

MAN AND MACHINE: THE LIMITS OF HUMANITY AND AI. PROLEGOMENA

Aldona MUSIAŁ-KIDAWA

Silesian University of Technology, Faculty of Organization and Management; aldona.musial-kidawa@polsl.pl,
ORCID: 0000-0001-9525-6348

Purpose: The aim of the publication is to analyze the limits of artificial intelligence in emulating the human mind and to show that, although AI can imitate cognitive processes, it remains far from fully capturing the specificity of human nature, which encompasses dimensions such as imagination, intuition, emotionality, and creative reflection.

Methodology: The publication is theoretical in nature and is based on an analysis of the relevant literature. The applied method allowed for an examination of the current state of knowledge by combining reflection on the concept of human nature with a critical interpretation of literature in the field of artificial intelligence research. The analysis focuses on comparing the functional capabilities of AI with the essence of the human mind and its elusive dimensions.

Findings: The publication highlights key issues related to the impact of AI on the condition of contemporary humans: the development of AI allows us to examine which aspects of human functioning can be formally captured and which remain elusive. Humans are not limited to information processing; they experience meaning, values, and responsibility, which go beyond the capabilities of AI. Analyzing these relationships leads to a deeper understanding of the differences and similarities between human and artificial intelligence.

Social consequences: The social consequences of the increasing significance and role of AI may include the need to revise the division of roles and responsibilities between humans and machines, the strengthening of the importance of ethical competencies, and the development of debates on regulations and the limits of artificial intelligence applications in social life.

Originality: The study indicates which aspects of human functioning can be replicated in computational systems and which remain elusive, treating artificial intelligence as a tool for exploring human nature.

Keywords: human, artificial intelligence, phenomenology, functionalism, subjective experience, human nature, human-machine relation.

Category of the paper: conceptual paper, viewpoint.

Introduction

The development of artificial intelligence systems in recent decades from classical symbolic algorithms to contemporary deep learning models and large language models (Kaplan, 2025) poses a question that has for centuries stood at the core of philosophical reflection: what, in fact, is a human being, and what distinguishes humans from machines? In an era when machines are capable of generating language, solving problems that require planning, and displaying apparent forms of creativity, the classic question “what makes us human” returns with renewed force and in a new context. It is no longer merely a problem of abstract philosophy; it is becoming a real epistemological, ontological, and ethical challenge (Harari, 2024). The reflections focus on the question of whether intelligence, understood as the capacity to process information, learn, reason, and adapt, constitutes the entirety of humanity, or whether there are aspects of human nature that cannot be reduced to it. In the context of contemporary AI systems, which can imitate human behavior at a functional level, the problem becomes particularly pressing (Chalmers, 2010). Is the simulation of thinking and understanding equivalent to genuine thinking? Does the ability to perform cognitive tasks imply participation in the same experience of consciousness and responsibility that is proper to human beings?

Meanwhile, the literature on artificial intelligence is dominated by analyses focusing on the technical capabilities of computational systems and their functional effectiveness in performing cognitive tasks. Much less frequently undertaken is an in-depth reflection on the ontological consequences of a functional understanding of intelligence, or on whether the ability to simulate intelligent behavior can be regarded as a sufficient criterion of humanity. This creates a research gap concerning the relationship between a functional conception of intelligence and the multidimensional structure of human subjectivity. Based on these gaps in the literature, the article proposes the hypothesis that the ability of artificial intelligence systems to simulate intelligent behavior does not exhaust the full structure of humanity, as it does not encompass the dimensions of phenomenal consciousness, embodied existence, and normative responsibility. The objective of the article is to analyze the limitations of a functionalist understanding of intelligence in the context of contemporary artificial intelligence systems, particularly large language models. The study has a philosophical-analytical character and is based on a conceptual analysis of key concepts such as intelligence, consciousness, intentionality, and subjectivity, as well as a comparison of the functional model of intelligence with the operational principles of contemporary AI systems and a critical reflection on the limitations of functionalism in capturing the full structure of humanity, highlighting dimensions that are irreducible to computational processes. The contribution of the article lies in demonstrating that the ability to simulate intelligent behavior does not exhaust the structure of the human mind. The analysis shows that dimensions such as phenomenal consciousness, embodiment, and normative responsibility remain irreducible to functional models of

information processing. In this way, the article contributes to a deeper philosophical reflection on the relationship between humans and artificial intelligence.

Cogito ergo sum... or not?

In the context of research on thinking machines, the discussion of what makes us human naturally focuses on the concept of intelligence. A fundamental question arises: is the capacity to “think” in a computational sense sufficient to regard a system as intelligent in the full, human sense? This problem has been considered in modern philosophy and cognitive science by both classical thinkers and contemporary researchers of artificial intelligence (Chwedeńczuk, 1995; Prescott, 2025). The beginning of the modern debate on “thinking machines” is associated with Alan Turing’s work *Computing Machinery and Intelligence* (Turing, 1950). Instead of asking the metaphysical question, “can machines think?”, Turing proposed to operationalize the problem in the form of the so-called imitation test. The criterion did not refer to the internal structure of the machine, but to its behavior: if, in the course of linguistic communication, a system proves indistinguishable from a human for a competent observer, then in a functional sense the system may be said to “think”. In the functionalist approach, mental states are defined by their role within a system: by input–output relations, patterns of information processing, and the function they perform in generating behavior (Dennett, 2016). In this sense, it does not matter what “material” the subject is made of whether biological or silicon-based. What matters is its functional organization. The Turing test shifts the centre of gravity from ontology to epistemology: instead of examining what thinking is “in itself”, we ask according to what criteria we can attribute it to someone. Intelligence becomes identified with the ability to produce contextually appropriate responses that are semantically coherent and pragmatically effective, such that an observer cannot distinguish the machine from a human. Contemporary language systems, especially generative models, are approaching the realization of this criterion in certain domains (Kaplan, 2025). In this sense, Turing’s project proved prescient, for intelligence appears as an engineering problem rather than a metaphysical one: functionalism, in Turing’s formulation, makes it possible to separate science from metaphysics. It becomes possible to study machine intelligence empirically: we measure effectiveness in performing cognitive tasks, generating coherent responses, or adapting to new problems: “We cannot rule out the possibility that one day some large generative network will begin to talk with us and we will be able to ask it for explanations” (Sejnowski, 2019, p. 151). However, this shift carries a significant philosophical consequence: if intelligence is taken as a sufficient criterion for humanity, then the human being ceases to be perceived through the lens of subjective perspective and individual experience, and instead begins to be evaluated according to measurable functions. To be human would mean nothing more than effectively processing

information, generating responses, and adapting to a symbolic environment. This perspective, however, leads to a far-reaching reduction, narrowing the concept of humanity to a set of cognitive competences that can be described in functional and operational terms. Functionalism assumes, often only implicitly, that everything essential to the mind can be captured in the structure of functional relations, that is, in the way processes are organized and interconnected. Consequently, consciousness can be interpreted as a particular type of computational organization, meaning as the result of operations performed on symbols, and subjectivity as the effect of integrating informational processes into a coherent whole. As Dennett observes: “Essentially, it does not matter whether a system is made of organic particles or silicon, as long as it performs the same task. Artificial hearts do not need to be made of organic tissue, so artificial brains need not be either - at least in theory. If all the control functions of a human wine taster’s brain can be reproduced using silicon circuits, the pleasure will be *ipso facto* reproduced along with them” (Dennett, 2016, p. 47). Understood in this way, the perspective shifts the focus from the ontological status of the person to their functional structure, leading to the conception of the human primarily as an information-processing system whose properties can be described, modeled, and reconstructed in formal terms. Such an approach may result in cognitive reductionism, in which humanity is reduced to intellectual capacities, while other dimensions, such as embodiment, emotionality, existential aspects, or normativity, are treated as secondary to information processing. From this perspective, the distinction between humans and advanced artificial intelligence would be quantitative, depending on the system’s level of complexity, the amount of data processed, or the speed of computation, rather than on essential qualitative differences. The human-machine boundary would thus become fluid and largely technical, practically reducible to the operational parameters of the system (Dennett, 2016).

Although the functional paradigm defines intelligence as the ability to effectively process information and generate appropriate responses, it reveals a number of critical tensions that highlight its limitations as a criterion for humanity. The first of these relates to the very process of understanding: the question arises whether a system that produces correct answers actually understands their meaning, or merely manipulates symbols according to formal rules. A classic illustration of this problem is Searle’s Chinese Room argument, which shows that even if a machine behaves intelligently from an external perspective, it lacks access to the semantics of the symbols, because there is no internal subject that genuinely understands. This issue will be explored in greater detail later in the publication. The second tension is ontological: the Turing test measures only the external indistinguishability of human and machine behavior, without resolving whether the system possesses any experience, intentionality, or first-person perspective. From Nagel’s point of view, even the most functionally complex systems cannot capture what it is “like” to be a conscious subject, because subjective experience cannot be reduced to behaviors observable from the outside (Nagel, 1997). This criterion is therefore pragmatic and epistemic in nature; it does not encompass or resolve metaphysical problems. The third tension concerns normativity: the functional understanding of intelligence focuses on

adaptive effectiveness, whereas humans are not limited to reacting to stimuli and optimizing actions; they are capable of questioning their own goals, refusing to act, acting against their own interest, or sacrificing themselves for the sake of values (Schaich-Borg, Conitzer, Sinnott-Armstrong, 2024). Reducing humanity to cognitive competences thus overlooks a crucial normative and existential dimension. Therefore, in the context of AI, the boundary between humans and machines is not merely a matter of technical parameters, system complexity, or the amount of data processed. It is philosophically rooted in the structure of subjectivity, in consciousness, intentionality, and normative responsibility, qualities that no machine, however advanced, possesses. As Putnam observes: “Admittedly, a machine may converse beautifully about, say, Polish landscapes. However, it would not be able to recognize a tree, an apple, a mountain, or a cow, a field, or a tower if placed before any of these things. There is no reason to believe that when a machine talks about apples, its words refer to actual apples, just as there is no reason to believe that an ‘outline’ of an ant refers to Winston Churchill” (Putnam, 1998, p. 309). Reducing these aspects to functional operations can create the mistaken impression that computational intelligence exhausts the essence of human being, whereas in reality they remain dimensions that cannot be translated into formulas, models, or algorithms, dimensions that make a human a conscious, moral, and self-aware subject. Consequently, although the functional paradigm has proven methodologically groundbreaking, enabling empirical research on artificial intelligence, its application as a criterion for humanity is insufficient. It does not engage with metaphysical questions that go beyond functional effectiveness and therefore overlooks the deeper dimensions of human existence: phenomenal consciousness, embodied existence, intentionality, and normative and moral subjectivity. Behavioral indistinguishability, then, guarantees neither understanding, consciousness, nor normative subjectivity. Efficient information processing and generation of appropriate responses may constitute a necessary condition, but they are insufficient for full human being. A profound insight in this context is provided by Kołakowski, who notes: “Sciences measure things without knowing what they measure; in performing cognitive acts, they are unable to grasp the acts themselves. They are incapable of spontaneously creating their own meaning and justifying their claims to objectivity” (Kołakowski, 2003, p. 32).

Man and machine

Since the functionalist conception of intelligence redefines the boundary between humans and machines, the fundamental question is whether its assumptions are borne out in actual artificial intelligence systems. Contemporary large-scale language models represent the most advanced realization of this paradigm, and therefore their analysis allows for an empirical examination of the philosophical consequences of functionalism.

The functionalist paradigm in the context of contemporary artificial intelligence systems gains a new perspective in the 21st century through large-scale language models (LLMs) such as GPT, PaLM, and LLaMA. These systems are based on transformer neural architectures, which enable the models to capture complex statistical dependencies present in data sequences. By analyzing vast text corpora, they identify linguistic regularities and extract recurring structural and semantic patterns. On this basis, the models generate predictions of subsequent tokens according to the context of an utterance, which in practice results in coherent, logically connected language sequences often interpreted as a sign of “understanding” (Kacprzak, 2024). At the same time, these systems demonstrate the ability to adapt to new tasks through procedures such as fine-tuning or prompt engineering, without the need to modify the underlying model architecture. This mechanism reflects the functionalist conception of intelligence, in which the essence of intelligence is determined by the cognitive function it performs, rather than the material substrate on which it is implemented. In practice, LLMs are capable of producing grammatically flawless statements, answering questions, and creating coherent narratives, which from the perspective of the Turing test can be interpreted as “intelligent behavior”. As Kaplan notes: “Generative artificial intelligence shakes the foundations of our sense of self and challenges our view of human uniqueness in much the same way the Copernican revolution altered our understanding of the solar system by placing the Sun at its center. Are we merely a stepping stone to a new form of non-biological life? Or are we simply getting better at constructing useful gadgets?” (Kaplan, 2025, p. 16). Technical analysis, however, reveals significant limitations of the functionalist approach to intelligence in the context of large language models. First, these models do not possess internal semantics; they operate solely on token statistics and correlations between them, without forming meaningful representations in a subjective sense. In Searle’s terminology, such a system “processes symbols” but does not understand them (Searle, 1995). Despite the impressive capabilities of contemporary LLMs, a crucial philosophical and technical limitation emerges: the absence of genuine understanding of meaning. Searle articulated this critique of functionalism in the context of functional intelligence through the Chinese Room argument. Imagine a language model that receives text in Chinese and generates correct responses using only statistical procedures and transformational rules. From the outside, the system’s behavior is flawless: it responds sensibly, coherently, and appropriately. Yet, it contains no semantic representation or subjective understanding of the language. The model manipulates symbols based on statistical patterns from training data, but it does not “know” what they mean, just as a person in the Chinese Room manipulating symbols according to instructions does not understand the language, even if their responses appear correct to an observer. Contemporary neural models confirm Searle’s fundamental thesis that simulating understanding is not equivalent to actual understanding, that the absence of intentionality and a first-person perspective limits the system to the syntactic rather than the semantic level, and that external evaluation of behavior cannot fully reveal the internal state of the model. The fact that responses are sensible does not prove

the existence of a subjective perspective. Searle argued that: “Equipping any artifact we build with a computer program is not sufficient for it to have mental states comparable to human ones. Such an artifact should, of course, have causal powers comparable to those of the human brain. Brain activities limited only to executing a computer program are not sufficient for the functioning of the brain to give rise to a mind” (Searle, 1995, p. 36). Consequently, the Chinese Room argument reveals a fundamental ontological gap, as it exposes the difference between simulation and realization. Simulating a cognitive process is not equivalent to its actual occurrence, just as simulating a physical phenomenon is not the phenomenon itself. A language model may reproduce the structures of reasoning, but it does not need to possess the states that, in humans, constitute the experience of understanding; it does not experience meanings, relate them to the world intentionally, or have a first-person perspective. This problem is illustrated by Chalmers, who analyzes a hypothetical system capable of reporting its own perceptual states. Such a system could claim that it “sees” an object and base its beliefs on this claim, even though it would be unclear whether it actually experiences anything. As Chalmers notes, it is more likely that the system would respond: “I know there is a red tricycle there because I see it”, and if asked how it knows that it sees it, it would reply: “I just see it” (Chalmers, 2010, p. 317). This description shows that the functional ability to report cognitive states does not entail the existence of phenomenal experience. A system may behave as if it has access to its own experiences, even though there is no basis for asserting that any experiences actually occur. Thus, the simulation of cognitive processes does not imply the existence of consciousness or a first-person perspective. A similar point is made by Sejnowski: “Although neural networks correctly solve the tasks assigned to them, we currently do not have an appropriate method that could explain how they arrive at the correct answers” (Sejnowski, 2019, p. 150). This gap is structural, not gradual. It does not concern the level of system complexity or the scale of its computational capacities, but the very kind of being we are dealing with. On one side is a functional system whose operations can be described in computational and behavioral terms; on the other is a subject capable of experiencing, understanding, and relating to meanings internally. Even perfect alignment of a model’s behavior with that of a conscious being does not eliminate this difference, because it concerns not what the system does, but what it is and whether it possesses a qualitative dimension of experience.

In this sense, the ontological gap indicates the limits of functionalist explanations of the mind. It shows that describing a system’s operations, even in extremely precise and complete terms, may not suffice to explain the existence of consciousness or subjective understanding. There remains a distinction between operational efficiency and phenomenal experience, a distinction that cannot be reduced to mere function, structure, or behavior. It is precisely this irreducible difference that forms the core of the ontological problem in the analysis of contemporary artificial intelligence systems. Another important limitation is the lack of intentionality and subjectivity. The cognitive capacities of models are largely determined by the scope and quality of the training data, and the systems themselves do not “think” or form

concepts independently of the available information, which restricts their cognitive autonomy. Furthermore, these models lack embodied experience; the absence of sensory input and interaction with the physical world means that their “situational awareness” is purely syntactic and contextual rather than phenomenal, that is, devoid of subjective experience. As Przegalińska notes: “Humans use body language, facial expressions. This is our interface. It gives us an advantage in discussion; it is in itself more persuasive than a computer block standing in front of you and presenting arguments” (Przegalińska, 2020, p. 109). Consequently, although language models can generate behaviors resembling intelligence, the question remains open as to whether these behaviors are sufficient to capture the fullness of human being or to draw conclusions about the nature of humanity. Analyzing contemporary large-scale language models (LLMs) in light of the functionalist paradigm, several key aspects can be observed. First, intelligence in this framework is perceived as irreducibly behavioral; it is evaluated based on manifested behaviors rather than by analyzing internal states, processes, or the structure of the system. Language models practically illustrate the potential of functionalism, as many behaviors traditionally considered signs of intelligence can be effectively replicated in digital systems. At the same time, a significant ontological gap remains: these models lack consciousness, intentionality, and the capacity for normative responsibility. They can simulate reasoning processes and generate meaningful responses, but they do not possess a first-person perspective or genuine understanding of meaning. As Schneider observes: “Currently, we face the question of consciousness - a kind of ‘hard problem’ of machine consciousness: whether information processing by AI will be associated with any internal experiences” (Schneider, 2021, p. 32). In this sense, LLMs confirm both the strengths and the limitations of the functionalist paradigm: they demonstrate that intelligence can be formalized and implemented in machines, yet at the same time they reveal that functionally intelligent behavior is not equivalent to humanity and is far from capturing the essence of the mind and human being. The way LLMs operate exposes the ontological limits of functionalism: intelligence in the sense of symbol processing does not provide subjective experience, normative subjectivity, or the capacity for creatively assigning meaning and value. In this context, one can refer to the embodied approach, which emphasizes an aspect of cognition unavailable to artificial intelligence: “The embodied approach to studying cognitive processes is based on the thesis that a condition for understanding the nature of cognition is to pay attention to the essential role played by the acting body, situated in a specific environment and interacting with it dynamically, online, and without mediation by representations” (Przegalińska, 2016, p. 178).

The analysis of the functionalist paradigm in relation to contemporary language systems reveals a subtle yet crucial problem: the simulation of behavior is not equivalent to the experiencing of consciousness. Machines can generate utterances suggesting an emotional state, report their own “experiences”, or respond contextually, both in Turing tests and in everyday interactions with LLMs. However, these actions remain entirely syntactic and formal; from the outside, they may appear as intelligent responses, but internally there is no subject that

actually experiences anything. Philosophically, this exposes a significant gap between the functional capacity to behave intelligently and phenomenal consciousness (Chalmers, 2010). The fact that a machine can “describe” emotional states or coherently report its “experiences” does not imply the existence of a subjective point of view. Subjectivity and experience are inextricably linked to a first-person perspective; without it, we observe only an external effect, a semblance of intelligence that is not fully human. In this sense, contemporary LLMs, despite their impressive ability to imitate behaviors and create first-person narratives, remain “black boxes” without a subjective interior. Their functioning may be functionally analogous to human thinking, but there is no guarantee that they actually “experience” anything. This raises the need for deeper reflection on the limits of artificial intelligence in the context of what defines human existence and allows us to speak of it as a unique form of being, especially as Harari notes: “Science concentrates all its claims in a single all-encompassing dogma, which states that organisms are algorithms, and life is data processing. Intelligence separates from consciousness. Devoid of consciousness but extremely intelligent algorithms may soon know us better than we know ourselves” (Harari, 2015, p. 445). Technological development provokes a shift in the boundaries of subjectivity and cognitive capacity. The question of the limits of artificial intelligence thus becomes, in essence, a question about the limits of humanity. The focus shifts from anthropology to modern technology. The question is old, but it arises in a new configuration: who is human, and in what sense do consciousness, subjectivity, and the capacity to assign meaning determine our uniqueness and remain the foundation for understanding our existence in terms irreducible to an algorithm. To speak of what truly makes us human, it is not enough to analyze cognitive abilities or external behaviors. One must consider the subjective dimension of experience, embodied existence, and normative responsibility (Przegalińska, 2016). Machines can simulate intelligence and generate linguistic descriptions of states and relations, but “thinking” in the functional sense does not imply being a subject that experiences. In this light, the Turing test and AI simulations remain extremely valuable research tools, but they do not resolve a question constitutive for the philosophical understanding of humanity: whether we are dealing with a being capable of lived experience, or merely a system reproducing its external forms. They rather reveal the limits of formal intelligence: limits that indicate that phenomenal consciousness and subjectivity remain a qualitatively distinct dimension from functional cognitive competencies (Przegalińska, Triantoro, 2025).

Conclusion

The development of artificial intelligence does not so much pose entirely new questions as it radically intensifies old ones, forcing us to reconsider them in light of new technological possibilities. Already in classical debates in the philosophy of mind, the relationship between function and experience, between behavior and lived experience, and between structure and meaning was already being examined (Searle, 1995). However, it is only with the emergence of systems capable of autonomously generating utterances, learning from data, and simulating decision-making processes that these questions cease to be purely speculative. They become experimentally tangible: machines are no longer merely hypothetical thought constructs, but real cognitive partners with whom humans interact, as Przegalińska aptly notes: “Chatbots and avatars defy the label of ‘computer program’; they are something much more, something that enters into an interesting, complex relationship with the identity and consciousness of the subject who chooses to engage in virtual reality” (Przegalińska, 2016, p. 189). In this sense, artificial intelligence functions as a kind of “epistemic mirror.” It allows us to observe which aspects of human functioning can be formally described and reproduced, and which elude modeling. Every success of AI systems in pattern recognition, language generation, or action planning provokes the question of whether a given capability was truly a constitutive feature of humanity or merely a functional manifestation of it. At the same time, every technological boundary, difficulties in capturing subjective experience, normative meaning, or the full existential context, points to potential dimensions of human uniqueness. The development of artificial intelligence thus operates as a tool of philosophical reduction: it systematically tests hypotheses about the nature of humans through attempts at their implementation. From this perspective, the question of humanity is no longer merely about what humans are, but also about what can be reconstructed, simulated, or replaced. The problem thus takes on a comparative form: it is not only about describing human nature, but about determining the ontological difference between biological beings and computational systems. Is this difference merely one of degree, arising from the level of complexity, or is it qualitative, concerning the very mode of existence? In other words, is a human being simply a particularly complex information-processing system, or a being whose structure encompasses dimensions irreducible to computation? This distinction directly leads to the tension between functional description and phenomenological description. Artificial intelligence models operate at the level of formal relations: symbol transformations, statistical structures, or vector representations. Meanwhile, human existence reveals itself not merely as a system of cognitive operations, but as a lived world of meanings, values, and references. Humans do not simply process information; they experience significance, suffer, desire, make decisions within a horizon of responsibility, and live in a world that is always already meaningful to them. Kołakowski observes that human existence has always developed in the context of symbols and myths, which give significance

to experience and shape our understanding of the world. Humans are not merely a collection of cognitive functions or biological mechanisms; their psychic and cultural life creates dimensions of meaning, value, and subjectivity that exceed the scope of simple data analysis (Kołakowski, 2005). If these aspects cannot be fully captured in functional terms, then computational intelligence, no matter how advanced, cannot exhaust the structure of humanity (Przegalińska, 2016). At the same time, the development of artificial intelligence forces a more precise clarification of what we actually mean by consciousness, embodiment, and normativity. These concepts cannot remain merely intuitive references to “what is human”, but require systematic philosophical analysis capable of engaging with cognitive science and artificial systems theory (Prescott, 2025). The stake is not merely defending human uniqueness, but developing an adequate model that allows us to understand the relationship between different forms of biological and artificial intelligence without reductionism, yet also without unjustified anthropocentrism. In this sense, a layered model of humanity is both analytical and methodological. It does not presuppose an insurmountable gulf between humans and machines, but systematically examines the levels at which similarity is possible and those at which qualitative differences emerge. This model allows computational intelligence to be treated as one level of description, a functional level that can be shared by both biological and artificial systems, but one that does not exhaust the full structure of human existence (Bennett, 2025). Ultimately, reflection on artificial intelligence becomes reflection on the limits of reduction. It asks not only what machines can do, but also what cannot be reduced to mere function. At this point, the problem of artificial intelligence becomes, in the deepest sense, an anthropological problem: it forces us to reconsider whether humanity is a functional, phenomenological, existential, or normative category, or perhaps a multilayered structure that no single description can fully capture. As Kołakowski notes: “One could say that in the entire universe, a human cannot find a well so deep that, leaning over it, they would not discover their own face at the bottom” (Kołakowski, 2000, p. 78). This serves as a starting point for further reflection, in which successive dimensions of humanity will be examined not only in isolation but also in their interrelations. Only through their integration can we grasp whether, and in what sense, it is possible to compare humans and artificial intelligence, and where the boundary lies between simulation and being.

References

1. Bennett, M. (2025). *Krótką historia inteligencji. Ewolucja, sztuczna inteligencja i pięć przełomów, które ukształtowały nasz mózg*. Zysk i S-ka.
2. Chalmers, D. (2010). *Świadomy umysł. W poszukiwaniu teorii fundamentalnej*. PWN.
3. Chwedeńczuk, B. (1995). *Filozofia umysłu*. Aletheia.

4. Dennett, D.C. (2016). *Świadomość*. Copernicus Center Press.
5. Harari, Y.N. (2015). *Homo deus. Krótka historia jutra*. Wydawnictwo Literackie.
6. Harari, Y.N. (2024). *Nexus. Krótka historia informacji od epoki kamienia do sztucznej inteligencji*. Wydawnictwo Literackie.
7. Kacprzak, A. (2024). *Prompt engineering i ChatGPT. Poradnik skutecznej komunikacji ze sztuczną inteligencją*. Helion.
8. Kaplan, J. (2025). *Generatywna AI*. PWN.
9. Kołakowski, L. (2000). Prawda i prawdomówność jako wartości kultury. In: *Kultura i fetysze*. PWN.
10. Kołakowski, L. (2003). *Husserl i poszukiwanie pewności*. Znak.
11. Kołakowski, L. (2005). *Obecność mitu*. Prószyński i S-ka.
12. Nagel, T. (1997). *Widok znikąd*. Aletheia.
13. Prescott, T. (2025). *Psychologia sztucznej inteligencji*. Wydawnictwo Uniwersytetu Łódzkiego.
14. Przegalińska, A. (2016). *Istoty wirtualne. Jak fenomenologia zmieniła sztuczną inteligencję*. Universitas.
15. Przegalińska, A., Oksanowicz, P. (2020). *Sztuczna inteligencja. Nieludzka, arcyłudzka*. Znak.
16. Przegalińska, A., Triantoro, T. (2025). *Przenikanie umysłów. Potencjał twórczy współpracy z AI*. AIBooks by CampuAI.
17. Putnam, H. (1998). Mózgi w naczyniach. In: *Wiele twarzy realizmu i inne eseje*. PWN.
18. Schaich-Borg, J., Conitzer, V., Sinnott-Armstrong, W. (2024). *Moralna AI. Czy bać się sztucznej inteligencji*. Prószyński Media Sp. z o.o.
19. Schneider, S. (2021). *Świadome maszyny. Sztuczna inteligencja i projektowanie umysłów*. PWN.
20. Searle, J. (1995). *Umysł, mózg i nauka*. PWN.
21. Sejnowski, T.J. (2019). *Deep learning. Głęboka rewolucja. Kiedy sztuczna inteligencja spotyka się z ludzką*. Poltext.
22. Turing, A. (1972). Maszyny liczące a inteligencja. In: E. Feigenbaum, J. Feldman (eds.), *Maszyny matematyczne i myślenie*. PWN.