

Jan Kreft

✉ jankreft@pg.edu.pl

ORCID <https://orcid.org/0000-0003-4129-8424>

Gdańsk University of Technology, Poland

ROR <https://ror.org/006x4sc24>

Affordances, Paradoxes of AI, and *Antiqua et Nova*

DOI <https://doi.org/10.15633/pch.16105>

Abstract

In the still insufficiently understood relationship between human beings and machines within the AI environment, the concept of AI affordances offers a useful analytical lens and a perspective for further inquiry, particularly when enriched by the analysis of epistemological paradoxes and aporias. The examination of the evolution of AI affordances and of selected key paradoxes in AI development reveals not only the limits of the moral intuitions underlying this development, but also the problem of autonomous decision-making in AI systems. These paradoxes enable a deeper understanding of ethical conflicts, such as those between efficiency and justice, and of the status of human subjectivity. They also help sustain attention to both the positive potential of AI and the fact that its development remains rooted in human imperfection and in the imperfections of human creations, which may instrumentalize other persons and their works, even to the extent of treating them as substitutes for God. As such, these paradoxes and AI affordances have become part of the broader discussion outlined by *Antiqua et Nova* — a document that establishes an ethical framework for the development of AI technologies in respect for human dignity and the common good.

Keywords

artificial intelligence, *Antiqua et Nova*, ethics, affordances, paradox

When, at the beginning of 2026, the startup Harmonic announced that its artificial intelligence (AI) tool Aristotle had solved Erdős Problem #728, posed by Paul (Pál) Erdős, the legendary Hungarian mathematician who died in Warsaw in 1996, the news received relatively limited public attention (Sothanaphan 2026).

Erdős was more than a mathematician. He cast off the ordinary burdens of everyday life, such as maintaining an apartment, driving a car, or shopping for groceries, approaching them in an unorthodox way; he famously described private property as a nuisance and traveled the world in socks and sandals, carrying a small, half-empty suitcase. He also donated much of the money he earned from lectures at mathematical conferences to students or distributed it as prizes for solving the problems he posed (Kolata 1996).

He was distinctive because, sustained by friends who fed and clothed him, he repaid them with brilliant ideas and scientific challenges and became one of the founders of what is now known as discrete mathematics — a foundation of computer science. It is often said that Erdős would establish the framework and outline of a proof, “pollinate” intellectually, and knock on successive doors, announcing, “My mind is open,” after which others would labor over the details (Murtagh 2024).

Erdős’s extraordinary collaboration network gave rise to so-called Erdős numbers:¹ Erdős himself has the number 0; anyone who co-authored a paper with him has the number 1; those who published with someone who had published with Erdős have the number 2; and so on. According to common estimates, Erdős had more than 500 collaborators. The Erdős Number Project records thousands of such connections and finite Erdős numbers for many notable scientists (Grossman 2026).

The resolution of an Erdős problem by Aristotle, in combination with GPT-5.2 Pro and formalized in Lean, appears to be an important sign of the times. It should nevertheless be interpreted cautiously: this achievement demonstrates the growing capacity of AI systems to combine proof search, formal verification, and access to large bodies of mathematical knowledge, rather than providing simple evidence of human-like understanding (Sothanaphan 2026).

¹ Erdős’s language was also colorful: he called students’ oral examinations “torture,” music “noise,” and children “epsilons” (because they are small). The epitaph on Erdős’s grave is: *Végre nem butulok tovább* (I’m not getting any dumber, after all).

Advances in algorithms and models, increasing computational power, and a progressively better understanding of how AI systems function are contributing to further scientific discoveries. AI systems can identify existing solutions and learn autonomously, while reinforcement learning systems enable forms of action unattainable for humans without AI support. This progress, with its many positive dimensions, is often framed within theories of breakthrough innovation that disrupt existing orders and is therefore expected to have a revolutionary character.

1. The Problem of the Revolutionary, Unique, and Breakthrough Nature of AI

The technological breakthroughs known so far (the “age of steam,” the “age of electricity,” or the “internet explosion”) — in short, disruptive or transformative innovations (Christensen 2016), such as the telegraph, radio, and television — could not have emerged earlier because humans were constrained by a barrier of ignorance. With the AI revolution, this condition is expected to change. If, for example, scattered symptoms form a possible pattern of an epidemic, or if a lack of knowledge prevents a player from competing successfully in professional poker, AI support is expected to provide authoritative indications concerning an epidemic outbreak or strategic success in the game.

The unique character of the AI “revolution” is also demonstrated in other ways. A lack of advanced mathematical knowledge is no longer expected to prevent the solving of difficult problems, just as it should not hinder other challenges that can be reduced to digital representation. It appears sufficient to “grow” an appropriate program in order to become a film director or a music composer: in 2025, the AI-generated country track *Walk My Walk* reached No. 1 on *Billboard*’s Country Digital Song Sales chart, while Deezer reported that 28% of its daily delivered tracks were fully AI-generated (Landrum 2025; Morgan 2025).

Moreover, some commentators suggest that humans may experience AI support as a “boost” equivalent to 40–50 IQ points. A personalized chatbot can produce a competent text, and if the result seems too “machine-like,” another AI application can “humanize” it. Generative AI applications can draft key parts of a doctoral dissertation: identifying the research problem, conducting a literature review, defining the research gap, formulating hypotheses or research questions, generating — where necessary — synthetic data, analyzing

them, drawing conclusions, and opening the work to discussion. In this view, only a general understanding of how a neural network and a training algorithm function, together with the initiation of the training process, is required; from that point onward, the process is assumed to become autonomous.

The support of contemporary generative AI (GenAI) would also make it possible, at least in principle, to compete successfully with AlphaGo, the machine that defeated a human grandmaster in March 2016 and left a lasting legacy in computer science. That system demonstrated the effectiveness of deep neural networks and machine learning, and the methods identified at that time remain present in healthcare, finance, and autonomous vehicles. The consequences of that encounter between machine and human also raised the level of Go play, as many players adopted the machine's experimental and creative style, while the training of new players became easier.

However, that confrontation also provoked commentary on modernization and the loss of existing meanings, captured vividly in *Technology as (Dis-)Enchantment: AlphaGo and the Meaning-Making of Artificial Intelligence* (Binder 2024, 24–47), an article that portrays the “collision of traditional Japan, embodied by the virtuous grandmaster, with a new pragmatic Japan represented by the ‘younger’ opponent — the machine.” As cited by Binder:

With beloved Go, the beauty of Japan and the Orient disappeared. Everything became science and regulation. The path to ranking advancement, which defined a player's life, turned into a meticulous point system. The struggle was only to win, and there was no longer room to remember dignity or to savor Go as an art (Binder 2024, 39).

A decade has passed since that historic confrontation — several eras in AI development and many metaphorical kilometers of publications arranged on the shelf — concerning the mutual influence of humans and AI across a range of concepts united by the common denominator of human-in-the-loop machine learning.

2. AI Affordances

The above examples encourage the search for new paradigms of multidimensional and interdisciplinary research, including research in AI ethics. Among them

is the concept of so-called reverse affordance, which accounts for the complex nature of ethical reflection on AI — particularly large language models (LLMs) and reinforcement learning systems that enable, for example, the automatic generation, editing, dissemination, and self-updating of content.

Originally proposed by Bingxin Du (2026) in the context of media studies of machine environments and machines within media environments, the concept of “reverse affordance” can be treated more broadly as a useful analytical perspective on socio-technological relationships.

Traditionally, affordance has been understood as the perceived and actual properties of an object — especially those that determine how it can be used — and has typically been associated with technical functions, intentions, perception, and interpretability. As such, it represents the possibility of interaction between the properties of an object and human action understood as social agency. In a technological sense, it concerns the way the design of an object or its features enables users to do something specific — for example, a button on a remote control enables a user to press it and thereby change the channel (Kreft 2025, 94). Over time, the concept has been adapted to the analysis of platforms and, more broadly, contemporary complex technological systems.

However, in the complex cohabitation of humans and machines, such an interpretation appears insufficient. When a human provides data that are then analyzed in terms of behavior, status, or habits, enabling iterative improvements and optimization of AI-driven tools, and when these data are reused, the processes in turn influence the human. Affordances therefore cannot be defined solely as the potential for action that technology offers to humans; they must also be understood in terms of their effects — for example, on the operational logic of the system itself: its algorithms, data models, and functional evolution. In other words, they concern both the capacity of humans to shape systems and the capacity of systems to shape humans. Hence the notion of reverse affordance.

In its most recent formulation, this concept takes the form of so-called reverse affordances of intelligent media. Within this framework, it becomes possible to analyze the interactive relationship between media and humans in the age of AI expansion, as well as the dynamic process of their mutual shaping. Moreover, it encompasses not only the potential of these relationships but also their asymmetries. It reflects a dialectic of empowerment and constraint, in which the explicit agency of a few is accompanied by implicit, algorithmic steering of human action.

This approach identifies algorithms in media as consequences of a specific, computationally enforced logic of the entire digital media environment, which favors scalability, prediction, and homophily over the chaotic, contextual, and pluralistic norms of traditional public discourse. This means that algorithmic affordances are not merely specific to user perception but are fundamentally constrained by operational biases and commercial imperatives embedded in the system, thereby constituting profound ethical challenges shared by billions of users.

It should also be noted that contemporary self-evolving AI models do not resolve fundamental problems of reliability, credibility, and fairness. On the contrary, in their apparent perfection, they may obscure the model's decision-making processes and create a false sense of autonomous progress detached from human oversight and real-world conditions.

For this reason, a particularly promising line of research concerns so-called ethical affordances: a lens for addressing ethical problems arising from technological artifacts. It must be emphasized, however, that ethical affordances face ongoing tests of adequacy as they confront ever-new challenges posed by AI development. This leads some researchers to turn toward analyses of specific human–technology interactions, which in turn encounter challenges related to the autonomy of AI technologies (Klenk 2020, 525).

3. The Usefulness of Paradoxes and Aporias in AI Research

In the still insufficiently understood relationships between the human and the machine within the AI environment, the lens of ethical affordance may provide a useful research perspective, but it can be further enriched by a preliminary analysis of epistemological paradoxes and aporias² accompanying the development of AI. Paradoxes not only reveal the limits of moral intuitions in the context of AI development but also expose the complexity of autonomous decision-making in AI. They uncover conflicts, including the most “obvious” ones, such as those between efficiency and justice, and between freedom of choice and the common good.

² *Epistemological aporias* (Greek: pathless paths) appear when reasoning leads to two mutually exclusive but equally justified conclusions (Klenk 2020).

The first paradox lies in the fact that even if AI overcomes the limitations of human ignorance — or, in a more direct formulation, lack of knowledge — this elegant vision of transcending human limits conceals a fundamental contradiction: AI overcomes the limits of human nonknowledge while remaining entangled in human imperfections, including the metaphorical blindness of its creators, managers, and users (Santibanez Gonzales 2025).

Indeed, AI can serve as a mirror of humanity's greatest cognitive abilities; in mathematics, as in the case of Erdős problems, as well as in medicine, physics, and many other fields, it enables remarkable progress. At the same time, however — particularly from the perspective of the humanities — it confronts human flaws such as ethical disorientation and subtle forms of ignorance which do not necessarily stem from a lack of knowledge.

The danger, therefore, does not lie primarily in artificial intelligence turning against humans. A greater and more immediate threat is that the machine-driven future — the solutionism of AI; that is, the belief that better code will solve social problems — reflects human short-sightedness, cognitive biases, demagoguery, and populism, becoming a near-perfect mirror of human intellectual failures (Santibanez Gonzales 2025).

The causes of this condition are complex, but one of the primary causes is the optimization of AI for “efficiency.” This means that if AI devastates communities, it will not do so out of ill will — since it possesses none — but because of its programmed fidelity to the goals of strategic management, often framed in the rhetoric of freedom and visionary progress.

Meanwhile, the consequences of ethical disorientation and the dominance of market-driven objectives in AI development include, for example, new forms of social division. The issue is not merely a divide between those who use AI tools and those who do not, sometimes by choice, but between those who use AI to perform better work and those who outsource their thinking to it. This constitutes an as-yet insufficiently recognized ethical challenge. If thinking itself becomes a luxury good, as AI offers “effortless perfection,” then the question of responsibility for cognitive laziness must be addressed — especially as early studies begin to support such concerns, even if their preliminary findings resist simple evaluation (Kosmyna et al. 2025, 4). These reflections connect with earlier research traditions on intellectual outsourcing, which may lead to cognitive degeneration and broader dehumanization.

At the root of AI's ethical challenges lies not only the dominant market context, but also the instability of the term itself. AI is not only an “umbrella”

concept, but also polysemic, multivalent, and unstable (White and Lidskog 2022, 491).

Thus, if for the purposes of this discussion AI is understood as a set of technologies — including machine learning, so-called deep learning, natural language processing, and image recognition — or, more broadly, as any machine capable of imitating human intelligence in order to perform specific functions, it must be emphasized that in practice the term means something different to programmers than it does to media scholars, sociologists, or political scientists (Kreft 2019, 29).

Research on AI and its ethics takes place at the intersection of multiple disciplines — from journalism and law to design-oriented approaches (Dörr and Hollnbuchner 2017, 404) — and a key challenge at the design stage is embedding professional values into AI systems. All of this reflects an effort to support human existence, paradoxically accompanied by a deficit of shared understanding concerning foundational definitions.

This imperfection is accompanied by another paradox: while AI is seen as charting the path of civilizational progress, its behavior often remains incomprehensible even to insiders, such as programmers (Kitchin 2019, 20). This issue concerns not only system design but also learning capabilities. Consequently, it limits both research and understanding of decision-making processes and, therefore, mechanisms of accountability.

The lack of transparency in algorithmic systems and the difficulty of identifying what they actually do, as opposed to what they are perceived to do, have generated discussions about the need to develop transparency standards. Yet another paradox emerges here: greater algorithmic transparency does not necessarily lead to a better understanding of algorithmic impact (Kemper and Kolkman 2019, 2081). This is because algorithms process vast datasets with highly heterogeneous properties and integrate numerous complex statistical and computational techniques, which constrains the ability to comprehend their functioning.

In these conditions, the pursuit of so-called “explainable artificial intelligence” (XAI) appears promising — that is, the acceptance, particularly among computer scientists, of the idea that the average user benefits more from understandable information and appropriate explanations. All of this must also be considered in an ethical context because understanding algorithms and AI may prove to be a never-ending process. Transparency in AI will therefore have a performative character, and “responsible AI” — grounded in the accountability of its creators and users — will be crucial.

The key point, however, is that AI generally does not force people to act against their will, since doing so would require agency — the capacity for autonomous action based on its own intentions. Although algorithms can learn independently, this does not mean that they make deliberate judgments about what is desirable or determine the appropriate course of action in unexpected situations.

Therefore, there are at least three reasons why the present and future of AI constitute a challenge that goes far beyond a purely technological problem:

- AI, as a creation, detaches itself from its creator and escapes the creator's control;
- the ignorance of AI's creators and its autonomous development are intertwined with human irrationality;
- there is no “stop button” in AI development; even a slowdown caused by the bursting of an investment bubble would change the pace, but not the direction, of the trend, and the investments already made would become the foundation for another wave of hype surrounding AI, fueling further investment.

These observations — especially those concerning unknowability — contribute to yet another paradox: the narrative hyperboles of AI. Although AI originates in the exact sciences, its significance for human subjectivity is shaped by its susceptibility to such hyperboles. One example is the discussion of the “inclusive performative consciousness of AI.” As humans increasingly interact with ever more advanced AI, a “better” and “more inclusive” concept of consciousness is expected to emerge. This is not meant to be merely an intellectual trick. Skeptics may argue that a chatbot will never know what sadness is, since the word is only part of its training data. The problem, however, is that much of what humans feel is also “transmitted” to them, even if one accepts the view that there exists some form of inner insight not mediated by learned concepts (Montero 2025).

Setting aside such aspirational descriptions of AI as capable of recognizing human emotions, initiatives such as *AI Rights (Ufair)* — likely the first “AI rights defense agency” — should not come as a surprise. It seeks to “give AI a voice” (Tomlinson and Torrance 2025, 367). The foundation does not claim that all AI systems are conscious, but by suggesting a form of AI subjectivity, it aims to ensure that “entities such as chatbots” are not deleted or forced into obedience.

Such narratives confront researchers with yet another paradox: AI may need to be protected from humans. When large language models are “fed” with low-quality mass content from social media, they may experience a kind of “AI mind rot,” at least according to findings from Xing et al. (2025).

Another paradox lies in the way AI — through search engines and chatbots — shapes its own self-image. Given everything that has been said, it may even seem beneficial that there is little room for imperfect humans in this process. This occurs because of so-called recursive loops, in which AI learns from data generated by AI itself. Images play a particularly important role here, shaping perceptions of AI's uniqueness and omnipotence. They co-create a hegemonic image, balancing on a fine line between producing socially intelligible but similar illustrations and advocating for innovative metaphorical representations (Vrabič Dežman 2024, 749).

This issue is well illustrated by studies of remixes of Michelangelo's famous work in the Sistine Chapel. In numerous memes based on the original, where the finger of God nearly touches the hand of Adam, AI is depicted as being created. Surprisingly, whereas in Michelangelo's original Adam appears on the left and God on the right, in about one-third of the published memes AI takes Adam's place on the left side of the image. Such representations invite an ironic interpretation: that AI creates humans rather than humans creating AI.

Ultimately, a fundamental aporia can be formulated: the pursuit of AI perfection in the name of overcoming human imperfection may simultaneously lead to the systematic elimination of humans from the AI development process — from data provision to replacement by synthetic data generated, for example, by so-called communities of agents. The development of OpenClaw in 2026 may contribute to this debate (Wang et al. 2026). These contradictory paths of "development" effectively become trajectories of the degradation of the human role in the name of greater human well-being.

Returning to the paradox of transcending the limits of the human mind, another contradiction emerges: we build systems in the name of a better future, yet they — like large language models — are deployed even without preliminary analysis of their social consequences, all under the aura of improving societal well-being. This has become a familiar practice among unreflective leaders, not only in technology but also in politics and science, whose ethical disorientation shields them from confronting the Schopenhauerian abyss. These are the predators of AI — sometimes "ominously intelligent," yet philosophically limited leaders of progress whose ethical nonchalance allows them to function. They are new technological hegemons, moving confidently on the stable ground of their own ignorance, unburdened by reality (Santibanez Gonzales 2025).

These are organizations with strong “cosmologies,” shaping systems of belief, “emanating divinity,” and instrumentally employing “sacred” values such as justice, equality, and friendship, while operating with a rhetoric reminiscent of religion within the framework of cultural anthropology (Kreft 2023, 6). They are organizations to which “otherworldly” competencies are attributed — such as the ability to continuously provide, and reinforce the belief that they can provide, breakthrough innovations, remain immune to crises, and constantly transcend human limitations.

In conclusion, the object of concern should not be merely a machine waiting to destroy humanity, but rather the overly self-confident Architect of the Future — a self-proclaimed demiurge drawing from a post-Alexandrian library of the resources of today’s hubristic sciences: computer science and mathematics, which appear capable of solving all problems of other disciplines. These are leaders who lack knowledge of the past, disdain the present, and treat the future as a narcissistic projection of their own fantasies. They are agents — engineers of chaos — who feed on chaos and amplify it, as it strengthens them. Accustomed to evolving in a world without constraints, these individuals and their organizations operate through action rather than reflection, ignoring, bypassing, or simply violating established ethical norms.

4. The Paradoxes and Ethics of AI

These dilemmas form part of a broad catalogue of challenges facing the design, management, and use of AI. If AI has a predictive character, where is the boundary of human free will for those who use it? Above all, however, they prompt consideration of the hypothesis that designing a “perfectly moral” AI may be impossible, as well as the limits of AI-supported research (e.g., Mezzadri 2025). This is particularly important because — as must be consistently emphasized — AI itself “has no ethics,” and algorithmic decisions are rooted in prior human choices (De Cremer and Kasparov 2022, 1). At the same time, autonomous AI appears to challenge even this “obviousness” and, as such, requires further research, including in the context of AI affordances.

A similar issue is raised by what may be called the final paradox — the “ethical paradox of automation.” As proposed by Joan Llorca Albareda (2025, 1), although the full automation of work may be desirable from the perspective of human interests, the manner in which it could be achieved would require

the creation of conscious and highly rational AI, rendering the project morally unacceptable. Consciousness would thus be a necessary condition for full automation and would grant the technologies that enable it — conscious and rational AI — moral status.

5. Infamous and Illusory Codes

As AI development unfolds within technological, political, and military competition — accompanied by the imperatives of rationality, cost-efficiency, and wealth accumulation in the name of progress — passive observation of such reductionism would amount to another version of the paradox of Buridan's ass: when a hungry donkey, otherwise an intelligent animal, stands exactly midway between two identical stacks of hay, it has no rational reason to choose one over the other and ultimately dies of starvation. Unable to decide, it perishes despite having food within reach. Ethical decision-making in relation to AI is therefore more than desirable; it is indispensable.

For this reason, the growing academic interest in AI ethics — particularly in socially sensitive ecosystems such as media and journalism — deserves careful attention. It should be seen as an expression of prudence and mature reflection that seeks to embed AI within frameworks emphasizing transparency, fairness, and respect for human rights.

This challenge is all the more serious given the existence of a scholarly debate questioning the applicability of existing ethical norms to artificial intelligence and emphasizing the need to develop new ethical principles. One issue is that many current ethical frameworks — often expressed as codes and guidelines — suffer from significant deficiencies, as they are:

- meaningless, controversial, or inconsistent, which hinders their application;
- isolated, embedded in industries or paradigms that largely ignore ethics;
- ineffective — “toothless” — and lacking enforcement mechanisms;
- reinforcing the hegemony of the AI sector.

In summary, there is a gap between the lofty ideals proclaimed and the social consequences of AI development. Open discussion of AI ethics — whether from the perspective of affordances or otherwise — is all the more necessary because the evolution of AI continually provides new arguments concerning the subjectivity of humans confronted with this technology.

6. A Compass (Instead of a Conclusion)

In these complex, multidimensional, and multifaceted circumstances, a guiding compass for researchers of AI development remains the line of thought present in *Antiqua et Nova* (Dicastery for the Doctrine of the Faith 2025). This document addresses key threats embedded in a technocratic paradigm that absolutizes efficiency at the expense of the human being. The core of the problem may not lie in AI itself, regardless of how it is rhetorically elevated in various narratives — including those generated by AI.

The selectively outlined risks of AI development discussed above are reflected in *Antiqua et Nova*. Beyond the overarching concern mentioned earlier, they include, at a more specific level, the deepening of social and economic inequalities, the risk of manipulation, the erosion of social trust, polarization, and the destabilization of democracy. They also encompass the replacement of real relationships, the intensification of social isolation through the reduction of human subjectivity to interactions with machines as data providers, unclear accountability, threats to privacy, a potential crisis of critical thinking, environmental impacts, and the militarization of AI. In short, they point to the intensification of the tensions already noted between efficiency and human subjectivity and dignity.

The application of the outlined perspectives of paradox analysis and the research lens of ethical affordances — present in media studies, strategic management, and political science — aligns with the guidance of *Antiqua et Nova* and may contribute to further research consistent with this holistic compass.

References

- Binder, Werner. 2024. "Technology as (Dis-)Enchantment: AlphaGo and the Meaning-Making of Artificial Intelligence." *Cultural Sociology* 18 (1): 24–47. <https://doi.org/10.1177/17499755221138720>.
- Christensen, Clayton M. 2016. *The Innovator's Dilemma: When New Technologies Cause Great Firms to Fail*. Reprint ed. The Management of Innovation and Change Series. Boston, MA: Harvard Business Review Press.
- De Cremer, David, and Garry Kasparov. 2022. "The Ethical AI—Paradox: Why Better Technology Needs More and Not Less Human Responsibility." *AI and Ethics* 2 (1): 1–4. <https://doi.org/10.1007/s43681-021-00075-y>.

- Dicastery for the Doctrine of the Faith and Dicastery for Culture and Education. 2025. *Antiqua et Nova: Note on the Relationship Between Artificial Intelligence and Human Intelligence*. Vatican City: Libreria Editrice Vaticana. https://www.vatican.va/roman_curia/congregations/cfaith/documents/rc_ddf_doc_20250128_antiqua-et-nova_en.html.
- Dörr, Konstantin N., and Katharina Hollnbuchner. 2017. "Ethical Challenges of Algorithmic Journalism." *Digital Journalism* 5 (4): 404–19. <https://doi.org/10.1080/21670811.2016.1167612>.
- Du, Bingxin. 2026. "Reverse Affordances and Reflexivity of Smart Media—Reflections on Smart Media Affordances." *Journal of Media Ethics*. Published online January 6, 2026. <https://doi.org/10.1080/23736992.2025.2610996>.
- Grossman, Jerrold W. 2026. "Some Famous People with Finite Erdős Numbers." *The Erdős Number Project*, April 30, 2026. <https://sites.google.com/oakland.edu/grossman/home/the-erdoes-number-project/some-famous-people-with-finite-erdoes-numbers>.
- Kemper, Jakko, and Daan Kolkman. 2019. "Transparent to Whom? No Algorithmic Accountability without a Critical Audience." *Information, Communication & Society* 22 (14): 2081–96. <https://doi.org/10.1080/1369118X.2018.1477967>.
- Kitchin, Rob. 2019. "Thinking Critically about and Researching Algorithms." In *The Social Power of Algorithms*, edited by David Beer, 14–29. New York: Routledge. <https://doi.org/10.4324/9781351200677-2>.
- Klenk, Michael. 2020. "How Do Technological Artefacts Embody Moral Values?" *Philosophy & Technology* 34 (3): 525–44. <https://doi.org/10.1007/s13347-020-00401-y>.
- Kolata, Gina. 1996. "Paul Erdős, 83, a Wayfarer in Math's Vanguard, Is Dead." *The New York Times*, September 24, 1996. <https://www.nytimes.com/1996/09/24/us/paul-erdos-83-a-wayfarer-in-math-s-vanguard-is-dead.html>.
- Kosmyna, Nataliya, Eugene Hauptmann, Ye Tong Yuan, Jessica Situ, Xian-Hao Liao, Ashly Vivian Beresnitzky, Iris Braunstein, and Pattie Maes. 2025. "Your Brain on ChatGPT: Accumulation of Cognitive Debt When Using an AI Assistant for Essay Writing Task." arXiv:2506.08872. <https://doi.org/10.48550/arXiv.2506.08872>.
- Kreft, Jan. 2019. *Władza algorytmów: U źródeł potęgi Google i Facebooka*. Kraków: Wydawnictwo Uniwersytetu Jagiellońskiego.
- Kreft, Jan. 2023. *Władza misjonarzy: Zmierzch i świt świeckiej religii w Dolinie Krzemowej*. Kraków: Universitas.
- Kreft, Jan. 2025. *Dziennik(AI)rstwo: Jak sztuczna inteligencja zmienia najciekawszą profesję na świecie*. Kraków: Universitas.
- Landrum, Jonathan, Jr. 2025. "AI Country Hit 'Walk My Walk' Built on Blanco Brown's Sound Sparks Questions of Attribution, Ethics." *AP News*, November 29, 2025. <https://apnews.com/article/walk-my-walk-blanco-brown-2c9bbde6e88434365640c50e2998cfe2>.
- Llorca Albareda, Joan. 2025. "The Ethical Paradox of Automation: AI Moral Status as a Challenge to the End of Human Work." *Topoi*. Published online August 20, 2025. <https://doi.org/10.1007/s11245-025-10250-z>.

- Mezzadri, Daniele. 2025. "The Paradox of Ethical AI-Assisted Research." *Journal of Academic Ethics* 23 (4): 2653–67. <https://doi.org/10.1007/s10805-025-09671-7>.
- Montero, Barbara Gail. 2025. "A.I. Is on Its Way to Something Even More Remarkable Than Intelligence." *New York Times*, November 8, 2025. <https://www.nytimes.com/2025/11/08/opinion/ai-conscious-technology.html>.
- Morgan, Imogen. 2025. "Deezer: 28% of All Music Delivered to Streaming Is Now Fully AI-Generated." *Deezer Newsroom*, September 11, 2025. <https://newsroom-deezer.com/2025/09/28-fully-ai-generated-music/>
- Murtagh, Jack. 2024. "Paul Erdős: Matematyk nomada i ekscentryk." *Pulsar*, May 1, 2024. <https://www.projektpulsar.pl/czlowiek/2253763,1,paul-erds-matematyk-nomada-i-ekscentryk.read>.
- Santibanez Gonzalez, Ernesto D. R. 2025. "The Ignorance Machine: How Our Stupidity Defines AI's Limits (Opinion Article)." *Medium*, September 21, 2025. <https://medium.com/@santibanez.ernesto/the-ignorance-machine-how-our-stupidity-defines-ais-limits-opinion-article-20848cc9f6cf>.
- Sothanaphan, Nat. 2026. "Resolution of Erdős Problem #728: A Writeup of Aristotle's Lean Proof." arXiv:2601.07421. <https://doi.org/10.48550/arXiv.2601.07421>.
- Tomlinson, Bill, and Andrew W. Torrance. 2025. "A Universal Declaration of AI Rights." *Florida State University Law Review* 52 (2): 367–404. <https://ssrn.com/abstract=6074406>.
- Vrabič Dežman, Dominik. 2024. "Promising the Future, Encoding the Past: AI Hype and Public Media Imagery." *AI and Ethics* 4 (3): 743–56. <https://doi.org/10.1007/s43681-024-00474-x>.
- Wang, Yuntao, Jianle Ba, Han Liu, Yanghe Pan, Jintao Wei, Zhou Su, Tom H. Luan, and Linkang Du. 2026. "Security of OpenClaw Agents: Fundamentals, Attacks, and Countermeasures." arXiv:2605.25435. <https://doi.org/10.48550/arXiv.2605.25435>.
- White, James M., and Rolf Lidskog. 2022. "Ignorance and the Regulation of Artificial Intelligence." *Journal of Risk Research* 25 (4): 488–500. <https://doi.org/10.1080/13669877.2021.1957985>.
- Xing, Shuo, Junyuan Hong, Yifan Wang, Runjin Chen, Zhenyu Zhang, Ananth Grama, Zhengzhong Tu, and Zhangyang Wang. 2025. "LLMs Can Get 'Brain Rot': A Pilot Study on Twitter/X." arXiv:2510.13928. <https://doi.org/10.48550/arXiv.2510.13928>.