

Challenges of Artificial Intelligence in Preserving Fundamental Human Values

تحديات الذكاء الاصطناعي في الحفاظ على القيم الإنسانية الأساسية

Ouahabi Naziha

University of Blida2 (Algeria),
n.ouahabi@univ-blida2.dz

Date of submission:02/04/2025 Date of final acceptance:04/06/2025 Date of publication :juine 2025

Abstract

This research aims to explore the challenges of artificial intelligence in preserving fundamental human values, with a focus on the impact of these technologies on justice, equality, and individual rights. The study showed that the use of artificial intelligence can lead to unexpected outcomes that negatively affect human values if these challenges are not properly addressed. The research also recommended the development of clear ethical and legal frameworks to ensure the responsible use of artificial intelligence technologies. By enhancing transparency and community engagement, a better balance can be achieved between technological innovation and the preservation of fundamental human values. These findings are crucial for understanding how to integrate artificial intelligence into our societies in a way that protects and enhances human values.

Keywords : Artificial Intelligence; Human Values; Social Responsibility;

ملخص:

يهدف هذا البحث إلى استكشاف تحديات الذكاء الاصطناعي في الحفاظ على القيم الإنسانية الأساسية، مع التركيز على تأثير هذه التقنيات على العدالة والمساواة وحقوق الأفراد. أظهرت الدراسة أن استخدام الذكاء الاصطناعي يمكن أن يؤدي إلى نتائج غير متوقعة تؤثر سلباً على القيم الإنسانية إذا لم يتم التعامل مع هذه التحديات بشكل مناسب. كما أوصى البحث بضرورة تطوير أطر أخلاقية وقانونية واضحة لضمان استخدام تقنيات الذكاء الاصطناعي بشكل مسؤول. من خلال تعزيز الشفافية والمشاركة المجتمعية، يمكن تحقيق توازن أفضل بين الابتكار التكنولوجي والحفاظ على القيم الإنسانية الأساسية. تعتبر هذه النتائج حيوية لفهم كيفية دمج الذكاء الاصطناعي في مجتمعاتنا بطريقة تحمي وتعزز القيم الإنسانية.

الكلمات المفتاحية: الذكاء الاصطناعي ؛ القيم الإنسانية ؛ المسؤولية الاجتماعية ؛

Auteur correspondant : Ouahabi Naziha

Introduction:

The rapid advancement of artificial intelligence (AI) has significantly transformed various aspects of our lives, ranging from healthcare to education, finance, and beyond. As AI technologies continue to evolve at an unprecedented pace, they are reshaping industries, influencing global economies, and even altering the way we interact with the world. However, alongside these remarkable innovations, AI's integration into our daily lives raises important questions regarding its impact on fundamental human values such as justice, equality, and individual rights. These values form the bedrock of our social systems and legal frameworks, and their preservation is essential to ensuring a just and equitable society. As such, it is crucial to explore how AI's rapid development aligns with these core principles and to identify the potential risks and benefits it brings to these values.

One of the central concerns in this discussion is how AI technologies, which are often driven by data and algorithms, may inadvertently perpetuate or even exacerbate existing social inequalities. For instance, AI systems can reflect the biases embedded in the data they are trained on, which may lead to discriminatory outcomes, especially in areas such as criminal justice, hiring practices, and healthcare access. These biases, whether intentional or unintentional, have the potential to undermine the very essence of fairness and equality, as they may disproportionately affect marginalized groups. Moreover, the use of AI in decision-making processes, such as automated sentencing in courts or algorithmic recruitment tools, raises significant concerns about accountability and transparency. Who is responsible for the decisions made by these AI systems? How can we ensure that AI operates in a way that is both transparent and aligned with ethical standards?

The primary question that emerges from these considerations is: **How can we develop and implement AI technologies in a manner that upholds and reinforces the core values of justice, equality, and individual rights?** This question lies at the heart of the ongoing debate surrounding the ethical implications of AI and the need for a regulatory framework that ensures responsible development and use of these technologies.

This study aims to address these critical issues by examining the influence of AI on human values and identifying the challenges it presents. Among the key concerns are data bias, privacy violations, lack of transparency in decision-making, and the potential for unintended consequences, such as the reinforcement of harmful stereotypes or the erosion of personal freedoms. The methodology used in this study includes a comprehensive review of existing literature, case studies of AI applications, and an analysis of ethical frameworks and regulations. By exploring these areas, the study seeks to provide a clearer understanding of both the opportunities AI offers and the risks it poses to fundamental human values.

The study is structured into several key sections. First, it delves into the complex relationship between AI and human values, exploring how these technologies interact with and challenge core ethical principles. Second, it analyzes the challenges posed by AI, particularly focusing on issues such as bias, privacy concerns, and transparency. Third, the study discusses the ethical and legal frameworks that can guide the responsible development and deployment of AI technologies. Finally, the study concludes with recommendations for future research and policy development, aimed at creating a balance between technological innovation and the preservation of essential human values.

Through this exploration, the study seeks to contribute to the ongoing discourse about AI's role in shaping our societies, and to ensure that as AI continues to advance, it does so in a way that promotes a fair, just, and ethical future for all.

1- Artificial Intelligence: A Theoretical Approach

Artificial Intelligence (AI) has evolved through significant phases, from its early computational models to modern systems capable of complex problem-solving and decision-making. This section explores the historical development, core concepts, characteristics, and diverse applications of AI, highlighting its transformative impact on various sectors.

A- Historical Evolution of Artificial Intelligence According to Ahmed Kazem, the historical development of artificial intelligence can be divided into three phases: (Kazem, 2022, pp. 5, 6)

- **First Phase:** This phase began after World War II, starting in 1950 with Shannon's research on chess, and ended in 1963 with the work of figures like Feigenbaum and Feldman. This stage was characterized by finding solutions to games and solving puzzles using computers, focusing on the development of search methods in spatial representations. It led to the creation of computational models based on three main factors:
 - Representing the initial state of the subject being examined (e.g., a chessboard at the start of a game).
 - Defining the conditions for perceiving the goal (e.g., defeating the opponent).
 - The set of rules governing the movement of the player's pieces on the board.

The first phase of AI, according to Ahmed Kazem, marks the foundational period of artificial intelligence, beginning in the 1950s after World War II. This era was largely focused on solving games and puzzles, particularly through research like Shannon's work on chess. The emphasis during this period was on developing computational models that could simulate problem-solving by focusing on spatial representations. AI systems were initially designed to mimic human strategies in games, with the goal of making decisions based on predefined rules. The computational models created in this phase were centered around three key components: defining the initial state of a scenario, setting a goal, and following a set of rules. This phase was critical in establishing the groundwork for future AI systems, though the models were relatively simple compared to modern AI's complexity.

- **Second Phase:** Known as the "Romantic" phase, it lasted from the mid-1960s to the mid-1970s. During this period, Minsky developed "Frames" to represent information, and other scholars worked on systems to understand English sentences, including stories and conversations. Researchers like Anonston and Brown summarized advanced words at the Massachusetts Institute of Technology, which involved studies on natural language processing, computer vision, robotics, and symbolic processing.

The second phase, often referred to as the "Romantic" phase, represents a pivotal shift in AI research from the mid-1960s to the mid-1970s. During this period, the focus expanded from simple problem-solving to more complex tasks such as understanding natural language and processing symbolic information. Minsky's development of "Frames" was a significant contribution, allowing for the representation of knowledge in a more structured and accessible way. Researchers also made strides in natural language processing, aiming to understand and interpret English sentences, stories, and conversations. This phase saw AI systems increasingly interact with human languages and complex data. The work done at institutions like MIT by scholars such as Anonston and Brown advanced these fields further, laying the groundwork for developments in computer vision, robotics, and symbolic processing. However, despite these breakthroughs, the technologies remained limited, and the phase ended with a realization that the initial promises of AI were still far from being fully realized.

- **Third Phase:** The modern phase began in the mid-1970s, marked by the emergence of various technologies that addressed many applications. This period is considered the golden age of AI, as it led to the development of many modern AI systems. The core techniques of

AI evolved, including symbolic modeling, list processing mechanisms, and programming techniques, which interacted with many scientific fields.

The third phase, which began in the mid-1970s, is often referred to as the "golden age" of AI due to the rapid advancements and applications that emerged during this time. AI saw a significant expansion beyond theoretical concepts to practical systems capable of addressing real-world problems. This period marked the evolution of core AI techniques, such as symbolic modeling, which helped machines understand and process complex data. The development of list processing mechanisms and advanced programming techniques also enabled AI to solve more sophisticated tasks. Additionally, the integration of AI with various scientific fields, including mathematics, neuroscience, and cognitive science, allowed for more diverse and powerful applications. This phase laid the foundation for many modern AI systems, influencing a range of industries and applications, from medical diagnostics to robotics. However, the complexity of these systems also led to new challenges, including issues with scalability and understanding the limits of AI's capabilities.

B- Concept of Artificial Intelligence: The term "Artificial Intelligence" consists of two parts: "intelligence" and "artificial." "Intelligence" is defined by Webster's Dictionary as the ability to understand new or changing circumstances. In a broader sense, intelligence refers to the ability to perceive, understand, and learn from new situations. (Kurzweil, Ray. 1990, p174) Thus, the keys to intelligence are perception, understanding, and learning. The word "artificial" refers to things that are man-made as opposed to those that occur naturally, which are not directly influenced by human intervention. Therefore, artificial intelligence refers to intelligence created by humans in machines or computers, the intelligence originally possessed by humans but transferred to machines. (Yassin, 2017, p. 179)

Artificial intelligence is defined as the science concerned with creating smart machines that behave as humans are expected to behave. (Rouwabeh, A., & Boudah, A. J., 2015, p. 203).

According to the "Petit Robert" dictionary, AI is part of computer science, aiming to simulate cognitive abilities to replace humans in performing tasks that require intelligence. Patterson. W. Dan defines AI as a branch of computer science studying and creating computer systems that exhibit some forms of intelligence. (Belhamou, F. Z., & Arzi, F., 2017, p. 66).

These systems are capable of drawing valuable inferences about a given problem and can understand natural language, perceive the world, and perform other tasks that require human-like intelligence. Minsky defines AI as the science that enables machines to perform tasks that require intelligence if executed by humans. (Sheikh, H., 2018, p. 82)

In general, AI is defined as "a set of efforts aimed at developing computer systems that can behave and think similarly to humans, learning natural languages, completing real tasks in an integrated manner, or using cognitive images and forms to rationalize physical behavior. These systems can store accumulated human experiences and knowledge in real-time and use them in decision-making." (Yassine, S. G., 2004, p. 165).

This conceptual definition provides a comprehensive definition of artificial intelligence (AI) by breaking down its two core components: "intelligence" and "artificial." It explains that intelligence involves perception, understanding, and learning from new circumstances, while "artificial" indicates that this form of intelligence is created by humans, not naturally occurring. The text emphasizes that AI systems are designed to replicate human-like behaviors, such as problem-solving and learning. Several definitions of AI are referenced, including those by experts like Kurzweil, Rouwabeh, and Patterson, highlighting the field's goal to create machines capable of performing tasks that would typically require human intelligence. Minsky's definition underscores AI's ability to mimic human cognitive functions. Additionally, the paragraph notes that AI systems can store and utilize human knowledge to make informed decisions, demonstrating AI's potential in

learning, language processing, and real-world task execution. The varying definitions emphasize AI's interdisciplinary nature and its growing role in simulating human cognition.

C- Importance of AI in Modern Society

Artificial Intelligence (AI) has become an integral part of modern society, influencing various sectors and significantly impacting how we live, work, and interact. AI technologies have evolved from theoretical concepts to powerful systems capable of solving complex problems, improving efficiency, and driving innovation across industries. One of the primary reasons AI is essential in modern society is its ability to handle large-scale data processing and analysis. AI can sift through vast amounts of data, extracting valuable insights that are often beyond the capabilities of humans. This has transformed sectors such as healthcare, where AI models can analyze medical images, predict patient outcomes, and assist in diagnosing diseases with a level of precision that enhances medical care (Brynjolfsson & McAfee, 2017, p. 45).

In the business world, AI has revolutionized industries by automating routine tasks, improving decision-making, and enabling more personalized consumer experiences. For instance, AI algorithms in e-commerce analyze consumer behavior to provide personalized recommendations, while chatbots and virtual assistants streamline customer service operations, reducing operational costs and improving efficiency (Chui et al., 2018, p. 23). Moreover, AI's impact on automation is not limited to consumer-facing industries. In manufacturing, AI-powered robotics and predictive maintenance technologies improve production line efficiency, reducing downtime and enhancing product quality (Ransbotham et al., 2017, p. 27).

AI also plays a crucial role in the advancement of scientific research and innovation. Machine learning algorithms can simulate complex systems, allowing researchers to predict outcomes in various fields, from climate science to drug discovery. AI's ability to simulate biological systems, for example, has sped up the process of drug development, leading to faster and more efficient creation of life-saving medications (VanderPlas, 2016, p. 54). Additionally, AI is crucial in improving education by offering personalized learning experiences. AI-powered platforms can adapt to a student's learning style, identifying areas where they need improvement and tailoring the curriculum to suit their individual needs (Mouza, 2020, p. 78).

Beyond economic and practical uses, AI is also transforming societal dynamics. For example, in smart cities, AI is used to optimize traffic flow, reduce energy consumption, and improve urban planning, contributing to more sustainable and livable environments (Vinuesa et al., 2020, p. 112). Furthermore, AI in cybersecurity has become a vital tool for protecting digital infrastructures, identifying and responding to security threats in real-time, which is particularly important in an era where cyber-attacks are increasingly sophisticated (Zhou et al., 2018, p. 34).

While AI presents numerous opportunities, it also raises concerns regarding job displacement and ethical issues. However, the positive contributions of AI to modern society are undeniable, as its potential to drive innovation, improve quality of life, and solve complex global problems remains vast and transformative.

D-Characteristics and Goals of Artificial Intelligence: The main characteristics of AI include: (Turban, E., 2021, p. 13)

- Thinking and perception
- Using intelligence to solve problems
- Learning or understanding from experience
- Acquiring knowledge and applying it
- Displaying creativity and imagination
- Handling complex situations
- Responding quickly and effectively to new circumstances
- Recognizing the relative importance of elements in a situation

- Dealing with incomplete and ambiguous information
- Supporting decision-making processes

The characteristics of Artificial Intelligence (AI) listed here underscore its multifaceted and dynamic nature. AI is defined not just by its ability to process data, but by its capacity to mimic human cognitive functions such as thinking and perception. These characteristics enable AI to solve problems, learn from experience, and apply acquired knowledge in new contexts, highlighting its adaptability. AI's creative and imaginative capabilities further demonstrate its potential to innovate, a crucial trait for tackling complex situations that may be beyond traditional computational methods. Another key feature is AI's ability to quickly respond to changing circumstances, making it ideal for fast-paced environments where real-time decision-making is essential. Additionally, AI excels at recognizing the relative importance of elements in a situation, which supports its ability to manage and prioritize tasks. Its skill in dealing with incomplete and ambiguous information makes AI resilient in scenarios where data may be uncertain or unclear. Finally, AI's support in decision-making processes is one of its most powerful applications, as it can analyze large volumes of data and suggest optimal solutions or strategies. These traits collectively reinforce AI's potential to transform various sectors, from healthcare to finance, by improving efficiency, accuracy, and innovation.

The goals of AI can be summarized in two main points: (Khalifa, E., 2017, p. 41)

- Enabling machines to process information in a manner similar to how humans solve problems, focusing on parallel processing, where several commands are executed simultaneously, similar to human problem-solving.
- Gaining a better understanding of human intelligence by delving into the brain to simulate it. The human nervous system and brain are the most complex organs and operate in a highly integrated and continuous manner to recognize things.

The goals of AI are centered around two key objectives: emulating human problem-solving and understanding human intelligence. First, AI seeks to enable machines to process information like humans by using parallel processing, allowing multiple tasks to be handled simultaneously, much like the human brain. This approach aims to improve efficiency and adaptability in decision-making and problem-solving. Second, AI endeavors to understand and simulate human intelligence by studying the brain's complex structure and function. The human nervous system is highly integrated and capable of recognizing patterns, which AI strives to replicate. By mimicking these cognitive processes, AI can enhance its own problem-solving abilities and provide insights into human cognition, potentially bridging the gap between artificial and biological intelligence. These goals represent the intersection of technology and neuroscience, driving advancements in both fields.

E-Types and Applications of Artificial Intelligence:

Artificial Intelligence (AI) can be classified into different types based on its capabilities, ranging from simple task-specific systems to highly advanced, autonomous machines. These types of AI have diverse applications across various industries, driving innovations in fields like healthcare, transportation, and finance. Understanding AI's types and applications is essential for recognizing its potential to transform society.

AI can be classified into three main types, ranging from simple reactions to perception and interaction: (Ben Tayeb, A., & Mahloul, Z., 2019, p. 7)

- **Types of AI:**
 - **Narrow or Weak AI:** The simplest form of AI, programmed to perform specific functions in a limited environment, reacting to certain situations and not functioning outside its designated environment.

- **Strong or General AI:** This AI can collect, analyze information, and accumulate experiences from various situations, enabling it to make independent decisions. Examples include self-driving cars and chatbots.
- **Superintelligent AI:** These models are still in experimental phases, aiming to simulate human abilities. There are two main categories: one focuses on understanding human thoughts and emotions influencing behavior, and the other represents a "theory of mind," where the AI can express its internal state, predict others' emotions, and interact with them. This represents the next generation of superintelligent machines.

The types of AI are classified based on their capabilities and levels of intelligence. Narrow or Weak AI is the most common form today, designed to handle specific tasks within a limited scope, such as voice assistants or recommendation systems. While it is highly efficient in its designated tasks, it lacks the ability to operate beyond its programmed function. On the other hand, Strong or General AI represents a more advanced form, with the ability to analyze diverse data and learn from experiences to make independent decisions, like self-driving cars or sophisticated chatbots. Superintelligent AI, though still in the experimental phase, aims to surpass human cognitive abilities and is envisioned to understand complex emotional and behavioral nuances, offering more advanced interaction with humans. This progressive evolution of AI promises groundbreaking applications, though it also raises questions about ethics and control.

- **Applications of AI:** AI is used in various fields such as military, industrial, economic, technical, medical, educational, and other services. Some of the key applications include: (Saudi Data and Artificial Intelligence Authority, 2023, p. 16).
 - Self-driving cars and drones
 - Robots (mechanical devices programmed to work independently, designed to perform tasks like moving parts, repairing underground wires, defusing bombs, manufacturing cars, etc.)
 - Nonlinear control systems, such as railway systems
 - Smart devices capable of performing mental operations, like checking industrial designs, monitoring processes, and decision-making
 - Cognitive simulation to test theories about the brain's functioning, such as recognizing faces, voices, handwriting, and memory management
 - Computer applications in medical diagnostics in clinics and hospitals, performing surgeries
 - AI systems for analyzing economic data, such as stock exchanges and trading systems
 - Google's research clusters for online computer searches
 - Natural language learning and automatic language translation systems
 - Expert systems that simulate expert decision-making by relying on logical operations to arrive at correct decisions or options
 - Smart homes, autonomous weapons, smartphones, smart TVs, and hundreds of other applications.

The applications of AI are vast and diverse, transforming nearly every sector of society. In the military, AI is used in autonomous drones for surveillance and targeted strikes, while robots are deployed for tasks like bomb disposal and precision manufacturing. The industrial sector also benefits from AI-driven robots for tasks such as assembling products and repairing infrastructure. Nonlinear control systems, like those found in railway systems, optimize operations and safety, making transportation more efficient. In the medical field, AI has revolutionized diagnostics, enabling quicker and more accurate identification of diseases, as well as performing surgeries with robotic precision. AI applications also extend to financial markets, where systems analyze vast

amounts of economic data for predictive modeling and decision-making. Additionally, AI powers search engines like Google, enabling more efficient online data retrieval. In language processing, AI systems such as Google Translate have improved communication across languages, while cognitive simulations enhance our understanding of the brain's functions. Moreover, expert systems have been developed to assist decision-making in specialized fields, mirroring the expertise of professionals. Smart homes, AI in smartphones, and autonomous vehicles are all examples of how AI is integrated into everyday life, further underlining its potential to improve both practical functionality and quality of life.

2- The Impact of Artificial Intelligence on Human Values and Human Rights

The rapid advancement of artificial intelligence (AI) presents significant challenges and opportunities in the realm of human values and human rights. As AI continues to evolve, it is crucial to address the ethical implications of its use, ensuring that technology aligns with societal needs while protecting fundamental human rights.

A-Key Principles of AI Ethics Concerned with Human Values

Humanistic principles and social and environmental benefits are essential in the age of artificial intelligence, as they guide the development of technology towards serving society. AI influences these principles by enhancing efficiency and improving the quality of life, but it also requires careful monitoring to ensure that technical considerations do not outweigh human values. Thus, the integration of these principles aims to strike a balance between innovation and societal needs.

The rapid development of AI brings both opportunities and challenges in preserving human values and protecting human rights. As AI becomes more integrated into various aspects of life, it is vital that ethical principles guide its application. While AI has the potential to improve efficiency and quality of life, it must not overshadow fundamental human values such as fairness, privacy, and dignity. By emphasizing humanistic and social benefits, AI development can serve society responsibly. However, continuous monitoring is necessary to ensure that technological progress remains aligned with the broader ethical and societal goals. Striking a balance between innovation and human values is essential for sustainable AI adoption.

First Principle: Humanity

The principle of humanity emphasizes the necessity of building AI systems using a fair and ethical approach based on human rights and fundamental cultural values to positively impact stakeholders, local communities, and contribute to achieving long-term and short-term goals for the benefit and prosperity of humanity. AI systems must be designed not to deceive, manipulate, or impose behaviors that do not empower, enhance, or increase human skills. Instead, they should adopt a design approach that focuses on enabling human choice and decision-making. This principle includes several key points: (Saudi Data and Artificial Intelligence Authority, 2023, p. 17).

- **Planning and Design:**

1. It is necessary to design and build models based on fundamental human rights, values, and cultural principles, applying them to decisions, processes, and functions of AI systems. AI model designers must adhere to this.
2. Identifying how the AI system will align with basic human rights and cultural and social values in Saudi Arabia, in addition to determining the necessary techniques and testing them, while also defining how the AI system will enhance human skills and abilities.

- **Data Preparation:**

1. To ensure that AI models adopt a structure and design focused on the principle of humanity, ethical data management practices must be adhered to, along with standards and regulations regarding data management in the Kingdom.
2. Data must be collected, classified, processed, and provided correctly to ensure respect for human rights and cultural and social values in society. (Saudi Data and Artificial Intelligence Authority, 2023, p. 17).

- **Building and Performance Measurement:**

1. Designers and engineers must prioritize building AI systems and algorithms that center on humans, facilitating the decision-making process and respecting human rights and cultural values in the Kingdom. Automated decisions resulting from AI systems should not operate independently without considering human rights and cultural and social values.
2. Designers and engineers should develop AI systems using ethical standards and train algorithms to achieve results that uplift humanity.

- **Application and Monitoring:**

1. Regular evaluations should be conducted on the AI system to ensure its results do not conflict with human rights and cultural and social values, to verify the accuracy of key performance indicators, and to monitor its impact on individuals or communities, ensuring continuous improvement of the technology.
2. AI model designers must implement mechanisms to assess AI systems based on cultural values and human rights, minimizing any negative or harmful outcomes resulting from the system's use. In case of any negative outcomes, the AI system's responsible party must identify areas needing intervention and apply corrective measures to improve system performance and results, with ongoing monitoring. (Saudi Data and Artificial Intelligence Authority, 2023, p.p 16, 17).

The principle of humanity in AI emphasizes the creation of systems that align with human rights and cultural values, focusing on ensuring that AI benefits society while empowering individuals. This principle underlines the importance of designing AI with fairness and ethics, ensuring that technology does not manipulate or diminish human autonomy. The planning and design phase requires developers to integrate fundamental human rights and values into the AI's functions and processes. This includes ensuring the system respects cultural norms and contributes positively to both short-term and long-term goals for society's prosperity. In the data preparation stage, ethical data management practices are vital, ensuring the data used respects human dignity and cultural values, and prevents biases from influencing outcomes.

Furthermore, the building and performance measurement phase calls for AI systems that support human decision-making, incorporating human rights and societal values into their design. Automated decisions should not be made in isolation from these considerations. Designers and engineers are tasked with creating AI systems that contribute positively to humanity's welfare, adhering to ethical standards to ensure the systems' results uplift human capabilities. Regular application and monitoring are essential to ensure that the AI system does not contradict human rights and cultural values. Continuous evaluations, accompanied by corrective actions when necessary, help maintain the system's alignment with ethical goals. This ongoing process ensures that AI serves humanity responsibly and contributes to societal progress, while addressing any negative impacts that may arise. By establishing these guidelines, AI development can better safeguard human dignity, prevent harm, and promote a future where technology enhances, rather than undermines, human rights and cultural integrity.

Second Principle: Social and Environmental Benefits

- Designers and engineers must prioritize building AI systems and algorithms that focus on humans, facilitating decision-making, and aligning with human rights and cultural values in the Kingdom. Automated decisions resulting from AI systems should not operate independently, without considering human rights and cultural values.
- Designers and engineers should develop AI systems using ethical standards and train algorithms to achieve results that benefit humanity.
- Regular evaluations should be conducted on the AI system to ensure its results do not conflict with human rights and cultural and social values, to verify the accuracy of key performance indicators, and to monitor its impact on individuals or communities, ensuring continuous improvement of the technology.

- AI model designers must implement mechanisms to evaluate AI systems based on cultural values and human rights to minimize any negative or harmful outcomes resulting from the system's use. In case of any negative outcomes, the responsible party must identify areas requiring attention and apply corrective measures to improve the system's performance and results, with ongoing monitoring. (Saudi Data and Artificial Intelligence Authority, 2023, p.p 17, 18).

The second principle, focusing on social and environmental benefits, emphasizes the need for AI systems to prioritize human welfare while adhering to cultural values and human rights. Designers and engineers are urged to create AI systems that facilitate human decision-making, not operating in isolation or without considering the broader societal impact. It is essential that these systems not only follow ethical standards but also aim to bring tangible benefits to humanity, such as improving quality of life, enhancing accessibility, or promoting environmental sustainability. Regular evaluations of AI systems are crucial to ensure that they do not conflict with societal values or undermine human rights. Key performance indicators must be carefully monitored to track the real-world impact of AI, both on individuals and communities. This ongoing evaluation allows for adjustments and improvements in the system, ensuring that technology evolves in a way that consistently benefits society.

Moreover, AI systems should be designed with the foresight to minimize any potential negative effects. If harmful outcomes are detected, the responsible parties must act swiftly to identify the problem and apply corrective measures. This principle encourages continuous monitoring, feedback loops, and active problem-solving to prevent unintended consequences. By focusing on social and environmental benefits, AI development can help mitigate risks, such as job displacement or privacy violations, while promoting solutions that contribute to the well-being of communities. It creates a framework where technology serves to support and enhance human lives, rather than causing harm or deepening existing inequalities. This approach fosters a sustainable and ethical model for AI's integration into society, reinforcing the importance of responsibility and foresight in the development of new technologies.

B- Ethical Considerations of Artificial Intelligence on Human Values and Human Rights

The integration of Artificial Intelligence (AI) into daily life raises significant ethical concerns, particularly in the context of human values and human rights. As AI systems become more autonomous and capable of performing tasks traditionally undertaken by humans, it becomes crucial to assess their impact on fundamental human rights such as privacy, freedom, and equality, as well as broader human values like dignity and autonomy. These ethical concerns intersect with legal, societal, and technological dimensions, all of which must be considered to ensure that AI technologies support rather than undermine human rights.

- Privacy and Surveillance

One of the most profound ethical concerns AI poses is its potential to violate privacy. AI systems, particularly those based on machine learning, require vast amounts of data to operate effectively. This data often includes sensitive personal information such as medical records, financial data, and social media activity. When this data is used without adequate protection, it can lead to violations of privacy, particularly if AI systems enable widespread surveillance (Zuboff, 2019, p. 45). The ethical principle of privacy, deeply rooted in human rights law, asserts that individuals should have control over their personal data and how it is used. In the case of AI-driven surveillance systems, such as facial recognition technology, there is the potential for governments or corporations to infringe on personal privacy without individuals' knowledge or consent, creating new challenges for privacy laws and protections (Greenleaf, 2018, p. 33).

- Bias and Discrimination

AI systems are often trained on historical data, which can inadvertently contain biases that reflect societal inequalities. If these biases are not addressed, AI can perpetuate and even exacerbate discrimination, particularly in areas like hiring, criminal justice, and credit scoring (O'Neil, 2016, p. 88). For instance, AI algorithms used in predictive policing have been shown to disproportionately target minority communities, while hiring algorithms may favor candidates of certain genders or

ethnicities. The ethical principle of equality, which is enshrined in the Universal Declaration of Human Rights, calls for individuals to be treated equally and without discrimination. Bias in AI undermines this fundamental right by reinforcing existing prejudices, leading to unfair outcomes and social harm. Efforts to eliminate bias from AI systems through more diverse data and algorithmic fairness are essential to ensure that AI supports rather than contradicts human rights principles (Binns, 2018, p. 71).

- Autonomy and Decision-making

AI also raises ethical concerns related to autonomy—the ability of individuals to make their own decisions. Autonomous AI systems, such as self-driving cars or AI-driven healthcare recommendations, can make decisions that significantly impact individuals' lives. For example, if a self-driving car is involved in an accident, who is responsible for the decisions made by the AI in navigating the situation? This leads to questions of accountability and the extent to which AI systems should make life-changing decisions without human oversight (Lin, 2016, p. 210). Ethical considerations emphasize that AI should enhance human autonomy, not limit it. Human dignity and autonomy are key principles in human rights law, particularly in relation to informed consent and decision-making. AI should be designed in ways that empower individuals to make decisions while ensuring that they maintain control over the processes affecting their lives.

Transparency and Accountability

Transparency and accountability are vital ethical concerns when it comes to AI's impact on human values. Many AI systems, particularly those utilizing complex algorithms, operate as “black boxes,” meaning their decision-making processes are often opaque even to the developers who created them (Lipton, 2016, p. 35). This lack of transparency can lead to mistrust in AI technologies, especially when decisions that affect people's lives—such as loan approvals, legal judgments, or health diagnoses—are made by machines. The ethical principle of accountability demands that individuals and organizations be responsible for the actions of AI systems, particularly when these systems cause harm. Legal frameworks, such as the European Union's General Data Protection Regulation (GDPR), are starting to address this issue by requiring greater transparency in automated decision-making processes and the ability for individuals to contest these decisions (Greenleaf, 2018, p. 47).

Human Rights in AI Development

The development and deployment of AI must be guided by principles that respect human rights, ensuring that AI technologies contribute to social good rather than harm. One of the ethical principles that must guide AI development is the principle of non-maleficence, which suggests that AI should not harm individuals or society. This principle is aligned with the international human rights framework, which calls for the protection of life, liberty, and security of persons (United Nations, 1948). The rapid development of AI systems should be accompanied by rigorous ethical oversight to ensure that these systems do not infringe on human rights, such as the right to freedom of expression, privacy, or protection from discrimination.

The Role of Ethics in AI Regulation

Regulatory bodies and ethicists are increasingly calling for robust AI governance frameworks to ensure that AI is aligned with human rights and values. Many advocate for global ethical standards that can guide the development and application of AI technologies. For example, UNESCO's AI ethics recommendations urge the development of AI systems that respect human rights, promote inclusivity, and are designed to reduce societal harms (UNESCO, 2021, p. 10). These frameworks should emphasize human-centered AI that supports dignity, equality, and freedom while safeguarding privacy and autonomy. Additionally, AI developers and organizations must collaborate with human rights experts to ensure the technologies they create respect fundamental freedoms and contribute to the well-being of society as a whole.

The ethical considerations surrounding Artificial Intelligence (AI) and its impact on human values and human rights are crucial as AI becomes increasingly integrated into society. Privacy concerns are paramount, particularly with AI systems requiring large amounts of personal data, which may lead to violations of privacy and the erosion of individual freedoms, especially with surveillance technologies like facial recognition. Additionally, AI systems are prone to perpetuating bias, reinforcing societal inequalities in areas such as hiring, criminal justice, and credit scoring. This highlights the need for algorithmic fairness and the ethical principle of equality. AI's potential to undermine human autonomy, such as in autonomous decision-making systems like self-driving cars, raises important questions about accountability and control. Transparency and accountability in AI development are necessary to build trust and ensure responsible usage, as opaque decision-making processes can lead to negative consequences. Ethical AI development should be guided by human rights principles, promoting dignity, equality, and freedom, while safeguarding privacy. Finally, AI regulations must evolve to include robust ethical standards, ensuring AI technologies are developed with respect for human rights and contribute positively to society.

C- Challenges of Legal Frameworks for Protecting Personal Data in AI Usage

Today, there is an increasing discussion on reinforcing AI ethics in light of the current critical situation. However, these ethics are often at odds with emergency situations where crisis management takes precedence over finding ethical solutions, given the differing perspectives on how crises should be handled—immediate management versus achieving sustainable solutions. This is evident today as the use of AI has expanded unprecedentedly, revealing the limitations of relying solely on legal frameworks and their inability to keep pace with the growing range of AI applications. Efforts in AI ethics and law have focused on establishing ethical frameworks and general principles, but these legal principles do not guide us when principles collide, such as balancing the principle of saving lives with respecting privacy and the protection of personal information. To reduce this conflict, a constructive dialogue must be established with various AI stakeholders, which requires preparation and effort. However, crises create a time pressure that works against this. We find ourselves between the reactive need for immediate solutions and the duty to thoughtfully interact with the reality of circumstances, considering potential and obstacles. But does this situation mean abandoning ethics in times of crisis? Quite the opposite: the current crisis emphasizes the need for enhanced efforts in the ethical and legal aspects of AI, striving to create solutions and regulations as rapidly as technological advancements in algorithms are being made. This requires invigorating the efforts of legal professionals and AI ethics experts to ensure that legal frameworks do not lag behind technological development but evolve in parallel to it, proactively anticipating challenges. (Al-Suhaili, Muhammad Al-Hadi, 2019, p. 49).

The challenges of legal frameworks in protecting personal data in the context of AI usage are increasingly pressing as the technology advances at a rapid pace. While AI offers vast potential, it often introduces conflicts between immediate crisis management and the ethical and legal considerations that typically govern technology. In emergency situations, there is a tendency to prioritize quick solutions, such as saving lives, which may conflict with the need to protect personal data and uphold privacy rights. This dilemma exposes the limitations of current legal frameworks, which often struggle to keep up with the evolving landscape of AI applications.

Moreover, the existing legal principles frequently fail to provide clear guidance when ethical principles clash, particularly when balancing privacy rights against the urgency of crisis management. For example, in cases like data collection for public health reasons, the legal and ethical challenges of managing personal data in AI systems become evident. The current framework may not address the nuances of such scenarios, requiring a more dynamic and flexible approach.

To address these challenges, a collaborative dialogue between legal professionals, AI ethics experts, and other stakeholders is crucial. However, in times of crisis, the pressure for immediate solutions often undermines the opportunity for thoughtful deliberation. This creates a tension between the need for reactive measures and the obligation to consider long-term implications for personal data and privacy.

Despite these challenges, the situation underscores the need for strengthening the ethical and legal dimensions of AI. The development of regulations and frameworks must evolve alongside technological advancements, ensuring that AI systems respect fundamental rights while also addressing immediate needs in crisis situations. By proactively anticipating the challenges posed by AI, legal professionals and ethics experts can help ensure that personal data is protected without stifling the potential of AI technologies.

Conclusion:

This research explores the profound impact of Artificial Intelligence (AI) on human values and human rights, addressing the critical challenges posed by the rapid development of AI technologies. The central problem under investigation is how AI can evolve while maintaining and upholding fundamental human rights, values, and ethical standards, particularly in areas such as privacy, transparency, and bias. The research hypothesizes that while AI holds immense potential for societal progress, its unchecked application can undermine core human values unless robust ethical and legal frameworks are established.

To investigate this issue, a comprehensive methodology was employed, involving an extensive review of existing literature, case studies, and an analysis of ethical frameworks and regulatory standards. This approach allowed for a thorough examination of the opportunities AI presents, as well as the risks it poses, especially when it comes to aligning technological advancement with human-centric principles.

The tools used for this research included qualitative analysis techniques, case study evaluations of AI applications, and an in-depth review of AI ethics literature. The findings suggest that without continuous monitoring, AI could inadvertently infringe on human rights and ethical standards. Therefore, it is imperative to establish dynamic frameworks that evolve in parallel with AI advancements, ensuring that technology benefits humanity while safeguarding essential values.

Ultimately, this research underscores the need for a balanced approach that fosters AI innovation while ensuring that its deployment does not compromise the ethical foundations of our society. Future research and policy development will be crucial in shaping a future where AI serves as a force for good, enhancing human dignity and rights rather than undermining them.

Here are some recommendations for addressing the challenges that artificial intelligence (AI) poses to fundamental human values:

1. **Develop a Legal and Ethical Framework:** Clear legislation should be established to regulate the use of AI, with a focus on protecting individual rights and ensuring accountability for decisions made by intelligent systems.
2. **Enhance Transparency:** Companies and developers should clarify how AI algorithms work, including how data is collected and processed, to ensure users understand the processes that affect their lives.
3. **Encourage Community Participation:** Civil society should be involved in the AI policy-making process by organizing dialogues and workshops to listen to the concerns and opinions of citizens.
4. **Education and Awareness:** Promote education about AI and its impact on human values through curricula and seminars to empower individuals to make informed decisions.
5. **Develop Responsible Technology:** Encourage researchers and developers to focus on creating technologies that support justice and equality, and work to reduce bias and discrimination in intelligent systems.
6. **Achieve a Balance Between Innovation and Ethics:** Companies and developers should maintain a balance between technological progress and ethical principles by adopting a responsible approach to AI development.
7. **Ongoing Monitoring:** Establish independent bodies to monitor the use of AI and assess its impact on human values, ensuring compliance with ethical and legal standards.

By implementing these recommendations, we can ensure that AI remains a tool for enhancing fundamental human values rather than threatening them.

-Bibliography:

A - Books:

1. Brynjolfsson, E., & McAfee, A., *The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies*, W. W. Norton & Company; 2017.
2. Chui, M., Manyika, J., & Miremadi, M., *The Next Generation of AI: How Machines Are Reshaping the World*, McKinsey Global Institute; 2018.
3. Mouza, C., *The Role of Artificial Intelligence in Education*, Springer; 2020.
4. O'Neil, C., *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*, Crown Publishing Group; 2016.
5. Zuboff, S., *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*, PublicAffairs; 2019.
6. Kurzweil, R., *The Age of Intelligent Machines*, MIT Press, Cambridge, Mass.; 1990.
7. Yaseen, S. G., *Decision Support Systems*, Dar Al-Manahej for Publishing and Distribution, Amman, Jordan; 2017.
8. Yaseen, S. G., *Decision Support Systems (2nd Edition)*, Dar Al-Manahej for Publishing and Distribution, Amman, Jordan; 2004.
9. Turban, E., *Management Support Systems: Decision Support and Expert Systems* (translated by S. Ali Sarour), Dar Al-Mareekh; 2000.

B - Theses:

- Kazem, A., *Artificial Intelligence*, College of Information Technology Publications, Imam Jaafar Al-Sadiq University, Iraq; 2012.

C - Newspaper Articles:

- VanderPlas, J., *Python Data Science Handbook: Essential Tools for Working with Data*, O'Reilly Media, p. 54; 2016.
- Lipton, Z. C., *The Mythos of Model Interpretability*, ACM Queue, p. 35; 2016.
- Greenleaf, G., *Global Data Privacy Laws: The Global Reach of GDPR*, Privacy Laws & Business International Report, p. 33; 2018.

D - Seminar Papers:

1. Khawald, A., & Thlaijiya, N., *Artificial Intelligence-Based Information Systems Between Theoretical Concepts and Practical Applications in Economic Institutions*, In the 10th National Conference on AI-Based Information Systems and Their Role in Economic Decision-Making, Skikda University, Algeria; 2012.
2. Ben Taib, A., & Mahloul, Z., *Applications of Artificial Intelligence and Its Role in Promoting the Digitization of Societies and the Shift Toward Smart Cities: The United Arab Emirates as a Model*, In the International Conference on Smart Cities in the Current Changing Context: Reality and Prospects, The Democratic Arab Center, Berlin, Germany; 2019.

E - Websites:

- UNESCO, *Ethical Artificial Intelligence: UNESCO's Recommendations*, <http://www.unesco.org> (consulted on 01/04/2025).

Appendix:

1. Al-Suheili, M. H. (2019). *Developments in Artificial Intelligence and the Implications for Protecting Fundamental Rights and Freedoms*. Islamic World Educational, Scientific and Cultural Organization; p. 49.
2. Belhamou, F. Z., & Arzi, F. (2017). *The Contribution of Expert Systems in Improving Decision-Making in Algerian Enterprises: A Case Study of ABRAS SPA*. *Revue*

- Maghrébine Management Des Organisations, 2(1). Abu Bakr Belkaid University, Tlemcen, Algeria; p. 66.
3. Binns, R. (2018). Fairness in Machine Learning: A Survey. *ACM Computing Surveys*, 51(3), p. 71.
 4. Brynjolfsson, E., & McAfee, A. (2017). *The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies*. W. W. Norton & Company, p. 45.
 5. Chui, M., Manyika, J., & Miremadi, M. (2018). *The Next Generation of AI: How Machines Are Reshaping the World*. McKinsey Global Institute, p. 23.
 6. Greenleaf, G. (2018). *Global Data Privacy Laws: The Global Reach of GDPR*. Privacy Laws & Business International Report, p. 33.
 7. Khalifa, I. (2017). The Impact of Increasing the Role of Smart Technologies in Human Daily Life. *Trends of Events Journal*, 20. Future Center for Research and Advanced Studies, Abu Dhabi, UAE; p. 41.
 8. Lin, P. (2016). Why Ethics Matters for Autonomous Cars. In *Autonomes Fahren* (pp. 179-191). Springer Vieweg, p. 210.
 9. Lipton, Z. C. (2016). The Mythos of Model Interpretability. *ACM Queue*, p. 35.
 10. Mouza, C. (2020). *The Role of Artificial Intelligence in Education*. Springer, p. 78.
 11. O'Neil, C. (2016). *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Crown Publishing Group, p. 88.
 12. Ransbotham, S., Kiron, D., & Prentice, P. (2017). Artificial Intelligence and Business Strategy: Insights for Industry Leaders. *MIT Sloan Management Review*, p. 27.
 13. Rawabeh, A., & Boudah, A. J. (2015). The Evolution of Credit Risk Assessment in Light of Artificial Intelligence Models. *Journal of Human Sciences*, 26(4). Mentouri University, Constantine, Algeria; p. 203.
 14. Saudi Data and Artificial Intelligence Authority. *Principles of Artificial Intelligence Ethics*, 2023; p. 16.
 15. Saudi Data and Artificial Intelligence Authority. *Principles of Artificial Intelligence Ethics*, 2023; p. 17.
 16. Saudi Data and Artificial Intelligence Authority. *Principles of Artificial Intelligence Ethics*, 2023; pp. 17-18.
 17. Sheikh, H. (2018). The Role of Artificial Intelligence in Managing the Electronic Customer Relationship for the Algerian Popular Loan (CPA). *Journal of Social and Human Studies*, 10(2). Hassiba Ben Bouali University, Chlef, Algeria; p. 82.
 18. UNESCO. (2021). *Ethical Artificial Intelligence: UNESCO's Recommendations*. United Nations Educational, Scientific and Cultural Organization, p. 10.
 19. UNESCO. *Artificial Intelligence and Education: Policy Guidelines for Policymakers*, 2021; p. 13.
 20. United Nations. (1948). *Universal Declaration of Human Rights*. United Nations, p. 5.
 21. VanderPlas, J. (2016). *Python Data Science Handbook: Essential Tools for Working with Data*. O'Reilly Media, p. 54.
 22. Vinuesa, R., Azizpour, H., Leite, I., et al. (2020). The Role of Artificial Intelligence in Achieving the Sustainable Development Goals. *Nature Communications*, p. 112.
 23. Zhou, J., Zhang, X., & Xu, L. (2018). *AI in Cybersecurity: The New Frontline*. IEEE Access, p. 34.
 24. Zuboff, S. (2019). *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. PublicAffairs, p. 45.