, Mengaldo G. Intelligence is physical: Energy, information, & a thermodynamically grounded theory of minds & machines. 2025. https://philsci-archive.pitt.edu/27741
4. Copernicus N. De revolutionibus orbium coelestium. Johannes Petreius, 1543.
5. Kuhn TS. The copernican revolution: planetary astronomy in the development of western thought. Cambridge: Harvard University Press; 1957.
6. Galilei G. Sidereus Nuncius. Johann Georg, Venice, 1610.
7. Darwin C. On the Origin of Species. London: John Murray; 1859.
8. Freud S. A difficulty in the path of psycho-analysis. In Strachey J, editor. The standard edition of the complete psychological works of sigmund freud. Hogarth Press; 1917. vol. 17, p. 135–144.
9. Floridi L. The logic of information: a theory of philosophy as conceptual design. Oxford: Oxford University Press; 2020.
10. Dennett DC. From bacteria to bach and back: the evolution of minds. New York: Norton and Company; 2017.
11. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Žídek A, Potapenko A, Bridgland A, Meyer C, Kohl SAA, Ballard A, Cowie A, Romera-Paredes B, Nikolov S, Jain R, Adler J, Back T, Petersen S, Reiman D, Clancy E, Zielinski M, Steinegger M, Pacholska M, Berghammer T, Bodenstein S, Silver D, Vinyals O, Senior A, Kavukcuoglu K, Kohli P, Hassabis D. Highly accurate protein structure prediction with AlphaFold. Nature. 2021;596(7873):583–9.
12. Yang Z, Du X, Li J, Zheng J, Poria S, Cambria E. Large language models for automated open-domain scientific hypotheses discovery. In Proceedings of ACL. 2024. p. 13545–13565
13. OpenAI. GPT-4 technical report. 2023. https://openai.com/research/gpt-4. Accessed 18 Jan 2026.
14. Ramesh A, Dhariwal P, Nichol A, Chu C, Chen M. Hierarchical text-conditional image generation with CLIP latents. 2022. arXiv preprint arXiv:2204.06125
15. Silver D, Huang A, Maddison CJ, Guez A, Sifre L, Van Den Driessche G, et al. Mastering the game of Go with deep neural networks and tree search. Nature. 2016;529(7587):484–9.
16. Lin Q, Zhu Y, Mei X, Huang L, Ma J, He K, et al. Has multimodal learning delivered universal intelligence in healthcare? A comprehensive survey. Inf Fusion. 2025;116:102795.
17. Freud S. Introductory lectures on psychoanalysis. London: Hogarth Press; 1917.
18. Chalmers DJ. Oxford: The character of consciousness. Oxford University Press; 2010.
19. Mao R, Chen G, Li X, Ge M, Cambria E. A comparative analysis of metaphorical cognition in ChatGPT and human minds. Cognitive Comput. 2025;17:35.
20. Rahwan I. Society-in-the-loop: Programming the algorithmic social contract. Ethics Inf Technol. 2019;21:5–14.