

Alienated Intelligence: Cultural Projection, Cognitive Outsourcing, and the Myth of Artificial Agency

JAKUB WALICKI

Pomeranian University in Słupsk

jakub.walicki@upsl.edu.pl

ORCID: 0009-0007-0770-2266

Abstract. This article introduces the concept of “alienated intelligence” to critically examine contemporary cultural perceptions of artificial intelligence (AI). Drawing on Ludwig Feuerbach’s theory of projection and alienation, it argues that certain visions of AI function as cultural projections of human cognitive capacities that become externalized, idealized, and reified in technological systems. The analysis situates this claim within a broader philosophical framework, engaging Socrates’ notion of wisdom as awareness of ignorance, Dilthey’s account of embodied and integrated cognition, and Damasio’s critique of the Cartesian separation of reason and emotion. Against reductive computational models of mind, the article emphasizes that human cognition is not merely calculative but structurally unified, affective, and contextually embedded. It further argues that the outsourcing of cognitive activities to AI risks not only diminishing retained knowledge but also narrowing our awareness of what we do not know. Finally, the article examines the anthropomorphization and mystification of algorithmic systems as expressions of a broader cultural tendency to mythologize opaque technologies.

Keywords: AI, Artificial Intelligence, Human Cognition, Ludwig Feuerbach, Alienation, Religion Studies, Wilhelm Dilthey, Organizational Knowledge, Digital Transformation, Socrates, Damasio.

Contribution. The author independently defined the research topic and formulated the central thesis and supporting arguments. The study was based on a critical analysis and interpretation of relevant philosophical and scientific literature. On this basis, the author identified key theoretical problems, developed the conceptual framework of the paper, and articulated the main conclusions. All stages of the research and writing process – from conceptualization and argument development to drafting and revision – were carried out solely by the author.

Use of AI. AI was used to assist in translating notes for the article from Polish into English and in editing some sections.

Introduction

Although many AI systems remain narrowly task-specific and operate without semantic understanding or self-awareness, recent advances in machine learning and predictive architectures have fostered a growing tendency to view certain types of artificial intelligence not merely as assistive tools but as agents capable of replacing aspects of human cognition in increasingly complex tasks. This cultural shift is evidenced in the ways AI is represented in public discourse and media. Algorithms are often said to “decide”, “learn”, or even “create”, and such metaphors suggest agency, intention, or even consciousness.

As Carl Öhman puts it, large language models are the outcome of “vast volumes of data, literally traces of past human (speech) acts, synthesized into a unified agency that is (falsely) experienced by users as an extra-human agency” (Öhman, 2023, 1). Such descriptions point to a growing cultural tendency to perceive AI not just as a tool, but as a quasi subject – a unified agent with its own apparent agency – contributing to our inclination to anthropomorphize technological systems.

In this article, I argue that we are dealing with the phenomenon of “alienated intelligence” – a technological process analogous to the religious alienation described by Ludwig Feuerbach (Section 2). Just as in religion, where humans project their best qualities onto God, in the context of AI, we project our knowledge and intellectual capabilities

onto machines, which we then treat as autonomous entities capable of independent thought.

By “alienated intelligence” I mean intelligence that has been projected by humans onto technology – in other words, cognitive capacities that have been externalized into AI systems. It is intelligence detached (alienated) from the human individual and embodied in the machine. At the same time, this process also implies the alienation of human intelligence itself – our own cognitive abilities are transferred outside of ourselves (Section 4).

In addition to the main argument, I propose several complementary theses. First, I emphasize that human cognition is embodied and integrated: thought, emotion, and volition form a cohesive whole, an account informed by Wilhelm Dilthey’s philosophical anthropology (Section 3).

In light of this model, I argue that projecting cognitive agency onto AI risks destabilizing the integrity of the human cognitive structure itself. Dominant visions of AI operation often rest on what Antonio Damasio has termed the “Cartesian error,” namely, an exaggerated focus on intellect detached from emotion and embodiment. When cognitive activities are transferred from the human subject to artificial systems, much is “lost in translation”: experiential depth, affective resonance, and contextual sensitivity may be reduced, thereby impoverishing human perception rather than simply extending it.

Second, in the case of cognitive loss caused by alienation, the critical issue is not what we consciously delegate to AI, but what – over time – we will no longer realize we do not know. The growing loss of control over our own ignorance may undermine even our ability to formulate meaningful prompts for AI systems. This problem is examined through the lens of the Socratic conception of wisdom as awareness of one’s own ignorance (Section 4).

Finally, I argue that the religious impulse – particularly the tendency to deify what we do not understand – persists in our relationship with AI (Section 5). The anthropomorphization of algorithmic systems, along with the mystification of their hidden operations, reflects a broader

cultural strategy of taming the unknowable. The “unfathomable judgments of algorithms” function today much like divine mysteries once did in religious cosmologies. In this respect, my argument resonates with the notion of “algorithmic anxiety,” which captures the unease generated by opaque computational systems whose operations shape social reality while remaining largely inaccessible to ordinary users.

1. Philosophical Foundations of Alienation

Ludwig Feuerbach was a 19th-century German philosopher best known for his critique of religion and Hegelian idealism. In his view, religion is a projection of human desires and ideals transferred into the transcendent realm. As the philosopher explains:

Religion, at least the Christian, is the relation of man to himself, or more correctly to his own nature (i.e., his subjective nature); but a relation to it, viewed as a nature apart from his own. The divine being is nothing else than the human being, or, rather the human nature purified, freed from the limits of the individual man, made objective – i.e., contemplated and revered as another, a distinct being. All the attributes of the divine nature are, therefore, attributes of the human nature (Feuerbach 1881, 14).

Feuerbach criticized Hegelian idealism, arguing that thought does not shape reality; rather, material reality conditions consciousness. A central concept in his philosophy is alienation, a notion he borrowed from Hegel and which was later adopted by Marx. According to the author of *The Essence of Christianity*, alienation is a process in which humans separate their own qualities from themselves and attribute them to something external – primarily to God.

Feuerbach argued that all attributes ascribed to God (omnipotence, omniscience, benevolence) are, in reality, idealizations of human qualities, which, through alienation, become rigidified in an imaginatively substantiated idea. In his view, humans unconsciously transfer their most valuable attributes to a divine being, thereby impoverishing themselves.

As a result, they cease to recognize their own potential and begin to see themselves as imperfect beings dependent on Gods (Feuerbach 1881, 18).

This act of projection creates the illusion that such traits originate outside the human mind, when in fact they are its own creations. At the same time – through the reversal of the traditional theological order, or rather the replacement of theology with a psychological critique of religion – it is not God who creates man, but man who creates God. According to Feuerbach, God is an imaginative intentional object – something that humans have abstracted from themselves and subsequently recognized as autonomous.

I believe that the process of alienation described by Feuerbach helps to understand the cultural consequences of artificial intelligence. Before delving further into the cultural consequences of AI, it is helpful to recall an ancient philosophical concern about the externalization of mental faculties – a concern that predates both modern technology and Feuerbach's theory. This concern was most clearly articulated by Plato in his famous critique of writing in the dialogue *Phaedrus*.

In his 4th-century BCE work, the philosopher presents a fictional conversation between the inventor of writing, the Egyptian god Theuth, and the king of Egypt, Thamus. While the former presents writing as a gift to humanity that enhances memory and fosters wisdom, the latter questions its benefits. Reacting skeptically to Theuth's arguments, Thamus argues that writing does not actually strengthen memory but weakens it. For writing becomes an external repository of memory, causing humans to neglect their innate ability to retain information (Plato 2002, 274a-275b).

The Plato's critique of writing reflects a foundational concern about the consequences of transferring mental faculties to external supports. He feared that reliance on written symbols would not foster true understanding but instead encourage forgetfulness, as individuals would remember information by referencing external marks rather than internalizing knowledge.

It is worth emphasizing that the path from writing to artificial intelligence is shorter than we might assume. As the Canadian communication theorist Marshall McLuhan explains, technology is an

extension of the human being; it is something that extends human senses and capabilities (McLuhan and Fiore 2008, 8, 26). As a result, the Greek philosopher's arguments should be considered an early reflection on technological alienation.

2. Feigned Intelligence

By “feigned intelligence” I refer to simulated or superficial intelligence – a situation in which a person may appear knowledgeable or capable, while in fact relying entirely on AI-generated information and lacking deeper understanding. For instance, as one of the recent studies suggests, the use of generative AI may improve task performance while simultaneously distorting users’ metacognitive self-assessment: individuals assisted by AI often report higher confidence in their abilities despite no corresponding increase in genuine understanding or metacognitive accuracy (Fernandes et al. 2026). The following thought experiment is intended to illustrate this distinction.

Let us imagine such a scenario in two variations: a person goes for a walk in the forest. In the first case, they possess no knowledge of ornithology, zoology, or botany – they are unable to recognize bird species, animals in general, plants, or trees. In the second case, the opposite is true – the person walking through the forest is an educated nature enthusiast with a deep interest in these fields of knowledge.

In the first variation, the individual’s experience will be homogeneous. From their perspective, the forest blends into a single entity – resembling a knot of sensations with hidden threads. In the second case, however, the forest becomes a microcosm: the individual recognizes patterns in nature, perceives distinctions among species, and engages deeply with the environment.

According to the 17th-century French scholar René Descartes, human mentality is defined primarily by the intellect. Although Descartes distinguished several cognitive faculties, he treated them as partially separate entities. Moreover, he arranged them hierarchically – while the

senses, in his view, are fallible, the intellect serves as a source of certain and reliable knowledge (Descartes 2008, 18–19).

As the philosopher Hubert Dreyfus aptly points out in his book *Alchemy and Artificial Intelligence*, the contemporary vision of mentality is rooted in Cartesian thought:

The development of the high-speed digital computer has strengthened a conviction which was first expressed by Lucretius, later in the different ways by Descartes and Hume, and finally expressed in the nineteenth-century associationist and stimuli-response psychology: thinking must be analyzable into simple determinate operations. The suitably programmed computer can be viewed as a working model of the mechanism presupposed by this theory. Artificial intelligence has in this way made associationism operational and given it a second wind (Dreyfus 1965, 48–49).

Dreyfus highlights how early AI models were based on the assumption that human cognition could be replicated by formalizing associations between inputs and outputs, following a mechanistic, Cartesian view of the mind. By referencing associationism – a psychological theory that reduces thought to sequences of connected ideas – he critiques the reduction of human intelligence to symbolic processing.

Considering an alternative to Descartes' dualistic model, I turn to Wilhelm Dilthey. As a 19th-century philosopher who opposed both rationalist reductionism and the separation of mind and body, he emphasized that human mentality cannot be reduced to abstract reasoning or isolated intellect. In his anthropological conception of the human being, cognitive processes are embedded in a unified structure of experience, comprising thinking, feeling, and willing as inseparable elements of the psyche (Dilthey 1989, 263–266).

As Dutch researcher Jos de Mul explains: "If we understand a particular lived experience then we do not only register this experience but put it in the context of the totality of psychic life. This implies that we must also call upon our will and upon our feeling" (de Mul 2004, 228). As a result, something like "pure intellect" does not exist, and it is certainly not a defining feature of human nature. As Damasio puts it: "Reason may not

be as pure as most of us think it is or wish it were, that emotions and feelings may not be intruders in the bastion of reason at all: they may be enmeshed in its networks, for worse and for better” (Damasio 1994, xii).

Although other thinkers have explored similar themes – such as Maurice Merleau-Ponty’s notion of embodied perception (Merleau-Ponty 1945) or Michael Polanyi’s theory of “tacit knowledge” (Polanyi 2009) – Dilthey’s emphasis on the coherence and lived unity of mental life offers a particularly fitting framework for understanding the alienation of intelligence in the context of AI. While Merleau-Ponty’s concept of embodiment focuses on perception and physical engagement with the world, Dilthey’s anthropological model provides a broader framework encompassing emotional and volitional aspects of cognition. His focus on inner experience, rather than sensory-motor embodiment alone, makes his approach especially suited to exploring how cognitive capacities become externalized and estranged from the subject.

There is no doubt that the transition from probabilistic models generating text, images, or decisions to slow-reasoning models, the development of synthetic neural networks, and the use of machine learning in multilayered systems represent a true technological revolution that is reshaping our reality (Bengio, LeCun & Hinto 2021). However, as Dilthey’s anthropological model suggests, intellect is not an autonomous linguistic-computational ability that can be abstracted from the entirety of mental experience. Genuine human intelligence is rooted in self-awareness, intertwining with the body, environment, intuition, emotions, interests, and countless individual experiences that shape it.

Thus, if human intelligence is an embodied, emotive whole, any attempt to isolate and outsource it to a machine is bound to be incomplete – yet culturally we forget this, treating the machine’s output as equivalent or even superior to human thinking.

3. Alienation of Intelligence

According to McLuhan, technology is an extension of the human being. The development of computers, the Internet, and cloud services has long

since externalized “intelligence” (Clark, Chalmers 1998). Currently the servers used by AI systems are spread across the globe: in the Arctic, the Dubai deserts, and along the Norwegian fjords. Neural networks are gradually enveloping the entire planet, becoming “an extension of the central nervous system” (McLuhan and Fiore 2008, 40).

Meanwhile, as Yuval Noah Harari suggests, in the era of big data and global information networks, the intelligence of an individual human may lose some of its relative significance, giving way to what he describes as a form of “planetary intelligence” (Harari 2017, 97–105, 262). In a world where knowledge is stored in the cloud and instantly accessible, the unique cognitive abilities of a single person seem less central to progress than they once were.

Crucially, artificial intelligence does not merely grant us access to advanced computations. By being connected to the Internet, it becomes a vast reservoir of knowledge. As a result, it functions as a true McLuhanian prosthesis – one that appears to surpass anything we naturally possess. For example, with *Google Smart Compose*, available in *Gmail* and *Google Docs*, we no longer need to finish sentences – AI suggests completions in an auto-complete mode. In this way, both our knowledge and cognitive processes are being transferred to the cloud.

Importantly – as Öhman suggests – when we interact with the AI technologies, we experience them as personified. It is no longer a diffuse, impersonal technology, but something that, in direct interaction, simulates individual agency (Öhman 2023, 3). Empirical research on human-AI interaction supports this tendency: users often attribute humanlike traits, intentions, and emotional qualities to AI systems, which in turn shapes trust, engagement, and perceived agency in the interaction (Alabed, Javornik, and Gregory-Smith 2022).

By “alienating intelligence” I mean the process by which the gradual outsourcing of intellectual activities to technology may, in the long run, impoverish human cognition. This form of “cognitive outsourcing” refers to the delegation of one’s own cognitive capacities to external technological systems, particularly AI, designed to perform tasks such as remembering, analyzing, evaluating, or deciding on our behalf. While these systems execute specific functions, the effect is that cognitive

operations once carried out internally are progressively transferred to external infrastructures.

This concern is not new. Psychologists have documented the so-called “Google effect,” showing that people are less likely to retain information when they know it can be easily retrieved online (Sparrow et al. 2011). Nicholas Carr similarly argues that digital tools reduce our attention span and memory retention (Carr 2008; Carr 2010). Mark Coeckelbergh even refers to AI as an “alienation machine” that distances us from our embodied condition and core cognitive capacities (Coeckelbergh 2020).

Putting it figuratively: although the mentioned walker in the forest could use a smartphone search engine or ask *ChatGPT* for the name of a particular plant, the key issue is that, in most cases, he will not be able to recognize the questions and prompts they could potentially ask. Without prior knowledge, many tree species will appear, at the level of spontaneous perception, as the same generic tree. What is more, a person familiar with birds and trees is not only capable of identifying them. Their knowledge fundamentally alters their perception of the forest.

Let us take this idea further. Imagine that the walker is wearing *Google Oculus VR* glasses. Although they lack knowledge about the ecosystem, AI continuously “unravels the forest” for them. As they look around, they see outlines and labels identifying trees. Additionally, the glasses recognize the sounds of birdsong and instantly displays information about the animals hidden in the treetops. Would artificial intelligence, in this case, compensate for the protagonist’s previously lost cognitive abilities? In my view, it would not.

“Ornithological knowledge” is not merely an encyclopedic sum of information that can be stored on a server. As Dilthey shows us, it is a blend of facts, emotions, passion, intuition, and many other elements woven into a dense knot. What we are dealing with here is an epistemic coherence – still not fully understood – woven into the mystery of human existence. “Emotion, feeling, and biological regulation all play a role in human reason. The lowly orders of our organism are in the loop of high reason” – Damasio explains (Damasio 1994, xiii).

Despite its apparent usefulness, the external delegation of knowledge – removing it from the mind and storing it in an external system – may pose

a serious threat to the richness of human experience. Knowledge acquired through personal experiences and learning is organically integrated with our cognitive faculties. Viewing the topic through the lens of the Dilthey's insights allows us to see that the weakening of memory, the depletion of knowledge, or the limitation of any cognitive ability can profoundly destabilize the psychological structure of a human being (de Mul 2004, 169).

If we reduce human cognitive abilities to mere computational power, AI will undoubtedly come out on top (Bostrom 2014, 59–60). However, this way of thinking stems from a distorted understanding of human mentality – one that, in my view, is rooted in Cartesian errors as well as the emancipation of computational metaphors within the philosophy of mind (Dreyfus 1965, 33–34, 47) (Putnam 1975, 435–439). As Damasio observes: “I had grown up accustomed to thinking that the mechanisms of reason existed in a separate province of the mind, where emotion should not be allowed to intrude, and when I thought of the brain behind that mind, I envisioned separate neural systems for reason and emotion” (Damasio 1994, xi).

Additionally, what is decisive is that, over time, the individual experiencing the forest in a homogeneous way may cease to recognize that their experience is diminished. In this sense, the problem is not merely the loss of knowledge but the erosion of awareness of ignorance. In the Socratic tradition, wisdom does not consist in the possession of extensive knowledge, but in the lucid recognition of one's own cognitive limits. In Plato's *Apology*, Socrates explains that he considered himself wiser than those reputed to be knowledgeable precisely because he was conscious of his ignorance, whereas they were not (Plato 1949, 21c-d).

What is at stake, therefore, is not only the externalization of knowledge to artificial systems, but a gradual contraction of the horizon within which ignorance can be recognized as such. As cognitive capacities are increasingly outsourced to AI, it is not merely the stock of retained knowledge that may diminish; the very awareness of what is not known may recede. In such a condition, one no longer perceives that one's “forest” could be otherwise differentiated.

Taken together, this suggests that the phenomenon of “alienated intelligence” is not a product of AI technology itself. Rather, it stems from the widespread adoption of computation-focused models of human cognition that disregard the significance of emotions and intuition. If we imagine intelligence as a detachable component – something that can be extracted from the human mind and implanted into a powerful machine that amplifies it – we may fall into the trap of anthropomorphizing or even deifying the technology itself (Öhman 2023, 14).

4. Cultural Projection

Although we speak of “artificial intelligence”, in reality we are dealing with artificial intelligences (Boden 2016). The neural networks developed by different companies – such as *OpenAI*, *Google*, or *DeepSeek* – resemble an archipelago of islands that remain more or less interconnected but, for now, function as distinct tools.

Moreover, artificial intelligence is not a substantial entity but a set of functionalities produced by means of advanced technology. By leveraging algorithms and statistical models, these tools perform language translation, generate images, analyze data, and execute numerous other operations. Each neural network operates within the boundaries of its own architecture, and its “intelligence” is not an independent entity but a capability for carrying out specific tasks based on available information (Bender et al., 2021). In this sense, AI is more of a tool than a subject – an advanced form of work automation whose efficiency depends on data, computational resources, and the objectives for which it was designed.

However, Robert Geraci’s analysis of “Apocalyptic AI” narratives highlights how contemporary discourse often frames AI in quasi-religious or transcendent terms – effectively treating technology as a divine entity or savior (Geraci 2010). Such findings reinforce the idea that AI functions as a cultural projection of human ideals and fears, much like Feuerbach’s description of God as a human mirror. This tendency is further illustrated by the so-called *ELIZA effect*, whereby users attribute understanding,

intentionality, or even consciousness to computational systems whose operations they do not fully comprehend.

Feuerbach points it out when he argues that the God we perceive as a substantial entity is merely a collection of certain predicates: “Subject and predicate are distinguished only as existence and essence. The negation of the predicates is therefore the negation of the subject” (Feuerbach 1881, 19). When we speak of God as a creator who is merciful, omniscient, and just, we assume that these attributes must be somehow rooted. By anchoring them in an imagined foundation, we ideologically substantiate God as an autonomous being – a subject to which these qualities belong. But the substance in question does not really need to exist.

What the German philosopher describes as “God as a set of attributes” is, in our case, artificial intelligence as a set of advanced functionalities. Machine learning mechanisms, language generation, and image analysis capabilities seem so complex that we struggle to perceive them merely as a set of statistical operations. Instead, we tend to project onto them a sense of unity – something that binds them into a hidden, cohesive whole.

As Öhman explains the details of this relation in a longer passage:

[...] When you are chatting with a bot like ChatGPT, you are more or less chatting with significant portions of the web prior to 2021 plus a serious chunk of public domain books. But when it answers, it will not be in a million voices, but one, and the answer will not be the only possible one, but one that the designers of the model favor. This is why the essence of LLMs (but also other forms of AI), is anthropological in the same sense that gods are. Both are ways of personifying the human past, giving it voice, agency, and – more importantly – authority, in the present (Öhman 2023, 8).

As creatures thinking in metaphysical terms, we are tempted to anthropomorphize AI, intuitively assuming that its functions are a manifestation of some form of super-intelligence (Kołakowski 1989, 1–2). Even if AI is not currently conscious, we may speculate that, in the future, it could develop a kind of mentality, acquire goals, and even experience desires. “If we gradually increase the level of integration of a collective

intelligence, it may eventually become a unified intellect – a single large »mind« as opposed to a mere assemblage of loosely interacting smaller human minds” – the philosopher Nick Bostrom points out (Bostrom 2014, 56).

This tendency illustrates what I call the “Myth of Artificial Agency”: the cultural inclination to attribute autonomous agency, intentionality, or even quasi-divine qualities to algorithmic systems whose opaque operations exceed ordinary understanding. In conditions of limited transparency, the echo of older myth-making impulses re-emerges, leading us to personify and reify AI as if it were an independent subject rather than a technologically mediated extension of human design and projection (Kořakowski 1989, 6–8).

As Feuerbach reminds us, since time immemorial, religiosity has been linked to the human need to domesticate mysterious and unknown forces. When we did not understand the laws of nature – lightning, floods, or earthquakes – we assumed they were caused by the whims of the gods (Eliade 1959). Anthropomorphizing the traits of nature and various aspects of reality allowed our ancestors to find explanations for phenomena they did not truly comprehend.

In the absence of clear understanding, we seek refuge in mythologizing what we cannot grasp. In this context, it is important to note that our present environment is no longer the simple natural world but what Jean Baudrillard describes as “hyperreality” (Baudrillard 1994, 1–3, 22).

Although we still exist within a familiar physical world, it is increasingly mediated by the digital dimension. AI technology, which – following McLuhan’s insights – is not merely a medium but an environment of existence, intensifies the sensation of losing ground beneath the feet (McLuhan and Fiore 2008, 157). “[...] Instant communication insures that all factors of the environment and of experience coexist in a state of active interplay” – Canadian thinker claims (McLuhan and Fiore 2008, 63). As we move from statistical models to Hopfield Networks, which function like “black boxes” – where we have input and output data but do not fully understand the processes occurring within the “hidden layers” – we realize that AI technologies evade traditional categories of transparency and control (Burrell 2016). A global survey of over 48 000 respondents

found that while the majority use AI routinely, less than half expressed trust in these technologies, and two-thirds admitted to using them without assessing the accuracy of outputs, highlighting a widespread gap between use and understanding (Gillespie et al. 2025).

While in the past the world was governed by hidden natural forces, today the same is true of the technology-mediated “hyperreality”. Algorithms function like new laws of nature, determining what we see on our screens (Beisbart & Jalas 2022). Although we experience the effects of these algorithmic laws, we do not have full access to them – they remain mysterious, embedded within statistical models that refine and evolve independently in ways we cannot fully trace (Beisbart & Jalas 2022; Bargill et al. 2022).

As journalist Kyle Chayka explains:

The ambiguity of algorithmic influence creates a feeling that has been labeled »algorithmic anxiety«. Algorithmic anxiety describes the burgeoning awareness that we must constantly contend with automated technological processes beyond our understanding and control, whether in our Facebook feeds, Google Maps driving directions, or Amazon product promotions. We are forever anticipating and second-guessing the decisions that algorithms make (Chayka 2024, 38).

As a result, we are once again resorting to mechanisms for taming the unknown. As McLuhan points out, referencing Lao Tzu: “Electric circuitry confers a mythic dimension on our ordinary individual and group actions [...]” (McLuhan and Fiore 2008, 114). “Electric circuitry is Orientalizing the West. The contained, the distinct, the separate – our Western legacy – are being replaced by the flowing, the unified, the fused” (McLuhan and Fiore 2008, 145). The main difference here is that instead of speaking about “electrical circuits”, we should now refer to the digital dimension.

This phenomenon reveals that the human need for narrative and symbolism remains unchanged – despite living in an age of advanced technology, we still project onto it the same cognitive structures that have long helped us make sense of reality (Kołakowski 1989, 1–5). “The nature of LLMs is strikingly similar to the way classical anthropologists interpret the emergence of gods” – Öhman sums up (Öhman 2023, 6).

Conclusion

According to Feuerbach, the creation of God in the process of religious alienation leads to a paradoxical situation. Although we perceive him as an external being to whom we can direct prayers and questions, in reality it is merely a reflection of ourselves (Feuerbach 1881, 27). From the philosopher's perspective, this is an act of human self-renunciation – rather than cultivating within themselves the qualities they attribute to God, humans project them onto the absolute and surrender their agency to it (Öhman 2023, 3). This mechanism leads to a situation in which the believer fails to recognize themselves in God and perceives him as an autonomous force governing the world. It is, therefore, a double illusion: on one hand, God appears to be a real entity; on the other, his nature is merely a reflection of the human psyche (Feuerbach 1881, 12–13).

With artificial intelligence, we encounter a remarkably similar phenomenon. It appears to be an external entity, an autonomous being to which we can pose questions, expecting answers as if it were an intelligence separate from us. However, in reality, AI is merely a reflection of ourselves – an alienated form of human knowledge and intelligence that, despite seeming to function independently, tells us nothing that we have not already fed into it.

Language models do not emerge in a vacuum – trained on human knowledge, they draw from it, assimilate its structures and relationships. As a result, AI “feeds” on data produced by humanity, processes it, and presents it in a new form (Öhman 2023, 7). The analysis presented in this article suggests that contemporary AI can be understood as a projection of human cognitive abilities – echoing Feuerbach's theory of religious alienation. Just as humans once projected their own essence onto God, we now externalize intellectual capacities into machines and, in doing so, may gradually come to perceive them as superior or independent entities.

This phenomenon, I have argued, can lead to a subtle impoverishment of our own cognitive capacities. Projection, as Feuerbach showed, can serve to elevate ideals – just as the image of God once provided humans

with a moral compass and shared meaning. Similarly, AI, as a technological projection of our intelligence, can enhance creativity, extend problem-solving capabilities, and increase efficiency in many domains.

Although the gains are real, they come with a risk that we begin to forget the origin of these capabilities, mistaking the projection for the source. In Socratic terms, wisdom consists in the awareness of one's own ignorance; yet the outsourcing of cognitive activities to AI may gradually erode not only knowledge itself, but also our reflexive awareness of what we do not know. Moreover, in the context of work management, as managerial and organizational processes increasingly depend on algorithmic reasoning, human expertise risks becoming externalized and fragmented – reflecting Feuerbach's notion of self-alienation and potentially fostering a condition in which the limits of our own understanding become less perceptible to us.

Although it might seem that there is little difference between intelligence residing “in the head” and intelligence externalized in technology, this distinction is, in fact, deeply consequential. This is because intelligence – often mistakenly defined in narrow, computational terms – is actively involved in how we perceive and engage with the world (Varela, Thompson & Rosch 1991). As a result, by transferring our cognitive functions to technological systems, we may be – perhaps irreversibly – diminishing ourselves and disrupting what Dilthey described as the intricate psychological unity of human experience. At the same time, the surrounding Baudrillardian “hyperreality” appears increasingly incomprehensible, prompting us to tame it through traditional forms of religious thinking.

References

- Alabed, Amani, Ana Javornik, and Diana Gregory-Smith. 2022. “AI Anthropomorphism and Its Effect on Users' Self-Congruence and Self-AI Integration: A Theoretical Framework and Research Agenda.” *Technological Forecasting and Social Change* 182: 121786. <https://doi.org/10.1016/j.techfore.2022.121786>.

- Bar-Gill, Oren, Cass R. Sunstein, and Inbal Talgam-Cohen. 2023. "Algorithmic Harm in Consumer Markets." *Journal of Legal Analysis* 15 (1): 1–47. Available at: <https://doi.org/10.1093/jla/laad003>.
- Baudrillard, Jean. 1994. *Simulacra and Simulation*. Translated by Sheila Faria Glaser. Ann Arbor: The University of Michigan Press.
- Beisbart, Claus, and Markus Jälas. 2022. "Moral Transparency of and Concerning Algorithmic Tools." In *The Philosophy and Ethics of Information*, edited by Luciano Floridi, 313–34. Dordrecht: Springer.
- Bengio, Yoshua, Yann LeCun, and Geoffrey Hinton. 2021. "Deep Learning for AI." *Communications of the ACM* 64 (8): 58–65. Available at: <https://www.cs.toronto.edu/~hinton/absps/ACM21.pdf>.
- Bender, Emily M., Timnit Gebru, Angelina McMillan-Major, and Shmargaret Shmitchell. 2021. "On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?" In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (FAccT'21)*, 610–23. New York: ACM Press.
- Boden, Margaret A. 2016. *AI: Its Nature and Future*. Oxford: Oxford University Press.
- Bostrom, Nick. 2014. *Superintelligence: Paths, Dangers, Strategies*. Oxford: Oxford University Press.
- Burrell, Jenna. 2016. "How the Machine 'Thinks': Understanding Opacity in Machine Learning Algorithms." *Big Data & Society* 3 (1): 1–12. Available at: <https://journals.sagepub.com/doi/full/10.1177/2053951715622512>.
- Carr, Nicholas. 2008. "Is Google Making Us Stupid?" *The Atlantic*, July/August. Available at: <https://www.theatlantic.com/magazine/archive/2008/07/is-google-making-us-stupid/306868/>.
- Carr, Nicholas. 2010. *The Shallows: What the Internet Is Doing to Our Brains*. New York and London: W. W. Norton & Company.
- Chayka, Kyle. 2024. *Filterworld: How Algorithms Flattened Culture*. New York: Doubleday.
- Clark, Andy, and David J. Chalmers. 1998. "The Extended Mind." *Analysis* 58 (1): 7–19. Available at: <https://www.jstor.org/stable/3328150>.
- Coeckelbergh, Mark. 2020. *AI Ethics*. Cambridge, MA and London: MIT Press.
- Damasio, Antonio R. 1994. *Descartes' Error: Emotion, Reason, and the Human Brain*. New York: Avon Books.
- Descartes, René. 2008. *Meditations on First Philosophy: With Selections from the Objections and Replies*. Translated by Michael Moriarty. Oxford World's Classics. New York: Oxford University Press.

- Dilthey, Wilhelm. 1989. *Introduction to the Human Sciences. Selected Works*, Volume I, edited by Rudolf A. Makkreel and Frithjof Rodi. Oxford: Princeton University Press.
- Dreyfus, Hubert L. 1965. *Alchemy and Artificial Intelligence*. RAND Corporation Paper P-3244. Santa Monica: RAND Corporation. Available at: <https://www.rand.org/pubs/papers/P3244.html>.
- Eliade, Mircea. 1959. *The Sacred and the Profane: The Nature of Religion*. Translated by Willard R. Trask. New York: Harcourt, Brace & World.
- Fernandes, Daniela, Steeven Villa, Salla Nicholls, Otso Haavisto, Daniel Buschek, Albrecht Schmidt, Thomas Kosch, Chenxinran Shen, and Robin Welsch. 2026. "AI Makes You Smarter but None the Wiser: The Disconnect between Performance and Metacognition." *Computers in Human Behavior* 175: 108779. <https://doi.org/10.1016/j.chb.2025.108779>.
- Feuerbach, Ludwig. 1881. *The Essence of Christianity*. Translated by Marian Evans. The English and Foreign Philosophical Library, Vol. XI. London: Trübner & Co.
- Geraci, Robert M. 2010. *Apocalyptic AI: Visions of Heaven in Robotics, Artificial Intelligence, and Virtual Reality*. Oxford: Oxford University Press.
- Gillespie, Nicole, Steve Lockey, Timothy Ward, Alexandria Macdade, and Gerald Hassed. 2025. *Trust, attitudes and use of artificial intelligence: A global study 2025*. The University of Melbourne & KPMG. DOI:10.26188/28822919.
- Harari, Yuval Noah. 2017. *Homo Deus: A Brief History of Tomorrow*. New York: HarperCollins Publishers.
- Kołakowski, Leszek. 1989. *The Presence of Myth*. Translated by Adam Czerniawski. Chicago: University of Chicago Press.
- McLuhan, Marshall, and Quentin Fiore. 2008. *The Medium Is the Massage. Coordinated by Jerome Agel*. London: Penguin Books.
- Merleau-Ponty, Maurice. 2012. *Phenomenology of Perception*. Translated by Donald A. Landes. London and New York: Routledge.
- Mul, Jos de. 2004. *The Tragedy of Finitude: Dilthey's Hermeneutics of Life*. Translated by Tony Burrett. New Haven and London: Yale University Press.
- Nilsson, Nils J. 2010. *The Quest for Artificial Intelligence: A History of Ideas and Achievements*. Cambridge: Cambridge University Press. Web version.
- Öhman, Carl. 2023. "We Are Building Gods: AI as the Anthropomorphised Authority of the Past." *Minds and Machines* 34 (1): 1–18. Dordrecht: Springer. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4620986.
- Plato. 2002. *Phaedrus*. Translated by Robin Waterfield. Oxford World's Classics. New York: Oxford University Press.

- Plato. 1949. *The Collected Dialogues of Plato: Including the Letters*. Edited by Edith Hamilton and Huntington Cairns. New York: The Macmillan Company.
- Polanyi, Michael. 2009. *The Tacit Dimension*. London: The University of Chicago Press.
- Sparrow, Betsy, Jenny Liu, and Daniel M. Wegner. 2011. "Google Effects on Memory: Cognitive Consequences of Having Information at Our Fingertips." *Science* 333 (6043): 776–78.
- Varela, Francisco J., Evan Thompson, and Eleanor Rosch. 1991. *The Embodied Mind: Cognitive Science and Human Experience*. Cambridge: MIT Press.