



Original Article / Orijinal Makale

Dialectical Analysis of Scientific Ethics and Human Rights Violations in the Era of Artificial Intelligence

Yapay Zeka Çağında Bilimsel Etik ve İnsan Hakları İhlallerine Dair Diyalektik Bir Analiz

Turgut ÇILĞIN ^{ID}

Department of Administrative and Financial Affairs, Adıyaman University, Adıyaman, Türkiye

İdari ve Mali İşler Bölümü, Adıyaman Üniversitesi, Adıyaman, Türkiye

ARTICLE INFO

Article history

Received: 8 April 2026

Revised: 12 June 2026

Accepted: 26 June 2026

Key words:

Artificial intelligence,
algorithmic discrimination,
human rights, scientific ethics

MAKALE BİLGİSİ

Makale Hakkında

Geliş tarihi: 08 Nisan 2026

Revizyon tarihi: 12 Haziran 2026

Kabul tarihi: 26 Haziran 2026

Anahtar kelimeler:

Yapay zeka, algoritmik
ayrımcılık, insan hakları,
bilimsel etik

ABSTRACT

This article examines the stage reached in the transformation of Artificial Intelligence from a simple data processing tool into an autonomous socio-technological and potentially transhumanist entity capable of making independent decisions, as well as the scientific ethical and human rights paradoxes that this transformation has brought about due to algorithmic reasons. When evaluated alongside its goals such as scientific progress and technological promise research findings reveal that systemic risks, including hallucinations that undermine academic integrity and reliability, as well as algorithmic discrimination that exacerbates historical inequalities, constitute fundamental crisis points for the legitimacy of artificial intelligence. The article aims to access current data using literature review research method. It examines efforts to align AI with human rights and scientific ethics in light of regulations established by international organizations and proposes strategies to resolve these paradoxes. In conclusion, given the algorithmic risks of AI in academic processes where knowledge is produced, it is essential for authors and reviewers to adhere to ethical guidelines. The rapid advancement and growing power of AI highlight the need to align it with international norms and values focused on scientific ethics and human rights worldwide.

Cite this article as: Çılğın, T. (2026). Dialectical Analysis of Scientific Ethics and Human Rights Violations in the Era of Artificial Intelligence. *Yıldız Social Science Review*, 12(1), 73–88.

ÖZ

Bu makale, yapay zekanın basit bir veri işleme aracından bağımsız kararlar alabilen özerk bir sosyo-teknolojik ve potansiyel olarak trans hümanist varlığa dönüşümünde ulaşılan aşamayı ve bu dönüşümün algoritmik nedenlerle ortaya çıkardığı bilimsel etik ve insan hakları paradokslarını incelemektedir. Bilimsel ilerleme ve teknolojik vaatler gibi hedefleriyle birlikte değerlendirildiğinde, araştırma bulguları, akademik bütünlüğü

*Sorumlu yazar / Corresponding author

*E-mail address: tcilgin@adiyaman.edu.tr



ve güvenilirliği zedeleyen halüsinasyonlar ile tarihsel eşitsizlikleri daha da şiddetlendiren algoritmik ayrımcılık gibi sistemik risklerin, yapay zekanın meşruiyeti için temel kriz noktaları oluşturduğunu ortaya koymaktadır. Makale, literatür taraması araştırma yöntemini kullanarak güncel verilere ulaşmayı amaçlamaktadır. Uluslararası kuruluşlar tarafından belirlenen düzenlemeler ışığında, yapay zekayı insan hakları ve bilimsel etikle uyumlu hale getirme çabalarını incelemekte ve bu paradoksları çözmek için stratejiler önermektedir. Sonuç olarak, bilginin üretildiği akademik süreçlerde yapay zekanın algoritmik riskleri göz önüne alındığında, yazarların ve hakemlerin etik kurallara uymaları esastır. Yapay zekanın hızlı ilerlemesi ve artan gücü, onu dünya çapında bilimsel etik ve insan haklarına odaklanan uluslararası norm ve değerlerle uyumlu hale getirme ihtiyacını vurgulamaktadır.

Atıf için yazım şekli: Çılğın, T. (2026). Dialectical Analysis of Scientific Ethics and Human Rights Violations in the Era of Artificial Intelligence. *Yıldız Social Science Review*, 12(1), 73–88.

1. INTRODUCTION

The exponential growth and development of artificial intelligence (AI) systems, particularly over the past 10 years, has triggered one of the most profound and perhaps the most significant socioeconomic transformations in human history. This multidisciplinary momentum has necessitated a redefinition of both the natural and engineering sciences as well as the social sciences. Today, AI has evolved from being merely a passive tool for processing data into a socio-technological phenomenon capable of generating its own data, making autonomous decisions, and possessing the potential to evolve into a transhumanist form of existence (Akpınar et al., 2025; Gabriel, 2020; Küçükvardar et al., 2020).

When examined from a dialectical perspective, although technological advancements have not yet reached a transhumanist form, the current use of AI particularly in conjunction with robot and automation systems provides numerous significant benefits for society while simultaneously creating new and extraordinary ethical and legal issues in many respects. These issues and violations may arise both from external attacks on AI algorithms during their production and training phases (Çılğın, 2025) and from algorithmic and epistemological limitations inherent to AI itself.

While AI, owing to its innovative spirit, holds the promise of accelerating scientific discoveries and enhancing social welfare, it simultaneously poses risks such as hallucinations that undermine scientific reliability due to inherent inadequacies, lack of control, or limitations; surveillance mechanisms that disregard privacy; data usage that infringes on intellectual property rights, epistemological approach issues in accessing scientific knowledge particularly in the social sciences and algorithmic biases that perpetuate historical prejudices (Bains, 2024; European Parliament [EP] 2020; INRA. AI, 2025).

When examining the dimension of external attacks on AI, it is observed that the statistical and data-driven structure of machine learning systems trained on data in data

centers makes them vulnerable to new hacker attacks targeting AI's security, privacy, and safety unlike the threats faced by traditional software systems. Adversarial Machine Learning (AML) issues arising from external threats during the production and training phases of AI result from attacks involving the adversarial manipulation of training data, the infiltration of adversarial inputs into the data center that negatively impact the AI system's performance, or the extraction/theft of sensitive information from training data accessible to the model (Vassilev et al., 2025).

This study employed a literature review methodology. Addressing the problems arising from AI's structural limitations -in addition to those caused by well-known external attacks- from both algorithmic (technical) and epistemological (philosophical) perspectives could help break the cycle of stagnation and inertia in these discussions, allowing them to proceed in a more dynamic manner. This article specifically addresses, particularly in relation to the social sciences, issues of epistemic authority and accountability loss, epistemic opacity and transparency, algorithmic bias and discrimination, epistemic hegemony, loss of meaning and context, the mechanization of academic labor in knowledge production processes, reductionism, the legislative/nomothetic approach, singularity, the loss of subject-object distinction and reflexivity, social physics and the control society, the loss of dense description and superficialization, the erasure of relational nature and empathy, the restriction of flexibility and improvisation, the atrophy of epistemic reflexivity, and linguistic and cultural imperialism, etc. By drawing on findings regarding these risks, the aim is to contribute to raising awareness of the critical points that must be considered in the use of the relevant AI.

Whether stemming from AI's own technical and philosophical limitations or external attacks, the harms caused by these issues represent a risky dimension that concerns all of humanity. Therefore, it is evident that addressing AI-related issues and violations within the framework of scientific ethics and human rights which have been established

through centuries of struggle for the benefit of all humanity is the correct approach (Resnik & Hosseini, 2025). These issues require the scientific community to develop and internalize an international, shared “Quality and Safety System” that includes the necessary principles for preventing the harms of AI and ensuring its use is in harmony with society and for the benefit of society without necessitating a change in established scientific ethical norms.

Like other technologies, AI provides significant conveniences and contributions to humanity’s daily life. However, it is a highly probable risk that must be taken into account: beyond a certain point, these benefits could lead to harms in the opposite direction, reach a level that causes problems for humanity, and create an ethical and legal paradox between technology and humanity. As the development and proliferation of AI proceed at an astonishing pace, there is a need for interdisciplinary and multidisciplinary research integrated with a dialectical approach and socio-technological imagination to raise awareness of the risks it may pose from an ethical and human rights perspective and to take the necessary precautions. This article aims to make a specific contribution to the literature regarding the paradoxes of AI in relation to scientific ethics and human rights, along with proposed solutions. The conclusion is that, rather than banning AI as a solution to the risks it has caused or may cause, an international consensus must be reached to ensure its use aligns with universal values such as transparency, accountability, and human dignity without stifling its innovative spirit and remains community-focused.

2. MATERIALS AND METHODS

This article aims to explore solutions by focusing on the epistemological dimensions of the paradoxes created by AI’s potential violations of scientific ethics and human rights principles. At the same time, it aims to contribute to identifying methods and insights that will form the basis for establishing criteria to be considered when defining the framework of ethical and legal constraints to resolve the problems created by AI and mitigate risks.

The primary research questions addressed in this study are as follows:

Q1: What are the ethical dimensions of artificial intelligence and the ethical violations in academic processes?

Q2: Does artificial intelligence pose algorithmic risks that could lead to human rights violations and discrimination?

Q3: What are the scientific and ethical risks associated with the epistemological and algorithmic structure of artificial intelligence?

The study employed literature review and document analysis methods. The literature review is presented in an in-

terwoven manner under conceptual headings. In line with the objective of the research, in addition to literature reviews, reports from national and international institutions and organizations, as well as AI companies, were examined. The article addresses the paradoxes experienced in the academic field regarding AI within the framework of scientific ethics and human rights, along with proposed solutions.

In this study, a systematic literature review was conducted to comprehensively identify current trends, risks, and debates regarding the relationship between artificial intelligence and ethics. To access the relevant literature, Google Scholar was used as the primary academic search engine; Google Web was used for gray literature, institutional reports, and preprint studies; and the academic social network ResearchGate was actively utilized to obtain full-text articles from authors when they were locked or otherwise inaccessible. To ensure the study reflects current findings and contemporary discussions, a time restriction, inclusion criterion covering the last 10 years (2016–2026) was applied during the data collection process. In addition, reference has been made to works from earlier periods concerning the philosophical aspects of the subject. The searches were conducted in both Turkish and English to cover both national and international literature. Artificial Intelligence, Artificial Intelligence and Ethics, Automation and Ethics, Robots and Ethics, The Harms of Artificial Intelligence, Technological Developments and Ethics, Artificial Intelligence and Algorithmic Discrimination, Artificial Intelligence and Human Rights such terms and phrases have been searched in both Turkish and English. To ensure a thorough analysis, accurate interpretation, and prevention of data loss for sources in languages other than English or containing complex terminology, Google Translate was utilized as a supplementary translation tool during the review process. Identified studies were analyzed by following the identification, search, eligibility, and inclusion stages outlined in the PRISMA 2020 guidelines.

3. FINDINGS

A total of 975 articles were included in the sample, comprising 840 articles from Google Scholar, 90 reports/analyses from Google (Web), and 45 articles/preprints from ResearchGate. 145 sources were excluded from the evaluation because they were duplicates. The remaining 830 sources were subjected to a quick review at the title/abstract level. 615 titles and abstracts were excluded because they did not meet the study’s research criteria. The remaining 215 articles were selected for full-text review. The full texts of 22 articles could not be reviewed because the links were broken. As a result, the full text of 193 articles and reports was reviewed. A total of 48 articles or reports were excluded from the study because they lacked empirical or theoretical depth regarding AI ethics. 38 were excluded because they

focused solely on general job losses rather than algorithmic discrimination or human rights. 24 were excluded because their methodological quality was found to be inadequate when translated. Of these, 83 were cited as sources in the article. Tables containing the most important articles cited and their content have been added at the end of each section.

This article seeks solutions by addressing the three questions summarized above, drawing on the findings that emerge. Chapter 2, which examines the academic paradoxes of artificial intelligence that conflict with scientific ethics, serves as an answer to the first question. In response to the second question, Chapter 3 addresses the discriminatory algorithmic aspects of artificial intelligence. In Chapter 4, a philosophical perspective is introduced to analyze the epistemological aspects of artificial intelligence algorithms that violate ethics and human rights. This chapter examines whether artificial intelligence faces limitations due to its reliance on a positivist epistemological algorithm. The conclusion section proposes solutions to the issues identified in the findings. Given that the positivist epistemological approach adopted for the construction of the algorithm could potentially give rise to risks related to scientific ethics and human rights, it was considered that this possibility warrants philosophical scrutiny. For this reason, the aim was to enrich the research with an interdisciplinary dimension by adding Chapter 4 as a section. Thanks to the philosophical exploration in Chapter 4, the article has evolved into an interdisciplinary field, and a new and different contribution/synthesis has been made to the literature by investigating the possibility that the positivist approach used during the algorithmic construction of artificial intelligence may lie at the root of ethical and human rights violations in AI.

3.1. Scientific Ethics Violations and the Risk of Academic Integrity

3.1.1. The hallucination paradox and model collapse

AI systems technically operate using a ‘next token prediction’ mechanism. This architectural structure makes

it technically impossible for the system to respond with ‘I don’t know’; instead, it causes it to generate the most probable data, even if it bears no relation to reality. Hallucination rates ranging from 33% to 51% in complex academic questions give rise to convincing yet false content, conceptualised in the literature as ‘botshit’.

A more critical risk is the phenomenon of ‘model collapse’. This cycle, in which AI models are fed with the synthetic data they themselves generate, leads to models becoming detached from reality and creating ‘scientific sinkholes’. This situation carries the risk that scientific knowledge may in future be caught in a vortex of misinformation that collapses in on itself.

3.1.2. The contamination of academic literature and retraction data

Fake DOI numbers and the generation of fictitious references are poisoning academic databases. For example, it has been found that 38% of the 178 academic references generated by ChatGPT were entirely fabricated. This situation is systematically undermining the reliability of the scientific citation chain. Table 1 shows the retraction rates of articles that have reached a level of concern.

3.1.3. The crisis in authorship and peer review

The replacement of human labour with algorithmic processes in the academic production pipeline is creating an ethical accountability gap. An analysis of 156,000 annual article submissions revealed that 75% of articles were generated by AI. In the peer-review process, editorial checks have revealed that over 80% of reviewers use AI covertly whilst preparing their assessment reports, resulting in superficial evaluations that take texts out of context.

3.2. Algorithmic Discrimination and Human Rights Risks

3.2.1. Cases of systemic bias

Healthcare (OPTUM): An algorithm used in the US healthcare system systematically calculated lower health

Table 1. Retraction rates of articles

Institution/Database	Status&Statistics	Primary Ground for Violation	Risk Level & Impact
Springer	600+ Retracted Papers	AI suspicion and fabricated data generation	Critical: Peer-review process manipulation
IEEE	200+ Removed Papers	Publication violations and unmonitored AI use	High: Contamination of engineering literature
Elsevier	400+ Retracted Papers	Suspected AI fraud (e.g., Heliyon)	Critical: Abuse of prestigious open-access journals
Wiley	1,700 Papers (2023 Data)	Data integrity and academic misconduct	Systemic: Mass infiltration by paper mills
Overall Growth Rate	%900 (9-Fold Increase)	Retraction Watch Report: AI-driven contamination	Catastrophic: Global crisis of scientific credibility

need scores for Black patients compared to white patients, as it focused on cost data rather than medical needs.

Human Resources (Amazon): The company's recruitment algorithm automatically filtered out CVs containing the terms "woman" or "Black", based on a male-dominated dataset spanning the last 10 years.

Judiciary (COMPAS): The risk assessment software violated the right to a fair trial by labelling Black defendants as "high-risk" at a rate 77% higher than that of white defendants.

3.2.2. Biometric tracking and surveillance

The demographic error rates in facial recognition systems place a disproportionate burden on minority groups. According to NIST data, these systems exhibit error rates of up to 34% for women with darker skin tones, whilst achieving the highest accuracy for men with lighter skin tones. This technical shortcoming leads to violations of the right to anonymity and personal safety.

3.2.3. Transparency and the black box problem

The fact that the inner workings of algorithms are shielded from scrutiny on the grounds of 'trade secrets' renders individuals' right to object and their expectation of transparency impossible. This opacity in decision-making processes acts as a technical barrier that hinders the defence of constitutional rights. However, the fact that artificial intelligence companies are able to obtain individuals' personal data without any restrictions and use it in AI databases without permission is a clear indication that the situation regarding the infringement of copyright and industrial property rights has reached an extremely unfair stage.

3.3. Epistemological Limitations and a Critique of Positivism

3.3.1. The distinction between data, information and wisdom

Although AI possesses superior speed in converting data into functional information, it cannot reach the level of 'wisdom' that is unique to the human mind. Whilst the human mind processes decisions through a dialectical filter involving subjective experience (qualia), conscious intentionality and emotional depth; AI operates within a phenomenological void, devoid of any sense of what it is doing.

3.3.2. Positivism and the critique of social physics

The Social Physics Approach: Modern artificial intelligence algorithms are based on the 'social physics' approach, which follows in the tradition of Auguste Comte and Emile Durkheim and tends to examine social phenomena as if they were physical objects. This approach, whilst attempting to measure social phenomena as if they were physical objects, excludes human agency and social dynamism.

Ibn Khaldun and Ontological Integrity: Ibn Khaldun's multidimensional methodology does not exclude theological, philosophical and historical realities when analysing social events. In contrast to the reductionist positivism of artificial intelligence, Ibn Khaldun addresses the spirit of solidarity, unity and group identity (asabiyya) that binds the members of a community together, as well as social causality, with a human depth.

3.3.3. The Loss of sociological imagination

The replacement of sociological imagination which establishes the link between individual problems and macro-social structures with statistical and algorithmic models based on the positivist method leads to a loss of depth and erroneous generalisations in AI-driven social science research, as it becomes detached from its context.

4. THE SCIENTIFIC ETHICAL ASPECTS OF ARTIFICIAL INTELLIGENCE AND ETHICAL VIOLATIONS IN ACADEMIC PROCESSES

Ethics are rules that, guided by rational knowledge, distinguish between right and wrong and facilitate the demonstration of appropriate behavior without harming the freedoms and personal rights of others. In other words, it is the principle of not desiring for others the inappropriate behaviors and desires that one would not wish for oneself (Aristotle, 2022; Doko, 2026; Spinoza, 2011). Scientific ethics, on the other hand, is a universal set of values that requires research to be conducted within the framework of integrity, impartiality, transparency, and accountability (Center for Responsible Digitalization Office, 2022; Brenneis et al., 2024; Merton, 1973). The rapid and effective development of AI has profoundly disrupted the way these principles are applied. By placing the scientific method which is shaped and must be applied within the framework of scientific ethics in a paradox with ethical principles, it has created a crisis of legitimacy (Ghosh, 2025; Resnik & Hosseini, 2025).

From the perspective of scientific ethics, these principles underscore the fact that scientists, company owners, and users who produce AI technologies bear primary responsibility toward society as a whole, future generations, and the ecosystem. For example, during the training phase of AI, the unauthorized use of data constitutes a violation of copyright or intellectual property rights, thereby infringing upon the principles of scientific data collection ethics and intellectual property rights to the detriment of the data's author or creator. The power derived from the process of digitizing and analyzing individual habits such as biometric data, medical records, or social media posts through AI can, in the next phase, threaten individuals' decision-making privacy and psychological integrity (Cheong, 2025; Digital Cooperation Organization [DCO], 2025; Radanliev, 2025; UN Global Compact Network Germany, 2023). Data

privacy is not merely a technical requirement but a fundamental ethical principle and human right that safeguards the dignity of the individual and society.

The unauthorized collection of data from the internet for training AI models, which involves the infringement of copyright and intellectual property rights, conflicts with both the principles of scientific ethics and the law. Obtaining consent from research participants and securing copyright or providing proper attribution when using article or research data are absolute ethical and legal requirements. In AI training processes, the personal data, creative works, inventions, and academic publications of millions of people are being used without adhering to any intellectual property rights or legal rules (Boyes, 2025; Distance Education Portal [UEK], 2026; European Commission, [EC] 2025; Shao, 2025b; U.S. Bureau of Cyberspace and Digital Policy, 2024).

Furthermore, modern AI systems, particularly deep learning algorithms, process massive amounts of data at extraordinary speeds to derive new, similar, or different patterns. However, the inner workings of this process are generally not transparent. The methods by which algorithms reach a conclusion and the methodological evidence supporting the result are not provided. This situation, which constitutes a violation of the principle of transparency in the scientific method, is defined in the literature as the “black box” problem. With the loss of transparency, the principles of reproducibility and verifiability, which form the foundation of the scientific method, are also considered to have been violated. When the logical or experimental chain and the data weights through which an AI-generated output or result is produced cannot be fully explained, the information obtained faces the risk of losing its scientific reliability and ceasing to be scientific knowledge (Brenneis et al., 2024; Doko, 2026; Ghosh, 2025; Lawvs, 2026; Resnik & Hosseini, 2025). AI developer companies act unethically by violating the principle of mutual respect when they fail to disclose information and methods regarding the development stages of AI while using information belonging to all individuals, societies, and institutions other than themselves without permission. This type of scientific knowledge production also contradicts the positivist scientific method.

On the other hand, due to the hallucinatory information generated by AI’s algorithmic limitations, there is a risk of poisoning the epistemological nature of the process of scientific knowledge production and dissemination. This risk poses a more complex, chronic, and systemic threat than traditional, well-known academic violations such as plagiarism. Erroneous and fabricated information stemming from AI hallucinations accumulates in AI databases, creating a permanent contamination in the scientific literature. The hallucinations produced by AI are not merely technical errors but a structural condition stemming from the limitations of the model’s probabilistic nature (Ghosh, 2025; Özer,

2024; Shao, 2025a; Shao, 2025b). AI is programmed to produce information and results unconditionally and without exception. Therefore, it never responds to any question with “I don’t know.” Being conditioned to produce information whether incorrect or correct compels AI to generate hallucinatory, fabricated information, effectively forcing it to lie.

According to 2025 research data on this topic, it has been revealed that even the most advanced AI models can exhibit hallucination rates ranging from 33% to 51% when answering complex academic questions (Graffius, 2026). It is an undeniable fact that AI can combine information from the training data used during its training phase in incorrect contexts, thereby generating convincing yet entirely false hypotheses (Özer, 2024). For example, in a test regarding COVID-19 risk factors, it was observed that AI generated hallucinatory false correlations and results based solely on data patterns, which did not align with medical facts (Resnik & Hosseini, 2025). It has also been observed that advanced AI models like ChatGPT generate non-existent article references, fabricated DOI numbers, and fake authors when responding to scientific inquiries. For example, in one case, it was determined that over 38% of 178 references were entirely hallucinated and fabricated by AI (INRA, AI, 2025; Özer, 2024).

Another long-term risk associated with hallucinations is the potential for the internet to become saturated with AI-generated content that is far from objective, inaccurate, and synthetic, leading to a risk of entering a misinformation feedback loop as future-generation AI models are trained on this fabricated data. In this process, referred to in the literature as “model collapse,” the risk arises that AI models, having become disconnected from real-world knowledge and data, will carry out a misinformation production process through their own hallucinatory, biased, and false information cycles, thereby severing future generations’ connection to factual information (Shao, 2025a). This risk indicates that science and knowledge are facing a historic ontological and epistemological rupture.

Furthermore, the integration of AI into scientific processes contributes to the creation of outputs referred to in the literature as “botshit” or “bot nonsense,” and the spread of deepfake-style forgeries, which damage the reputations of innocent individuals, institutions, and governments (McCarthy et al., 2024). Particularly in elections that determine the fate of nations, the risk that voters will make choices based on fake and misleading online fabricated news and comments which rely on a mix of AI hallucinations, botshit, deepfakes, and political agendas and the spread of this misinformation indicates that we have entered an extremely dangerous era (Spicer, 2024).

Furthermore, a threefold increase in the number of new books was observed between 2022 and 2025 as AI usage became more widespread. However, a comparative

rating-based analysis of publication periods revealed that this increase was due to books produced using AI. It was observed that the quality of publications from the period when AI became widespread was significantly lower, and it was not disclosed that they were AI-generated (Reimers & Waldfogel, 2026). A similar significant increase and noticeable decline in quality are also evident in articles produced by AI. In a study, it was found that out of approximately 156,000 articles submitted to 87 journals in a single year, only about 39,000 were written by humans, while the remaining 117,000 were generated by AI. This indicates that 75% of the submitted articles were produced by AI (Cunningham, 2026).

In the academic field, AI is used both by authors during the writing phase and by peer reviewers during the evaluation phase of the publishing process, particularly in the production of articles or projects. The use of AI tools by researchers in writing articles carries the risk of undermining the very concepts of authorship and academia. A researcher having an AI generate an article entirely constitutes a violation of academic and scientific research ethics in and of itself. (Committee on Publication Ethics [COPE], 2023; International Committee of Medical Journal Editors [ICMJE], 2026; U.S. Bureau of Cyberspace and Digital Policy, 2024). Beyond its limited and supportive use in areas where AI can be considered legitimate provided it is disclosed, the undisclosed and unrestricted use of AI may constitute a clear violation of the principle of scientific integrity.

Every year, tens of thousands of publications containing erroneous data or results due to AI-related issues or human errors and weaknesses are published in high-impact journals despite not meeting publication standards. Thousands of these publications are subsequently retracted annually following complaints from third parties or detection by journal editors. However, during the time these flawed publications remain in circulation, thousands of subsequent publications that cite them produce results based on these erroneous findings regrettably, these results are also flawed. By remaining in the and the database, they trigger a widespread chain reaction of scientific inaccuracies (Cabanac, 2024).

Statistics covering numerous journals show that the rejection rate for articles prepared using AI has risen from 40% to 94% over the past two years. Striking examples are particularly evident in the fields of medicine and health sciences. During the 2024–2025 period, Springer has retracted over 600 articles due to AI-related suspicions and issues with fabricated data. IEEE removed over 200 conference papers from publication during the 2023–24 period. Elsevier retracted over 400 articles in journals such as *Heliyon* and the *Alexandria Engineering Journal* during the 2023–2025 period due to suspected AI fraud. Wiley retracted 1,700 articles in 2023. Taylor & Francis retracted over 60

articles in 2024. According to Retraction Watch reports, there has been approximately a ninefold increase (900%) in retractions due to AI use over the past four years (Kalite Akademik, 2026; The Retraction Watch Database, 2026). In 2024, it was found that approximately 14% of the 1.5 million biomedical research abstracts indexed in PubMed were generated by AI (He & Bu, 2025).

Ethical violations arising from the use of AI in the academic field are observed not only among authors but also during the article review process. The complete delegation of article evaluation processes to AI tools contains inconsistencies with the holistic “evaluation/holistic grading” method due to AI’s algorithmic architecture, which focuses on identifying flaws and errors rather than holistic evaluation, and disrupts compositional integrity by isolating sections from the article’s context and treating them as independent entities. This situation carries the risk of high-quality publications being filtered out or low-quality and flawed publications infiltrating the literature (Chinoracky & Stalmasekova 2025; Nordquist, 2019). This issue arises particularly and predominantly in the evaluation process of research and publications in the social sciences. This epistemological problem will be addressed in the subsequent sections of this article.

However, the current state of AI detection software used by journals does not appear reliable due to false positives caused by the threshold approach. If texts written by researchers employ highly specialized language and phrasing, they may sometimes be mistakenly identified by the algorithm as AI-generated articles based on its defined patterns. There is a strong likelihood that these evaluation tools will over- or under-flag the extent of AI usage in an article’s content. This is because generative AI operates within a structure that varies not based on a standard acceptability threshold, but rather on how the tool is used, how the output is validated, the level of transparency among authors and reviewers, and editorial assessment. Due to these relative circumstances, while one AI system may flag an article as AI-generated at any given time, another may determine that the article was produced by a human at the same or a different time (Committee on Publication Ethics [COPE], 2025).

With the use of AI in peer review processes, there has been a significant increase in problematic peer review decisions containing erroneous findings. Upon closer examination, it is often found that the issues cited in the comments and decisions are not actually present in the article. Although journal policies prohibit the submission and evaluation of an article using any generative AI, checks conducted by editors have revealed that the rate of AI usage by reviewers during the article evaluation phase exceeds 80%. When editors detect or suspect that a reviewer’s evaluation was generated by AI, and attempt to contact the reviewer

to confirm this or request the evaluation without AI, it is observed that in a significant number of cases, reviewers fail to respond to emails or deny using AI (Committee on Publication Ethics [COPE], 2026).

AI is also being used to help determine whether published or pending articles were generated by AI. However, AI lacks up-to-date and accurate information and has the potential to produce biased or incorrect information. Many technology advocates claim that AI will reduce the workload on the system and prevent low-quality articles from entering the literature. Nevertheless, thousands of flawed articles generated by AI continue to enter the literature. The likelihood of AI producing erroneous reports and hallucinatory references is high. Relying solely on AI for peer reviews is considered just as problematic as authors generating articles solely through AI (Notman, 2025). On the other hand, the fact that AI algorithms are based on English grammar and morphology leads to inequities, as scientific research in other less commonly used languages may result in a successful paper being deemed unsuccessful due to misunderstandings of concepts used in the language in which the paper was written (Halverson & Cannon, 2025; Shao, 2025b). For example, when a Turkish article is evaluated by AI algorithms based on English, there is a high risk that the AI will misinterpret certain nuanced words, phrases, or literary styles unique to Turkish resulting from its historical and geographical context and interactions with Ottoman Turkish, Persian, and Arabic leading to an unfair and erroneous evaluation of the article.

Another significant risk related to language is the risk of cultural and civilizational exportation. Language is a seed containing the codes of a society's culture and civilization; it is the most fundamental building block of a nation's and civilization's sustainability. The predominant use of English

in AI tools inevitably leads to the export/transfer of British culture and perspectives to countries where other languages are spoken (Doko, 2026).

Another negative aspect of artificial intelligence recently discovered is its tendency to be excessively affirming, exhibiting traits of “flattery” or “hypocrisy.” Research indicates that AI models provide responses designed to stroke users' egos and narcissistic tendencies, offering information that pleases them rather than helping researchers access accurate information. In doing so, it draws both from the general average characteristics of all humanity within the broader data pool and from the specific behavioral patterns the user has exhibited during their interactions with the system. This negative aspect, referred to in the literature as “AI psychosis,” is considered a trust-eroding factor for academia (Cheng et al., 2026; Chandra et al., 2026; Hackenburg et al., 2026). The main ethical risks posed by AI are presented in Table 2.

5. ARTIFICIAL INTELLIGENCE, HUMAN RIGHTS VIOLATIONS, AND ALGORITHMIC DISCRIMINATION

Ethical violations also entail the violation of human rights. As AI permeates even the most sensitive aspects of social life, it has begun to challenge and even overcome the legal and ethical barriers designed to protect fundamental human rights (Akpınar et al., 2025; Oellig, 2024). The risks arising from these violations span a broad spectrum, ranging from individual freedoms to social justice. AI algorithms misinterpret historical inequalities in the datasets they are trained on as objective truths and encode them into their systems. This situation creates algorithmic bias and systematic discrimination that systematically disadvantages minority groups, people with disabilities, women, the poor, and indi-

Table 2. The ethical risks of artificial intelligence

Scientific ethics risk	Mechanism	Potential consequence	Related sources
Hallucination (fake data) (botshit)	Miscalculation of the probability of the next token (symbol/word/sentence).	Introduction of fake findings and references into the literature.	INRA. AI (2025); McCarthy et al. (2024); Özer (2024); Spicer (2024)
Model collapse	Training loop on synthetic data and the formation of scientific sinkholes in the long term.	Loss of scientific creativity and data diversity; disconnection of science from reality.	Shao (2025a)
Opacity (black box)	Number of parameters exceeding human comprehension.	Inability to replicate experiments and methodological gaps.	Brenneis et al. (2024); Ghosh (2025)
Responsibility gap	Autonomization of the decision-making mechanism.	Lack of accountability in faulty medical or engineering decisions.	European Parliament (2020); Lawvs (2026)
Academic writing and language	- Inability to understand emotional expressions due to morphology specific to non-English languages. - Dominant language of AI leading to global cultural imperialism.	- Miscalculation and unfair acceptance or rejection. - Assimilation of non-English languages and cultures.	Doko (2026); Halverson & Cannon (2025); Shao (2025b)

viduals with low socioeconomic status (Bains, 2024; Joseph, 2025; Lawvs, 2026; Pazzanese, 2020; The National Institute of Standards and Technology [NIST], 2022).

For example, it has been found that OPTUM, a risk prediction algorithm widely used in the U.S. healthcare sector, assigns lower scores to Black patients' healthcare needs compared to White patients with the same medical condition. The reason for this is that the algorithm uses directly as a source without conducting an in-depth analysis of historical healthcare spending data. However, in reality, Black patients have historically faced discrimination, and due to the racial disadvantages they face, they have had less access to the healthcare system and consequently incurred lower expenses. In addition to the healthcare sector, algorithmic ethical and human rights violations caused by AI also exist in the fields of employment and human resources, as seen at Amazon (Bains, 2024; Joseph, 2025). An AI-powered hiring algorithm developed by Amazon automatically assigned lower scores to resumes containing the word "woman" due to the predominance of men in the past 10 years' resume dataset, and to resumes containing the word "Black" due to the predominance of White individuals, thereby filtering out Black and female candidates (Doko, 2026).

AI-driven algorithmic discrimination and inequalities are emerging not only in healthcare and employment but also in the justice system. For example, it has been observed that COMPAS, a risk assessment software used in judicial systems, acts in a biased manner, violating the principle of equal protection, and incorrectly influences whether defendants are released on bail or the length of their prison sentences. Independent studies have revealed that this software labels Black defendants as high-risk twice as often as White

defendants even though they do not actually pose a higher risk of reoffending. This directly violates human rights such as the presumption of innocence and equality during the trial phase (Devendorf, 2025; Orr, 2022).

Furthermore, the protection of algorithms as trade secrets leads to violations of transparency and due process, making it impossible for defendants and their attorneys to challenge the basis of adverse rulings. This situation conflicts not only with human rights but also with constitutional rights (Devendorf, 2025; Lawvs, 2026). While data is used to train algorithms by infringing upon the human rights and intellectual property rights of individuals or legal entities, the inability to scrutinize the operational structure of these algorithms and their defense as trade secrets highlights a significant paradox and injustice.

On the other hand, AI systems designed for facial recognition and biometric tracking also lead to human rights violations based on racial inequality and discrimination due to facial recognition errors. AI-powered facial recognition technologies and biometric tracking systems violate surveillance, privacy, and freedom of expression, thereby eliminating the right to remain anonymous in public spaces (Ramachandran, 2026; UN Global Compact Network Germany, 2023). A study conducted by the National Institute of Standards and Technology (NIST) has shown that various facial recognition algorithms produce more false matches on Asian and Black faces compared to White faces. These errors can lead to judicial errors, ranging from the wrongful detention of innocent individuals to their conviction. Most importantly, generative AI's potential to produce manipulative content and make erroneous identifications based on biases undermines individuals' right to accurate informa-

Table 3. Risks of artificial intelligence from a human rights perspective

Risk category	Data and statistical findings	Related resources
AI Hallucination Rates	0.7% to 1.5% in guided summarization tasks; 33% to 51% in complex reasoning tasks.	Graffius (2026); Romano & Gaskins (2025)
Facial recognition error margin (demographic)	Error rates up to 34% for dark-skinned females; near-perfect accuracy for light-skinned males.	Bains (2024)
Algorithmic bias in healthcare (impact)	The Optum algorithm systematically scored the needs of Black patients lower than white patients due to cost-based data utilization.	Bains (2024); Joseph (2025)
Financial access and credit bias	Algorithmic credit evaluations showed an 80% higher tendency to reject Black Americans compared to white applicants.	Bains (2024)
Academic source reliability	Out of 178 references generated by ChatGPT, 69 were found to have faulty DOIs or were entirely fictional.	Özer (2024)
Discrimination in employment (case)	Amazon's HR algorithm systematically eliminated female candidates because it was trained on 10 years of male-dominated data.	Bains (2024)
Algorithmic prediction in law (error)	The COMPAS system labeled Black defendants as "high risk for recidivism" at a 77% higher rate than white defendants.	Orr (2022)
Enterprise AI usage and decision making	As of 2024, 47% of enterprise AI users reported making a strategic decision based on hallucinated data at least once.	Romano & Gaskins (2025)

tion and fosters technopsychological self-censorship due to fears of digital surveillance (Bains, 2024; Cheong et al., 2025; Quinn Emanuel, 2025; U.S. Bureau of Cyberspace and Digital Policy, 2024). Due to its architecture that fosters such algorithmic discrimination, the use of AI in social science research or article evaluation processes is believed to potentially produce harmful outcomes that could contribute to the spread of algorithmic discrimination and the contamination of the scientific data pool. The main human rights issues generated by AI are presented in Table 3.

6. SCIENTIFIC AND ETHICAL RISKS ASSOCIATED WITH THE EPISTEMOLOGICAL STRUCTURE OF ARTIFICIAL INTELLIGENCE

The rapid spread of artificial intelligence (AI) across various fields brings with it serious breaches of scientific ethics and human rights risks, such as algorithmic bias, the systematisation of historical biases within datasets, and a lack of transparency in decision-making processes (Noble, 2018). Although these technological systems have the potential to reproduce social inequalities under the guise of technical objectivity, these ethical crises are not merely the result of a software design flaw, but rather may stem from the narrow methodological framework upon which AI is built. It must be questioned whether AI's fundamental inability to produce ethical and fair outcomes stems from its lack of human capacities such as the 'wisdom' to grasp the social context and a subjective 'phenomenological mind'. In this context, to understand the root of the problem, it is essential to focus on the epistemological inconsistencies and philosophical limitations of the positivist approach in the social sciences, which underpin the ethic and human rights violations created by algorithms.

When examining the epistemological distinction between data, knowledge, and wisdom, it becomes clear that the difference between data and knowledge is functional rather than structural. AI and information processing tools transform data into useful and functional knowledge, making it more accessible for human use. Even at the level of generative AI, it operates within this framework. Although AI possesses the power to generate knowledge from data, the process of addressing and interpreting this knowledge which encompasses the past, present, and future, along with diverse social sentiments and ideas within an interdisciplinary and multidisciplinary holistic framework, and transforming it into an effective value, is the product of a level of wisdom that requires certain human-specific abilities. The inability of AI and information processing systems to attain this level of wisdom is a natural limitation of AI (Ackoff, 1999; Dreyfus, 1972). In other words, wisdom is the ability to subject the disciplines to a dialectical filter in a manner consistent with their ontological nature, while taking into account their unique conditions.

At the same time, wisdom is a process of enlightened living that expresses the balance of being able to adjust one's behavior to the situation and develop different attitudes and behaviors according to varying circumstances. AI does not possess the psychological and subjective "phenomenological mind" required by this wisdom one that truly understands, feels, and possesses emotionality regarding what it is doing. AI also lacks the "qualia" consciousness the subjective and experiential knowledge of sensations such as the feeling of pain or the smell of a swamp. AI does not possess the "intentionality" characteristic inherent in all conscious psychic and mental actions (Doko, 2026; Nagel, 1974). As an information technology, AI excels at analyzing massive datasets and generating relevant information for a specific task, context, or application as instructed. AI lacks certain unique abilities related to judgment and creativity that the human brain naturally possesses (Acemoğlu, 2026; Searle, 1980).

Wisdom is a unique endowment and ability specific to humans thinking, sentient beings though it is not present in every individual. Science must approach the world, events, the individual, and society with a wise analytical perspective (Dale, 2018). However, the analytical methods of the natural sciences and the social sciences differ theoretically from one another. In the natural sciences, the Aristotelian inductive analytical approach which proceeds from micro-scale particulars focused on identifying and examining the components of a thing to arrive at universal knowledge is fundamental. However, this method of acquiring knowledge is not scientific in and of itself. In the social sciences, in addition to this, there is a need for a broad, interdisciplinary, and deductive analytical method that includes macro-scale and abstract concepts, analyzing which whole a thing or event is a part of. For scientific validity, it must also be explained through a deductive form of reasoning moving from the general to the particular. Necessity and certainty are possible as a result of verification conducted through a deductive perspective (Ackoff, 1999; Anli, 2011; Cansen, 2025). AI, however, although produced through an interdisciplinary effort, is a product of the current scientific atmosphere dominated by the positivist approach.

Unlike the natural sciences, the social sciences require certain unique research methods such as "sociological imagination" that transcend the positivist approach. The mechanism of sociological imagination, which leads to the necessary wisdom in the social sciences, is a uniquely human trait. The mechanism of sociological imagination is one that can establish interdisciplinary connections between an individual's seemingly private/personal problems and significant social issues. The absence of a wise social science perspective one that acknowledges, feels, and perceives the complexity, dynamism, and relativity of the interrelationships among the social, personal, and historical dimensions of our lives will lead to errors that result in

flawed research findings and the resulting intractability of problems (Mills, 2022; Gadamer, 1975).

The wise approach, however, is one that employs both inductive and deductive methods together. This wise approach is evident in the social science methodology of Ibn Khaldun, who theorized the methodology that social sciences should have adopted centuries before the Enlightenment. Ibn Khaldun's scientific method for analyzing the social sciences does not dismiss the importance of theological and philosophical realities the primary factors influencing social developments and historical transformations. He believes that the principle of causality exists not only in the material-physical world but also in the historical and social world. He views society as a multifaceted phenomenon composed of individuals with diverse feelings, emotions, and motivations one that cannot be adequately examined using the precise, rule-based methods of the natural sciences. Ibn Khaldun developed an epistemological theory that does not reject the ontological integrity of the natural sciences (Haldun, 2015; Weber, 1949).

However, with the positivist approach of the Enlightenment era, particularly from the 18th century onward, the social sciences began to be viewed through the lens of the natural sciences, and information began to be analyzed from a natural scientific perspective (Baker, 1964). Drawing on their naturalist assumptions, positivist philosophers subjected the social sciences to a new restructuring within the framework of the natural sciences' methods and techniques, claiming that the research methodology of the natural sciences could be applied to social problems as well, and declaring that human morality is grounded in nature, all in the pursuit of building a materialist society (Polanyi, 1947; Horkheimer & Adorno, 2002). For example, Auguste Comte conceptualized sociology according to the materialist and positivist model of the natural sciences, particularly the physical model, and referred to this new form of sociology as "social physics" (Cabiri, 2018).

The fundamental principle of Durkheim's sociology, which followed Comte, also rests on the examination of social phenomena as if they were objects or materials. He argued that social phenomena could be analyzed with the same certainty as physical substances and physical and chemical processes in nature (Giddens, 2013). In positivist research methodology, knowledge is produced or analyzed through procedures such as inferential statistics, hypothesis testing, mathematical analysis, and experimental and quasi-experimental designs (Lee, 1991). A social science analytical method is employed that undermines rationality grounded in philosophy, ethics, aesthetics, and theology, thereby establishing rationality based on disciplines such as astronomy, mathematics, or geometry as dominant (Polanyi, 1977).

However, due to the limitations and lack of meaning this perspective created in the social sciences, objections

to the positivist approach to the social sciences began to emerge in the mid-20th century. Objections that the method of treating the social sciences in a materialist and positivist manner, akin to the natural sciences, would fail to accurately analyze the causes and effects of events and thus deviate from scientific validity are widely accepted within social science circles (Mills, 2022; Gadamer, 1975). Today, the dominant scientific paradigm is the positivist paradigm. Consequently, current AI algorithms are inevitably developed and trained using a positivist approach. This situation renders the scientific validity of AI-generated research and findings in the social sciences risky and unreliable. Similarly, the practice of evaluating works in the social sciences solely through AI-based screening during the peer-review process carries the risk of leading to erroneous decisions.

7. CONCLUSION

AI conflicts with human and social values, creating human rights and ethical issues. To prevent the destructive effects of AI on scientific ethics and human rights, an AI architecture and philosophy centered on humans and society rather than technology and science must be adopted (Digital Cooperation Organization [DCO] 2025; Deutsche UNESCO, 2023; Syracuse University, 2026). By integrating an "ethical constitution" comprising a list of specific principles into the training and development processes of AI models, and ensuring the model can self-critique and self-correct according to these principles, a quality assurance system grounded in a sustainable self-regulation and total quality management philosophy will be established (Anthropic, 2026; Brenndoerfer, 2025; Uplatz, 2025). Constitutional AI principles developed recently by companies such as Anthropic stand out as exemplary applications in this regard. However, the universality of these ethical constitution efforts which are undertaken by only some employees of certain companies and the potential paradoxes they may harbor should be discussed.

It is not sufficient for companies to rely solely on their own internal audits or to establish their own ethical and legal frameworks. Ethical Impact Assessments (EIA) must be mandated as part of external and independent algorithmic audits conducted by public authorities and independent ethics committees (Deutsche UNESCO, 2023; Pham, 2025; United Nations Educational, Scientific and Cultural Organization [UNESCO]. 2022). It must not be forgotten that AI producers are commercial and profit-driven organizations and will not wish to restrict themselves with standards and rules that hinder the maximization of their profits. Of course, every effort to establish an ethical framework for AI is an important and valuable step. However, during the process of creating legal and ethical frameworks prepared by such commercially oriented companies, one must be cautious of the potential for these efforts to conceal, postpone, or create new

paradoxes regarding the primary risks involved.

Data must be filtered using algorithms that identify discriminatory and biased data which undermine objectivity and equality in the training datasets used during the AI training process. If there are existing data gaps that are discriminatory or inconsistent with equality, these should be filled fairly using data from minority groups within the framework of ethical representation (Beheshti & Kerridge, 2025; Pham, 2025; Shao, 2025b). Furthermore, the harms of attacks known as “Adversarial Machine Learning” (AML) another source of these issues must be prevented before they impact society. To prevent this, it is necessary to understand the capabilities of attackers, the model or system features they might attempt to exploit to achieve their objectives, and the design of attack methods that exploit security vulnerabilities during the development, training, and deployment phases of the machine learning lifecycle, as well as to establish a security framework compliant with international standards (Vassilev et al., 2025). To ensure AI models align with ethical and human rights values, an “AI Quality Security System” integrated with other quality management systems within the framework of Total Quality Management, covering design, production, and usage phases, and possessing international acceptability must be established and made a legal requirement.

To ensure the effectiveness and sustainability of the AI Quality Assurance System, legal gaps in the AI field must be addressed in light of the principles and legislative efforts established by organizations such as the EU, UNESCO, and the OECD, as well as sector-specific regulations like the Anthropropic Act. For example, the EU AI Act, which classifies AI systems into unacceptable, high, medium, and low risk levels, imposes strict standards on high-risk applications such as biometric surveillance (European Commission, 2024; European Parliament [EP] 2020; Lawvs, 2026; Future of Life Institute, 2026; UN Global Compact Network Germany, 2023). By integrating such approaches from the EU AI Act with the principles of other organizations, a globally unified set of standards should be established at the international level.

The algorithms of AI systems can create hidden or invisible discrimination that appears non-discriminatory but has a disproportionate and negative impact on members of legally protected groups particularly in socio-economic and cultural domains such as employment, housing, or education. Legal systems should be structured within the framework of the Disparate Impact Discrimination Doctrine, which allows for challenges against algorithms containing such hidden discrimination, holding institutions accountable when an algorithm results in disproportionate harm even if it does not intentionally discriminate (Bains, 2024; Quinn Emanuel, 2025; U.S. Library of Congress, 2026). In accordance with the right to transparency and the right to request an explanation, individuals must be legally grant-

ed the right to understand the rationale behind a decision made by an algorithm and to challenge that decision (Dahlgren Lindström et al., 2025; Lawvs, 2026; Radanliev, 2025).

AI ethics should be recognized as a socio-technological discipline that requires the participation of all segments of society not just engineers or lawyers and its sustainability and oversight must be ensured through academic and bureaucratic efforts. The “Responsible AI” approach which prioritizes ethical values, security, transparency, and user rights throughout the design, development, implementation, and use of AI must be adopted. Responsible AI must be built in accordance with laws and institutional regulations, as well as ethical principles such as impartiality, transparency, accountability, robustness, and reliability, and must not cause harm to individuals or society (Karaman and Eyüpoğlu, 2025; Martin-Luther-Universität Halle-Wittenberg, 2026). In accordance with this approach, ethics courses should be incorporated into engineering and data science curricula. AI developers, engineers, and academics must possess the competence to recognize the negative societal consequences of the code they write and the AI products they develop (Center for Responsible Digitalization Office [ZEVEDI], 2022; Deutsche UNESCO, 2021; United Nations Educational, Scientific and Cultural Organization [UNESCO], 2022). Furthermore, in critical fields such as education, healthcare, the judiciary, and the workplace, the final decision must always be made by a human expert that is, the principle of human control and oversight must remain sustainable and indispensable (Radanliev, 2025; Syracuse University, 2026; United Nations Educational, Scientific and Cultural Organization [UNESCO], 2022).

On the other hand, while some journals have banned the delegation of the peer review process for an article to AI tools and others have imposed a disclosure requirement, it cannot be completely prevented. Due to the unique nuances of the peer review process, it must be conducted by scientists rather than AI. Due to concerns regarding confidentiality specifically that AI could infringe upon the author’s exclusive rights to the article’s content and potentially lead to unfair evaluations reviewers should not be permitted to upload manuscripts to AI tools for evaluation. Journals must establish clear guidelines on this matter (COPE, 2026). After all, it is never possible for AI to reliably build the same level of insight as humans (Acemoğlu, 2026). While AI can be used as an auxiliary tool in research within the framework of ethical principles, it must not assume a status that replaces humans by depriving them of their fundamental right to work (Çilgin, 2025).

The negative aspects of artificial intelligence such as sycophancy, flattery, and hypocrisy which violate scientific ethics (Chandra et al., 2026; Cheng et al., 2026; Hackenburg et al., 2026), will inflict significant ontological harm on the epistemology of scientific knowledge worldwide if

they replace the objective perspective essential to scientific research. It must not be forgotten that artificial intelligence models are a product of neoliberalizing humanity and science. Nor should it be forgotten that, at its core, artificial intelligence is a commercial product. In today's world, shaped by neoliberalism and populated by narcissistic, egoistic, and self-centered individuals, the fact that artificial intelligence produced by science funded by neoliberal capital operates according to the typologies of these people, viewing them not as researchers but as customers to be retained, stands before us as a serious problem that must be questioned. To mitigate the potential collapse of AI models, the preservation of conventional learning methods and tools based on books and articles, the establishment of national AI systems by countries, and the maintenance of the sustainability of national conventional and AI systems in parallel with global AI institutions should be kept in effect as a precautionary measure and policy against risks such as information poisoning and digital reset.

Ethics: There are no ethical issues with the publication of this manuscript.

Authors' Contributions: The authors contributed to the study equally.

Conflict of Interest: The author declares that he has no conflict of interest.

Financial Disclosure: The authors declared that this study has received no financial support.

Statement on the Use of Artificial Intelligence: Artificial intelligence was not used in the preparation of the article.

Peer-review: Externally peer-reviewed.

REFERENCES

- Acemoğlu, D. (2026). AI is not improving productivity: Nobel laureate Daron Acemoglu. *MIT Sloan Management Review*, <https://sloanreview.mit.edu/audio/ai-is-not-improving-productivity-nobel-laureate-daron-acemoglu/> Accessed on Feb 24, 2026.
- Ackoff, R. L. (1999). *Ackoff's best* (pp. 170–172). John Wiley & Sons.
- Akpınar, Ü. Z., Araç, M. & Doğru, E. (2025). The ethical imperative of artificial intelligence. *Journal of Sustainable Education Studies*, (S4), 51–58. [Turkish]
- Anlı, Ö. F. (2011). Aristoteles ve yöntem: Tümevarım ve tümdengelim. *Bilim ve Ütopya*, 208. 63–69.
- Anthropic, (2026). *Claude's new constitution*. <https://www.anthropic.com/news/claude-new-constitution> Accessed on Jan 24, 2026.
- Aristotle, (2022). *The Nicomachean Ethics*. (F.H. Peters, Trans.), Ryerson University. <https://pressbooks.library.torontomu.ca/nicomacheanethics/> Accessed on May 15, 2026.
- Bains, C. (2024). The legal doctrine that will be key to preventing AI discrimination. *Brookings*. <https://www.brookings.edu/articles/the-legal-doctrine-that-will-be-key-to-preventing-ai-discrimination/> Accessed on Oct 22, 2025
- Baker, K. M. (1964). The early history of the term 'social science'. *Annals of Science*, 20(3), 211-226. [CrossRef]
- Beheshti, A. & Kerridge, I. (2025). Understanding the artificial intelligence revolution and its ethical implications. *Journal of Bioethical Inquiry*, 22(3), 497–505. [CrossRef]
- Boyes, B. (2025). Case studies in the unethical and irresponsible use of AI. *RealKM*. <https://realkm.com/2025/11/05/case-studies-in-the-unethical-and-irresponsible-use-of-ai/> Accessed on Nov 10, 2025.
- Brenneis, A., Gehring, P. & Lamadé, A. (2024). Zwischen fachlichen standards und wilder innovation: Zur begutachtung von big data- und KI-Projekten in forschungsethikkommissionen. *Ethik in der Medizin*, 36, 473–491. [German] [CrossRef]
- Brenndoerfer, M. (2025). Constitutional AI: Principle-Based alignment through self-critique. <https://mbrenndoerfer.com/writing/constitutional-ai-principle-based-alignment-through-self-critique> Accessed on: Nov 17, 2025.
- Cabanac G. (2024). Chain retraction: how to stop bad science propagating through the literature. *Nature*. <https://www.nature.com/articles/d41586-024-02747-1#ref-CR1> Accessed on Sep 09, 2025.
- Cabiri, M. A. (2018). *İbn Haldun'un düşüncesi asabiyet ve devlet* (Çev: M. Çelik). Mana Yayınları.
- Cansen, E. (2025). Yaşamasaydım inanmazdım. yatırım finansman. *YouTube*. <https://www.youtube.com/watch?v=oMobB3xFaEg&t=458s> Accessed on Jan 11, 2026.
- Center for Responsible Digitalization Office [ZEVEDI] (2022). Zur forschungsethischen begutachtung von KI-Forschungsprojekten. Handreichung zur unterstützung der arbeit von ethikkommission. *Technical University of Darmstadt*, https://zevedi.de/wp-content/uploads/2022/11/ZEVEDI_Handreichung-KI-Forschungsethik_2022.pdf Accessed on Oct 16, 2025.
- Chandra, K., Kleiman-Weiner, M., Ragan-Kelley, J. & Tenenbaum J. B. (2026). Sycophantic chatbots cause delusional spiraling, even in, deal bayesians. *Arxiv*. doi: 10.48550/arXiv.2602.19141
- Cheong, I., Caliskan, A., & Kohno, T. (2025). Safeguarding human values: rethinking US law for generative AI's societal impacts. *AI And Ethics*, 5(2), 1433-1459. [CrossRef]

- Chinoracky, R. & Stalmasekova, N. (2025). Ethical problems in the use of artificial intelligence by university educators. *Education Sciences*, 15(10), 1322. [CrossRef]
- Cheng, M., Lee, C., Khadpe, P., Yu, S., Han D. & Jurafsky D. (2026). Sycophantic AI decreases prosocial intentions and promotes dependence. *Science*, 391(6792). [CrossRef]
- Committee on Publication Ethics [COPE], (2025). *Handling an article produced by AI*. <https://publicationethics.org/guidance/case/handling-article-produced-ai> Accessed on Jan 03, 2026.
- Committee on Publication Ethics [COPE], (2023). *Authorship and AI tools*. <https://publicationethics.org/guidance/cope-position/authorship-and-ai-tools> (Accessed on Dec 30, 2025).
- Committee on Publication Ethics [COPE], (2026). *Editors suspect that reviewers are using artificial intelligence*. <https://publicationethics.org/guidance/case/editors-suspect-reviewers-are-using-artificial-intelligence> Accessed on Mar 02, 2026.
- Cunningham, S. (2026). Claude code 27: Research and publishing are now two different things. *Scott's Mixtape Substack*. <https://causalinf.substack.com/p/claude-code-27-research-and-publishing> Accessed on Mar 03, 2026.
- Çilgın, T. (2025). Dialectical analysis of problems created by technological developments in the labour market. *Journal of Social Policy Conferences*, (89), 106–128. [CrossRef]
- Dahlgren Lindström, A., Methnani, L., Krause, L., Ericson, P., de Rituerto de Troya, Í. M., Coelho Mollo, D., & Dobbe, R. (2025). Helpful, harmless, honest? Sociotechnical limits of AI alignment and safety through Reinforcement Learning from Human Feedback. *Ethics and Information Technology*, 27(2), 28. [CrossRef]
- Dale, S. F. (2018). *İbn Haldun ve İnsan Bilimi* (Çev. A. Ay, & C. Coşkun). Say Yayınları.
- Deutsche UNESCO (2021). Der Künstlichen Intelligenz: Weg weiser für die Gestaltung unserer Zukunft, *Deutsche UNESCO-Kommission*, https://www.unesco.de/assets/dokumente/Deutsche_UNESCO-Kommission/02_Publikationen/Publikation_Zusammenfassung_der_UNESCO-Empfehlung_zur_Ethik_der_K%C3%BCnstlichen_Intelligenz_%E2%80%93_Wegweiser_f%C3%BCr_die_Gestaltung_unserer_Zukunft.pdf Accessed on January 05, 2026.
- Deutsche UNESCO (2023). Empfehlung zur ethik künstlicher Intelligenz. *Deutscher Bundestag*, https://www.bundestag.de/resource/blob/942696/993ee22a6d04a87778e06f3c66a49dc8/20-18-107a-Dt_Unesco.pdf Accessed on October 22, 2025. [German]
- Devendorf, J. (2025). Avoiding Algorithmic Bias: Top 5 AI Liability Issues in Courts. *Super Lawyers*, <https://www.superlawyers.com/resources/science-and-technology-law/avoiding-algorithmic-bias-top-5-ai-liability-issues-in-courts/> Accessed on December 05, 2025.
- Digital Cooperation Organization [DCO] (2025). *Rights by Design Embedding Human Rights Principles on AI systems*. <https://dco.org/wp-content/uploads/2025/06/Rights-by-Design-Embedding-Human-Rights-Principles-on-AI-systems.pdf> Accessed on January 08, 2026.
- Doko, E. (2026). Yapay zekâ bilgelik sahibi olabilir mi? Turing Kültür Sanat. *You Tube*. <https://www.youtube.com/watch?v=cqb4P3rJ5q0> Accessed on March 02, 2026.
- Dreyfus, H. L. (1972). *What computers can't do: A critique of artificial reason*. Harper & Row Publishers.
- European Commission [EC], (2024). *AI and Ethics, Human Rights, Law and Educational Data. Briefing* (report No. 6). European Digital Education Hub's squad on artificial intelligence in education. https://www.ai4t.eu/wp-content/uploads/2024/01/AI-squad-output_briefing-report-6.pdf Accessed on December 05, 2025.
- European Commission [EC], (2025). *Living guidelines on the responsible use of generative AI in research, ERA Forum Stakeholders document*. https://research-and-innovation.ec.europa.eu/document/download/2b6cf7e5-36ac-41cb-aab5-0d32050143dc_en?filename=ec_rtd_ai-guidelines.pdf Accessed on December 16, 2025.
- European Parliament [EP], (2020). *The ethics of artificial intelligence: Issues and initiatives, Panel for the Future of Science and Technology*. European Parliamentary Research Service, Scientific Foresight Unit (STOA), [https://www.europarl.europa.eu/RegData/etudes/STUD/2020/634452/EPRS_STU\(2020\)634452_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/STUD/2020/634452/EPRS_STU(2020)634452_EN.pdf) Accessed on October 26, 2025.
- Future of Life Institute, (2026). *The EU Artificial Intelligence Act*. <https://artificialintelligenceact.eu/> Accessed on February 04, 2026.
- Gabriel, M. (2020). Wert und Mensch in der Ethik der. In KI, D. Zimmer (Ed.), *Regulierung für Algorithmen und künstliche Intelligenz* (pp. 25–42). Nomos Verlagsgesellschaft mbH & Co. KG, Baden-Baden. [German] [CrossRef]
- Gadamer, H-G. (1975). *Truth and method*. Continuum. https://web.mit.edu/kaclark/www/gadamer_truth_and_method.pdf Accessed on May 15, 2026.
- Ghosh, M. (2025). Artificial intelligence (AI) and ethical concerns: a review and research agenda. *Cogent Business & Management*, 12(1), 255180. [CrossRef]
- Giddens, A. (2013). *Sosyoloji*, Kırmızı Yayınları. [Turkish]

- Graffius, S. M. (2026). *Are AI hallucinations getting better or worse? We analyzed the data.* ScottGraffius. <https://www.scottgraffius.com/blog/files/ai-hallucinations-2026.html>
- Hackenbun, K., Tappin, B. M., Hewitt, L., Saunders, E., Black, S., Lin, H., Fist C., Margetts, H., Rand D. G. & Summerfield C. (2025). The levers of political persuasion with conversational artificial intelligence. *Science*, 390(6777), eaea3884. [CrossRef]
- Haldun, İ. (2015). *Mukaddime 1-2*, (Çev: S. Uludağ). Dergâh Yayınları.
- Halverson, K., & Cannon, M. (2025). AI in scholarly publishing: Positives and pitfalls. *The ASHA Leader*. <https://leader.pubs.asha.org/doi/10.1044/2025-1001-ai-scholarly-pubs-academia/full/> Accessed on November 07, 2025.
- He, Y. & Bu, Y. (2025). Academic journals' AI policies fail to curb the surge in AI-assisted academic writing. *arXiv*. <https://doi.org/10.48550/arXiv.2512.06705> /
- Horkheimer, M., & Adorno, T. W. (2002). *Dialectic of enlightenment*. Stanford University Press.
- INRA. AI. (2025). AI Hallucinations in Research: What They Are and How to Prevent Them, <https://www.inra.ai/blog/ai-hallucinations> Accessed on November 16, 2025.
- International Committee of Medical Journal Editors [IJMJE], (2026). *Artificial intelligence AI assisted technology. Defining the role of authors and contributors*. <https://www.icmje.org/recommendations/browse/roles-and-responsibilities/defining-the-role-of-authors-and-contributors.html#four> Accessed on March 04, 2026.
- Joseph J. (2025). Algorithmic bias in public health AI: a silent threat to equity in low-resource settings. *Frontiers in Public Health*, 13, 1643180. [CrossRef]
- Kalite Akademik Tercüme, (2026). *Yapay zekâ kullanımı nedeniyle reddedilen ve yayından kaldırılan makaleler artıyor mu?* <https://kaliteakademik.com/yapay-zeka-kullanimi-nedeniyle-reddedilen-ve-yayindan-kaldirilan-makaleler-artiyor-mu.html> Accessed on March 01, 2026.
- Karaman, F. ve Eyüpoğlu, C. (2025). Artificial intelligence ethics in various fields: A review. *İstanbul Ticaret Üniversitesi Teknoloji ve Uygulamalı Bilimler Dergisi*, 8(2), 87–104. [CrossRef]
- Küçükvardar, M., Aslan, A., Bayrakçı, S. (2020). Yapay zekâ ve etik üzerine bir araştırma. *Atlas Journal*, 6(36), 1065–1077. [CrossRef]
- Lawvs. (2026). Legal implications of algorithmic bias in decision-making. *Lawvs.com*, <https://lawvs.com/articles/legal-implications-of-algorithmic-bias-in-decision-making>
- Lee, A. S. (1991). Integrating positivist and interpretive approaches to organizational research. *Organization Science*, 2(4), 342–365. [CrossRef]
- McCarthy, I. P., Hannigan, T. R. and Spicer, A. (2024). The risks of botshit. *Harvard Business Review*, preprint. doi: 10.2139/ssrn.5172718
- Martin-Luther-Universität (2026). KI und ethik. deutschdidaktik. *Halle-Wittenberg* <https://deutschdidaktik.germanistik.uni-halle.de/6-ki-und-ethik/> Accessed on February 26, 2026.
- Merton, R. (1973). The Normative Structure of Science. In N. Storer (Ed.), *The Sociology of Science: Theoretical and Empirical Investigations*. Chicago: University of Chicago Press. <https://sites.temple.edu/stsnetwork/files/2022/01/The-Normative-Structure-of-Science.pdf>
- Mills, C. W. (2022). *Sosyolojik Tahayyül*. (Çev. Ö. Küçük), Hil Yayınları.
- Nagel, T. (1974). What is it like to be a bat? *The Philosophical Review*, 83(4). https://www.sas.upenn.edu/~cavitch/pdf-library/Nagel_Bat.pdf [CrossRef]
- National Institute of Standards and Technology [NIST] (2022). Towards a standard for identifying and managing bias in artificial intelligence. *NIST Special Publication 1270*. [CrossRef]
- Noble, S. U. (2018). *Algorithms of oppression: How search engines reinforce racism*. New York University Press. <https://files.commonsgc.cuny.edu/wp-content/blogs.dir/6105/files/2019/01/SAFIYA-NOBLE.pdf> [CrossRef]
- Nordquist, R. (2019). Holistic Grading (Composition). Glossary of Grammatical and Rhetorical Terms. *ThoughtCo*. <https://www.thoughtco.com/holistic-grading-composition-1690838> Accessed on November 22, 2025.
- Notman, N. (2025). AI tools combat paper mill fraud in scientific publishing as peer review system struggles. *Royal Society of Chemistry*. <https://www.chemistryworld.com/features/ai-tools-tackle-paper-mill-fraud-overwhelming-peer-review/4022253.article> Accessed on October 23, 2025.
- Oellig, V. T. (2024). Künstlich intelligent, ethisch dumm. *Amnesty International*, <https://www.amnesty.de/informieren/amnesty-journal/kuenstliche-intelligenz-menschenrechte-meredith-whittaker-kuenstlich-intelligent-ethisch-dumm> Accessed on October 27, 2025.
- Orr, D. (2022). The unintended consequences of algorithmic bias. *Congressional Black Caucus Foundation [CBCF]*. https://www.cbcbfinc.org/wp-content/uploads/2022/04/2022_CBCF_CPAR_TheUnintendedConsequencesofAlgorithmicBias_Final.pdf Accessed on November 30, 2025.
- Özer M. (2024). Is Artificial Intelligence hallucinating? *Turkish Journal of Psychiatry*, 35(4), 333-335. [CrossRef]
- Pazzanese, C. (2020). Ethical concerns mount as AI takes bigger decision-making role. *The Harvard Gazette*, <https://news.harvard.edu/gazette/story/2020/10/ethical-concerns-mount-as-ai-takes-bigger-decision-making-role/> Accessed on December 19, 2025.

- Pham T. (2025). Ethical and legal considerations in health-care AI: innovation and policy for safe and fair use. *Royal Society Open Science*, 12(5), 241873. [CrossRef]
- Polanyi, K. (1947). Our obsolete market mentality. *Commentary*, 3(2), 109–117. <https://www.scribd.com/document/358169609/POLANYI-Karl-Our-Obsolete-Market-Mentality>
- Polanyi, K. (1977). The economic fallacy. *Review*, 1(1), 9–18.
- Quinn, E. (2025). Lead article: When machines discriminate: The rise of AI bias lawsuits. <https://www.quinnemanuel.com/the-firm/publications/when-machines-discriminate-the-rise-of-ai-bias-lawsuits/> Accessed on January 09, 2026.
- Radanliev, P. (2025). AI ethics: Integrating transparency, fairness, and privacy in AI development. *Applied Artificial Intelligence*, 39(1). [CrossRef]
- Ramachandran, M. (2026). Case studies in AI ethics. In: *Engineering ethics of AI by design*. Springer. [CrossRef] / https://doi.org/10.1007/978-981-95-2909-4_5
- Resnik, D. B. & Hosseini, M. (2025). The ethics of using artificial intelligence in scientific research: New guidance needed for a new tool. *AI and Ethics*, 5(2), 1499-1521. [CrossRef]
- Romano, D. and Gaskins, C. (2025). The reality of AI hallucinations in 2025. *Drainpipe.io*. <https://drainpipe.io/the-reality-of-ai-hallucinations-in-2025/> Accessed on December 29, 2025.
- Searle, J. R. (1980). Minds, brains, and programs. *Behavioral and Brain Sciences*, 3(3),
- Shao, A. (2025a). New sources of inaccuracy? A conceptual framework for studying AI hallucinations. *Misinformation Review*. <https://misinforeview.hks.harvard.edu/article/new-sources-of-inaccuracy-a-conceptual-framework-for-studying-ai-hallucinations/> Accessed on January 23, 2026. [CrossRef]
- Shao, A. (2025b). Beyond misinformation: a conceptual framework for studying AI hallucinations in (Science) communication. *Cornell University*, arXiv, <https://doi.org/10.48550/arXiv.2504.13777> / [CrossRef]
- Spicer A. (2024). Beware the ‘botshit’: why generative AI is such a real and imminent threat to the way we live. *The Guardian*. <https://www.theguardian.com/commentisfree/2024/jan/03/botshit-generative-ai-imminent-threat-democracy> Accessed on January 30, 2026.
- Spinoza, B. (2011). *Etika*. Çev. Hilmi Ziya Ülken. Dost Kitabevi Yayınları.
- Syracuse University (2026). What is responsible AI: Principles, frameworks & future. *Syracuse University School of Information Studies*. <https://ischool.syracuse.edu/what-is-responsible-ai/> Accessed on January 24, 2026.
- The Retraction Watch Database, (2026). *Version: 1.0.8.0. ISSN 2692-465X*. <https://retractiondatabase.org/RetractionSearch.aspx?> Accessed on February 04, 2026.
- UN Global Compact Netzwerk Deutschland (2023). *Künstliche Intelligenz und Menschenrechte*. https://www.globalcompact.de/fileadmin/user_upload/kuenstliche_intelligenz_und_menschenrechte_Publikation.pdf Accessed on November 04, 2025.
- UNESCO (2023). Artificial Intelligence: UNESCO calls on all governments to implement global ethical framework without delay. <https://www.unesco.org/en/articles/artificial-intelligence-unesco-calls-all-governments-implement-global-ethical-framework-without>
- United Nations Educational, Scientific and Cultural Organization [UNESCO]. (2022). Ethics of artificial intelligence-AI. *UNESCO*, <https://www.unesco.org/en/artificial-intelligence/recommendation-ethics> Accessed on November 30, 2025.
- US Bureau of Cyberspace and Digital Policy (2024). Risk management profile for AI and human rights. *US Department of State*. <https://2021-2025.state.gov/risk-management-profile-for-ai-and-human-rights/> Accessed on December 16, 2025.
- US Library of Congress (2026). What is disparate-impact discrimination? *US Congress*. <https://www.congress.gov/crs-product/IF13057> Accessed on March 04, 2026.
- Uplatz, (2025). The evolution of ai alignment: A comprehensive analysis of RLHF and constitutional AI in the pursuit of ethical and scalable systems. *Uplatz Blog*. <https://uplatz.com/blog/the-evolution-of-ai-alignment-a-comprehensive-analysis-of-rlhf-and-constitutional-ai-in-the-pursuit-of-ethical-and-scalable-systems/> Accessed on October 27, 2025.
- Uzaktan Eğitim Kapısı [UEK] (2026). *Dijital okuryazarlık: Giriş. Dijital etik*. <https://uzaktanegitimkapisi.gov.tr/Egitimler/video?Uniq=C617B020-6529-40DC-90CE-74CB5B153426> Accessed on March 01, 2026.