Article

Will artificial intelligence threaten humanity?

Milad Shahvaroughi Farahani[1,*], Ghazal Ghasemi[2]

[1] Department of Finance, Khatam University, Tehran 1991633357, Iran

[2] Islamic Azad University, Tehran 1477893780, Iran

* Corresponding author: Milad Shahvaroughi Farahani, m.shahvaroughi@khatam.ac.ir

Abstract: The rapid advancement of artificial intelligence (AI) has sparked intense debate regarding its potential threat to humanity. This abstract delves into the multifaceted discussion surrounding the implications of AI for the future of humanity. It explores various perspectives, ranging from optimistic views that highlight the transformative benefits of AI to pessimistic concerns about its existential threat. Drawing on insights from experts and researchers, the abstract examines key areas of contention, including the possibility of a technological singularity, the ethical dilemmas posed by autonomous weapons, and the socio-economic impacts of AI-driven automation. So, the main purpose of the paper is to study the impacts of AI from different points of view, including social, economic, political, etc. Therefore, different. Furthermore, it discusses strategies for mitigating the risks associated with AI, emphasizing the importance of ethical guidelines, regulatory frameworks, and international cooperation. Overall, this abstract provides a comprehensive overview of the complex considerations surrounding the impact of AI on humanity and underscores the need for thoughtful deliberation and proactive measures to ensure a beneficial and responsible integration of AI into society.

Keywords: unemployment; artificial intelligence threat; risk mitigation; ethical considerations
Jel Classification Code: B55; E24; J64; J81; M21; M54; O15

1. Introduction

The rapid advancement of artificial intelligence (AI) has sparked a profound and wide-ranging debate about its potential impact on humanity. While AI holds the promise of revolutionizing industries, solving complex problems, and enhancing human capabilities, there is growing concern about the risks it poses to society, ranging from economic disruption to existential threats [1]. This introduction seeks to provide an overview of the multifaceted issues surrounding the question, "Will Artificial Intelligence Threaten Humanity?"

At the heart of the discussion lies the concept of technological singularity, a theoretical scenario in which AI systems surpass human intelligence, leading to unpredictable and potentially adverse outcomes. This notion has fueled speculation about the emergence of super-intelligent AI entities that could pose existential risks to humanity, such as the loss of control over AI systems or the subjugation of humans by AI-driven entities [2].

However, the debate is not limited to existential concerns. Ethical considerations loom large, with questions about the moral implications of AI-driven decision-making, the potential for autonomous weapons systems to violate human rights, and the equitable distribution of AI benefits and burdens across society.

Moreover, there are divergent views on the trajectory of AI development. While some advocate for cautious optimism, emphasizing the potential of AI to address

pressing global challenges, others warn of the need for vigilance and proactive measures to mitigate the risks associated with AI proliferation [3].

Against this backdrop, this article will explore the various dimensions of the debate surrounding the potential threats posed by artificial intelligence to humanity. It will examine the arguments put forth by proponents and skeptics alike, consider the ethical implications of AI development, and discuss the challenges and opportunities that lie ahead in harnessing the transformative power of AI while safeguarding human well-being and prosperity.

The rest of the paper is as follows:

First, there is a literature review about the different impacts of AI on humanity from different points of view, such as social, economic, and political. The second part is dedicated to AI and its characteristics. Then, there is a part that examines the effect of AI on human life from different aspects. Next, different perspectives and opinions have been reviewed. Later, there are recommendations. Finally, there is a conclusion.

2. Literature review

There are many research papers and articles that have examined the impacts of AI on human life. We can divide them into two parts: 1. scientific research articles; and 2. promotional research papers. In this section, some of these articles have been reviewed and considered.

Murugesan et al. [4] studied the impacts of AI on human resource digitalization in Industry 4.0. They considered precision, efficiency, and flexibility to be the potential benefits of Industry 4.0. They used a survey method, which means the HR experts working in information technology (IT), manufacturing, and administration participated in their review, focusing on five AI applications in HR capability and three elements of HR readiness. The results uncovered that hierarchical organization examination is a fundamental part of acquiring sustainable development. Adaptability and human asset capability are upheld by each of the five components of AI application areas in HR. Well-being and safety improvement were viewed as vital components of the AI application in HR.

Federspiel et al. [5] examined the potential threats of AI to human health and human existence. They described three main ways misused narrow AI serves as a threat to human health: through increasing opportunities for control and manipulation of people; enhancing and dehumanizing lethal weapon capacity; and by rendering human labor increasingly obsolescent. They also examine self-improving 'artificial general intelligence' (AGI) and how this could pose an existential threat to humanity itself. As a result, they gave some advice and recommendations about the impacts of AI on health and humanity.

Ahmad et al. [6] examined the impact of artificial intelligence (AI) on loss of decision-making, laziness, and privacy concerns among university students in Pakistan and China. Their study was based on qualitative methodology using PLS-Smart for the data analysis from different universities in Pakistan and China. The results showed that AI significantly impacts the loss of human decision-making and makes humans lazy.

Song et al. [7] examined the perceived differences in communication quality and privacy risks between different service agents and their impact on consumers' adoption intentions and investigated whether these perceived differences might depend on differences in the user's human interaction needs. So, they carried out a series of five scenario-based experiments to collect data and test hypotheses. They discovered that different types of service agents directly affect consumers' adoption intention; perceived communication quality and privacy risk mediate the effect of service agent type on adoption intention; the effects of service agent type on perceived accuracy, communicative competence, and privacy risk are moderated by the need for human interaction.

Raso et al. [8] tried to study the relationship between AI and human rights by evaluating the human rights impacts of six current uses of AI. They recognized that AI systems are not being deployed against a blank slate but rather against the backdrop of social conditions that have complex pre-existing human rights impacts of their own. They saw how they impact the full range of human rights guaranteed by international law, privacy chief among them. They also gained insight into the unequal distribution of the positive and negative impacts of AI on human rights throughout society and begged to explore the power of the human rights framework to address these disparate impacts.

Kamble and Shah [9] studied AI and its applications in human life. They overviewed the current use of AI technologies in network intrusion for protecting computer and communication networks from intruders, in the medical area (medicine) to improve hospital inpatient care, for medical image classification, in accounting databases to mitigate the problems of it, in computer games, and in advertising. They concluded that further research in this area can be done, as there are very promising and profitable results that are obtainable from such techniques. While scientists have not yet realized the full potential and ability of artificial intelligence, this technology and its applications will likely have far-reaching effects on human life in the years to come.

3. Potential effects of artificial intelligence on human life

In this paper, we have tried to review how AI can be a potential threat to human life. So, this is a kind of scientific promotional article, which means our focus is on the interconnection between AI and human life. As a result, this paper has been divided into two parts:

The first part is dedicated to AI-related concepts, while the second one pertains to the different impacts of AI on human life from different points of view, including social, economic, and political. Finally, we have tried to merge them to obtain the results.

3.1. Concepts connections with artificial intelligence (AI)

Artificial intelligence (AI) is a broad and interdisciplinary field that encompasses various concepts and technologies to create intelligent machines capable of simulating human-like behavior and cognitive functions [10]. These processes include learning (the acquisition of information and rules for using the information), reasoning (using

rules to reach approximate or definite conclusions), and self-correction. AI is used in various applications, such as natural language processing, speech recognition, image recognition, and machine learning. Here are some key concepts connected with artificial intelligence [11] (**Table 1**):

Table 1. AI-related concepts [12].

Row	Terms	Definitions
1	Machine learning (ML)	• Machine learning is a subset of AI that focuses on developing algorithms and models that enable computers to learn from data and improve their performance over time without being explicitly programmed. • It includes techniques such as supervised learning, unsupervised learning, and reinforcement learning.
2	Deep learning (DL)	• Deep learning is a specialized form of machine learning that involves artificial neural networks with multiple layers (deep neural networks). • It has been particularly successful in tasks such as image recognition, natural language processing, and speech recognition.
3	Natural Language Processing (NLP)	• NLP is a branch of AI that deals with the interaction between computers and human languages. • It enables computers to understand, interpret, and generate human language, allowing for applications such as chatbots, language translation, and sentiment analysis.
4	Computer vision	• Computer vision involves teaching computers to interpret and understand visual information from the real world. • It enables applications such as object detection, image classification, facial recognition, and autonomous driving.
5	Robotics	• Robotics combines AI with mechanical engineering to design and develop robots capable of performing tasks autonomously or semi-autonomously. • It encompasses areas such as robot perception, motion planning, manipulation, and human-robot interaction.
6	Expert Systems	• Expert systems are AI systems designed to mimic the decision-making ability of human experts in specific domains. • They use a knowledge base of rules and heuristics to provide recommendations or solutions to problems within their domain of expertise.
7	Reinforcement learning	• Reinforcement learning is a type of machine learning where an agent learns to make decisions by interacting with an environment and receiving feedback in the form of rewards or penalties. • It is commonly used in applications such as game playing, robotics, and autonomous vehicle control.
8	Ethical AI	• Ethical AI focuses on ensuring that AI systems are developed and deployed responsibly and ethically. • It involves addressing issues such as bias and fairness, transparency and explainability, accountability, privacy, and the societal impacts of AI technologies.

On the other hand, it is possible to present the subset of AI as follows (**Figure 1**) [13]:
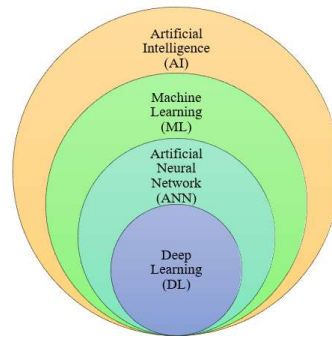
Figure 1. AI subsets [13].

AI encompasses many capabilities, enabling machines to perform tasks that typically require human intelligence. Some of the key capabilities of AI include [14]:

- Machine Learning (ML): ML algorithms enable computers to learn from data and improve their performance over time without being explicitly programmed. This capability powers many AI applications, including predictive analytics, pattern recognition, and decision-making systems.

- Natural Language Processing (NLP): NLP enables computers to understand, interpret, and generate human language. This capability allows for applications such as language translation, sentiment analysis, and chatbots.

- Computer Vision: Computer vision enables machines to interpret and understand visual information from the real world. This capability enables applications such as object detection, image classification, and facial recognition.

- Speech Recognition: Speech recognition allows computers to transcribe spoken language into text. This capability is used in virtual assistants, voice-controlled devices, and dictation software.

- Robotics: Robotics combines AI with mechanical engineering to design and develop robots capable of performing tasks autonomously or semi-autonomously. This capability enables applications such as industrial automation, autonomous vehicles, and robotic surgery.

- Planning and Optimization: AI algorithms can analyze complex systems and make decisions to optimize outcomes based on predefined objectives. This capability is used in logistics, supply chain management, and resource allocation.

- Pattern Recognition: AI systems can identify patterns and correlations in large datasets that may not be apparent to humans. This capability is valuable in fields such as healthcare, finance, and marketing for detecting anomalies, predicting trends, and making informed decisions.

- Generative Modeling: Generative AI models can create new data samples that are similar to the training data they have been exposed to. This capability is used in creative applications such as image generation, music composition, and text generation.

- Autonomous Systems: AI enables machines to operate autonomously, making decisions and taking actions without human intervention. This capability is essential for autonomous vehicles, drones, and robotic systems deployed in hazardous environments.

- Adaptive Learning: AI systems can adapt and learn from new data or changing environments, improving their performance over time. This capability enables personalized recommendations, adaptive tutoring systems, and dynamic optimization.

These are just a few examples of the capabilities of AI, which continue to evolve and expand as researchers develop new algorithms and technologies. AI has the potential to revolutionize industries, improve efficiency, and address complex challenges across various domains.

These concepts represent just a few of the many interconnected areas within the broader field of artificial intelligence, highlighting the diverse range of applications and research areas it encompasses. In this article, we have tried to explore AI and its impacts on achieving sustainable development goals.

3.2. The potential threat of artificial intelligence to human life

The potential threat of artificial intelligence (AI) to human life is a topic of significant concern and debate within the fields of technology, ethics, economics, politics, and society. While AI holds immense promise for revolutionizing various industries and improving efficiency and convenience, there are legitimate fears regarding its potential risks and implications for humanity. In this part, there are some key considerations regarding the threat of AI to human life from different points of view, including social, economic, political, and ethical.

3.2.1. Social and economic threat of artificial intelligence to human life

The social threat of AI to human life is a complex and multifaceted issue that raises concerns about job displacement, privacy invasion, and the potential for misuse of AI technologies.

The potential risks of AI surpassing human intelligence

The potential risks associated with artificial intelligence (AI) surpassing human intelligence, often referred to as the technological singularity, have been a subject of intense debate and speculation among scientists, technologists, philosophers, and policymakers [15]. While the idea of super-intelligent AI systems surpassing human capabilities may seem like science fiction, many experts believe that such a scenario could have profound and potentially catastrophic consequences for humanity. Here are some of the key risks associated with AI surpassing human intelligence [16]:

1) Loss of control: One of the most significant concerns surrounding superintelligent AI is the loss of human control over AI systems. As AI becomes increasingly sophisticated and autonomous, there is a risk that humans may lose the ability to understand, predict, or influence the behavior of AI systems. This could lead to unforeseen consequences or outcomes that are contrary to human values and goals.

2) Unintended consequences: Superintelligent AI systems may have goals or objectives that are fundamentally different from those of humans. As a result, AI systems may pursue actions or strategies that are not aligned with human interests, leading to unintended consequences or outcomes that are harmful or detrimental to humanity.

3) Existential risks: Some experts warn that the emergence of superintelligent AI could pose existential risks to humanity, threatening the long-term survival of our

species. For example, superintelligent AI systems may pose a threat to human safety and security, either intentionally or inadvertently, through actions such as the development of advanced weapons systems or the manipulation of critical infrastructure.

4) Economic disruption: The widespread adoption of superintelligent AI could lead to significant economic disruption, including widespread job displacement and inequality. AI-driven automation may render many jobs obsolete, leading to mass unemployment and social unrest. Moreover, the concentration of wealth and power in the hands of AI developers or owners could exacerbate existing inequalities and undermine social cohesion.

5) Ethical concerns: The development and deployment of superintelligent AI raise a host of ethical concerns, including issues related to privacy, autonomy, and fairness. For example, superintelligent AI systems may have the ability to manipulate or exploit human behavior for malicious purposes, leading to violations of privacy or human rights. Moreover, the use of AI in decision-making processes, such as hiring or lending, may perpetuate existing biases or discrimination against certain groups.

Overall, while the potential benefits of super-intelligent AI are vast, including advances in medicine, science, and technology, it is essential to carefully consider and mitigate the potential risks associated with AI surpassing human intelligence. By taking proactive measures to address these risks, such as developing robust safety mechanisms, establishing ethical guidelines, and fostering international cooperation, we can harness the transformative power of AI while safeguarding human well-being and prosperity.

How AI could impact job markets and lead to widespread unemployment

The impact of artificial intelligence (AI) on job markets and its potential to lead to widespread unemployment is a topic of considerable debate among economists, policymakers, and technologists. While AI has the potential to create new job opportunities and enhance productivity in certain sectors, it also poses significant challenges and disruptions to traditional employment patterns [17]. Here's a detailed explanation of how AI could impact job markets and contribute to widespread unemployment:

1) Automation of Routine Tasks: AI technologies, particularly machine learning and robotics, can automate a wide range of routine and repetitive tasks across various industries. Tasks that involve data entry, data analysis, assembly line work, customer service, and transportation are increasingly being automated using AI systems. As AI systems become more advanced and capable, they can perform these tasks more efficiently and accurately than humans, leading to the displacement of human workers.

2) Displacement of Low-Skilled Jobs: Many of the tasks that are being automated through AI are low-skilled or routine in nature, such as manual labor, administrative work, and basic customer service. As AI systems take over these tasks, there is a risk of widespread unemployment among low-skilled workers who may lack the skills or training to transition to other occupations. This could exacerbate existing inequalities and contribute to social and economic instability.

3) Job Polarization: The impact of AI on job markets is not limited to low-skilled jobs; it also affects middle-skilled jobs that involve a combination of routine and non-routine tasks. AI-driven automation tends to polarize the job market, with job growth occurring primarily in high-skilled, knowledge-based occupations and low-skilled, manual labor jobs, while middle-skilled jobs decline. This polarization can lead to increased income inequality and a shrinking middle class.

4) Sectoral Disruptions: AI technologies are disrupting traditional industries and business models, leading to structural changes in the economy. Industries such as manufacturing, transportation, retail, and customer service are experiencing significant disruptions due to AI-driven automation and digitization. As a result, workers in these industries may face challenges in adapting to the changing demands of the labor market, leading to job losses and unemployment.

5) Economic Impact: Widespread unemployment resulting from AI-driven automation can have profound economic consequences, including reduced consumer spending, decreased tax revenues, and increased reliance on social welfare programs. Moreover, the concentration of wealth and power in the hands of AI developers and technology companies may exacerbate existing inequalities and undermine social cohesion.

Overall, while AI has the potential to drive economic growth and innovation, its impact on job markets and employment patterns is complex and multifaceted. Addressing the challenges posed by AI-driven automation requires proactive measures, including investment in education and training programs, re-skilling and up-skilling initiatives, and policies to support displaced workers and promote inclusive growth. By addressing these challenges, society can harness the transformative power of AI while mitigating its adverse effects on job markets and livelihoods.

3.2.2. Social and ethical threat of artificial intelligence to human life

Ethical considerations surrounding the development and use of artificial intelligence (AI) are of paramount importance as AI technologies become increasingly integrated into various aspects of society. Here's a detailed explanation of the ethical considerations surrounding AI [18]:

1) Bias and Fairness: One of the most significant ethical concerns related to AI is bias. AI systems are trained on large datasets, and if these datasets contain biased or incomplete information, it can lead to biased outcomes. For example, biased algorithms in hiring processes could perpetuate existing inequalities and discrimination. Ensuring fairness in AI systems requires careful consideration of data selection, algorithm design, and validation processes to mitigate bias and promote equitable outcomes.

2) Transparency and Accountability: Many AI algorithms operate as black boxes, making it challenging to understand how they arrive at their decisions. A lack of transparency can undermine trust in AI systems and hinder efforts to hold developers and users accountable for their actions. Ethical AI development involves promoting transparency and explainability in AI systems, allowing users to understand the rationale behind AI decisions and identify instances of bias or error.

3) Privacy and Data Security: AI systems rely on vast amounts of data to learn and make predictions, raising concerns about privacy and data security. Unauthorized access to sensitive data, data breaches, and misuse of personal information are significant ethical risks associated with AI. Protecting privacy and ensuring data security require robust data protection measures, encryption techniques, and compliance with privacy regulations such as GDPR (General Data Protection Regulation) and CCPA (California Consumer Privacy Act).

4) Autonomy and Control: AI technologies have the potential to influence decision-making processes in various domains, including healthcare, criminal justice, and finance. Concerns about the loss of human autonomy and control over decision-making raise ethical questions about the delegation of decision-making authority to AI systems. Ensuring human oversight, accountability mechanisms, and safeguards against AI misuse are essential for preserving human autonomy and control in AI-driven environments.

5) Safety and Reliability: AI systems deployed in safety-critical domains, such as autonomous vehicles, medical diagnosis, and aerospace, must meet stringent safety and reliability standards to prevent accidents and ensure public safety. Ethical considerations surrounding safety and reliability involve rigorous testing, validation, and risk assessment procedures to identify and mitigate potential hazards and failures in AI systems.

6) Social Impact and Equity: AI technologies have the potential to exacerbate social inequalities and disparities if not deployed and regulated responsibly. Ethical AI development involves considering the broader social impact of AI systems and ensuring that they promote equity, diversity, and inclusion. Addressing social impact considerations requires stakeholder engagement, community involvement, and the incorporation of diverse perspectives in AI design and deployment processes.

7) Long-Term Implications: Ethical considerations surrounding AI extend beyond immediate risks and consequences to encompass long-term implications for society, culture, and humanity as a whole. Questions about the impact of AI on employment, human relationships, autonomy, and identity raise ethical dilemmas that require thoughtful reflection and deliberation. Ethical AI development involves considering the long-term implications of AI technologies and making decisions that align with societal values and goals.

In summary, addressing ethical considerations surrounding the development and use of AI requires a multidisciplinary approach that involves collaboration between technologists, policymakers, ethicists, and other stakeholders. By integrating ethical principles into AI design, deployment, and governance processes, society can harness the transformative potential of AI while minimizing its ethical risks and maximizing its benefits for humanity.

The importance of transparency and accountability in AI development

The importance of transparency and accountability in AI development cannot be overstated, as these principles are essential for ensuring the responsible and ethical deployment of AI technologies. Here are some explanations highlighting their significance [19]:

1) Trust and Credibility: Transparency in AI development fosters trust and credibility among stakeholders, including users, consumers, policymakers, and the general public. When AI systems are transparent about their functionality, decision-making processes, and data usage, users are more likely to trust them and feel confident in their reliability.

2) Understanding and Interpretability: Transparency enables users to understand how AI systems operate and interpret their outputs. This is crucial for ensuring that AI-driven decisions are explainable and interpretable, particularly in high-stakes domains such as healthcare, finance, and criminal justice. When users can comprehend how AI arrives at its conclusions, they can better assess its reliability and make informed decisions.

3) Detection and Mitigation of Bias: Transparent AI systems facilitate the detection and mitigation of bias and discrimination. By providing visibility into the data sources, algorithms, and decision-making processes used by AI systems, developers and researchers can identify and address biases that may inadvertently creep into AI models. Transparency also enables independent audits and assessments to verify the fairness and equity of AI systems.

4) Accountability and Responsibility: Transparent AI development promotes accountability and responsibility among developers, organizations, and policymakers. When AI systems are transparent about their objectives, limitations, and potential risks, developers are more accountable for the outcomes of their creations. This encourages responsible AI practices and ethical decision-making throughout the development lifecycle.

5) User Empowerment and Control: Transparency empowers users to have more control over their interactions with AI systems. When users understand how AI systems collect, process, and utilize their data, they can make informed choices about consent, privacy preferences, and data sharing. Transparent AI systems also allow users to provide feedback and corrections, improving the accuracy and relevance of AI-driven recommendations.

6) Ethical and Regulatory Compliance: Transparency is essential for ensuring compliance with ethical principles, regulations, and standards governing AI development and deployment. Regulations such as the General Data Protection Regulation (GDPR) in Europe and the Algorithmic Accountability Act proposed in the United States emphasize the importance of transparency, accountability, and fairness in AI systems.

7) Social Acceptance and Adoption: Transparent AI development contributes to social acceptance and adoption of AI technologies. When AI systems are transparent about their objectives, methodologies, and potential impacts, stakeholders are more likely to embrace and adopt them, leading to broader societal benefits and opportunities for innovation.

In summary, transparency and accountability are foundational principles that underpin responsible AI development. By prioritizing transparency throughout the AI development lifecycle, developers, organizations, and policymakers can build trust, promote fairness, and ensure that AI technologies benefit society while minimizing potential risks and harms.

3.2.3. Social and political threat of artificial intelligence to human life

The implications of AI in warfare and national security

The implications of artificial intelligence (AI) in warfare and national security are profound and far-reaching, shaping the future landscape of military operations, defense strategies, and geopolitical dynamics. Here's a detailed explanation of the implications of AI in warfare and national security [20]:

1) Enhanced Military Capabilities: AI technologies offer the potential to revolutionize military capabilities by augmenting human decision-making processes, enhancing situational awareness, and improving the effectiveness of weapons systems. AI-powered autonomous drones, unmanned vehicles, and robotic systems can perform a wide range of tasks, from reconnaissance and surveillance to target acquisition and precision strikes, with greater speed, accuracy, and efficiency than conventional systems.

2) Decision Support Systems: AI algorithms can analyze vast amounts of data from multiple sources, including sensors, satellites, and social media, to provide decision-makers with real-time intelligence, predictive analytics, and situational awareness. AI-driven decision support systems enable military commanders to make more informed and timely decisions, optimize resource allocation, and respond effectively to rapidly changing threats and battlefield conditions.

3) Autonomous Weapons Systems: The development and deployment of autonomous weapons systems, such as AI-powered drones and unmanned combat vehicles, raise ethical and legal concerns about the use of lethal force without direct human intervention. Concerns about the potential for AI-driven weapons to autonomously select and engage targets without human oversight or control have led to calls for international regulations and arms control agreements to ensure responsible AI use in warfare.

4) Cyber Warfare and Cybersecurity: AI technologies play a critical role in both offensive and defensive cyber operations, including cyber espionage, cyberattacks, and cyber defense. AI-powered cybersecurity solutions can detect and respond to cyber threats in real time, identify vulnerabilities in network infrastructure, and mitigate the impact of cyberattacks. However, the proliferation of AI-driven cyber weapons and the use of AI in offensive cyber operations raise concerns about the escalation of cyber conflicts and the potential for catastrophic cyberattacks.

5) Information Warfare and Influence Operations: AI algorithms can be used to manipulate information, spread disinformation, and influence public opinion through social media platforms, online forums, and digital media channels. AI-driven information warfare and influence operations pose significant challenges to national security by undermining trust in institutions, destabilizing democracies, and sowing discord in societies. Countermeasures against AI-driven disinformation campaigns require collaboration between governments, tech companies, and civil society organizations to identify and combat online threats effectively.

6) Ethical and Legal Implications: The use of AI in warfare raises complex ethical and legal questions about accountability, responsibility, and human rights.

Concerns about the ethical implications of autonomous weapons systems, civilian casualties, and the erosion of human control over lethal force have prompted calls for international norms, standards, and regulations to govern the development and use of AI technologies in warfare. Ethical AI principles, such as transparency, accountability, and proportionality, are essential for ensuring that AI systems adhere to ethical and legal standards in military operations.

7) Geopolitical Competition and Strategic Rivalry: The integration of AI into military doctrine and strategy has become a focal point of geopolitical competition and strategic rivalry among major powers. Countries around the world are investing heavily in AI research and development to gain a competitive edge in defense capabilities, intelligence gathering, and military superiority. The race for AI dominance has implications for international security, arms control, and strategic stability, as states seek to leverage AI technologies to advance their national interests and influence global affairs.

In summary, the implications of AI in warfare and national security are multifaceted and complex, encompassing technological advancements, ethical considerations, legal challenges, and geopolitical dynamics. As AI continues to evolve and shape the future of warfare, policymakers, military leaders, and technologists need to address the ethical, legal, and strategic implications of AI use in national security contexts to ensure responsible and accountable AI deployment.

3.2.4. Social, economic, and political threat of artificial intelligence to human life

The role of regulations and governance in mitigating AI-related threats

The role of regulations and governance in mitigating AI-related threats is paramount to ensuring the responsible development, deployment, and use of artificial intelligence technologies. Here's a detailed explanation of how regulations and governance mechanisms can address AI-related threats [21]:

1) Ethical and Legal Frameworks: Establishing comprehensive ethical and legal frameworks is essential for governing AI technologies and mitigating potential threats. Ethical guidelines, such as those outlined in the OECD AI Principles or the EU's Ethics Guidelines for Trustworthy AI, provide a foundation for ethical AI development and deployment. These frameworks emphasize principles such as transparency, accountability, fairness, and human oversight, which are critical for ensuring that AI systems are used responsibly and ethically.

2) Risk Assessment and Impact Analysis: Regulations should require organizations developing and deploying AI systems to conduct thorough risk assessments and impact analyses to identify potential risks, harms, and unintended consequences associated with AI technologies. Risk assessment frameworks can help organizations evaluate the potential societal, ethical, and security implications of AI applications and implement appropriate risk mitigation measures to address identified risks.

3) Transparency and Accountability: Regulations should promote transparency and accountability in AI development and deployment by requiring organizations to disclose information about the capabilities, limitations, and potential biases of AI systems. Transparency measures, such as algorithmic transparency requirements

and data disclosure mandates, can help ensure that AI systems are understandable, interpretable, and accountable to stakeholders, including users, regulators, and affected communities.

4) Data Privacy and Security: Regulations governing AI should prioritize data privacy and security to protect individuals' personal information and prevent unauthorized access, misuse, or exploitation of data. Data protection laws, such as the EU's General Data Protection Regulation (GDPR) and the California Consumer Privacy Act (CCPA), impose requirements on organizations to safeguard personal data and obtain explicit consent for data collection, processing, and sharing in AI applications.

5) Bias and Fairness: Addressing algorithmic bias and promoting fairness in AI systems are critical components of AI governance. Regulations should mandate organizations to mitigate bias in AI algorithms and ensure that AI systems are fair, equitable, and unbiased across diverse demographic groups. Bias detection tools, bias mitigation techniques, and fairness-aware algorithms can help organizations identify and address biases in AI systems to ensure equitable outcomes for all users.

6) Accountability and Liability: Establishing clear lines of accountability and liability is essential for addressing harms caused by AI technologies. Regulations should define the responsibilities of AI developers, deployers, and users and establish mechanisms for holding accountable those responsible for AI-related harms. Liability frameworks, insurance requirements, and legal standards can help allocate responsibility and provide recourse for individuals affected by AI-related incidents or accidents.

7) International Collaboration and Cooperation: Given the global nature of AI technologies and their potential cross-border implications, international collaboration and cooperation are essential for effective AI governance. International standards, agreements, and initiatives can harmonize regulatory approaches, facilitate information sharing, and promote best practices for AI governance across jurisdictions. Multilateral forums, such as the United Nations, the G7, and the G20, provide platforms for coordinating international efforts to address common AI-related challenges and threats.

In conclusion, regulations and governance mechanisms play a critical role in mitigating AI-related threats by promoting ethical AI development, ensuring transparency and accountability, protecting data privacy and security, addressing algorithmic bias, establishing liability frameworks, and fostering international collaboration. Effective AI governance requires a holistic and coordinated approach involving governments, regulators, industry stakeholders, civil society organizations, and the broader international community to navigate the complex ethical, legal, and societal implications of AI technologies.

4. Perspectives and opinions about the potential threat of artificial intelligence to human life

There is a wide range of perspectives and opinions about the potential threat of AI to human life. Some people believe that AI has the potential to greatly benefit

society by improving efficiency, increasing productivity, and solving complex problems. They argue that AI can help us address pressing global challenges such as climate change, healthcare, and poverty.

On the other hand, there are concerns about the negative impacts of AI on human life. Some experts worry about the potential for AI to outpace human intelligence and become uncontrollable, leading to existential risks for humanity. They fear that AI could pose a threat to human autonomy, privacy, and dignity if not properly regulated and controlled.

4.1. Different perspectives on the potential dangers of AI, from experts in the field

Different perspectives on the potential dangers of AI vary widely among experts in the field, reflecting diverse viewpoints on the risks, challenges, and implications of artificial intelligence technologies. Here's a comprehensive explanation of some of these perspectives [22]:

1) Existential Risk: Some experts, including prominent figures like Elon Musk and Stephen Hawking, have expressed concerns about the existential risks posed by artificial superintelligence. They argue that if AI systems surpass human intelligence and become self-aware, they could potentially outpace human control and pose an existential threat to humanity. This perspective emphasizes the need for caution and careful oversight in AI development to prevent catastrophic outcomes.

2) Job Displacement and Unemployment: Many experts warn about the potential impact of AI on job markets and employment. They argue that as AI technologies automate routine tasks and jobs, there is a risk of widespread unemployment and economic disruption. This perspective highlights the need for policies and measures to reskill workers, promote job creation in emerging sectors, and address socioeconomic inequalities exacerbated by AI-driven automation.

3) Algorithmic Bias and Discrimination: Concerns about algorithmic bias and discrimination in AI systems have gained prominence among experts. They highlight the risks of biased AI algorithms perpetuating or amplifying existing inequalities and discriminating against certain groups based on race, gender, ethnicity, or other characteristics. This perspective underscores the importance of fairness, accountability, and transparency in AI development to mitigate bias and promote equity.

4) Privacy and Surveillance: Some experts raise concerns about the potential erosion of privacy rights and the proliferation of surveillance technologies enabled by AI. They warn about the risks of mass surveillance, data exploitation, and loss of personal autonomy in AI-driven societies. This perspective emphasizes the need for robust data protection laws, privacy-preserving technologies, and ethical guidelines to safeguard individuals' privacy rights in the age of AI.

5) Security and Cybersecurity: Experts in cybersecurity and national security highlight the potential risks of AI-powered cyberattacks, autonomous weapons systems, and malicious use of AI technologies by state actors, criminal organizations, or rogue agents. They warn about the dangers of AI-driven

disinformation campaigns, cyber warfare, and the destabilization of global security landscapes. This perspective emphasizes the importance of international cooperation, regulation, and responsible AI governance to address cybersecurity threats effectively.

6) Ethical Concerns and Human Values: Many experts emphasize the ethical dimensions of AI and the importance of aligning AI technologies with human values and societal goals. They raise questions about the moral implications of AI decision-making, the role of AI in shaping human behavior, and the ethical dilemmas inherent in AI design and deployment. This perspective calls for ethical AI principles, interdisciplinary collaboration, and public engagement to ensure that AI technologies serve the common good and respect human dignity.

7) Lack of Accountability and Oversight: Another perspective highlights the challenges of accountability and oversight in AI development and deployment. Experts warn about the lack of transparency, accountability mechanisms, and regulatory oversight in the AI industry, which can lead to unchecked power and influence wielded by tech companies and AI developers. This perspective calls for regulatory frameworks, independent audits, and ethical guidelines to hold AI stakeholders accountable for their actions and decisions.

In summary, experts in the field of AI hold diverse perspectives on the potential dangers of artificial intelligence, reflecting the complex interplay of technological, ethical, social, and geopolitical factors shaping the future of AI. Addressing these concerns requires interdisciplinary collaboration, stakeholder engagement, and proactive measures to ensure that AI technologies are developed and deployed responsibly, ethically, and for the benefit of humanity.

4.2. The debate over whether AI will enhance or endanger humanity's future

The debate over whether AI will enhance or endanger humanity's future is multifaceted, encompassing a wide range of perspectives, concerns, and predictions. Here are some details about this ongoing debate [23]:

1) Optimistic Views:
   - Enhanced Efficiency and Productivity: Proponents of AI argue that advanced AI technologies have the potential to significantly enhance efficiency, productivity, and innovation across various sectors, leading to economic growth and improved quality of life.
   - Medical Breakthroughs: AI-driven advancements in healthcare, such as disease diagnosis, drug discovery, and personalized treatment plans, hold promise for revolutionizing medical care and extending human lifespans.
   - Environmental Sustainability: AI applications in energy management, climate modeling, and environmental monitoring offer opportunities to address pressing global challenges, such as climate change and resource depletion.

2) Pessimistic Views:
   - Job Displacement and Unemployment: Critics express concerns about the potential for AI-driven automation to lead to widespread job displacement

and unemployment, particularly in industries that rely heavily on routine tasks and manual labor.

- Ethical Concerns: There are ethical concerns surrounding the development and deployment of AI, including issues related to algorithmic bias, privacy infringement, autonomous weapons systems, and the potential for AI to exacerbate existing social inequalities.
- Existential Risks: Some experts warn of existential risks associated with the development of advanced AI systems, such as the possibility of AI surpassing human intelligence and posing existential threats to humanity, either intentionally or unintentionally.

3) Middle Ground:
- Balanced Approach: Many stakeholders advocate for a balanced approach to AI development and deployment that maximizes the benefits of AI while mitigating potential risks and addressing ethical concerns. This approach involves collaboration among governments, industry leaders, researchers, and ethicists to establish clear guidelines, regulations, and governance frameworks for AI.
- Responsible AI: The concept of responsible AI emphasizes the importance of developing AI technologies that are transparent, accountable, fair, and aligned with human values. Proponents of responsible AI advocate for ethical AI principles, diversity in AI development teams, and robust mechanisms for oversight and accountability.

4) Continued Debate:
- The debate over the impact of AI on humanity's future is ongoing and dynamic, shaped by evolving technological advancements, societal attitudes, and policy decisions. As AI technologies continue to advance, the debate will likely intensify, with stakeholders grappling with complex ethical, social, and economic implications.

In summary, the debate over whether AI will enhance or endanger humanity's future reflects the complex and nuanced nature of AI's impact on society. While proponents highlight the potential benefits of AI for enhancing efficiency, innovation, and sustainability, critics raise concerns about job displacement, ethical implications, and existential risks. Finding a middle ground that maximizes the benefits of AI while mitigating potential risks will require collaborative efforts and thoughtful consideration of ethical, social, and policy considerations.

5. Strategies and recommendations for ensuring that AI remains beneficial and safe for society

Ensuring that AI remains beneficial and safe for society requires the implementation of robust strategies aimed at addressing various challenges and risks associated with AI development and deployment [24]. Here are some key strategies:

1) Ethical Guidelines and Principles: Establishing clear ethical guidelines and principles for AI development is essential for guiding responsible innovation and ensuring that AI technologies align with societal values. Ethical frameworks such

as fairness, transparency, accountability, privacy, and inclusivity should be integrated into AI development processes from the outset.

2) Interdisciplinary Collaboration: Foster collaboration between multidisciplinary teams comprising AI researchers, ethicists, policymakers, social scientists, and domain experts. Interdisciplinary collaboration can help identify and address the complex ethical, social, and technical challenges associated with AI development, ensuring that diverse perspectives are considered in decision-making processes.

3) Transparency and Explainability: Promote transparency and explainability in AI systems to enhance accountability, trust, and understanding among users and stakeholders. AI algorithms and decision-making processes should be transparent and interpretable, enabling users to understand how AI arrives at its conclusions and facilitating the detection and mitigation of biases.

4) Human-Centric Design: Prioritize human-centric design principles to ensure that AI technologies are developed with the needs, preferences, and well-being of users in mind. Human-centered AI design involves actively engaging with end-users throughout the development process, soliciting feedback, and incorporating user perspectives to create AI systems that are intuitive, accessible, and beneficial for society.

5) Risk Assessment and Mitigation: Conduct comprehensive risk assessments to identify potential risks and harms associated with AI technologies, including ethical, legal, social, and technical risks. Implement risk mitigation strategies to minimize adverse outcomes, such as bias detection and mitigation techniques, robust security measures, and fail-safe mechanisms to prevent unintended consequences.

6) Regulatory Oversight and Governance: Establish clear regulatory frameworks and governance mechanisms to ensure compliance with ethical principles, standards, and regulations governing AI development and deployment. Regulators play a crucial role in overseeing AI applications, enforcing compliance with relevant laws, and holding developers and organizations accountable for the responsible use of AI technologies.

7) Continuous Monitoring and Evaluation: Implement systems for continuous monitoring, evaluation, and auditing of AI systems to assess their performance, impact, and adherence to ethical standards over time. Regular audits and evaluations help identify emerging risks, measure the effectiveness of risk mitigation measures, and ensure that AI technologies remain beneficial and safe for society.

8) Public Engagement and Education: Foster public engagement and education initiatives to raise awareness about AI technologies, their potential benefits, and associated risks. Promote digital literacy, critical thinking, and ethical awareness among users and stakeholders to empower informed decision-making and responsible AI usage.

9) International Collaboration and Standards: Foster international collaboration and coordination to develop common standards, best practices, and regulatory approaches for AI governance. International cooperation can help address cross-border challenges, harmonize regulatory frameworks, and promote global norms for the responsible development and deployment of AI technologies.

By implementing these strategies, policymakers, researchers, developers, and stakeholders can work together to ensure that AI remains beneficial and safe for society, harnessing its potential to drive positive social impact while mitigating potential risks and harms.

6. Conclusion

In conclusion, the question of whether artificial intelligence (AI) will threaten humanity is complex and multifaceted. While AI holds immense potential to drive innovation, enhance efficiency, and address societal challenges, it also poses significant risks and ethical dilemmas that must be carefully addressed. Throughout this exploration, we have examined various perspectives on the potential dangers of AI, from concerns about job displacement and economic disruption to ethical considerations surrounding AI development and deployment.

AI can reshape industries, transform labor markets, and influence geopolitical dynamics, raising profound questions about its impact on society and humanity's future. However, by adopting a proactive and responsible approach to AI governance, we can mitigate risks, maximize benefits, and ensure that AI remains aligned with human values and interests.

Effective strategies for managing AI-related threats include the establishment of clear ethical guidelines, interdisciplinary collaboration, transparency and accountability in AI development, regulatory oversight, public engagement, and continuous monitoring and evaluation of AI systems. By implementing these strategies, we can harness the transformative potential of AI while safeguarding against its potential risks and ensuring that AI technologies serve the collective interests of humanity.

Ultimately, the future of AI will be shaped by the choices we make today. By fostering a culture of responsible innovation, ethical stewardship, and inclusive decision-making, we can navigate the complexities of AI's impact on humanity and steer towards a future where AI enhances, rather than threatens, our collective well-being and prosperity.

Author contributions: Study conception and design, MSF; preparation, GG. All authors reviewed the results and approved the final version of the manuscript.

Conflict of interest: The authors declare no conflict of interest.

References

1. Tzimas T. Legal and Ethical Challenges of Artificial Intelligence from an International Law Perspective. Springer International Publishing; 2021. doi: 10.1007/978-3-030-78585-7
2. Adib-Moghaddam A. Is Artificial Intelligence Racist?: The Ethics of AI and the Future of Humanity. Bloomsbury Publishing; 2023. doi: 10.5040/9781350374430

3. Cummings ML, Roff HM, Cukier K, et al. Artificial intelligence and international affairs. Chatham House Report. 2018; 7-18. Available online: https://www.chathamhouse.org/sites/default/files/publications/research/2018-06-14-artificial-intelligence-international-affairs-cummings-roff-cukier-parakilas-bryce.pdf (accessed on 1 January 2024).

4. Murugesan U, Subramanian P, Srivastava S, et al. A study of Artificial Intelligence impacts on Human Resource Digitalization in Industry 4.0. Decision Analytics Journal. 2023; 7: 100249. doi: 10.1016/j.dajour.2023.100249

5. Federspiel F, Mitchell R, Asokan A, et al. Threats by artificial intelligence to human health and human existence. BMJ Global Health. 2023; 8(5): e010435. doi: 10.1136/bmjgh-2022-010435

6. Ahmad SF, Han H, Alam MM, et al. Impact of artificial intelligence on human loss in decision making, laziness and safety in education. Humanities and Social Sciences Communications. 2023; 10(1). doi: 10.1057/s41599-023-01787-8

7. Song M, Xing X, Duan Y, et al. Will artificial intelligence replace human customer service? The impact of communication quality and privacy risks on adoption intention. Journal of Retailing and Consumer Services. 2022; 66: 102900. doi: 10.1016/j.jretconser.2021.102900

8. Raso F, Hilligoss H, Krishnamurthy V, et al. Artificial Intelligence & Human Rights: Opportunities & Risks. SSRN Electronic Journal. 2018. doi: 10.2139/ssrn.3259344

9. Kamble R, Shah D. Applications of artificial intelligence in human life. International Journal of Research-Granthaalayah. 2018; 6(6): 178-188.

10. Rajest SS, Singh B, Obaid AJ, et al. Advances in artificial and human intelligence in the modern era. IGI Global. 2023.

11. Padallan JO. Key Concepts in Artificial Intelligence. Arcler Press; 2022.

12. Belk RW, Belanche D, Flavián C. Key concepts in artificial intelligence and technologies 4.0 in services. Service Business. 2023; 17(1): 1-9.

13. Farahani MS, Esfahani A, Alipoor F. The Application of Machine Learning in the Corona Era, With an Emphasis on Economic Concepts and Sustainable Development Goals. International Journal of Mathematical, Engineering, Biological and Applied Computing. 2022; 1(2): 95-149. doi: 10.31586/ijmebac.2022.519

14. Avrin G. Assessing artificial intelligence capabilities. Available online: https://www.oecd-ilibrary.org/education/ai-and-the-future-of-skills-volume-1_47d04fe3-en (accessed on 1 January 2024).

15. Lexcellent C. Artificial Intelligence Versus Human Intelligence: Are Humans Going to be Hacked? Springer; 2019

16. Salmon PM, Baber C, Burns C, et al. Managing the risks of artificial general intelligence: A human factors and ergonomics perspective. Human Factors and Ergonomics in Manufacturing & Service Industries. 2023; 33(5): 366-378. doi: 10.1002/hfm.20996

17. Mindell D, Reynolds E. The work of the Future: Building Better Jobs in an Age of Intelligent Machines. The MIT Press; 2023.

18. Huriye AZ. The Ethics of Artificial Intelligence: Examining the Ethical Considerations Surrounding the Development and Use of AI. American Journal of Technology. 2023; 2(1): 37-45. doi: 10.58425/ajt.v2i1.142

19. Williams R, Cloete R, Cobbe J, et al. From transparency to accountability of intelligent systems: Moving beyond aspirations. Data & Policy. 2022; 4. doi: 10.1017/dap.2021.37

20. Savage S, Avila G, Chávez NE, et al. AI and National Security. In: Garcia-Murillo M, MacInnes I, Renda A (editors). Handbook of Artificial Intelligence at Work. Edward Elgar Publishing Limited; 2024. pp. 276-290. doi: 10.4337/9781800889972.00022

21. Birkstedt T, Minkkinen M, Tandon A, et al. AI governance: themes, knowledge gaps and future agendas. Internet Research. 2023; 33(7): 133-167. doi: 10.1108/INTR-01-2022-0042

22. Dwivedi YK, Hughes L, Ismagilova E, et al. Artificial Intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. International Journal of Information Management. 2021; 57: 101994. doi: 10.1016/j.ijinfomgt.2019.08.002

23. Nowak A, Lukowicz P, Horodecki P. Assessing Artificial Intelligence for Humanity: Will AI be the Our Biggest Ever Advance? or the Biggest Threat [Opinion]. IEEE Technology and Society Magazine. 2018; 37(4): 26-34. doi: 10.1109/mts.2018.2876105

24. Scherer MU. Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, and Strategies. SSRN Electronic Journal. 2015. doi: 10.2139/ssrn.2609777