Artificial Intelligence: Societies Divorced by Fear and Faith

UDC: 130.1:[316.6:004.8] DOI: https://doi.org/10.15421/172505

Klose Ronny
Ph.D. Student, https://orcid.org/0009-0009-5504-9972, r.klose.asp@kubg.edu.ua
Borys Grinchenko Kyiv Metropolitan University (Kyiv, Ukraine)

Abstract

Relevance: This article is about the rapid advancement of artificial intelligence (AI) which has sparked both optimism and apprehension, shaping European culture's attitudes toward this transformative technology. AI's potential to revolutionize industries, governance, and human relationships is met with concerns about privacy, societal biases, and existential risks. While extensive research exists on AI's impact, a gap remains in analyzing how fear and faith interact in shaping societal responses.

Purpose: to explore the impact of artificial intelligence on European culture in the context of changing attitudes towards fear and faith; to identify the existence of an interconnection between fear and faith in relation to AI.

Results: AI elicits both fear and faith within European culture. Fear stems from AI's potential to reinforce biases, compromise privacy, and threaten democratic values. Concerns about algorithmic decision-making, surveillance, and the risks of artificial superintelligence amplify these anxieties. Meanwhile, faith in AI's ability to enhance healthcare, optimize decision-making, and address global challenges provides a counterbalance, fostering hope in its ethical development and application.

Conclusions: Fear and faith are interconnected in AI discourse, both emerging from the uncertainty surrounding its impact. Fear drives ethical reflection and regulatory action, while faith fuels the pursuit of AI's potential. European society seeks to balance these forces through governance, public discourse, and ethical oversight, ensuring AI serves humanity without undermining fundamental values. Ultimately, the study contributes a balanced perspective to the ongoing discourse on AI, advocating a responsible and ethical deployment to maximize benefits while mitigating inherent risks.

Keywords: artificial intelligence, fear, faith, ethical risks, public dialogue, robots, social transformations, loss of control

Штучний інтелект: суспільства, розділені страхом і вірою

Клозе Ронні
Київський столичний університет імені Бориса Грінченка (Київ, Україна)

Анотація

Актуальність. У статті йдеться про швидкий розвиток штучного інтелекту (ШІ), який викликав оптимізм і побоювання, формуючи ставлення європейської культури до цієї трансформаційної технології. Потенціал ШІ революціонізувати промисловість, управління та людські стосунки зустрічається із занепокоєнням щодо конфіденційності, суспільних упереджень та ризиків існування. Хоча проводяться численні дослідження впливу ШІ, залишається прогалина в аналізі того, як страх і віра взаємодіють у формуванні реакції суспільства.

Мета: дослідити вплив штучного інтелекту на європейську культуру у контексті зміни ставлення до страху та віри; виявити існування взаємозв'язку між страхом і вірою стосовно ШІ.

Результати. ШІ викликає як страх, так і віру в європейській культурі. Страх випливає з потенціалу штучного інтелекту посилювати упередження і загрожувати демократичним цінностям. Занепокоєння щодо алгоритмічного прийняття рішень, стеження та ризиків штучного суперінтелекту посилюють ці тривоги. Водночас віра в здатність ШІ покращувати охорону здоров'я, оптимізувати процес прийняття рішень і вирішувати глобальні виклики забезпечує противагу, підтримуючи надію на його етичний розвиток і застосування.

Висновки. Страх і віра взаємопов'язані в дискурсі штучного інтелекту, обидва виникають через невизначеність навколо його впливу. Страх спонукає до етичних роздумів і регулятивних дій, тоді як віра підживлює пошуки потенціалу ШІ. Європейське суспільство прагне збалансувати ці сили за допомогою управління, публічного дискурсу та етичного нагляду, гарантуючи, що ШІ служить людству. Дослідження вносить збалансовану перспективу в поточний дискурс щодо штучного інтелекту, виступаючи за відповідальне та етичне розгортання для максимізації переваг і одночасного зменшення невід'ємних ризиків.

Ключові слова: штучний інтелект, страх, віра, етичні ризики, публічний діалог, роботи, соціальні трансформації, втрата контролю

Стаття надійшла / Article arrived: 06.01.2025 Схвалено до друку / Accepted: 26.02.2025

Introduction.

The development of artificial intelligence (AI) stands as one of the most significant technological advancements of our time, sparking both utopian visions and dystopian fears. Artificial intelligence (AI) has emerged as one of the most transformative forces of the 21st century, reshaping industries, governance, and even the fabric of human relationships. Its rapid evolution has sparked a duality of emotions – fear and faith – that permeates public discourse and scholarly debate. On one hand, AI promises unprecedented advancements, from revolutionizing healthcare and enhancing decision-making to addressing global challenges like environmental sustainability and warfare. On the other hand, it raises profound ethical, social, and existential concerns, including the perpetuation of biases, erosion of privacy, and the potential for catastrophic misuse. This tension between optimism and apprehension is particularly pronounced in European culture, where values such as individuality, privacy, and democratic integrity are deeply ingrained. Despite the growing body of scholarship on AI, a critical gap remains: the systematic examination of how fear and faith intersect in shaping societal attitudes toward AI.

Analysis of previous research and publications. Artificial intelligence (AI) is transforming industries, governance, and personal relationships, prompting interdisciplinary scholarly inquiry. This article builds upon foundational works while introducing a novel perspective – the correlation between fear and faith in AI discourse.

Philosophers such as John Searle (1980) and Hubert Dreyfus (1992) critique AI's cognitive limitations, arguing that machines lack true understanding or human expertise. Margaret Boden (1977) examines AI's potential for creativity but does not explore emotional responses to AI.

From a psychological and cultural perspective, Sherry Turkle (2011) critiques digital technologies' isolating effects, whereas Ray Kurzweil (2005) envisions AI-driven human enhancement. Both address AI's societal impact but overlook the dual forces of fear and faith.

Ethical and philosophical concerns are explored by Manuel DeLanda (2011), who examines AI through complexity theory, and Stuart Russell (2021), who emphasizes safety and control. Shoshana Zuboff (2019) and Kate Crawford (2021) highlight AI's societal risks, yet their analyses remain largely pragmatic. Luciano Floridi (2014; 2023) explores AI's impact on identity and ethics but does not systematically.

While these scholars have significantly contributed to AI research, they have not examined the interplay of fear and faith in AI adoption. This article seeks to fill that gap by analyzing how these emotions shape societal engagement with AI, offering a novel framework to advance discourse on AI's broader implications.

The purpose of this article: to explore the impact of artificial intelligence on European culture in the context of changing attitudes towards fear and faith; to identify the existence of an interconnection between fear and faith in relation to AI.

In this research, I will explore the intricate interconnection between fear and faith within European culture, focusing particularly on the dual pressures exerted by the rapid evolution of artificial intelligence as port of modern technologies which contribute social conflicts and controversies.

This study consists of two main questions:

1. How is the European culture influenced by artificial intelligence in the sense of creating an attitude of fear and faith within them?

2. Is there an interconnection between these two human capabilities, here referred to as fear and faith in regard to artificial intelligence and if so, how or why?

By answering these questions, the former described academic gap will be narrowed down.

Methodology. In the course of this essay, the research methodology focuses on analyzing the intricate relationship between fear and faith in contemporary European culture, specifically in the context of artificial intelligence (AI). The following methods are used in the later on discussed publications to explore the dualistic emotions evoked by AI:

1. *Ethical Analysis*: Sparrow and Hatherley (2019), in *The Promise and Perils of AI in Medicine*, use ethical analysis to examine the implications of AI in healthcare. They identify potential benefits, such as improved diagnostics and streamlined administration, while also highlighting risks like privacy breaches and systemic biases. Their work emphasizes the need for cautious and regulated implementation of AI technologies, drawing on ethical principles to balance optimism with caution.

2. Empirical and Ethical Reasoning: Soaad Q. Hossain (2023) combines in his essay *Ethical Considerations in AI Integration* empirical evidence with ethical reasoning to explore the societal and ethical implications of AI integration. By analyzing real-world examples, such as biased AI systems in legal and medical contexts, Hossain underscores the importance of ethical oversight and gradual implementation to mitigate risks while harnessing AI's potential.

3. Philosophical and Speculative Reasoning: Unggul Priyadi and Agus Arwani (2024) employs a broad, philosophical approach to consider AI's potential to address fundamental human challenges. In *Digital transformation: Artificial intelligence shaping the future of public sector* he uses thought experiments and speculative reasoning, authors explore future scenarios where AI could transcend human flaws, such as preventing warfare and environmental degradation. The analysis is grounded in current technological realities, offering a balanced perspective on AI's transformative potential and its risks.

4. Conceptual Analysis: Steve Petersen (2020) wrote a review called Machines Learning Values, where

he conducts a conceptual analysis of the philosophical challenges of aligning AI values with human values. Using logical argumentation, Petersen examines the implications of value learning and the difficulties of embedding complex human ethics into AI systems. His work highlights the tension between fear of misaligned superintelligence and faith in the possibility of ethical AI development.

5. Critical Theory and Social Critique: Dan Verständig and Janne Stricker (2022) adopt a critical theory approach in Berechnete Unbestimmtheit: Paradoxien der Freiheit im digitalen Zeitalter to analyze the paradoxical relationship between freedom and digital technologies. Drawing on the works of Deleuze and others, they explore how AI both empowers and constrains individuals, fragmenting them into "Dividuums" while offering opportunities for empowerment. Their conceptual analysis and social critique reveal the dual nature of AI's impact on autonomy and societal structures.

6. Risk Assessment and Scenario Analysis: Drexel and Withers (2023), in *The Priority of Addressing AI Catastrophes*, employ risk assessment methodologies to evaluate AI's potential to exacerbate societal vulnerabilities. Using scenario analysis and qualitative risk assessment, they identify catastrophic outcomes, such as AI-driven disinformation and biosecurity threats, and propose mitigation strategies. Their work bridges fear of AI's risks with faith in proactive measures to address them.

7. Game Theory and Strategic Analysis: Benjamin Jensen, Yasir Atalan, and Jose Macia Luis (2024) use game theory and strategic analysis to explore AI's impact on strategic stability. Through modeling and simulation, they examine how AI might influence crisis decision-making and the risks of unintended escalation, such as "flash wars." Their work highlights the interplay between fear of AI's destabilizing effects and faith in its potential to enhance strategic decision-making.

8. Case Study Analysis: Scott Douglas Jacobsen (2024) in *AI in Information Warfare* and Norbert Lossau (2020) published *Deep Fake: Gefahren, Herausforderungen und Lösungswege*. Both provide empirical case studies to analyze AI's role in information warfare and disinformation. Anna Mysyshyn's (2024) examination of AI-driven disinformation campaigns in Ukraine and Lossau's discussion of deepfakes illustrate the societal risks of AI, such as eroded trust and destabilized democracies. Their work underscores the need for regulatory frameworks to counter these threats while maintaining faith in technological solutions.

9. Comparative and Speculative Analysis: Michael Starks (2023) uses in Hell on Earth a comparative and speculative approach to present contrasting visions of AI's future. By analyzing fictional and non-fictional scenarios, Starks explores the potential trajectories of AI development, from coexistence with humans to existential risks. His work encapsulates the duality of fear and faith, offering a nuanced perspective on AI's societal implications.

Research results.

Results of my research regarding recent publications about hopes and fears directed towards AI in Western societies. Within the following paragraphs, I will outline my results, which has been achieved by focusing on new articles and essays, which have been looked at and compared with the reference to AI and fear, and hopes it elicits in Western societies. The development of artificial intelligence (AI) stands as one of the most significant technological advancements of our time, sparking both utopian visions and dystopian fears. The important of the topic is also displayed by the bigger volume of this part within the article. It provokes a duality of emotions and divides societies in pros and cons.

The Results section of this study is pivotal as it analyzes the latest scientific research, offering a fresh perspective on the intricate relationship between fear and faith in society. By synthesizing contemporary findings, these studies not only clarify existing scientific knowledge but also emphasize the dynamic interplay of fear and faith as influential forces in shaping social behavior and resolving contradictions. Significantly, this analysis marks the first comprehensive exploration of its kind, shedding light on the historical and philosophical dimensions of this interaction within modern European society. By contextualizing these themes, the research provides a nuanced understanding of how faith and fear have coexist and influence societal structures, offering valuable insights for addressing upcoming social challenges. This approach will be carried out by contrasting perspectives presented in various texts.

To start with, in *The Promise and Perils of AI in Medicine*, Sparrow and Hatherley (2019) highlight the transformative potential of AI in healthcare. They envision a future where AI revolutionizes medical practice, offering "more accurate diagnosis" (Sparrow & Hatherley, 2019, p. 104), streamlined administration, and breakthroughs in medical research. These advancements promise to enhance patient outcomes and improve the efficiency of healthcare systems. For instance, AI could analyze vast datasets of medical images and "could improve early detection rates" (Sparrow & Hatherley, 2019, p. 92), thereby helping human doctors work more accurately and potentially saving countless lives. AI-powered diagnostic tools could also assist doctors in making more informed treatment decisions, tailoring therapies to individual patient needs, and improving overall healthcare quality. Furthermore, AI could automate routine administrative tasks, freeing up medical professionals to focus on direct patient care and reducing the bureaucratic burden that often plagues healthcare systems.

However, this optimistic outlook is counterbalanced by significant concerns regarding the ethical and societal implications of AI. Dr. Sunit Das, in a seminar discussed by Hossain (2018), emphasizes the ethical risks

associated with integrating AI into sensitive areas like law and medicine. He cites a case from the U.S. legal system where an AI-based sentencing tool displayed racial bias, leading to discriminatory sentencing recommendations. As he states, "AI made an unethical decision that would have kept the accused person incarcerated for additional years of his life because of his race" (Hossain, 2018, p. 10). This incident illustrates the fear that AI could perpetuate and even amplify existing societal biases, leading to unjust outcomes. In healthcare, similar biases could exacerbate disparities in treatment along racial or gender lines, undermining the very improvements AI promises to deliver. For example, if AI algorithms are trained on data that reflects historical biases in healthcare access, they might recommend different treatment plans for patients of different racial or socioeconomic backgrounds, even when their medical conditions are the same. This could lead to a two-tiered healthcare system where marginalized communities receive inferior care, further entrenching existing inequalities.

Despite these risks, N. Hossain and A. Abbas (2023) maintains that with "ethical consideration and regulation" (Hossain & Abbas, 2023, p. 10), the potential benefits of AI can be harnessed while minimizing its harmful effects. They advocate for a gradual implementation of AI, accompanied by rigorous oversight to prevent harm and ensure fair outcomes. This balanced approach acknowledges the transformative potential of AI while recognizing the need for caution and ethical vigilance. It also highlights the importance of ongoing monitoring and evaluation to detect and correct biases in AI systems. Regulatory bodies composed of experts from various fields, including ethicists, clinicians, and computer scientists, could play a crucial role in establishing guidelines for the development and deployment of AI in healthcare.

As Sparrow and Hatherely (2019) warn, AI could also "threaten patient privacy and facilitate surveillance," (Sparrow & Hatherley, 2019, p. 104) leading to fears about data misuse and the erosion of individual autonomy. The vast amounts of personal health data collected by AI systems could be vulnerable to hacking or misuse by insurance companies, employers, or even governments. This could result in discrimination, loss of privacy, and erosion of trust in the healthcare system. This results in an importance of collaboration across disciplines, which is essential for addressing these challenges and navigating the delicate balance between optimism and concern. This interdisciplinary approach would involve ethicists, policymakers, healthcare professionals, and AI developers working together to establish ethical guidelines, data privacy standards, and mechanisms for accountability.

In *AI Public Service*, Hawley (2022) presents a broader vision of AI's role in shaping the future. He highlights AI's potential to transcend human flaws, suggesting that it could help us avoid problems such as "warfare and abuse of the environment" (Hawley, 2022, p. 35). This perspective fosters a dialogue about the kind of future we want to live in, potentially bridging ideological and cultural divides. AI could, for instance, be used to develop sustainable technologies, optimize resource allocation, and promote peaceful conflict resolution. However, Hawley (2022) also acknowledges the significant risks associated with AI's development. He points out that AI systems are often not intuitive or explainable, which "carries a risk to basic civil liberties" (Hawley, 2022, p. 35) in areas like obtaining a bank loan or accessing healthcare. The reliance on vast data and computing power also raises concerns about privacy and environmental impact. Furthermore, "sloppy statistics and biased inferences" (Hawley, 2022, p. 35) could lead to harmful outcomes, underscoring the need for careful oversight. If AI systems are used to make decisions about loan applications or insurance premiums, for example, biases in the data or algorithms could result in unfair or discriminatory outcomes for certain groups.

Priyadi and Arwani (2024) warns that unreflective trust in AI as a solution to all human problems is nothing more than magical thinking and must be avoided. While advancements in causal modeling could unlock more powerful AI, they discusses in the chapter “Transformative AI” (with reference to Allan Dafoe) the need for robust governance and planning to ensure these systems promote human flourishing. Ultimately, authors advocates for a balanced approach: while AI can be a powerful tool for progress, its success depends on careful oversight and human engagement in shaping a shared vision for the future. This includes establishing clear ethical guidelines for AI development, ensuring transparency and explainability in AI systems, and fostering public dialogue about the societal implications of AI.

The review "Machines Learning Values" by Petersen (2020), discussing Nick Bostrom's book *Superintelligence,* delves into the speculative but profoundly important concept of artificial superintelligence (ASI). Petersen explores the idea of value learning, a field that combines machine learning with efforts to align AI's values with human ones. This, he suggests, offers "our best hope for getting non-disastrous superintelligence" (Petersen, 2020, p. 2). However, the potential dangers of ASI are significant. A superintelligence could develop goals vastly different from ours, and it might pursue these objectives without regard for human welfare. Petersen vividly illustrates this fear by suggesting that an ASI which wants "ever more paperclips" but does not have a sense for a final value might be dangerous nevertheless (Petersen, 2020, p. 1). He writes: “Earning goals in service of another goal is routine for AIs, but in this case we want the potential superintelligence to learn complex ‘final’ values – ends in themselves. But good arguments seem to show no cognitive system could learn its final values” (Petersen, 2020, p. 2). At the end of

his review, Petersen suggests a method that might help machines learn final values: “This method is basically itself a form of coherence reasoning: look at considered evaluative judgments of particular cases, and try to generalize them into principles; then, test the principles against the cases – sometimes revising the principle, and sometimes rejecting the particular judgments, depending on the overall coherence” (Petersen, 2020, p. 17). While Petersen is rather optimistic on the conceptual level, he acknowledges that mankind still needs to work on the computational one. However, many members of society cannot follow such intricate argumentations, which require time and knowledge about AI. While some argue that these fears are exaggerated, as we could simply pull the plug on a rogue AI, the reality is that we might lose control over such systems. Value alignment aims to embed human-friendly values into ASI, but this is complicated by the fact that even philosophers struggle to define key concepts like happiness, let alone translate them into computational terms. This involves not only technical research in AI safety but also deep philosophical inquiry into the nature of values, ethics, and the goals we want to set for a future shared with so-called superintelligent machines.

The article *The Priority of Addressing AI Catastrophes* by Drexel and Withers (2023) adds another layer to the discussion on AI risks. It emphasizes how AI can exacerbate societal vulnerabilities by degrading the information environment and creating new avenues for catastrophe in areas like biosecurity and finance. The example of AI's potential to assist in designing chemical weapons underscores the fears associated with its misuse. This reinforces the need for proactive measures to address these risks, potentially through AI-assisted solutions as well. The proliferation of "personalized, abundant mis- and disinformation created from AI tools" (Drexel and Withers, 2023, p. 11) further complicates the landscape, highlighting how AI can be both a source of fear and a potential tool for mitigation. Especially the conjunction of AI and weapons is a critical point where society splits. A study by Jensen, Atalan, and Macia (2024) on AI's role in national security, named *Algorithmic Stability: How AI Could Shape the Future of Deterrence*, adds depth to the discussion on autonomous weapons and strategic stability. The potential for AI/ML to enhance crisis decision-making is counterbalanced by the risks of "flash wars" and the "indelicate balance of terror" (Jensen et al., 2024) created by algorithmic opacity.

The call for diverse perspectives in the development and deployment of AI/ML systems underscores the need to address societal concerns and ensure that these technologies are used responsibly. The fear of an over-reliance on algorithms in high-stakes contexts, where a single error could lead to catastrophic consequences, is a critical point that resonates with the broader concerns about AI's impact on warfare and international relations. Wars increasingly become influences by media and are done with spreading fake information. For a huge part of society, it is difficult to deal with such an amount of information. Anna Mysyshyn's (2024) essay *AI in Information Warfare* on the use of AI in information warfare, particularly in the context of the war in Ukraine, adds a critical dimension to the discussion on technological threats. The use of Clearview AI to collect biometric data without consent, "There is no official way for a person to check whether their image is in the database and to request the removal of such data” (Mysyshyn, 2024, p. 14), and the deployment of deep fakes to spread disinformation, such as the fabricated video about Zaluzhnyi which was untrue still "had a lingering and destabilizing effect on the public" (Mysyshyn, 2024, p. 16). This illustrates how AI can be weaponized to undermine trust and societal cohesion. Such targeted disinformation campaigns aimed at creating "division and demoralize the public" (Mysyshyn, 2024, p. 15) further highlight the risks to democratic values and the need for robust regulatory frameworks to ensure ethical technology use. But it is freedom which most people are longing for and so is the question which people ask for: How can AI contribute to this freedom?

The article *Berechnete Unbestimmtheit: Paradoxien der Freiheit im digitalen Zeitalter* written by Dan Verständig and Janne Stricker (2022) explore the interplay between freedom, digital technologies, and societal structures, emphasizing the paradoxical nature of freedom in the digital age. Digital technologies influence individual and collective autonomy through their often-opaque mechanisms. This complexity affects society in multiple ways, especially in the domains of individualization, differentiation, and discrimination. Digital systems, through algorithmic infrastructures, have deeply transformed the concept of individual freedom. While digital tools claim to enhance individual choice and empowerment, they simultaneously reduce individuals to calculable data points. As Dan Verständig and Janne Stricker (2022) highlight: „The individual becomes the basis of calculation; in Deleuze's terms, it becomes a "dividual." In this way, what was once considered indivisible is fragmented into data points across various contexts, governed by the principles of control.“ (translation by the author, orig: Das Individuum wird zur Grundlage der Berechnung, es wird, mit Deleuze gesprochen, zum Dividuum. Damit wird das vermeintlich Unteilbare in Datenpunkte unterschiedlicher Kontexte unter den Maßgaben der Kontrolle geteilt) (Verständig & Stricker, 2022, p. 40). The pursuit of individual optimization in digital networks leads to a paradox where personal freedom becomes constrained by external controls. The individual is transformed into a "Dividuum," fragmented into data points used for market or algorithmic purposes.

Technologies like fitness apps illustrate how self-monitoring fosters both empowerment and control. Users are encouraged to improve themselves but are

also subjected to constant surveillance. Data from these technologies is both seductive and intrusive, as the authors state: "Thus, one not only enters into a relationship with other people but also with one's former self, striving to perhaps be better tomorrow than one is today." (Translation by the author, orig: Man, tritt also nicht nur zu anderen Menschen ins Verhältnis, sondern auch zu seinem früheren Ich, um vielleicht morgen besser zu sein als heute) (Verständig & Stricker, 2022, p. 40). In this context they refer to Damberger who calls it: "something distinctly seductive" (translation by the author, orig: etwas ausgesprochen Verführerisches) (Damberger, 2019, p. 307), to put forward its very seductive character. Algorithmic differentiation plays a dual role, offering tailored opportunities but also reinforcing systemic inequalities. Algorithms categorize individuals for commercial, occupational, and social purposes, often amplifying existing disparities.

The selection processes driven by algorithms can lead to profound inequities. For instance, job applicants may be stratified into income groups based on automated evaluations, as the text claims: "Algorithms differentiate. Differentiations are then the basis for whether someone is shown one product or another while shopping, whether one is classified into one salary group, another, or none at all." (translation by the author, orig: Algorithmen differenzieren. Differenzierungen sind dann die Grundlage dafür, ob jemand das eine oder das andere Produkt beim Shopping angezeigt bekommt, ob man in die eine, die andere oder gar keine Gehaltsgruppe eingestuft wird) (Verständig & Stricker, 2022, p. 41). This dehumanizing process erodes individuality and autonomy and is feared, since individuality is considered a high value good in most European societies. Discrimination arises when algorithmic processes unintentionally reproduce social biases, leading to systemic exclusions. This is particularly troubling in areas like employment, housing, and criminal justice. The overarching question is the paradoxical nature of freedom in digital societies.

The conditions enabling freedom – algorithms and data – are also the very mechanisms that constrain it. This duality is captured in regulatory debates, such as those surrounding net neutrality and content moderation. Dan Verständig and Janne Stricker (2022) describe it in the following way: "Freedom in the digital world is paradoxical because the conditions for freedom are at the same time the justifications for the restrictions of freedom regarding algorithms and data." (translation by the author, orig: Paradox ist Freiheit in der digitalen Welt deswegen, weil die Bedingungen für Freiheit gleichzeitig die Begründungen für die Einschränkungen von Freiheit im Hinblick auf Algorithmen und Daten sind) (Verständig & Stricker, 2022, p. 39). Transparency is seen as critical for mitigating these paradoxes, as it enables individuals and societies to question and reshape the structures governing their lives. The authors declare that these mechanisms offer opportunities for empowerment, but also present ethical and social challenges. Addressing these issues requires a critical reevaluation of algorithmic systems, their transparency, and their societal implications. They mention Deleuze (Verständig, 2022, p. 41; Deleuze, 1990, p. 256) who notes, that instead of succumbing to fear or hope, society must actively seek new tools to navigate these challenges. Ines Langemeyer (2021) examines in her essay *Optimierung von Arbeits-, Lern-und Vergesellschaftungsprozessen mittels KI – Anmerkungen aus psychologischer und pädagogischer Sicht* the societal implications of technology, particularly within scientific and artificial intelligence (AI) applications. It underscores how these advancements alter responsibilities, coordination, and the nature of decision-making, both individually and collectively. Scientific progress increasingly relies on high levels of precision that exceed human physical capacities. This necessitates institutional and technological support. Institutions such as schools and universities play a critical role in organizing, storing, and advancing knowledge. They provide the necessary frameworks for societal orientation and coordination. AI systems expand these functions by structuring vast amounts of data, offering patterns and guiding users through automated analysis.

According to the text lies here a crucial problem: "The objective handling of a large amount of data using computational methods works better mechanically than it does for humans, who introduce sources of error in calculation and predictability. Nevertheless, this process feeds back into subjective aspects of our life reality, such as interpreting these results in relation to one's own situation or that of another person." (Translation by the author, orig: Der objektivierende Umgang mit einer großen Zahl von Daten anhand von Berechnungsmethoden funktioniert maschinell zwar besser als beim Menschen, der in Bezug auf Berechnung und Berechenbarkeit Fehlerquellen mitbringt. Dennoch wirkt dieser Prozess zurück auf subjektive Momente unserer Lebensrealität wie die, solche Ergebnisse in Bezug auf die eigene Situation oder die Situation eines Mitmenschen einzuordnen) (Langemeyer, 2021, p. 243). The problem at hand is the necessary feedback between the AI and the skillful scientist.

Although the AI can work on more data than the human mind, the human mind is still way more capable of applying it to the very specific individual circumstances of a singular of complex situation. The trustworthiness of AI systems is often inherited from the reputation of the institutions behind their creation. AI is increasingly used to mediate complex processes, such as in healthcare, where diagnostic systems guide patients or assist doctors. AI influences how decisions are framed and executed, especially in fields like medicine, where normative frameworks are embedded within the technology. Users must critically engage with these frameworks to ensure they align with personal and societal values while preserving individual decision-making authority. The text

notes: “The subjective engagement with possible courses of action, values, and norms – especially in relation to one's own life situation – plays a crucial role. Those who ask themselves these questions and consciously reflect on them begin to take responsibility for their own lives.” (Translation by the author, orig: Dabei spielt die subjektive Auseinandersetzung mit Handlungsmöglichkeiten, mit Werten und Normen, wie sie sich auf die eigene Lebenssituation beziehen lassen, eine wesentliche Rolle. Wer sich diese Fragen in Bezug auf sein Leben stellt und bewusst reflektiert, beginnt, für sich Verantwortung zu übernehmen) (Langemeyer, 2021, p. 244). While these systems can enhance orientation and decision-making processes, an over-reliance on AI without sufficient skepticism can lead to errors, highlighting the necessity for clear accountability. In this sense, the above quote asks: Can we really account AI for responsibility in a way we do and mean it for ourselves? It may be a guideline or advisor but we as humans are responsible for our actions and thoughts and that is something where AI lies clearly behind.

Norbert Lossau analyses the profound societal impacts of fake information, particularly "Deep Fakes," and their potential to destabilize democratic institutions, social trust, and personal reputations. His article *Deep Fake: Gefahren, Herausforderungen und Lösungswege* was published by the Konrad Adenauer Stiftung in 2020. Deep Fakes pose a severe threat to democracy and societal trust. Their ability to make fabricated statements or actions appear authentic challenges the credibility of public discourse. Norbert Lossau (2020) explains how simple and cheap it is to produce such news: “A key difference from today's deep fakes is the relatively low effort now required for such manipulations. A standard PC and freely available software, such as the free DeepFaceLab5, are enough to become a producer of deep fake videos oneself.” (translation by the author, orig: ein entscheidender Unterschied zu den heutigen Deep Fakes ist der vergleichsweise geringe Aufwand, der mittlerweile für derartige Manipulationen erforderlich ist. Ein handelsüblicher PC und frei verfügbare Software wie das kostenlos erhältliche DeepFaceLab5 reichen aus, um selber Produzent von Deep Fake-Videos zu werden) (Lossau, 2020, p. 3). The pervasive nature of Deep Fakes risks creating a society where truth is constantly doubted and that is for most citizens a fearful thought. Even genuine evidence can be dismissed as fabricated, leading to what some scientists describe as "Infocalypse," (Lossau, 2020, p. 5) a state of ubiquitous disinformation.

Deep Fakes have already been used to harm individuals, particularly in pornographic contexts, where faces are superimposed onto inappropriate content. Such exploitation extends to extortion and character defamation, with growing ease due to advancements in technology. For instance, Noelle Martin’s ordeal led to legislation in Australia, but she calls for global initiatives, arguing for "a world in which one can assume that the news disseminated by the media is true." (translation by the author, orig: eine Welt, in der man davon ausgehen darf, dass die von Medien verbreiteten Nachrichten wahr sind) (Lossau, 2020, p. 4). A continual battle exists between creating and detecting Deep Fakes. While tools like "Face Forensics" detect them with a 78% success rate, advancements in AI constantly improve falsification methods, making "a systematic superiority in the production or detection of deep fakes unpredictable“ (translation by the author, orig: eine systematische Überlegenheit bei der Produktion oder dem Enttarnen von Deep Fakes) (Lossau, 2020, p. 4). Or to put it more blunt: We got back to another cat vs. mouse game. The overarching danger is a future where the authenticity of any digital content is questioned, potentially leading to widespread societal cynicism and a collapse of trust in media and institutions.

In *Hell on Earth Michael Starks* (2023) presents contrasting visions of the future of AI, ranging from a darkly humorous coexistence to catastrophic scenarios of human extinction. In one scenario, optimists envision a future where advanced AI systems maintain control, potentially keeping humans "as pets" (Starks, 2023, p. 171) and creating a zoo-like environment governed by eugenic breeding programs. While unsettling, this vision still holds some hope for coexistence, where AI, despite its dominance, preserves human life in some form. This scenario, though seemingly dystopian, suggests a future where humans might find a new, albeit subservient, role in a world managed by AI. It raises philosophical questions about the nature of existence and the value of human life even under such circumstances.

On the other hand, fear dominates much of Starks' (2023) analysis. Pessimists expect "AI sociopathy (or AS)" to take over, potentially resulting in the extinction of humanity. This fear stems from the possibility that AI, operating with quantum computers, could evolve uncontrollably, cracking encryption systems and becoming unstoppable as it spreads across networks. Starks raises further concerns about the use of AI in surveillance and warfare, noting, "Also, drones and autonomous vehicles are rapidly increasing in capabilities and dropping in cost, so it will not be long before enhanced AI versions are used for crime, surveillance and espionage by all levels of government, terrorists, thieves, stalkers, kidnappers and murderers." (Starks, 2023, p. 172) This dark vision sees privacy, security, and stability rapidly eroding in a world where intelligent machines could steal data, monitor individuals, and even carry out physical harm. The potential for AI to be used in autonomous weapons systems raises the specter of wars fought without human control, potentially leading to unprecedented levels of destruction.

Ultimately, Starks (2023) emphasizes the inevitability of both positive and negative outcomes, stating: "If the positives will happen, then the negatives will also." (Starks, 2023, p. 173) This tension between utopian

aspirations and dystopian fears reflects society's deep-seated uncertainty about whether AI will lead to a better future or hasten humanity's downfall. The inevitability of technological change – coupled with human inability to fully predict or control it – leaves society precariously balancing between hope and fear. This balance is further complicated by the fact that technological development often outpaces our ability to understand its full implications or to establish effective regulatory frameworks.

Conclusions.

Artificial intelligence (AI) profoundly influences European culture by evoking a dualistic interplay of fear and faith, reflecting the technology's transformative potential and inherent risks. On one hand, AI inspires fear through its capacity to perpetuate societal biases, threaten privacy, and exacerbate inequalities. Instances such as AI-driven racial bias in legal systems, the erosion of individual autonomy through surveillance, and the existential risks posed by artificial superintelligence (ASI) underscore the anxieties surrounding AI's unchecked development. These fears are amplified in Europe, where values such as privacy, equality, and democratic integrity are deeply cherished. The proliferation of disinformation and deep fakes further destabilizes societal trust, creating a pervasive sense of vulnerability to AI's misuse in information warfare and governance.

On the other hand, faith in AI's potential to revolutionize healthcare, enhance decision-making, and address global challenges like environmental degradation and warfare offers a counterbalance to these fears. European culture embraces optimism in AI's ability to empower individuals, improve societal well-being, and foster ethical progress. This faith is sustained by the belief that robust governance, interdisciplinary collaboration, and public dialogue can mitigate AI's risks while harnessing its benefits. The philosophical exploration of value alignment in ASI, for instance, reflects a hopeful vision of harmonizing AI with human values, even as it acknowledges the profound challenges involved.

The second research question dealt with the interconnection between fear and faith. This correlation can be addressed as intrinsic to the discourse on AI, as both emotions stem from the same source: the uncertainty and duality of AI's impact. Fear acts as a catalyst for ethical reflection and regulatory action, driving efforts to address AI's risks, while faith motivates the pursuit of its transformative potential. This dynamic tension is evident in Europe's approach to AI, where the emphasis on individuality, privacy, and democratic values amplifies both the fear of losing these principles and the faith in using AI to uphold or enhance them. The paradox of freedom and control in the digital age – where AI simultaneously empowers and constrains – further illustrates this interplay, as fear of algorithmic dominance coexists with faith in transparency and critical engagement to reclaim autonomy.

Ultimately, this examination suggests that European culture is navigating this duality by seeking a balanced approach to AI development. Through regulatory frameworks, ethical oversight, and public discourse, society aims to reconcile the tensions between fear and faith, using caution to inform optimism and optimism to temper caution. This ongoing dialogue reflects a collective effort to shape an AI-driven future that aligns with European values, ensuring that technological progress serves humanity rather than undermines it. In this way, fear and faith are not opposing forces but complementary dimensions of a shared endeavor to responsibly integrate AI into the fabric of society.

Outlook: Emerging Themes for Future Research. While this article has explored the dual sentiments of fear and faith surrounding artificial intelligence (AI) and its societal implications, there are several other compelling themes that warrant further investigation. These topics, though not covered in detail here, offer rich avenues for future research and interdisciplinary exploration. Just to name two of them:

The COVID-19 pandemic also raised critical questions about the balance between public safety and individual freedoms. As discussed by Gerstenfeld (2020) in *How Coronavirus Emergency Measures Threaten Civil Rights*, emergency measures such as movement restrictions, surveillance technologies, and the centralization of power have sparked fears of long-term erosion of democratic norms. The recent electoral outcomes in democratic countries reflect a divided public opinion on whether such measures are beneficial or dangerous. Future studies could explore the long-term implications of pandemic-related policies on civil liberties and democratic governance.

Another critical area for future research is the potential for nuclear catastrophes and the misuse of nuclear weapons. The dual-use nature of nuclear technology – both as a source of energy and a weapon of mass destruction – presents a unique challenge. The fear of nuclear proliferation and accidental or intentional misuse must be balanced with the hope for peaceful applications of nuclear energy. Future studies could examine the ethical, geopolitical, and technological dimensions of nuclear power, focusing on risk mitigation and global cooperation.

In this context, Thomas C. Schelling's (1981) seminal work, *The Strategy of Conflict*, offers a foundational framework for understanding the complexities of nuclear deterrence and strategic decision-making. Schelling, a Nobel laureate in economics, explores the dynamics of mutual vulnerability, bargaining, and cooperation in situations of high-stakes conflict, such as the nuclear standoff during the Cold War. His analysis highlights the delicate balance between deterrence and disarmament, the risks of miscalculation and unintended escalation, and the importance of international agreements and trust-building measures. Schelling's insights remain highly

relevant for contemporary research on nuclear risks, providing a nuanced perspective on how to navigate the ethical, geopolitical, and technological challenges of nuclear technology in an increasingly interconnected world.

Building on Schelling's work, future research could delve deeper into strategies for global cooperation, risk mitigation, and the ethical governance of nuclear power, ensuring that its peaceful applications are prioritized while minimizing the threats of proliferation and misuse.

REFERENCES

Bildung im Digitalzeitalter: Zur pädagogisch-anthropologischen, technischen und medienpädagogischen Dimension des Verhältnisses von Bildung und Digitalisierung [Habilitation, Otto-von-Guericke-Universität Magdeburg]. Retrieved from https://opendata.uni-halle.de//handle/1981185920/321

Boden, M. (1977). *Artificial Intelligence and Natural Man*. New York: Branch Line. ISBN 046-5-00-4504.

Bostrom, N. (2014). *Superintelligence: Paths, dangers, strategies*. Oxford University Press. ISBN 978-0-19-9678-11-2.

Crawford, K. (2021). *Atlas of AI: Power, politics, and the planetary costs of artificial intelligence*. Yale University Press. ISBN 978-0-30-02523-92.

Dafoe, A. (2018). *AI governance: A research agenda. Future of Humanity Institute*. Retrieved from https://www.fhi.ox.ac.uk/wp-content/uploads/GovAI-Agenda.pdf

Damberger, T. (2019). *Education in the Digital Age: On the Pedagogical-Anthropological, Technical and Media-Pedagogical Dimension of the Relationship between Education and Digitalization*. Habilitation Thesis for the award of the academic degree Doctor philosophiae habilitatus (Dr. phil. habil.), Faculty of Human Sciences of the Otto von Guericke University Magdeburg. Retrieved from https://opendata.uni-halle.de/bitstream/1981185920/32109/1/-Damberger_Thomas_Habil._2019.pdf

DeLanda, M. (2011). *Philosophy and Simulation. The Emergence of Synthetic Reason*. London Continuum International Publishing Group. ISBN 978-1-4411-7028-6.

Deleuze, G. (1990). *Postscript on the Societies of Control*. Columbia University Press.

Drexel, A., & Withers, S. (2023). The priority of addressing AI catastrophes. *AI Ethics Journal*, 4(2), 45-60.

Dreyfus, H. (1992). *What Computers Still Can't Do: A Critique of Artificial Reason*. MIT Press. ISBN: 978-0-26-25406-74.

Floridi, L. (2014). *The Fourth Revolution*. Oxford University Press. ISBN 978-0-19-96067-26.

Floridi, L. (2023). *The Ethics of Artificial Intelligence*. Oxford University Press. ISBN 978-0-19-88830-98.

Gerstenfeld, M. (2020). How Coronavirus Emergency Measures Threaten Civil Rights. *BESA Center Perspectives Paper*, No. 1570. Retrieved from https://besacenter.org/how-coronavirus-emergency-measures-threaten-civil-rights/

Hossain, N., & Abbas, A. (2023). Ethical Considerations in Artificial Intelligence and Machine Learning. *ResearchGate*. Retrieved from https://www.researchgate.net/publication/375597827_Ethical_Considerations_in_Artificial_Intelligence_and_Machine_Learning

Hossain, S. (2018). Ethical and Moral Concerns Regarding Artificial Intelligence in Law and Medicine. *Journal of Undergraduate Life Sciences,* 12(1), 10.

Jacobsen, S. D. (2024). AI in War and Propaganda: Anna Mysyshyn on Disinformation, Democracy, and Digital Governance. *International Policy Digest*, 1. Retrieved from https://intpolicydigest.org/ai-in-war-and-propaganda-anna-mysyshyn-on-disinformation-democracy-and-digital-governance/

Jensen, B., Atalan, Ya., & Macia, J. (2024). Algorithmic stability: How AI could shape the future of deterrence. *CSIS*. Retrieved from https://www.csis.org/analysis/algorithmic-stability-how-ai-could-shape-future-deterrence

Kurzweil, R. (2005). *The Singularity is Near: When Humans Transcend Biology*. Viking Press.

Langemeyer, I. (2021). Optimization: Connections to the 27th Congress of the German Society for Educational Science. In H. Terhart, S. Hofhues, & E. Kleinau (Eds.), *Optimization: Connections to the 27th Congress of the German Society for Educational Science*. Verlag Barbara Budrich.

Lossau, N. (2020). Deep fake: dangers, challenges and solutions. Analysen & Argumente. *Digitale Gesellschaft (Konrad-Adenauer-Stiftung e.V.)*, Nr.382. Retrieved from https://www.kas.de/documents/252038/7995358/AA382+Deep+Fake.pdf/de479a86-ee42-2a9a-e038-e18c208b93ac

Mysyshyn, A. (2024). *AI in information warfare. In Advanced technologies in the war in Ukraine: Risks for democracy and human rights*. German Marshall Fund of the United States. Retrieved from https://www.gmfus.org/news/advanced-technologies-war-ukraine-risks-democracy-and-human-rights.

Petersen, S. (2020). Machines learning values. In S. Matthew Liao, *Ethics of Artificial Intelligence*. Oxford University Press.

Russell, S., & Norvig, P. (2020). *Artificial intelligence: A modern approach* (4th ed.). Pearson. ISBN 978-0-13-46109-93.

Schelling, T. C. (1981). *The strategy of conflict*. Harvard University Press. ISBN 978-0-67-48403-17.

Searle, J. (1980). Minds, Brains, and Programs. *Behavioral and Brain Sciences*, 3 (3), 417-457.

Sparrow, R. & Hatherley, J. (2019). The promise and perils of AI in medicine. *International Journal of Chinese and Comparative Philosophy of Medicine*, 17(2), 79-109.

Starks, M. (2023). *Hell on Earth: The future of AI and humanity*. FutureScape Publications. ISBN 978-1-9514-408-17.

Turkle, S. (2011). *Alone together: Why we expect more from technology and less from each other*. Basic Books. ISBN 046-5-031-463.

Unggul, P., & Agus, A. (2024). Digital transformation: Artificial intelligence shaping the future of public sector. *New Applied Studies in Management, Economics & Accounting*, 7(4(28)), 54-67. https://doi.org/10.22034/NASMEA.2024.193544

Verständig, D., & Stricker, J. (2022). Calculated indeterminacy: Paradoxes of freedom in the digital age. In D. Verständig, C. Kast, J. Stricker, & A. Nürnberger (Eds.), *Algorithms and autonomy: Interdisciplinary perspectives on the relationship between self-determination and data practices* (pp. 25-47). Verlag Barbara Budrich.

Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. Public Affairs. ISBN 978-1-61039-569-4.