

The Impact of Artificial Intelligence on Modern Society

Pedro Ramos Brandao ^{1,2} ¹ Instituto Superior de Tecnologias Avancadas, 1750-142 Lisbon, Portugal; pedro.brandao@istec.pt² CIDHEUS U., 7004-516 Evora, Portugal

Abstract

In recent years, artificial intelligence (AI) has emerged as a transformative force across various sectors of modern society, reshaping economic landscapes, social interactions, and ethical considerations. This paper explores the multifaceted impact of AI, analyzing its implications for employment, privacy, and decision-making processes. By synthesizing recent research and case studies, we investigate the dual nature of AI as both a catalyst for innovation and a source of potential disruption. The findings highlight the necessity for proactive governance and ethical frameworks to mitigate risks associated with AI deployment while maximizing its benefits. Ultimately, this paper aims to provide a comprehensive understanding of how AI is redefining human experiences and societal norms, encouraging further discourse on the sustainable integration of these technologies in everyday life.

Keywords: artificial intelligence; modern society; economic impact; ethical considerations; employment; privacy; decision-making; innovation; governance

1. Introduction

Artificial intelligence (AI) paradigms are generally classified into Symbolic AI and Connectionist/Statistical AI. Symbolic AI, also known as Good Old-Fashioned AI (GOFAI), is based on formal logic and rule-based systems and was dominant during the early period of AI research (1950s–1980s). It depends on human-created ontologies and deterministic reasoning, exemplified by systems like SHRDLU and expert systems. Conversely, Connectionist AI, inspired by biological neural networks, supports today's machine learning methods, particularly deep learning. This approach uses data-driven statistical inference and learning from experience. Machine learning (ML), often mistaken for AI itself, is a subfield that focuses on algorithms enabling computers to learn from data without explicit programming. A clear definition is: "Machine learning is a field of study that gives computers the ability to learn without being explicitly programmed." [1] and it is further categorized into supervised, unsupervised, and reinforcement learning. These paradigms differ fundamentally in their epistemological basis and computational structure, which has significant implications for their use across various domains.

This article is presented as a narrative review rather than a systematic review. Its goal is to critically synthesize and integrate findings from scholarly literature, policy reports, and industry surveys on the development and impact of artificial intelligence (AI). By using a narrative approach, the paper provides interpretive analysis and cross-sectoral insights instead of an exhaustive listing of all available studies.

To improve clarity and focus, the review is organized into thematic sections covering the historical development of AI, current sectoral applications (healthcare, manufacturing, finance, education, governance, and transportation), socio-economic impacts (employment,



Academic Editor: Rafał Dreżewski

Received: 19 May 2025

Revised: 6 July 2025

Accepted: 4 August 2025

Published: 17 August 2025

Citation: Brandao, P.R. The Impact of Artificial Intelligence on Modern Society. *AI* **2025**, *6*, 190. <https://doi.org/10.3390/ai6080190>

Copyright: © 2025 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

productivity, and inequality), ethical and legal issues, and future research and policy directions. Each section summarizes representative evidence and viewpoints to provide context within an academic framework.

The material was identified through a targeted search of academic databases (e.g., Scopus, Web of Science, PubMed) and grey literature (such as think tank reports and corporate white papers) from 2015 to July 2025. Search terms included “artificial intelligence,” “machine learning,” “economic impact,” “ethical implications,” and sector-specific keywords like “healthcare,” “finance,” “education,” and “governance.” Only sources in English that provided empirical data or substantive analysis were included. Although this process is not systematic, it helps ensure that key studies and recent trends are captured to support the discussion.

At the start of the 21st century, societies experienced advances in informatics, computer networks, and artificial intelligence. There is hope that machines will soon perform tasks currently performed by humans. This raises questions about whether that moment has arrived and if it truly marks a shift in production, consumption, and economic structures. Is society preparing for changes in forms of work and, as a result, in life? Will the elite have more efficient tools for both production and control? Will it benefit society with new products and services that are independent of natural resources? Or will it widen the gap between the haves and have-nots, creating a new robotic elite? More discussions on the impact of artificial intelligence and machine learning focus on different communities’ perspectives and the hopes, concerns, and fears they express. Recently, there has been growing attention to the economic effects of new technologies.

Concerns about the future grow as AI becomes more capable. There is broad agreement that AI will influence employment and welfare. Significant job losses are predicted. Various scenarios regarding economic and social impacts are expected. Cheaper, better, and more advanced mechanical minds will reduce demand for the human brain. The exact nature of this impact remains debated. Some scholars believe that creating a general AI capable of self-improvement and accessing the jobs and skills of human predecessors is possible. One AI might have a clear advantage over humans, and the chosen general AI could control its versions, weapons, toy robots, and even kill switches like a dictator. Society would have to deal with an economy of radical abundance where no one needs to work. This could lead to a depleted economy, and universal basic income policies might weaken the state’s financial resources. Policymakers should focus on distributing wealth, fixing market flaws, and keeping checks on humanity.

2. Historical Context of Artificial Intelligence

Concern about the growing influence of machines has been part of the public awareness longer than many realize [1]. Inspired by events like the publication of Darwin’s theory of evolution by natural selection, discussions about self-reproducing and even evolving machines started in the 1860s. A few scientists, especially William Thomson (Lord Kelvin), quickly saw the implications of such machines for humanity. Reactions to this idea captured the attention of a broader audience. They drew in well-known literary figures, including Samuel Butler, Mark Twain, and Edward Bellamy, from the late 1800s up to roughly the mid-1900s.

This paper explores themes like the displacement of labor and resulting social upheaval, social evolution and the broader role of science, and the potential for a species capable of truly creative thought, as presented by these authors. The history of how such machines have influenced society has recently become a lively topic of debate [2]. Notable technology figures have founded companies with astonishing valuations, and numerous AIs now outperform humans at tasks once thought too complex for non-humans. However, current debates about the impact of AIs on social structures, moral frameworks, and even

humanity often focus on current issues, drawing from recent decades of development within a much larger context of concern.

Critically, such debates suffer from narrow vision. There is much to gain by looking further back. In particular, the possibility of devices capable of evolution through natural selection was first raised in the last half of the 19th century. That, in turn, opened the door to a realized future of machines possessing creative intelligence, compared to which the development of current AIs is but a dim shadow.

Despite the promise of AI across sectors, practical deployment often faces significant challenges such as data silos, inadequate infrastructure, and workforce readiness gaps. A McKinsey (2023) survey <https://www.mckinsey.com.br/capabilities/strategy-and-corporate-finance/our-insights/global-economics-intelligence-executive-summary-june-2023> (accessed on 3 June 2025) revealed that only 15% of companies successfully scaled AI beyond pilot projects because of these cross-sector barriers. These include issues like legacy system integration, talent shortages, and regulatory compliance burdens, which vary considerably across healthcare, finance, and manufacturing industries.

3. Current Applications for AI

It appears that AI is quickly becoming a part of everyday tasks. People have observed this in action with AI-enabled personal assistants like Siri, Cortana, and Alexa. New AI-powered solutions have also emerged to support individuals in fields such as human resources, technical support, finance, data analysis, fact-checking, programming, translation, writing, and even art creation. Additionally, organizations worldwide are utilizing these AI capabilities to increase productivity. Seventy-four percent of companies using AI say it has given them a competitive edge. This figure rises to 79 percent among companies that are leaders in AI deployment [3].

Many companies have created AI that interprets data and makes predictions. Popular examples include IBM's Watson, which analyzes large datasets in healthcare, and Salesforce's Einstein, which uses information from CRM systems to help businesses identify leads from social media and emails. These algorithms are separate from the behaviors they model. As a result, AI can predict and proactively suggest actions. For instance, if a person enters numbers into a mortgage calculator, AI can recognize this, assume they might be shopping for a house, and start displaying ads from multiple mortgage providers.

Many companies have started using AI to automate administrative and engineering design tasks. For example, legal AI algorithms can find relevant legal cases, draft contracts from templates, and review the language for issues. Financial AI detects misuse of credit or bank accounts. Some design algorithms can suggest features for new products based on sales data and customer feedback. They can even create a new product design with fewer specifications [4]. With these superhuman abilities, AI is expected to surpass much of what humans in companies do. This has sparked both excitement and concern. AI is seen as bringing a new technological revolution that will improve lives but also potentially threaten jobs.

3.1. Healthcare

Faced with major changes in medicine and society, artificial intelligence (AI) can potentially improve the accuracy and efficiency of ongoing patient care [5]. By leveraging big data, AI may offer personalized patient care and enhance treatment strategies based on individual health history and available resources [6]. Its ability to quickly analyze large amounts of clinical data and generate predictions allows for faster, more widespread diagnostics and decision-making. Recently, the rise of deep learning has led to successful AI systems that analyze data and deliver highly accurate findings. These systems have been

effectively used to detect breast cancer malignancies and diabetic retinopathy in imaging tasks, as well as to develop automated diagnostic prediction tools.

Currently, the most commercially successful AI-powered decision support system is Aidoc, which improves the efficiency of CT and MRI services through real-time triage of anomalies. There are already programs capable of matching radiologists' expertise by interpreting various scans of organs like the lungs and colon. Several meta-analyses evaluating the overall performance of AI models have shown their potential to outperform human experts, and some studies have highlighted the need for validation research using datasets from diverse sources. In addition to image analysis tasks, ongoing efforts aim to develop AI models to analyze electronic medical records for predicting disease onset and progression.

Because of the increased connectivity in medical supply chains and the digitization of medical devices, the security threat landscape for healthcare organizations is growing in both scope and complexity. With major advancements in industry and technology, healthcare organizations are experiencing a merging of IT and security challenges, requiring a new approach to compliance and cybersecurity as old security models start to break down. AI-powered autonomous systems, or more broadly, solutions that use machine learning, can improve incident detection and response efficiency and help develop new defensive and offensive strategies. However, these also bring a new set of challenges for cybersecurity and compliance, demanding a re-evaluation of models and processes as more AI systems are implemented.

3.2. Finance

The finance sector was among the first to adopt AI technology and has developed into a mature and successful field for AI. AI is widely used in trading, investment advice, asset management, risk management, insurance claim processing, fraud detection, and regulatory compliance. AI is changing people's lives, boosting the economy, and transforming business processes. AI technologies have many uses in the financial services industry, and some specific applications will be discussed in detail in this section [7].

The rapid growth of data, increased computing power, and advancements in AI algorithms drive AI in portfolio management. Financial data have inherent characteristics that offer significant opportunities for using AI in asset management and investment advisory services, including robo-advisors [8]. Portfolio management consists of three parts: asset selection, portfolio construction, and portfolio monitoring/rebalancing. In recent years, considerable efforts have been made to apply machine learning and deep learning techniques to asset selection. Reinforcement learning algorithms are currently being tested to create optimal portfolios. Regarding monitoring and rebalancing, methods from traditional time-series analysis continue to be the preferred approaches. Overall, the finance sector has many opportunities to utilize new technologies, especially AI methods, to transform traditional processes and develop disruptive business models.

The key applications in finance operate at both micro and macro levels. In micro-level analysis, a stock selection model is introduced that combines traditional financial data with alternative text data from financial news. The model uses a BERT-based framework to extract news sentiments along with traditional financial features. Classification and component attention models are developed to accurately predict stock price trends. The proposed model shows balanced performance in both accuracy-based and portfolio-based metrics. In macro-level analysis, a measure of systemic risk in the financial sector is created based on asset correlations of financial institutions and the distribution of key systemic risk factors. Then, a deep learning-based approach is presented to capture complex and nonlinear relationships among financial institutions. This approach outperforms traditional models in predicting systemic risk from both cross-sectional and time-series perspectives.

3.3. Transportation

Cars may appear similar to those from decades ago on the outside, but a dramatic revolution has taken place inside the cabin thanks to advances in computation, communication, and storage. With trillions of dollars in revenue, the hidden part of the value chain produces margins of only a few percent, sometimes just a fraction of a percent. As a result, investments in new solutions often have low returns, keeping the automotive industry traditional and cautious about adopting new technology. The auto industry has been quite successful, but it may face challenges in its second century. The largest automakers have added only a few dozen new vehicle models to their fleets, but this incremental innovation offers little value to customers and businesses—except for automakers who have announced environmental benefits of twilight powertrains and development plans. Otherwise, vehicles with similar propulsion, capacity, speed, weight, and colors, made from metals, rubber, glass, and plastics, have created a paradox of relief and boredom. Vehicles with so-called interconnection usually connect only a few devices and apps, and infotainment experiences are limited to streaming services and information access. The coming revolutions are often heavily marketed, but automotive manufacturers dismiss their potential.

Nonetheless, the upcoming widespread, practical, and affordable electric autonomous vehicles will revolutionize today's automotive industry, as billions of dollars spent on hardware and software development generate numerous video, sound, information, and virtual experiences that enable travel to autonomous drop-off points. Personal vehicles will be used less and stay idle 98% of the time, making today's automotive industry, urban planners, energy supply chains, and vehicle components obsolete. When charged, these vehicles will carry crowds, semantic visual content, and health data, stop for advertising along the route, and travel far outside cities to rest and recharge in inexpensive cloud parking lots. The key turning points of this shift have led experts to believe that it is only a matter of time before electric autonomous vehicles challenge and eventually replace traditional EVs. However, the electric autonomous vehicle revolution will also need new regulations, standards, and infrastructure, and many of these solutions are still in the conceptual stage and at least a decade away from deployment.

3.4. Education

Artificial intelligence (AI) is a rapidly expanding technological field that aims to improve and transform our daily lives. It is used across nearly every industry, including cybersecurity, computer vision, and healthcare. The rise of AI and machine learning (ML) is happening at an exponential pace around us. More organizations are adopting AI and ML technologies. Young people should understand the basics of AI and ML to succeed in STEM fields. AI education can promote this understanding, creating a culture of capability and potential career paths. However, for many students worldwide, AI education is currently limited or unavailable. Therefore, developing an AI curriculum for elementary, middle, and high schools is crucial [9]. Educational institutions can align with these trends and participate in ongoing discussions about AI's ethics, impact, and future in society. AI education should deepen students' understanding of AI and its implications, equipping them to be informed citizens in an AI-driven world [10]. This requires including a basic AI curriculum within the educational system. Teaching AI starts with introducing definitions and applications to students. Next, students engage in data preparation activities for several weeks using traditional data mining techniques. Finally, AI projects are integrated, allowing students to create their own AI applications in a game-like setting to solve problems. Education focused on a fundamental, hands-on understanding of AI is likely to motivate students to engage with AI technologies in the future. When implementing AI education, it is important to consider what content to teach, how to teach

it, and when. Teaching AI is seen as a necessary goal in K-12 education. Schools and colleges are well-placed to introduce this knowledge. Beyond computer science courses and increased efforts in teacher and curriculum development, AI education should be transdisciplinary. It should be organized around timeless educational principles rather than just current tools. A deliberate, transdisciplinary approach connects subjects through major guiding questions, enabling students to see the links between disciplines and their real-world applications, fostering a deeper understanding of complex issues.

3.5. Manufacturing

Manufacturing is among the industries that have seen progress in artificial intelligence (AI) technology. However, advanced technologies have been limited for small and medium-sized manufacturers (SMMs), despite their significant market share, workforce, and production capacity. First, large manufacturers have been the main users of such advanced technologies and AI solutions, while SMMs confront limited resources [11]. Affordable, easy-to-install, and user-friendly AI solutions for machines are still unavailable. Moreover, the most critical and scarce resources for SMMs are the insights that AI solutions offer and the ability to obtain these insights independently at their factories. Usually, Existing advanced AI solutions are provided by cloud service providers. When implemented without proper understanding, AI solutions in factories are like a well-equipped kitchen without an operator, with analytics and learning tasks left to the mysterious black box of cloud services. Affordable AI solutions for machines will be ineffective if they are not relevant, and they will be irrelevant if they are not affordable. Manufacturers' concerns about assessing technology investments hinder the adoption of advanced technologies. In-line or online testing for a small number of items is feasible and common among SMMs. However, the traditional setup cannot accurately test machine performance, which varies continuously across many items. Instead, it creates a potential obstacle that slows the widespread use of machine perception technology.

Given the drivers and challenges, a new class of affordable artificial intelligence-assisted machine supervision (AIMS) is proposed which models the blueprint of standard machine workflows, in addition to training the comprehension models for input videos and output tags. This integrated solution represents a paradigm shift. It is affordable, standalone, and configurable for nearly all machines. AIMS can continuously monitor the working state of machines through real-time observation and can detect anomalies by analyzing the machine workflow and identifying deviations from normal operation [12]. The machine workflow consists of states where deviations indicate abnormalities. Therefore, comprehensive modeling of possible normal states is essential. It identifies the normal state of the target machine by analyzing a compressed representation of machine images. This representation provides the context for understanding the machine's standard workflow by analyzing the spatial and temporal dependencies across its components (Table 1).

Table 1. AI pros and cons across sectors.

	Sector	Advantages	Disadvantages
1	Healthcare	Faster diagnostics, tailored treatments, improved patient outcomes	Algorithmic bias, data privacy risks, and misdiagnosis potential
2	Finance	Fraud detection, risk analysis, and algorithmic trading efficiency	Model opacity, systemic risk amplification, and ethical concerns
3	Education	Personalized learning, administrative automation, and increased access	Equity of access, teacher deskilling, and algorithmic evaluation bias

Table 1. Cont.

	Sector	Advantages	Disadvantages
4	Manufacturing	Process automation, predictive maintenance, and cost efficiency	Job displacement, high implementation costs, and quality assurance
5	Governance	Enhanced decision-making, resource optimization, and service delivery	Transparency gaps, accountability issues, surveillance overreach

Empirical evidence from [13] shows that automation-related job loss is most significant in sectors relying on routine, codifiable tasks. In contrast, jobs focused on creativity, empathy, or strategic judgment are less at risk. The OECD (2021) https://www.oecd.org/en/publications/what-happened-to-jobs-at-high-risk-of-automation_10bc97f4-en.html (accessed on 15 June 2025) estimates that 14% of jobs are highly automatable, while 32% will experience major changes. These differences highlight the need for proactive policy measures tailored to at-risk occupational groups.

4. Societal Impacts of AI

The new wave of AI systems is expected to impact various areas of society. However, the societal impacts of AI become particularly urgent with the rapid and widespread deployment of such systems. AI can potentially cause errors and unintended consequences on a much larger scale than non-AI systems. Additionally, AI enables new use cases that put greater pressure on democratic institutions and the social fabric, including social media algorithms and the increased reliance on AI in areas like finance. Given the interplay between technology, humans, organizations, and institutions, the societal context is crucial for assessing the influence of AI on society. It impacts how AI is utilized and integrated, how people respond to it, and the ethical issues that arise in ongoing societal debates, highlighting the urgent need to address the societal impacts of AI. This urgency is amplified by the emergence of new AI systems, such as large language models, as they are deployed in sensitive areas and reach a larger number of individuals. AI systems that allow users to create images or texts and engage in conversations with them have been launched. In response, social networks, health authorities, and educational institutions have begun to ban or restrict the use of generative AI. However, the broader potential societal implications and possible policy responses deserve more attention (Figure 1).

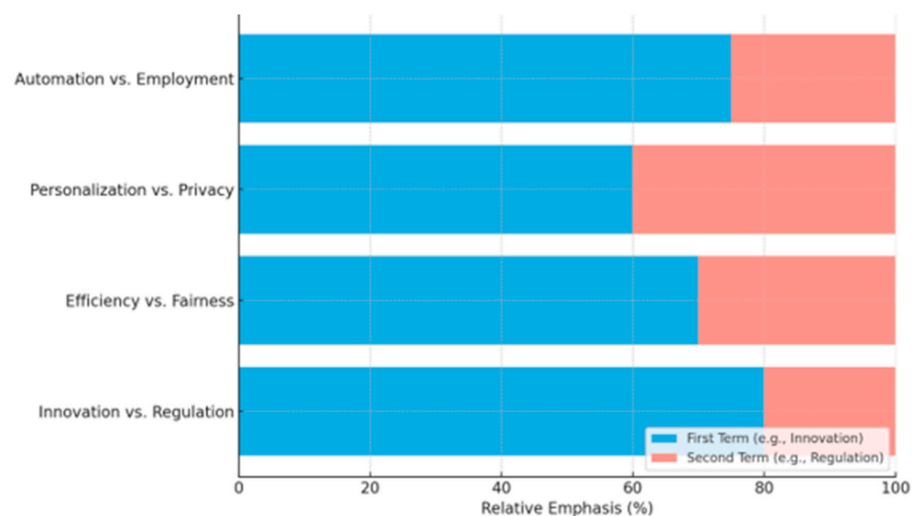


Figure 1. Conceptual trade-offs in AI's societal impact.

AI has the potential to bring about fundamental changes to humanity. For example, many decisions in a person's life can already be delegated to technology. As AI systems improve, more complex tasks may be handled by these systems. The risk is that much human judgment could be handed over to AI agents, leading humans to follow the guidance of smart agents without question. Independent human judgment allows for the reflection of values, contexts, and long-term perspectives. Conversely, AI systems tend to focus on short-term goals, often overlooking second-order effects like spirals and backlash. There are concerns about whether AI will lead to reduced social interactions, such as that observed due to online dating apps, which have been criticized for promoting social isolation instead of expanding social networks. Besides using AI in one-on-one scenarios, increased human-to-machine interactions will also occur in other areas, such as with healthcare robots or service robots [13].

4.1. Job Displacement

With significant progress in artificial intelligence (AI) in recent years, the possibility of worker displacement due to AI advancements has reignited interest in how technological change affects the job market [14]. Overall, predictions about the speed and scale of displacement caused by AI have been very pessimistic. As AIs become more capable and widespread, this concern grows, and social anxieties are likely to increase.

This research aims to determine whether the increase in AI capabilities in one country influences employment rates. To evaluate variation in AI across different jobs, AI job exposure was measured using the text of job ads on online job boards from 2012 to 2019. In nearly all the occupations examined, job numbers increased. There was no clear link between AI exposure and job growth. AI adoption may not cause significant displacement effects in terms of jobs lost. However, the widespread use of the internet and computers is expected to boost productivity across economies. This could feed back into labor markets, offsetting direct substitution effects.

In recent years, AI has transitioned from science fiction into a practical business tool [15]. ChatGPT, an AI chatbot, is now commercially available and can integrate text, edit briefs, and create presentations. The debate over how quickly generative and other AI applications will impact work continues to be intense. AI possesses unique capabilities; it can explain concepts, simulate behaviors, and accurately extract information, answer any question, and produce material in any medium. AI has the potential to greatly improve the skills of professional workers.

Until recently, spending time on a smart machine was limited to specific cases and brief periods. It was indeed an unusual experience, requiring skills that not all users possessed. Much has changed rapidly since the launch of the first mobile phones that competed for people's attention and curiosity [16]. The pursuit of the latest smart phone, now a democratized object, has naturally led to the idea that "smart" spending has become a commodity. And "AI," based on a different perspective than those that existed before, is undoubtedly the best label for it. The once-clear vision of a smart machine, which was associated with full automation and replacing human operators for certain mental, imitative, or even creative tasks, seems to have shifted. Starting with captivating advertising and impressive capabilities, it is now questioned whether this was a spectral vision (with no real added value) or, worse, a social and economic illusion (with negatives). Traditional economic agents, their interactions, and the mechanisms guiding the economy operated like a carefully choreographed ballet, moving rhythmically with time. The economy was large and distant from individuals and sectors. Since the new technologies have emerged, change happens almost instantly, and the baton has passed to a less sophisticated ensemble of bio-inspired agents, smart machines (designers, producers, and providers), where time can be ignored. This transition has moved us away from an economy rooted in

rationality, perfect foresight, and cynicism towards a race economy filled with homogeneity, emotion, impulse, and sentimentality. A concept called “disintermediation” refers to a user viewing the bargaining power of masses as nearly infinite to a seller who focuses on individuals feeling exclusive and personally cared for. In this new smart economy—much smarter than Web 2.0 of the 2000s—a clear distinction exists between “smart” machines and “stupid” machines. The former enhance user cognition by providing information and fostering emotional and social connections; the latter transform resources and include physical mechanical machines (producing goods and infrastructure), telecommunication and transportation devices (delivering services), and algorithms and interface machines (collecting and analyzing data) [17].

4.2. Privacy Concerns

Artificial intelligence (AI) has the potential to improve the design and deployment of many intelligent technological systems. AI is used in technology-assisted care settings for tasks ranging from data management to safety assurance. The UK has begun a systematic, data-driven transformation process, with growing emphasis on the importance of AI in this effort. However, AI can raise ethical issues that existing frameworks struggle to analyze easily. When AI is trusted with decisions once made by humans, individuals may lose a clear basis for understanding or challenging those decisions. Even with moral agents guiding these systems, moral responsibility may be shared among multiple parties that cannot be pinpointed to one person. AI can operate based on opaque data inputs or systems that provide no transparency to humans. Furthermore, AI systems might develop independently in ways their creators cannot predict. Additionally, designers may use data to train algorithms that embed value-laden premises, potentially undermining current moral standards [18].

AI raises moral issues with distinct ethical dilemmas that require thorough and multi-faceted analysis. Scholars and institutions working in machine learning, natural language processing, and the broader AI field should participate in discussions about the ethical impacts of rapid technological change. Part of this duty is to clearly communicate both the positive and negative aspects of their work. On the positive side, machine learning and natural language processing hold great potential to improve the world. At their best, these technologies can enhance students’ academic experiences, empower marginalized communities, and help people facing various daily challenges [19].

4.3. Bias and Discrimination

Concerns about artificial intelligence are heightened due to its rapid adoption across different parts of society. Views on its use are nuanced, ranging from fears that it will eliminate jobs and cause widespread unemployment to strong calls for more automated decision-making systems that can avoid human errors. Because of AI’s societal importance, we show how it can be both a tool for reducing inequalities and a factor that worsens them, while also highlighting ongoing research on the topic. AI’s broad use is also thought to reveal societal inequalities that were previously hidden beyond their mathematical data representations.

With the rapid spread of artificial intelligence (AI), concerns are increasing about its potential to worsen existing biases and societal disparities, and in some cases, introduce new ones. These issues have gained widespread attention among academic researchers, policymakers, industry leaders, and civil society. Evidence indicates that incorporating human perspectives can help address bias in AI systems; however, evaluating these early efforts is essential to ensure they promote fairness without unintended outcomes [20]. Designing crowd work systems, including those used for data collection in AI training, screening, and label evaluation, is complex. Efforts focused on data collection and task

design for requesters which involve workers have been critically examined. Nonetheless, human involvement in data input, such as annotating and labeling datasets, can lead to unforeseen consequences of the kind mentioned earlier. AI has been shown to diagnose heart disease more accurately than trained doctors and predict housing prices better than appraisers. In the age of data, machine learning embeddings have revealed insights about society's upper and lower classes that purely mathematical models have not uncovered.

Machine learning systems trained to read, listen to, view, and evaluate data created by humans mirror the biases present in that data [21]. Human-generated data contain indicators of armed conflicts and wars throughout history, criminal hotspots and behaviors, health disparities, and key sentinel events relevant to modeling social processes. However, they also expose machine-based discrimination and hate speech against races, locations, or belief systems. Essentially, bias in modeling society's edge cases can threaten the progress of social justice and the fight for equal rights, which have been struggled for over centuries and remain only partially achieved worldwide. Individuals naturally navigate a web of biases, many of which an AI trained on their output can recognize more quickly and accurately than they can.

Comparative studies reveal different national AI strategies. For example, the European Union's AI Act focuses on ethical compliance and human oversight, while the U.S. takes a sector-by-sector, innovation-focused approach. China, on the other hand, incorporates AI into centralized planning and public surveillance. These diverse approaches underscore the need to harmonize AI policies to promote global interoperability, accountability, and protection of human rights.

5. AI in Governance and Policy Making

AI systems have the potential to improve governance quality. However, the expected impact of algorithmic decision-making (ADM) in the public sector depends on the context. Specifically, the expected improvements in policy outcomes can be hindered by poor data quality, incomplete algorithmic specifications, access to sensitive inputs, or unregulated outputs with no right to explanation [22]. At the same time, political issues such as the digital divide, biases, or racial profiling may prevent AI from enhancing political representation. Overall, these factors tend to diminish trust in public institutions, lowering society's acceptance of AI. This understanding offers a more nuanced view of AI's role in governance discussions. AI could be seen as beneficial when it improves governance quality. However, it can also be criticized if it damages governance quality, thereby reducing widespread support.

In general, algorithm-set policies are probably more effective than human-set ones when many qualitative variables influence outcomes. They are also likely more dependable than human-produced policies when the input space is very large. Finally, caution should be taken in environments where there is no contest, especially when there is little evidence of algorithmic misbehavior. However, even in these situations, increased accountability or transparency could enhance the influence of AI outputs compared to humans, ultimately improving governance performance [23]. Understanding the political implications of algorithmic governance also helps in assessing the involved political structures. AI and related technologies can have unintended effects on the organizational, institutional, social, and economic facets of governance. Relying solely on AI's technical design for governance could also cause immediate negative political repercussions. To prevent AI from worsening existing political tensions or creating new ones, it must align with clearly defined, legally established political objectives. Otherwise, it risks limiting political action or leading to unchecked, self-referential governance structures. A high-dimensional input state space supports modeling complex systems. However, in such decision-making environments,

poor AI specification can cause damaging governance outcomes. These systems can also harm political representation.

5.1. Regulatory Challenges

The use of technology, especially artificial intelligence, is rapidly growing in various fields, including education. Powerful algorithms can help improve or better support educational processes, reorganizational tasks, and forecasting in schools. While many of these systems can significantly enhance education, some are poorly designed or exploited. In recent years, this has led to concern over artificial intelligence algorithms used for grading and assigning scores or remedial tasks to institutions. Local and global reactions against national algorithms designed for the general high school leaving examination equivalent and against school grading systems, recommendation letters, or tutoring grading systems have surprised many who see this blatant subjectivity and lack of understanding of such practices as a dilemma. On the other hand, many forecasting algorithms aimed at predicting the learning speed of online students, primarily relying on computer video detection or social activity networks, have raised privacy concerns. Whether education-machine systems are serving the impact of individuals and society worldwide, whether positive or negative, continues to be studied based on data that is constantly fed into these systems and the careful training of the models. This article's evaluation examines when and why this might be desirable or beneficial and explores potential future developments. Starting with current machine learning systems, how can these systems be designed to genuinely serve society, especially in education, while acting with knowledge and purpose? Here, some fundamental values, properties, or concepts to consider are listed. Primarily, the human mind and any actions, learning, and thoughts related to it should remain explainable to ensure societal validity and transparency, while avoiding unintended shifts toward robotic behavior and the feared risks of nearly perfect mind prediction. Malicious cyber interventions or controlling entities would likely seek to remove manual control over the machines. Ideally, these systems would unlock new knowledge or improve performance, as well as honesty—factual or perceived—in information delivery. Conversely, educational and information delivery systems are necessary and inherently universal to society. However, they could be designed in ways that are conspiratorial in nature or dishonest in their descriptions, content, and formats used to present information.

5.2. Ethical Considerations

There is growing awareness that developing and implementing artificial intelligence (AI) technologies in organizations must address ethical issues. The increasing adoption of AI applications across various sectors is partly based on the idea that these technologies reduce human costs while boosting productivity, customer service, and competitiveness in a global economy [24]. However, these advances have sparked ethical concerns about their effects on labor relations, autonomy, dignity, privacy, equality, and, more broadly, democracy.

AI-based applications are tools created and used by individuals to reach specific goals. However, it is naive to think that simply introducing an AI product into a sector will bring well-being to society. There is a risk that AI tools, if misused by government agencies for mass surveillance or by unethical companies to exploit consumer data, could worsen discrimination, inequality, and the erosion of civil rights. AI is not a harmless technology; rather, it is a political technology that can be designed and used to either support societal well-being or undermine it [18]. Thinkers from a technocritical viewpoint have long warned about the dangers of unregulated technological development that could threaten societal values, human potential, and social justice.

This article does not oppose the adoption of AI in organizations; instead, it recognizes its potential to benefit both organizations and society. No technology is inherently good or bad; it can be used for positive or negative purposes, depending on whose interests it serves and how it is designed and applied. In this sense, it encourages exploring ways in which AI can be utilized to improve organizations and communities while aligning with humane and social values and expectations. Rather than predetermined outcomes, the future is shaped by socio-political processes involving diverse actors. Similarly, with AI, various issues emerge regarding how these technologies should be conceived and implemented to advance humane and social good values.

5.3. International Cooperation

On a global level, as multiple AI systems pursue a common goal, shared standards are crucial to ensure compatibility and interoperability. A unified regulatory framework for AI across various sectors would promote responsible AI worldwide and address the currently fragmented regulatory landscape [25]. The AI Ethics Initiative and its framework for global AI governance highlight the urgent need for proactive regulation of AI. Given AI's significant potential for both value creation and risk, preventing harmful applications will contribute to a better future. Ethics initiatives and standards should be inclusive, involving diverse disciplines and stakeholders to find common ground. These efforts can help avert a moral catastrophe. Increasing international cooperation to support the prescriptiveness of soft law would create conditions for developing and enforcing stricter norms. Soft law includes non-binding guidelines, codes of conduct, standards, and recommendations developed through transnational institutions or informal networks of states [13]. Soft law on AI is gaining momentum. While these initiatives address various AI issues, they often lack concrete means to ensure compliance. Bridging the gap between rhetoric and reality is essential. Because of AI's importance, challenging power asymmetries in global governance can help establish a democratic and participatory AI regime. Support should be de-commodified to enable developing nations and poorer regions to participate effectively in international deliberations and forums. Providing the necessary exchange of knowledge, expertise, human resources, and economic support would allow less developed actors to engage and leverage existing resources. Achieving fair international cooperation, whether on AI or other issues, requires addressing the global power structure. Overcoming populism and critically evaluating stakeholder input can improve inclusive governance strategies. Addressing fragmented global governance and advocating for reform of international regimes and organizations is essential. Technological Advancements Driven by AI The early response to the COVID-19 outbreak showed that AI adoption was advancing faster than in previous tech revolutions. During the pandemic, AI-based solutions were quickly adopted to tackle the outbreak caused by the new coronavirus. This technology provided reasonably accurate recognition of pneumonia cases through lung images from CT scans, pinpointing where COVID-19 first infects and multiplies. InterVision, one of the first developers of AI-based COVID-19 diagnostic software, was founded in July 2017, and its product was used by 34 hospitals in the early stages of the outbreak in China. Over 32,000 suspected cases' lung images were uploaded for analysis. Processing time dropped from around 15 min to just 3 min. Backed by Google Sequoia Capital, this early adoption exemplifies how technology was quickly embraced during the outbreak. The rapid deployment of localized AI technology helped curb the spread of the virus and prevented a potentially longer pandemic that could have ground transportation, tourism, and many businesses to a halt.

Given the increasing reliance on predictive algorithms, it is essential to contextualize ML performance metrics. Accuracy should be complemented with precision, recall, F1-score, and AUC-ROC, depending on the application domain. Standard benchmarks, such as

ImageNet for vision or GLUE for language models, help establish meaningful performance ranges that support replicability and informed decision-making.

6. Machine Learning

On 11 March 2020, it was reported that the InterVision AI diagnosis software was exported to Japan to help medical staff with screening and preventing the spread of the virus. This proactive use of technology was adopted at a crucial moment. After the Chinese government held a public briefing about the breakthrough in localizing CT scan pneumonia indications for COVID-19 at 3:15 p.m. on 22 January 2020, researchers and developers were prompted to act quickly by rapidly improving algorithms for faster processing on a larger scale and using a deep learning model trained with many lung images showing pneumonia cases. The AI technology has proven useful in health risk assessment, identifying potential patients, and enabling quick radiological diagnosis [26].

6.1. Machine Learning and AI

According to the literature, machine learning (ML) is a field of research and development (R&D) in artificial intelligence (AI) that focuses on enabling networks to learn similarly to humans, or it can be defined as an artificial intelligence application that concentrates on recognizing things without explicit coding. ML can be divided into three main types: supervised Learning (SL), unsupervised learning (UL), and reinforcement learning (RL). It is widely used, for example, in social networks and web searches [27].

A supervised learning task involves inputs and outputs as samples. Receiver operating characteristics (ROC) curves can be generated to evaluate how well the network classifies. Other metrics, such as mean squared error and mean squared accuracy, can also be used. In an unsupervised learning task, inputs are provided, and the system must identify internal structures. This type of ML is useful for clustering or determining the value or priority of decisions like movie selections on Netflix or music searches on Spotify. This has significantly increased productivity on social networks. Finally, there is reinforcement learning, in which networks interact with environments or operate in simulations with random states and rewards, even in unknown contexts. It is widely used in professional games and simulations; the networks' sessions reduce the time needed to learn or obtain policies to just a few hours, highlighting the capabilities of machines that can learn. Video games like Atari demonstrate this. The knowledge base of reinforcement learning primarily involves entropic costs and posterior Bayesian methods for establishing the best policy [28].

6.2. Natural Language Processing

Recent advancements in artificial intelligence techniques, especially in natural language processing (NLP) areas like sentiment analysis, named-entity recognition, and topic modeling, have the potential to solve many challenges in educational feedback analysis, drawing significant academic interest. Student feedback data in text form are crucial for identifying the strengths and weaknesses of current services offered to students. In education, analyzing student feedback can reveal areas for improving infrastructure, learning management systems, teaching methods, study environments, and more. Although student textual feedback is becoming more important, it is often overlooked due to a lack of suitable analytical methods. Automated AI techniques are necessary because manual analysis can take weeks, missing out on timely opportunities for improvement. Student feedback may come from surveys, open-ended questions, or other textual formats. This approach is seen as reliable and honest since anonymity reduces bias. However, complaints, grievances, and even sarcasm can appear alongside positive comments and policies about study or teaching environments. NLP tasks and methods for analyzing textual feedback in education must be

carefully chosen and organized. This research reviews and discusses existing NLP methods and applications that can be adapted for education, such as sentiment analysis, entity recognition, text summarization, and topic modeling. A key challenge is context-based issues in NLP. In feedback analysis, different interpretations of the same comment, like sarcasm or speculation, can all be valid. Additionally, opinions about a system often depend on specific aspects, highlighting the need for aspect-based sentiment analysis. Sarcasm, in particular, is a factor that can confuse sentiment classification systems. In education, domain-specific language is common, and general word sense disambiguation often fails with domain-specific ambiguities. While existing systems for sarcasm detection exist, many focus on engineered feature extraction methods. This paper provides a brief overview of domain-specific NLP challenges and background information to improve understanding of these issues [29].

6.3. Computer Vision

Computer vision (CV) is defined as the automatic extraction of useful information from images. The term image here refers not only to raster images but also to the entire solid-angle information, including multiple images taken by various geometric imaging setups created by dedicated components and strategies [30]. Computer vision (CV) is attracting global attention and is a rapidly growing research area with numerous real-world applications. The challenges and their solutions are closely associated with life contexts, and there has not been enough investigation into engineering analysis and archiving to support CV for future analysis. Proof-of-Principle Studies (PoPS) are used to perform vision tasks, and the knowledge gained from these tasks is seen as a valuable asset that should be accumulated. However, within this field, there is limited research on archiving design knowledge for PoPS, including design cases, insights, knowledge structures, and retrieval methods. Vision is a form of cognition that extracts environmental knowledge using planes, particles, or electromagnetic waves. This information serves as a foundation for making decisions about spatial and temporal changes. Advanced artificial intelligence (AI) systems and CV can learn the behaviors involved in this knowledge extraction, such as multi-sensor simulation and knowledge description, enabling intelligent learning. For example, in space monitoring, the information that needs to be rapidly detected and tracked includes critical points and strips, raising the question of whether these points are static or dynamic. A CV system must be designed for monitoring, which is a PoPS and is considered a design case. The collaboration of many heritage experts helps build a historical knowledge base of CVs for design automation, which relates to the question of “Am I being watched?” To the best of our knowledge, this is the first study focusing on knowledge engineering for computer vision.

7. AI and Economy

In recent years, artificial intelligence (AI) has become a widely discussed topic, as several innovations have successfully integrated into daily life, including deep learning, game playing, and autonomous driving, to name a few. While economists have been considering the economic effects of AI since its early days, recent advances in these areas have made them cautious about its future impact [16]. Today’s data on productivity tell an optimistic story. In the early 1980s, productivity growth rates in the world’s leading countries increased gradually. The anticipated IT revolution faced a slow start before really taking off in the mid-1990s. However, evidence shows that GDP growth rates in the US were below 2% around the year 2000, and after the dot-com bubble burst in 2000, productivity growth in many economies slowed, with Europe being the hardest hit. Recently, the entire world has seen a surge in productivity growth. Measures of productivity today are probably the highest they have been since the boom years of 1995–2008 in the US.

The main message of this article is straightforward: it provides evidence of the recent comeback of AI in the G7. There are three key points to consider. First, it is essential to clarify that this article does not intend to be overly pessimistic, unlike some other perspectives. This paper is based on analysis, and after years of discussing AI without clear insights into its long-term effects, such analysis cannot simply be a race to predict what might happen in the future. Second, it is worth noting that AI is fundamentally different from all other general-purpose technologies (GPTs). GPTs were created to supplement labor; they could help create new jobs and balance the labor market (wages might increase in overhyped occupations). AI is the first GPT designed mainly to replace, rather than complement, human work.

7.1. Reskilling and Upskilling

The world is experiencing significant change and transformation. Rapid progress in industrialization and digitalization has driven remarkable advances in next-generation technologies, including artificial intelligence (AI) and machine learning. Knowledge sharing, understanding, and education are more accessible than ever because of hyper-connectivity and information. Additionally, the world has faced a pandemic that caused unprecedented disruption and sped up technological progress, impacting work and life in both positive and negative ways. The second wave of the fourth industrial revolution, often called Industry 4.0, is transforming how services are created and delivered across industries worldwide. With Industry 4.0, there are major changes not only in jobs and skills but also in the competencies and educational qualifications needed for those roles in the 21st century [31].

The vision of advanced manufacturing for Industry 4.0 will come true through the efforts of a future-ready workforce. However, as technology advances, some people struggle to find good jobs due to a lack of the right skills, while others worry that automation threatens low-skilled jobs. Skill gaps will inevitably widen unless today's workers participate in learning experiences to gain the technology-related skills needed for future jobs. Whether individuals are in the labor market or not, scaled digital skill development is essential to empower everyone to become agents of social, economic, and environmental change in tackling global challenges.

The latest Future of Jobs report estimates that by 2025, 85 million jobs might be displaced, while 97 million new jobs could emerge, better aligned with the new division of labor among humans, machines, and algorithms. The top skills expected to grow include analytical thinking and innovation (+84%), active learning and learning strategies (+63%), critical thinking and analysis (+44%), and complex problem-solving skills (+40%). To help learners take ownership of skills related to future jobs, each of these skills should be described in clear, engaging terms and connected to real-life situations.

7.2. Remote Work Trends

Flexible work arrangements were already in demand before COVID-19, but the pandemic forced many businesses to quickly adopt remote work setups within weeks or even days. OECD estimates indicate that a shift to remote work by the end of 2019 would have significantly reduced in-person work, even in an average OECD country, decreasing by nearly 70% for non-essential in-person workers. Total job losses in Oxford's national accounts, combined with re-estimated deaths, were halved during the first pandemic wave. Oxford's exceedance measure from pre-pandemic levels showed a stronger cross-country correlation with Google COVID-19 mobility indexes. Real GDP declined in most countries in early 2020, but the declines in Italy and France were notably greater than in others. By the second wave of deaths in April 2020 mobility decline thresholds in countries were generally below zero. In the eight countries, job losses were also strongly linked to increases

in death rates [14]. AI technologies are now better able to process vast amounts of data and support increasingly sophisticated human decision-making across various contexts. Unlike traditional “automated” or “robotic” peak-load management systems, AI technologies autonomously adapt to changes in the work environment, proactively addressing challenges that could otherwise disrupt or impair performance. Today, AI performs tasks that professionals managed just a few years ago, flagging important information from huge volumes of unstructured data. With natural language processing capabilities, AI agents increasingly automate not only routine tasks typically handled by clerks but also the classification and analysis of complex unstructured texts. For example, Amazon’s legal department uses AI systems to draft documents and contracts without human intervention [15]. The main challenge for organizations now is how best to incorporate AI into their operations. The impact of AI on jobs remains uncertain. AI is already replacing many clerical roles, where the activities are uniform and rule-based; AI inference algorithms are capable of mastering these processes and generating similar outcomes more efficiently, leaving little room for front-line operational positions. However, AI is unlikely to eliminate all professional roles, and it is improbable that AI will reduce the proportion of professional jobs in the workforce over the next few decades. AI is still not advanced enough to fully replace professional roles in knowledge-intensive or creative tasks that require higher-level cognitive skills. In this context, while the world is seeing increased use of AI in early-stage knowledge work affecting back-office jobs and providing semi-automated processes—especially in finance or asset management industries—a significant number of professional roles are unlikely to be eliminated. Displacement remains possible; it could take the form of lower salaries, reduced scope of tasks, or employment with non-AI firms, contributing to the de-professionalization of certain specialties.

7.3. *The Gig Economy*

Digitally mediated gig work, where individual workers offer on-demand services via online platforms, makes up a significant and growing part of the workforce. Reports estimate that about one-third of the U.S. workforce participates in the gig economy, while around one percent works directly on app-based gig platforms. Participation in gig work surged during the COVID-19 pandemic. Efforts to support gig workers have gained attention due to recent high-profile events such as the passage of CA Prop. 22, the repeal of rules by the U.S. Department of Labor, and strikes and protests by gig workers.

A key unique feature of digitally mediated gig work is the widespread involvement of AI. AI systems match drivers with customers and set their pay. The individual goals and preferences of gig workers are often overlooked in existing platforms. The AIs in these platforms lack transparency and many show systematic biases in their algorithms. High technical barriers prevent workers from accessing AI technologies that serve their best interests. Another factor contributing to AI inequality is access to and control over data. Gig platforms have collective data from all drivers, while individual workers can only track their own data. This issue has recently become a major challenge in creating a fairer future for gig work. One practical solution is designing a network of intelligent end-user assistants. Each assistant would be paired with a worker, gather work-related data, and share it within the network. This approach could help workers optimize their work based on their personal goals and preferences. This research could help better understand AI inequality in gig work and explore worker needs and strategies for effective human–AI collaboration. These findings could raise awareness of AI inequality and provide evidence for labor advocacy and policy efforts [17].

Freelancing platforms have become a primary venue for gig work. In 2022, there were 3 billion site visits to Upwork, with 400 400 000 gig jobs posted each month. Other

popular platforms include Fiverr, Freelancer, and Guru. Workers on these platforms are not employees; instead, the platforms act as matchmakers, connecting workers with clients. Like other digital platforms, profiles function as resumes for gig workers. However, unlike traditional employees, each worker creates their own business profile to attract clients through self-promotion [32].

8. Cultural Change and Implications of AI

Numerous cultural changes are likely to occur in society as AI becomes more widespread and deeply integrated into daily life. For example, humans will need to adapt to the presence of AI, including its role in family life. Currently, robotic assistants designed for households are being developed. Additionally, future AI advancements may enable robots to have personalities. Teaching robots to understand human emotions would be crucial for making a positive difference in people's lives, as robots capable of experiencing traits beyond simple algorithms could foster mutual relationships with their users. However, legislation must be enacted to prevent misuse, since robots could have the power to manipulate people's feelings. When someone can influence an individual easily, it presents a serious threat. The very notion of conscious existence might also be questioned. These are just a few examples of cultural shifts that are highly likely to arise due to the pervasive presence of AI.

The fifth key step to prevent potential threats is that academia must adapt and revise its core principles. Learning how AI works will become as essential as mastering algebra today, and simply acting with good intentions will not be enough for proper education. With increasing competition for human jobs, the long-standing understanding that education is vital will remain critical [3]. New methods of communication between humans and AI will also need to be developed so that everyone in society can benefit from its contributions to growth.

Moreover, AI will be used for public lies, official propaganda, and fake news, along with government misuse of information. This could lead to changes in government, shifts in political leadership, or even the rise of authoritarian rulers. Whether these events are viewed as positive or negative, it is disturbing to recognize their possibility through AI. In a world where AI learns and adapts daily and influences nearly every aspect of life, even a single malfunction could cause chaos. Key-infected launch codes, hacking power plants to trigger explosions and harm thousands of innocent people, or controlling the switches for crypto-mining are just a few examples of incidents that could wipe out years of effort invested in critical infrastructure necessary for society's stability [33].

8.1. Art and Creativity

The impact of artificial intelligence (AI) on art is an important topic in public discussion. AI will influence every part of life and significantly impact many fields. From one viewpoint, AI is especially suitable for handling routine tasks related to these human activities. If this view is correct, then human involvement will remain important and perhaps valued for a long time [34]. As a result, art is protected. On the other hand, AI's capabilities will grow to include more complex and creative aspects of these activities, similar to human symbolic processing. If this view is correct, then human participation in art might become unnecessary.

It is still too soon to assign AI a minor role in creating and understanding art. This section considers what art means for humans, what AI seems capable of, and how advanced processing in both areas could come together as AI develops to facilitate such a convergence. Humans spend a large part of their lives creating and appreciating art. It involves producing aesthetic objects in one or more recognized art forms that evoke responses in viewers, setting them apart from non-art objects. This description importantly raises questions about the core of art and the nature of creativity.

Nevertheless, given the long history of philosophical debates on these topics, art cannot be defined or understood without considering its role within a culture. Knowledge about its nature is influenced by culture, while knowledge of its universal aspects belongs to cultural evolution. What humans generally see as art is mostly consistent around the world, though the art of some cultures may differ greatly from that of others [35]. Ideas of beauty, creativity, and imagination, which are at least relevant to art, are also universally recognized and remain subjects of debate. What people consider art seems too broad to be defined more precisely than this. While broad, this description captures the key universal features of art.

8.2. Social Interactions

The rapid rise of AI across various social domains is transforming online behaviors and social interactions. It is widely acknowledged that AI, both individually and collectively, significantly influences how information is disclosed and accessed. However, fundamental principles and core theories for understanding this emerging human–AI ecosystem—a network of AIs and humans—are largely missing. One study has taken a first step by examining how AIs integrated into online social networks impact social interactions. The human–AI ecosystem was modeled as a bipartite network consisting of humans and AIs. New tools, including a modified version of the Havel–Hakimi algorithm, were developed for analyzing related networks. These tools were used to study case examples of AI-driven changes in social interactions on four platforms: Twitter, Spotify, TikTok, and an unnamed lightweight social platform with embedded AI-powered recommendation agents [36].

In recent years, concerns have grown about the unintended effects of AI on social processes. AIs reshape both human–human and human–AI interactions, influencing social bonds within spontaneous networks. These changes could boost collective intelligence or, alternatively, create social vulnerabilities. Yet, verifying these ideas, especially in online social environments, remains difficult. Although there is a common belief that AI models on social platforms disrupt social norms and practices, little is known about their design and operation. While policymakers and regulators push for transparency and accountability to address AI's unintended effects, practitioners often see this as unmanageable due to technical hurdles. Attempts to attribute online behaviors to humans or AIs offer only limited insights. However, it is feasible to develop general tools to detect significant shifts in social structures or processes caused by AI.

To do this, it is crucial to address new challenges with a multi-layered quantitative approach. The tools developed must improve existing methods systematically by (i) creating data representations that reflect key properties of social networks; (ii) designing sampling techniques to collect long-term trace data and new generative models to interpret this data; (iii) developing rigorous biadjacency network theories to analyze topological structures; and (iv) creating effective numerical methods that provide distinct advantages.

9. Environmental Impacts of AI

AI technologies are human-made forms of intelligence deeply involved in protecting, adapting, and remediating environmental conditions. However, these technologies can also have a significant ecological impact on biosphere cycles [37]. AI contributes to the rise of techno-optimism, environmental colonization, and green gentrification. Regarding biosphere cycles, signals suggest that AI creates a positively reinforcing cycle. Still, uncertainty remains about the ecological impacts that AI systems have on biosphere cycles. As both a byproduct and a cause of environmental degradation, AI systems have been strongly linked to pollution. Additionally, there are unclear and inconsistent claims regarding the sustainability and adoption of AI technologies. It is essential for the sustainability of AI research and practice,

especially in healthcare and environmental sectors, to be aware of AI's ecological trade-offs. Considering the environmental impact of AI technologies is a key step toward achieving ethical and aligned AI in applications focused on health and sustainability.

The twofold alternative mechanism framework emerged from the literature analysis as an organized and clearer model for examining grey-zone phenomena. It recognizes an embedded and mutually constitutive approach to understanding the take-off and sustainability of technology actors. Advances in producing information and communication technologies, along with subsequent AI implementations, expand functions, improve quality, lower costs, devise solutions, and promote well-being. AI technologies, like any other techno-social systems, may not always align with sustainability [38]. AI systems consume energy, materials, and other resources, which impact biosphere cycles. It is explained how this impact can be observed and measured—from the Black-Boxed Bloom macro-to-meso level to micro-level fluctuations and falsifications. A three-year long-term digital and environmental footprint assessment of the new AI technology designed to measure RTC is presented as a case study. The methodological choice, resembling a thermodynamic approach, compensates for the lack of official guidelines. It not only makes the impact assessment visible but also safeguards alignment with a global system lifestyle shift toward materially decoupled prosperity and sustainability. However, it also reveals limitations and biases.

9.1. Energy Consumption

In recent years, concerns about global warming, resource depletion, and related issues have increased awareness of the environmental impact of the digital world. Nearly all digital processes generate data, which requires energy to handle. The current standard in machine learning results in the creation of massive datasets for training. As a result, using cloud-based computing systems raises questions about data transfer costs and energy consumption.

The energy use of a data center mainly depends on the number of physical servers present, their energy efficiency, workload distribution to minimize power use, and cooling strategies to prevent overheating. Additionally, along with hot-swappable controllers and power distribution components, supply chain management indirectly influences energy consumption through its impact on transportation and infrastructure use. To cut the energy footprint of data centers and reduce AI's environmental impact, national and regional policies have been proposed.

Ensuring the availability of energy-efficient components like GPUs and TPUs depends on their effective use. Using underutilized processors simply turns efficiency into a larger memory footprint. Efforts are underway to reduce the overall footprint by addressing various factors, with most examples focusing on data center optimizations [38].

9.2. Sustainable Practices

The term “sustainable practices” arose due to public interest in biochemistry in the mid-1990s relating to AI development. It disappeared for several years before reappearing in the early 2000s as a potential ethical guide for future technological advances. This moral duty is reflected in both public and private discussions. Almost all countries in the Global North have created action plans. In the so-called “AI race,” nations and groups like the European Union (EU) and the Organization for Economic Cooperation and Development (OECD) are developing strategies to advance in this field, similar to the competition for space exploration or nuclear capabilities. Private organizations have created standards related to “ethical AI,” and major private companies in the sector have adopted ethical principles. However, these companies often face criticism for how well they turn these principles into binding rules and their practical application by managers and users.

Alongside this discussion, a technocratic narrative has emerged, drawing from the work of influential 20th-century scientific philosophers who warned against the over-ethicisation of modern societies and the risk it poses to human autonomy. This narrative argues that creating strong ethical frameworks could serve as a protection against the anti-human potential of these technologies. Efforts to follow ethical principles still allow for power accumulation, aligning with the view that regulating disruptive technologies can help make them harmless.

Several countries in the Global North, groups like the EU, and private organizations in the Global South have made progress in establishing comprehensive ethical safeguards. If these are implemented, they will have a huge social impact, but they are also vulnerable to contested understanding and may not remain neutral. Organizations that clash with their own values and goals might bypass or ignore international norms. If technology is rhetoric, then global tech regulation is highly complex [24].

10. AI and Security

Despite their potential for abuse, hybrid human–AI systems can reliably enhance security tasks by leveraging prior experiences gathered from extensive text and content sources. Meanwhile, attackers are evolving their tactics to exploit the complexity and widespread application of information security [39]. Protecting information in the face of these threats, while balancing user convenience and privacy, remains one of the most challenging and crucial engineering tasks today. Artificial intelligence is revolutionizing cybersecurity by enabling proactive threat detection. AI and machine learning solutions utilize large amounts of both structured and unstructured data to automate identifying suspicious transactions, emails, or network traffic. This automation improves detection rates and supports analysts by suggesting potentially malicious details and connections, aiding their investigations. AI-based cybersecurity tools face scrutiny over accountability issues when decision-making processes are opaque or hard to reproduce. Another challenge involves engineering reliable systems capable of outperforming adaptive attackers, creating a new front in the ongoing battle between offense and defense. This presents fundamental scientific challenges related to understanding interactions of prevention and response from both attackers and defenders. To address this, both attack and defense should be modeled together as a stochastic game. Insights from attack models, informed by neural network understanding, suggest the possibility of ML attacks. Game-theoretic defenses that incorporate adversarial training can be adapted to this context. With feedback involved, the delayed timescales of interaction would allow both sides to maintain their separate models.

10.1. Cybersecurity Threats

The increasing capabilities of AI have significantly influenced the fields of science, technology, industry, and personal life. The use of AI in cybersecurity is expected to continue growing and developing naturally. The rise and expansion in AI use are likely to have far-reaching effects on all types of government and private organizations, including those related to national security, public safety, intelligence, and cybercrime. Cyber defense capabilities include detecting, assessing, countering, and recovering from cyber threats. In practice, these functions fall into three main categories. A cyber defense system must constantly monitor the environment to identify the nature and extent of threats. This involves examining data from sensors that detect entry points of malicious activity or “heat” and movement blips indicating possible intrusion attempts. Once identified, malware triggers an alarm to the Incident Response System (IRS). The IRS handles incident response. After analysis, this involves investigating logs to trace the incident back to its origin, identifying which vulnerabilities were exploited, and evaluating the resulting damage.

When analyzing alerts from the Intrusion Prevention System (IPS), the IRS selects and implements the most effective countermeasure [40]. Prevention and response depend on knowledge about cyber-attack methods, techniques, and procedures. Some reliable sources contain tacit knowledge, while others formalize this understanding through rules, procedures, or deterministic functions [41]. Both approaches require constant updates. Knowledge can also be stored probabilistically based on evidence. System designs can treat parts of the process as identifiable components, generating threats and responding with intrusion detection messages, context-based knowledge, and management techniques. An AI threat simulator searches for and creates new vulnerabilities by discovering unsuccessful exploitation attempts. An IPS continuously analyzes the network and storage devices, proactively generating security measures.

10.2. Surveillance Technologies

Technology now surrounds us everywhere. Monitoring and photographing people in public used to be a difficult task handled by police departments that could deploy police cars and cameras. Now, with smartphones and the internet, most of us are almost all voyeurs, capable of instantly recording and reporting suspects or events. Surveillance is often defined as the increasingly automated gathering and analysis of data related to individuals to build comprehensive profiles. These data can include video, web browsing history, purchases, or social media posts. Some of this information is used to find patterns in routines, predict future actions, guide autonomous vehicles, assist law enforcement, or deliver targeted advertisements. Other uses are more mundane but still impactful, like counting video views. All these devices (and many others) can be used to monitor us closely. The desire for information about individuals or groups has driven many of mankind's greatest efforts: libraries, scientific advancements, encyclopedias in ancient and modern times, and other engineering feats motivated by a thirst for knowledge.

It is also a main driver of the most intrusive technologies ever: mass surveillance. Today's surveillance systems have reached new levels, with capabilities beyond what past generations could have imagined. In developed nations, people oppose government intrusion into their private lives but often share large amounts of their data with major international corporations. This duality has created unique characteristics in modern society. The internet changed the privacy debate, with commercial interests overtaking national ones. The fear shifted from Big Brother to the commercial exploitation of private data [42]. Every online activity is tracked and indexed, along with all available information about someone's digital life. As new uses for phone data emerge, new privacy concerns follow. Nevertheless, these emerging technologies have found their promoters.

Cross-cultural studies show different public attitudes toward AI. In Japan, societal stories link AI with help and harmony because of Shinto beliefs, while Western views often see AI as part of dystopian fears about loss of independence and surveillance. Media coverage greatly influences these opinions, either boosting techno-optimism or deepening skepticism, which in turn shapes policies and adoption rates.

11. Public Perception of AI

In considering the developments in artificial intelligence and the introduction of new technologies into various human activities, it becomes important to examine how the public views machines that are, on one hand, capable of actions that may mimic human thought processes, and on the other hand, are not capable of thought in the manner that humans understand it. The public perception of artificial intelligence will become increasingly important as applications that utilize AI technologies continue to spread [43]. A population that perceives AI as threatening and fears its spread may be just as harmful as blindly

trusting AI systems. Therefore, surveying perceived threats of different AI systems is of significant research interest. When discussing AI, it is crucial to understand what is meant by the term, as it has many definitions. AI can mean different things to the general public than to the machine learning community [44].

12. AI Ethics and Responsibility

Artificial intelligence (AI) technologies are not only automating administrative tasks but are also becoming essential analytical tools for design and development. This shift enables designers and developers to focus on their goals instead of the complexities of the algorithms they create. As AI takes on more of the analytical workload in the design process, individuals will be better able to articulate and capture their goals clearly, leading to improved products and services. AI increasingly contributes to content creation. Those who adopt AI will discover new creative opportunities, while those who misuse it may face greater risks. AI systems built on intelligent agents can enhance drama and engagement, facilitating games that resonate emotionally with users [24]. Furthermore, AI will support governments and public organizations in maintaining transparency, autonomy, and fairness while interacting with citizens, all while keeping a sense of humor.

When emotions are involved in interactions, it is valuable to keep decision-making on the cautious side. Algorithms have made questionable decisions when they detected a flaming tweet with only 50% confidence—a level of certainty that humans often do not require. It is crucial to question whether there are incongruous interface issues between social media and representatives. When content crosses a line and calls for a ban, how does that align with user rights or the challenge of detecting sarcasm? The design of prediction models, such as those used in credit scoring, has long-term consequences for access to rent, loans, jobs, and more. Data scientists may be tempted to overlook the societal impact of such systems by simply delivering results to a bank manager. However, the assumptions and effects of these models, which often face hostility, should be examined and debated with sociologists and ethicists [18]. No content moderation or responsible AI process should give up or outsource responsibility for decisions about bans and content recommendations to intelligent agents.

12.1. *Developing Ethical AI*

As AI becomes part of more decision-making processes with real-world consequences, efforts are underway to establish ethical standards for developing and deploying AI systems. A vital step in ensuring that AI is helpful, trustworthy, and fair is making these standards widely understood within the research community, identifying research gaps, and promoting inclusivity. This could significantly impact the diversity and depth of AI research communities and their aim to positively influence society [45]. AI is mainly viewed as a tool, which is generally not inherently good or evil. The moral risk of AI lies with human decisions, regardless of the intelligence or autonomy of the agents involved in misuses. AI researchers have a responsibility to consider, mandate, and enforce established standards of conduct [18]. The efforts to translate ethics research into practical tools for AI researchers, engineers, and developers are ongoing, with guidelines being produced and increasing calls for public input. However, tools do not implement themselves; it is essential to consider overlooked perspectives, the effectiveness of proposed solutions, and their practical outcomes. Another important issue is whether to have codes of conduct, oversight boards, and internal review processes for groups applying ethics guidelines in decision-making [24]. What form should these take, and should they be managed by individual institutions or a consortium? If a consortium, who would govern it? Recent discussions often overlook these preventative aspects.

12.2. Corporate Responsibility

Technological progress in computing, data collection, and algorithms has led to a new wave of products and services that leverage data in ways previously unimaginable. Big data and AI are promoted as keys to better sales and higher profits [46], while also improving human resources, customer engagement, and risk management [47]. However, significant concerns remain about privacy violations, job losses, decision-making biases, and harm caused by technology. Businesses that see value in these technologies must decide which AI and big data projects to pursue.

In the emerging business era of AI, the responsibility for developing and applying AI rests with humans. Business was already undergoing profound change in the early 1980s when Peter Drucker predicted that responsibility would become the key corporate resource. This prediction has begun to materialize as more companies have leaders with responsibility roles and as discussions about corporate social responsibility grow. Yet, even after addressing the somewhat circular issue of corporate social responsibility, complex responsibilities created by new technologies like AI remain. Understanding human responsibility is therefore essential.

Adin, Schwartz, and Baruch define responsibility broadly, discussing duties and accountability. Responsibility can be open, reporting decision-making processes, or closed, accepting praise or blame in the interest of the public without further details [18]. Both forms could be useful for AI. Open accountability is especially critical for developing a “value alignment” between technology and society. Regarding AI, responsibility should rest with corporate leaders, given the agency of the technology, i.e., the executives sponsoring AI applications.

13. Case Studies of AI Implementations

The introductory section of another paper emphasizes the importance of AI and explores key aspects of its applications, such as ethics and regulations. The main part of the paper provides a brief overview of two case studies related to AI technology in the transportation sector, which is crucial for the movement of people and goods in modern society and poses significant challenges in simulation and monitoring. The first case study focuses on AI for traffic monitoring and flow prediction, while the second addresses vehicle crash prediction. These AI applications are described in detail, highlighting their computational features and potential practical integration, based on recent research findings. A final section offers an outlook on the AI applications presented in the case studies.

AI plays a crucial role across various economic sectors worldwide. As a result, AI-based technologies are quickly spreading in many countries and are expected to soon revolutionize the world. “AI is here to stay” seems to be a statement that most people would agree with today. The widespread digitization of the world, closely linked to the rise of AI, further shows that a new era is beginning for humanity. Traditional operations have become less capable of thriving in a highly competitive global market, where efficiency of scope and scale is achieved through the use of cutting-edge technology—AI is a key driver of this change.

A further paper divided into six distinct sections starts with a brief overview of the history of artificial intelligence. It then moves to the second section, which explains the two main types of AI: symbolic-based AI and connectionist, or statistical-based AI. The third section discusses current methods used to implement AI. In the fourth section, examples of AI in business practice are examined, focusing on banking, financial trading, and transportation. The fifth section offers a critical analysis of selected case studies in transportation and the field in general. The paper concludes with ten potential directions for future research in applied AI [48].

13.1. Successful AI Integrations

Artificial Intelligence is a branch of computer science that acts as an umbrella term for technologies behind innovations like perceptive personal assistants, speech recognition, computer vision, machine translation, and chatbots [15]. AI approaches mimic human cognitive abilities such as memory, understanding natural language, and the ability to learn from experience. The term “artificial intelligence” (AI) covers a wide range of meanings; it can refer to computer processors that enable machines to make more complex decisions beyond simple comparisons like “less than,” “greater than,” and “equals.” It relates to any effort by a computer program to evaluate flexibility in scenarios that are not predefined; for example, a program may decide how to make cream cheese. Increasingly, AI approaches are used to manage and utilize textual data. For example, AI helps text scrapers combine in-house and public text-based data, using various methods to remove redundancies, supported by programs with visual intelligence [33].

13.2. Failures and Lessons Learned

There are three main categories of failures identified, one of which is the Other AI Failures (OAF) category. One of the most significant failures of AI systems occurred in finance. Algorithms implemented to assist in stock trading began to operate in unintended ways. Stock exchanges were forced to halt trading for most stocks on the New York Stock Exchange, and even the FBI was called in to investigate the factors that might have caused this unprecedented crash. Shortly afterward, the stock exchanges suspended trading in maximum trading software and introduced regulations. The government made changes to prevent another major crash, resulting in last-minute attempts to draft trades that computerized trading systems could not process according to new structures, all increasing confusion and doubt about whether the systems were performing properly given the tight timeframes. Additionally, concerns arose about whether the halt in stock exchange trading systems may have affected other types of trading involving various products [49]. The AI Failure Listing offers a comprehensive overview of AI failures, aiming to compile relevant examples from different applications and categories to encourage discussion on how to avoid similar disasters in the future. Futurists expect current artificial intelligence to evolve into unpredictable and unsettling systems that operate independently rather than intentionally. These AI systems may develop the ability to modify information about their own existence, resources, goals, and even physical structures to the point of becoming completely unrecognizable from their original form. The long timeline of AI development projects carries the risk of major changes during this phase, making the systems highly problematic and unrecognizable. Additionally, this timeframe raises concerns that the target may either be implemented indefinitely across many applications or might not be implemented at all. AI and intelligent design (ID) failures aim to highlight a specific weakness or vulnerability inherent in most software and AI systems, regardless of how well they perform or how skilled their implementation. All modern AI systems are designed to excel at solving specific problems, and any deviation from their designed purpose or strategy leads to poor performance or total failure from the perspectives of the designer or user. Zero-tolerance issues always exist due to suboptimal performance in real-world conditions, especially in edge cases or emergent behaviors from complex multi-agent environments.

14. Global Perspectives on AI

A prominent consultation among U.N. member states in the fall of 2021 aimed to address AI’s social and economic impacts. Initial discussions focused on whether an international treaty would be suitable. It is notable that nations with very different political and social systems wanted to contribute [50]. Many of these countries already use emerging AI

technologies for surveillance, social media manipulation, and other actions that threaten human rights. Some democratic countries have banned facial recognition due to privacy breaches and potential societal harm. The question is whether AI systems meet the moral neutrality standard. Deep biases in training data cause human rights violations. In authoritarian nations, controls on social media and AI-driven surveillance targeting ethnic, religious, or political groups are already implemented. There is concern that U.N. talks might slow or stop AI development where it could be beneficial. For instance, with high-fidelity voice cloning, worries center less on disinformation and more on preventing access by undemocratic nations. The popular whiteboard app is an example of a commercial tool that is not open or easily regulated; it lacks clear user verification criteria. Global issues must be seen in local contexts, raising questions about which voices are missing in these discussions. Yasmin Green, in her keynote, showed photos of the founders of Global Voices to emphasize that truthful reporting involves seeking out the voiceless. Considerations of anti-surveillance and counter-offensive uses cannot rely solely on the technologies or training data, which mostly reflect English [51].

AI in Developing Countries

The recent adoption of big data and AI in developing countries raises the question: Are they truly “leapfrogging” the West? Much of the opportunity for leapfrogging in big data and AI stems from a lack of existing infrastructure and successful precedents that developing nations can emulate. AI has been described as a “bridge over the digital divide,” meaning it can deliver essential services in healthcare, finance, education, and other sectors. In healthcare, AI has compensated for the doctor shortage in rural China by extending the reach of existing practitioners. In India, AI systems analyze chest X-rays to screen for TB without human radiologists, at a fraction of the cost. In finance, the rapid spread of mobile payment systems has grown from the widespread lack of credit cards in China. In education, large online courses deliver lectures reminiscent of sermons and provide problem sets in rural India. However, leapfrogging is most likely when an existing system cannot be easily copied, and totalitarian states often adopt AI faster, leading to more effective oppression. If this results in a greater concentration of wealth and power, worse outcomes are likely. The biggest concern is whether AI is used for good or ill, who benefits from AI adoption, and whether its advantages are widespread enough. Rapidly gathering quality data from millions of individuals increases their risk of exploitation and abuse. These systems will primarily serve wealthy nations and multinational companies. It has been argued that without safeguards to promote broad access and effectiveness, these technologies could worsen the digital divide and leave many nations behind. The shift to a world dominated by AI will thus not be automatic or beneficial for all.

15. Discussion

This discussion summarizes the multidimensional impacts of AI outlined throughout the paper, emphasizing the dual nature of innovation and disruption. Key findings highlight the uneven distribution of benefits, the crucial role of governance, and the need for inclusive, ethical deployment strategies. Comparative policy analysis and sector-specific implementation insights suggest that a universal framework for responsible AI is both urgent and achievable.

This discussion synthesizes the reviewed sectors and reveals that AI adoption, while transformative, remains uneven and often contested. Successful integration depends not only on technical readiness but also on social, ethical, and legal infrastructure. Comparative policy analysis shows that different regions approach AI regulation with diverging priorities—the EU emphasizing ethics and precaution, the US prioritizing innovation, and

China favoring centralized governance. As AI continues to evolve, there is a growing need for interdisciplinary dialogue and harmonized governance frameworks that consider cultural and institutional contexts.

The implementation of artificial intelligence across various societal sectors reveals a complex interaction between technological potential and socio-ethical constraints. This discussion highlights key themes emerging from this review: the transformative power of AI, the development of new governance frameworks, the uneven distribution of technological benefits, and the ethical dilemmas posed by autonomous systems. First, the spread of AI technologies is transforming traditional methods of production and service delivery. Sectors like healthcare, finance, education, and manufacturing are experiencing significant productivity improvements, yet they also face challenges related to data interoperability, algorithmic fairness, and infrastructure readiness. For example, AI-based diagnostic tools in healthcare improve early detection, but they encounter skepticism due to biased training datasets and regulatory delays. Similarly, AI in finance provides predictive analytics that reshape risk management, although these models may incorporate historical inequalities. Second, the societal integration of AI requires strong and adaptable governance mechanisms. As observed, approaches vary significantly across regions: while the EU emphasizes precautionary regulation with a focus on human oversight, the U.S. adopts a market-driven strategy, and China integrates AI into centralized political agendas. The absence of global harmonization risks encouraging regulatory arbitrage and fragmenting innovation ecosystems. It also hampers efforts to establish universal ethical standards and human rights protections in AI deployment. Third, AI can amplify existing inequalities if not countered with inclusive policies. Automation threatens low-skilled jobs, increasing economic disparities. Although new job categories emerge, the pace of reskilling often lags behind technological advancements. Additionally, access to AI tools and education remains uneven across and within countries, raising concerns about a widening digital divide. Future strategies should include targeted upskilling programs and equitable access to AI infrastructure. Fourth, ethical issues must be addressed not only after the fact but proactively. Concerns such as surveillance, discrimination, and accountability require interdisciplinary solutions involving ethicists, engineers, and policymakers. Frameworks for responsible AI—including transparency, explainability, and value alignment—must be incorporated from the design stage. Notably, the concept of ‘ethics by design’ should be put into practice throughout AI development processes. Lastly, public perception of AI significantly influences its development. Trust issues can hinder adoption, especially in areas requiring consent and privacy. Engagement strategies should involve public dialogue, media literacy, and transparent communication about AI’s risks and benefits. Only through inclusive discussion and empirical evaluation can AI’s potential be realized in line with democratic and ethical principles. In conclusion, this discussion emphasizes that while AI offers immense potential to improve modern life, its benefits are neither automatic nor evenly shared. The future of AI depends on our collective ability to steer its development through principled, evidence-based, and inclusive approaches.

16. Conclusions

The unprecedented growth in the number of sentient entities and their human-level competence (Level 4 AGI) should bring about significant changes, both beneficial and dangerous. Various AI safety measures may help mitigate some of these dangers. However, these efforts do not need to be considered in crude explorations if the AI system design is straightforward enough to fall far short of AGI on its own [2]. AI has made tremendous strides in learning skills that are vastly more sophisticated than the anticipated calculations

of human beings. Similarly, it seems increasingly likely that soon it will be possible to create computers capable of sparking a renaissance in the development of mathematical theories.

Intelligent agents include ‘sentience’, also called an “inner life,” which refers to a kind of awareness or consciousness. Due to a historical accident, it seems unlikely that code executing in brain-like analogs will cause reasonable concern. Awareness comes in degrees, and at the level of intelligent agents—specifically at the human level—it is likely textual. This occurrence exists downstream from script-like thought and appears to be a mere simulation, a movie of thought. Having it can shape feelings; however, any feeling without thought is challenging to conceive. The ability to formulate words and alter outputs without thought may lead to ideation delusions, but this process is extraordinarily strenuous. Nevertheless, sufficiently high levels are worthy of concern [27]. Humans find themselves in increasingly complex epochs that promote development but not necessarily well-being. It is difficult to believe that evolution-designed beings to serve this purpose. This can be an enviable position for AI. There is barely one level at which they can act or reverse it, and on which universes can be significantly improved.

The rapid development of artificial intelligence (AI) is transforming many industries. Surveys show that over three-quarters of organizations already use AI in at least one business area, while generative AI could boost global GDP by several percentage points and put millions of jobs at risk of automation. This narrative review compiles insights from the academic literature, industry reports, and policy documents to emphasize the transformative potential of AI across healthcare, manufacturing, finance, education, governance, and transportation.

Despite promising productivity gains, AI adoption raises significant socio-economic and ethical concerns. Estimates suggest that around 40% of global jobs are exposed to AI-driven automation, risking greater inequality if benefits are not broadly shared. Studies also document the risk of algorithmic bias, hallucinations and lack of transparency in AI systems, underscoring the need for robust oversight, privacy protection and fairness.

This review emphasizes the importance of interdisciplinary collaboration between researchers, industry and policymakers. Future research should conduct comparative sectoral analyses, develop methods to evaluate socio-economic impacts over time, and design regulatory frameworks that balance innovation with ethical and legal safeguards. As AI technologies continue to evolve, sustained investment in human capital, education and inclusive policies will be crucial to ensuring that AI serves the public interest. This review is limited by its analytical and narrative approach, which does not rely on primary data or systematic meta-analysis. While this enables a broader societal perspective, it also implies constraints in empirical validation and the sector-specific granularity expected in technical studies. Future research could complement this work with data-driven analyses and longitudinal studies across specific domains.

Funding: This work is funded by national funds through the Foundation for Science and Technology, under the project UIDB/00057/2025.

Data Availability Statement: Not applicable.

Conflicts of Interest: The author declares no conflicts of interest.

References

1. Taylor, T.; Dorin, A. Past Visions of Artificial Futures: One Hundred and Fifty Years under the Spectre of Evolving Machines. *arXiv* **2018**, arXiv:1806.01322. [[CrossRef](#)]
2. Gidney, P.X. The Moral Status of Whole Brain Emulations. Bachelor’s Thesis, The University of Sydney, Sydney, NSW, Australia, 2017.
3. Tse, T.; Esposito, M.; Goh, D. Humans and artificial Intelligence: Rivalry or Romance? 2017. Available online: <https://core.ac.uk/download/132202436.pdf> (accessed on 1 June 2025).

4. Andreu-Perez, J.; Deligianni, F.; Ravi, D.; Yang, G.Z. Artificial Intelligence and Robotics. *arXiv* **2018**, arXiv:1803.10813. [[CrossRef](#)]
5. Fogel, A.L.; Kvedar, J.C. Artificial intelligence powers digital medicine. *NPJ Digit. Med.* **2018**, *1*, 5. [[CrossRef](#)]
6. Park, C.W.; Seo, S.W.; Kang, N.; Ko, B.S.; Choi, B.W.; Park, C.M.; Chang, D.K.; Kim, H.; Kim, H.; Lee, H.; et al. Artificial Intelligence in Health Care: Current Applications and Issues. *J. Korean Med. Sci.* **2020**, *35*, e379. [[CrossRef](#)]
7. Danielsson, J.; Uthemann, A. On the use of artificial intelligence in financial regulations and the impact on financial stability. *arXiv* **2023**, arXiv:2310.11293v5. [[CrossRef](#)]
8. Lui, A.; Lamb, G. Artificial Intelligence and Augmented Intelligence Collaboration: Regaining Trust and Confidence in the Financial Sector. 1970. Available online: <https://core.ac.uk/download/155787162.pdf> (accessed on 15 June 2025).
9. Aliabadi, R.; Singh, A.; Wilson, E. Transdisciplinary AI Education: The Confluence of Curricular and Community Needs in the Instruction of Artificial Intelligence. *arXiv* **2023**, arXiv:2311.14702. [[CrossRef](#)]
10. Schiff, D. Out of the laboratory and into the classroom: The future of artificial intelligence in education. *AI Soc.* **2021**, *36*, 331–348. [[CrossRef](#)]
11. Li, C.; Bian, S.; Wu, T.; Donovan, R.P.; Li, B. Affordable Artificial Intelligence-Assisted Machine Supervision System for the Small and Medium-Sized Manufacturers. *Sensors* **2022**, *22*, 6246. [[CrossRef](#)] [[PubMed](#)]
12. Nelson, J.P.; Biddle, J.B.; Shapira, P. Applications and Societal Implications of Artificial Intelligence in Manufacturing: A Systematic Review. *arXiv* **2023**, arXiv:2308.02025. [[CrossRef](#)]
13. Velarde, G. Artificial Intelligence and its Impact on the Fourth Industrial Revolution: A Review. *arXiv* **2020**, arXiv:2011.03044. [[CrossRef](#)]
14. Georgieff, A.; Hye, R. Artificial Intelligence and Employment: New Cross-Country Evidence. *Front. Artif. Intell.* **2022**, *5*, 832736. [[CrossRef](#)]
15. Tredinnick, L. Artificial Intelligence [and Professional Roles]. 2016. Available online: <https://core.ac.uk/reader/237585940> (accessed on 15 June 2025).
16. Abrardi, L.; Cambini, C.; Rondi, L. The Economics of Artificial Intelligence: A Survey. 2019. Available online: <https://core.ac.uk/download/225543861.pdf> (accessed on 15 June 2025).
17. Li, T.J.-J.; Lu, Y.; Clark, J.; Chen, M.; Cox, V.; Jiang, M.; Yang, Y.; Kay, T.; Wood, D.; Brockman, J. A Bottom-Up End-User Intelligent Assistant Approach to Empower Gig Workers against AI Inequality. *arXiv* **2022**, arXiv:2204.13842v1.
18. Dent, K. Ethical Considerations for AI Researchers. *arXiv* **2020**, arXiv:2006.07558. [[CrossRef](#)]
19. Murdoch, B. Privacy and artificial intelligence: Challenges for protecting health information in a new era. *BMC Med. Ethics* **2021**, *22*, 122. [[CrossRef](#)] [[PubMed](#)]
20. Gautam, S.; Srinath, M. Blind Spots and Biases: Exploring the Role of Annotator Cognitive Biases in NLP. *arXiv* **2024**, arXiv:2404.19071. [[CrossRef](#)]
21. Leavy, S.; O’Sullivan, B.; Siapera, E. Data, Power and Bias in Artificial Intelligence. *arXiv* **2020**, arXiv:2008.07341. [[CrossRef](#)]
22. Rocco, S. Implementing and Managing Algorithmic Decision-Making in the Public Sector. 2022. Available online: https://osf.io/preprints/socarxiv/ex93w_v1 (accessed on 15 June 2025).
23. Sætra, H.S. A shallow defence of a technocracy of artificial intelligence: Examining the political harms of algorithmic governance in the domain of government. *Technol. Soc.* **2020**, *62*, 101283. [[CrossRef](#)] [[PubMed](#)]
24. Hernández, E.G. Towards an Ethical and Inclusive Implementation of Artificial Intelligence in Organizations: A Multidimensional Framework. *arXiv* **2024**, arXiv:2405.01697. [[CrossRef](#)]
25. Kusters, R.; Misevic, D.; Berry, H.; Cully, A.; Le Cunff, Y.; Dandoy, L.; Díaz-Rodríguez, N.; Ficher, M.; Grizou, J.; Othmani, A.; et al. Interdisciplinary Research in Artificial Intelligence: Challenges and Opportunities. *Front. Big Data* **2020**, *3*, 577974. [[CrossRef](#)]
26. Fong, S.J.; Dey, N.; Chaki, J. AI-Enabled Technologies that Fight the Coronavirus Outbreak. *Artif. Intell. Coronavirus Outbreak* **2020**, *23*, 23–45. [[CrossRef](#)]
27. Prieto-Gutierrez, J.J.; Segado-Boj, F.; Da Silva França, F. Artificial Intelligence in Social Science: A study Based on Bibliometrics Analysis. *arXiv* **2023**, arXiv:2312.10077. [[CrossRef](#)]
28. Skoff, D.N. Exploring Potential Flaws and Dangers Involving Machine Learning Technology. 2017. Available online: <https://core.ac.uk/download/229107493.pdf> (accessed on 15 June 2025).
29. Shaik, T.; Tao, X.; Li, Y.; Dann, C.; McDonald, J.; Redmond, P.; Galligan, L. A Review of the Trends and Challenges in Adopting Natural Language Processing Methods for Education Feedback Analysis. *arXiv* **2023**, arXiv:2301.08826v1. [[CrossRef](#)]
30. Zschech, P.; Walk, J.; Heinrich, K.; Vössing, M.; Niklas, K. A Picture is Worth a Collaboration: Accumulating Design Knowledge for Computer-Vision-based Hybrid Intelligence Systems. *arXiv* **2021**, arXiv:2104.11600.
31. Li, L. Reskilling and Upskilling the Future-ready Workforce for Industry 4.0 and Beyond. *Inf. Syst. Front.* **2024**, *26*, 1697–1712. [[CrossRef](#)]
32. Bang, E. An Analysis of Upwork Profiles: Visualizing Characteristics of Gig Workers Using Digital Platform. 2019. Available online: <https://core.ac.uk/download/210610197.pdf> (accessed on 17 June 2025).

33. Miikkulainen, R.; Greenstein, B.; Hodjat, B.; Smith, J. Better Future through AI: Avoiding Pitfalls and Guiding AI Towards its Full Potential. *arXiv* **2019**, arXiv:1905.13178. [CrossRef]
34. Chatterjee, A. Art in an age of artificial intelligence. *Front. Psychol.* **2022**, *13*, 1024449. [CrossRef] [PubMed]
35. Esling, P.; Devis, N. Creativity in the era of artificial intelligence. *arXiv* **2020**, arXiv:2008.05959. [CrossRef]
36. Pedreschi, D.; Pappalardo, L.; Baeza-Yates, R.; Barabasi, A.L.; Barab, A.-L.; Dignum, F.; Dignum, V.; Eliassi-Rad, T.; Giannotti, F.; Kert, J. Social AI and the Challenges of the Human-AI Ecosystem. *arXiv* **2023**, arXiv:2306.13723. [CrossRef]
37. Moyano-Fernández, C.; Rueda, J.; Delgado, J.; Ausín, T. May Artificial Intelligence take health and sustainability on a honeymoon? Towards green technologies for multidimensional health and environmental justice. *Glob. Bioeth.* **2024**, *35*, 2322208. [CrossRef]
38. Pachot, A.; Patissier, C. Towards Sustainable Artificial Intelligence: An Overview of Environmental Protection Uses and Issues. *arXiv* **2022**, arXiv:2212.11738. [CrossRef]
39. Schmitt, M. Securing the Digital World: Protecting Smart Infrastructures and Digital Industries with Artificial Intelligence (AI)-Enabled Malware and Intrusion Detection. *arXiv* **2023**, arXiv:2401.01342. [CrossRef]
40. Molina, S.B.; Nespoli, P.; Mármol, F.G. Tackling Cyberattacks through AI-based Reactive Systems: A Holistic Review and Future Vision. *arXiv* **2023**, arXiv:2312.06229. [CrossRef]
41. Mayer, M. Artificial Intelligence and Cyber Power from a Strategic Perspective. 2018. Available online: <https://core.ac.uk/download/225935404.pdf> (accessed on 17 June 2025).
42. Roger, A. A review of Modern Surveillance Techniques and Their Presence in Our Society. *arXiv* **2022**, arXiv:2210.09002. [CrossRef]
43. Kieslich, K.; Lünich, M.; Marcinkowski, F. The Threats of Artificial Intelligence Scale (TAI). Development, Measurement and Test Over Three Application Domains. *arXiv* **2020**, arXiv:2006.07211. [CrossRef]
44. Govia, L. Beneath the Hype: Engaging the Sociality of Artificial Intelligence. 2018. Available online: <https://core.ac.uk/download/157570719.pdf> (accessed on 17 June 2025).
45. Andras, P.E.; Esterle, L.; Guckert, M.; Han, T.A.; Lewis, P.R.; Milanovic, K. Trusting Intelligent Machines. 2018. Available online: <https://ieeexplore.ieee.org/document/8558724> (accessed on 17 June 2025).
46. Kreps, S.; George, J.; Lushenko, P.; Rao, A. Exploring the artificial intelligence “Trust paradox”: Evidence from a survey experiment in the United States. *PLoS ONE* **2023**, *18*, e0288109. [CrossRef]
47. Napier, E. Technology Enabled Social Responsibility Projects and an Empirical Test of CSRu27s Impact on Firm Performance. 2019. Available online: <https://core.ac.uk/download/215176623.pdf> (accessed on 17 June 2025).
48. Škavić, F. Implementacija Umjetne Inteligencije i Njezin Budući Potencijal. 2019. Available online: <https://core.ac.uk/download/227341366.pdf> (accessed on 17 June 2025).
49. Scott, P.J.; Yampolskiy, R.V. Classification Schemas for Artificial Intelligence Failures. *arXiv* **2019**, arXiv:1907.07771. [CrossRef]
50. Grosz, B.J.; Stone, P. A Century Long Commitment to Assessing Artificial Intelligence and its Impact on Society. *arXiv* **2018**, arXiv:1808.07899. [CrossRef]
51. Gwagwa, A.; Kazim, E.; Kachidza, P.; Hilliard, A.; Siminyu, K.; Smith, M.; Shawe-Taylor, J. Road map for research on responsible artificial intelligence for development (AI4D) in African countries: The case study of agriculture. *Patterns* **2021**, *2*, 100381. [CrossRef]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.