Ethical Standards for the Design and Development of Artificial Intelligence Applications in the Digital Age

■ **Prof. Dr. Mohamed Fawzy Waly**

Professor of Educational Technology, Faculty of Education
Former Dean, Faculty of Computers and Information,
Damanhour University

■ Abstract:

The field of artificial intelligence (AI) is witnessing rapid advancement, accompanied by a growing range of applications across various domains of life. This evolution raises essential and continually evolving questions regarding how to develop and utilize AI applications in an ethical, responsible, and socially beneficial manner. The ethics of AI applications refer to a set of principles and values that guide the responsible development and deployment of these technologies, aiming to maximize their benefits while minimizing potential risks.

This study seeks to investigate the most critical ethical standards that should govern the design and development of AI applications in the digital era. The research culminated in a proposed list of ethical standards for designing and developing such applications. This list comprises fifteen key standards, including responsibility, accountability, transparency, fairness, privacy, and environmental sustainability.

The study recommends the integration of these standards and their respective indicators into the design and development processes of AI applications. These ethical guidelines serve as a compass to ensure optimal benefits from AI while mitigating associated risks. Furthermore, the research advocates for additional studies to explore the implications of applying these ethical standards from the perspectives of both developers and users.

Keywords

Ethical standards, Design and development of AI applications, Artificial intelligence strategy, Anti-corruption strategy.

Introduction

Artificial Intelligence (AI) has become an integral part of our daily lives, witnessing widespread application and continuous advancement. LeCun, Bengio, and Hinton (2015) define AI as a set of technologies that enable computers to understand and interact with the surrounding world intelligently—through data processing, decision-making, and the development of predictive models.

Similarly, Stone et al. (2016) describe AI as a broad field aimed at developing systems capable of simulating human cognitive abilities such as learning, reasoning, problem-solving, decision-making, and creativity. Ahmed Al-Arabi (2021) further emphasizes that AI encompasses multiple disciplines, including machine learning, natural language processing, computer vision, and robotics. The goal is to develop intelligent systems that can operate autonomously and make informed decisions. Russell and Norvig (2021) describe AI as a branch of computer science focused on designing and building intelligent systems capable of acting independently in complex and uncertain environments.

In summary, AI represents one of the most significant modern technological trends. It is a contemporary field of computer science that seeks to understand the nature of human intelligence, monitor attention and responses, and simulate intelligent behavior to provide advice, solve problems, or assist in decision-making (Sreenu,2023).

AI holds immense potential to reshape society. The widespread implementation of intelligent systems contributes to economic gains; however, it also raises a range of ethical challenges. Most researchers and developers agree on the importance of developing AI in a trustworthy manner that serves the collective good of society. Figure 1 below illustrates the various purposes for which AI applications are used (Berendt, 2019; Cath et al., 2018; European Commission, 2020; Floridi et al., 2018; Jobin et al., 2019).

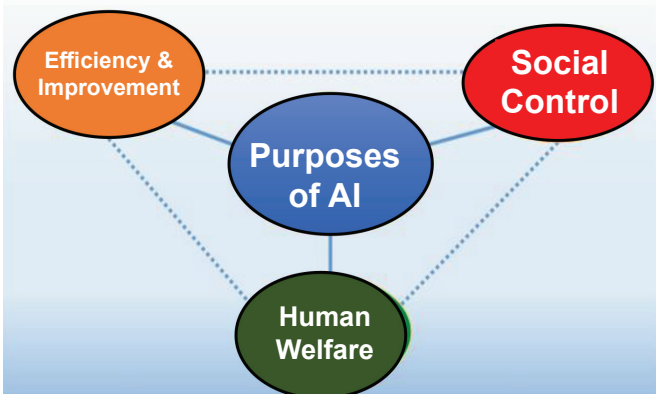

Figure (1): Purposes of Artificial Intelligence (Berendt, 2019)

The design and development of AI applications necessitate overcoming significant ethical challenges. Although a substantial body of research exists on the technical development of AI applications, there remains a considerable gap in studies that

clearly define the ethical standards necessary for their appropriate use (Shin & Park, 2019).

Striking a balance among ethical principles during the design of AI systems is critically important, yet achieving all ethical objectives simultaneously is often difficult—if not impossible. While AI applications are primarily intended to reduce user burdens and facilitate various processes, the use of personal data can provoke significant concerns, such as violations of privacy (Marwick & Lewis, 2017; Shin & Park, 2019).

A team of experts on artificial intelligence—the High-Level Expert Group on Artificial Intelligence (AI HLEG) appointed by the European Commission—emphasized the importance of meeting seven key requirements to ensure the trustworthiness of AI applications. These requirements are outlined by Shneiderman (2020) as follows:

- Human agency and oversight
- Technical robustness and safety
- Privacy and data governance
- Transparency
- Diversity, non-discrimination, and fairness
- Societal and environmental well-being
- Accountability

Similarly, the Organisation for Economic Co-operation and Development (OECD, 2021) recommended five ethical principles to guide the design and implementation of AI applications:

- Inclusive growth, sustainable development, and well-being
- Human-centered values and fairness
- Transparency and explainability
- Robustness, security, and safety
- Accountability

Jobin et al. (2019) conducted a comprehensive review of 84 global studies on AI ethics and identified eleven overarching ethical principles. Among these, five—transparency, fairness and equity, non-maleficence, responsibility, and privacy—were referenced in more than half of the studies analyzed. The principles of beneficence, freedom, and autonomy were identified in 41 studies each, while others such as trust, sustainability, dignity, and solidarity appeared in approximately one-third of the reviewed literature.

While these ethical principles are theoretically desirable, their practical implementation poses significant challenges. As Köbis et al. (2021) note, the design of AI applications often makes it difficult—if not impossible—to fulfill all ethical considerations simultaneously.

Ethics and legal frameworks are two fundamental pillars of the regulatory landscape governing the development and use of AI applications. Ethical guidelines ensure that intelligent systems align with core human values such as justice, transparency, and accountability. Meanwhile, legal regulations establish binding rules and procedures to govern AI use and mitigate risks such as privacy violations and algorithmic bias. Bryson (2018) stresses that integrating ethical and legal dimensions is crucial for developing responsible AI that serves society in a safe and equitable manner.

It is worth emphasizing that ethical engagement with AI applications is not solely the responsibility of developers and researchers. Rather, it is a collective responsibility involving users, companies, governments, organizations, and society as a whole. Through collaboration and adherence to shared ethical principles, AI can become a positive force that contributes to building a better future for all.

In this context, governments, academic institutions, and technological organizations have increasingly sought to establish policies and ethical guidelines to ensure the responsible design, development, and deployment of AI. Numerous organizations, including the European Union and the OECD, have issued frameworks to guide the ethical development of AI systems (Jobin et al., 2019). Egypt has also taken a pioneering step by launching its National AI Strategy to maximize the societal and economic benefits of AI applications. However, challenges remain in translating these ethical principles into practical action. This calls for concerted efforts among researchers, policymakers, and developers to ensure that AI systems align with fundamental human values (Mittelstadt et al., 2016).

Research Problem

In the digital age, certain artificial intelligence (AI) applications suffer from a lack of ethical consideration, which can lead to numerous societal challenges. Among the most concerning are the reinforcement of bias and discrimination against specific social groups, violations of privacy, and the misuse of personal data. In many instances, user data may be collected without explicit consent, leading to serious risks such as digital theft, extortion, and breaches of personal privacy. Additionally, AI technologies can be exploited to disseminate disinformation, including the generation of fake news and manipulated videos. Moreover, the overreliance on AI may reduce direct human interaction, contributing to increased social isolation.

The key implications of the current state of AI application usage can be summarized as follows:

- **Data Bias**: Pre-existing biases within datasets can increase the likelihood of biased AI model outputs, resulting in unfair outcomes such as racial or socio-economic discrimination.

- **Privacy and Security**: Mishandling sensitive data can lead to privacy breaches or data leaks.
- **Data Quality and Integrity**: Incomplete, outdated, or low-quality data can lead to unreliable model predictions and poor decision-making.
- **Algorithmic Bias**: Beyond data, biased algorithm designs can produce unjust or discriminatory results if not developed with rigorous scrutiny.
- **Automation Bias**: Users may place excessive trust in AI recommendations, wrongly assuming higher accuracy than is warranted.
- **System Failures**: AI systems may malfunction or behave unpredictably in dynamic environments, particularly when confronted with unforeseen scenarios.
- **Scalability Issues**: As AI systems grow in scope, maintaining their efficiency and quality across diverse environments becomes increasingly complex.
- **Adversarial Attacks**: AI systems—particularly in domains like cybersecurity—are vulnerable to attacks that manipulate inputs to produce incorrect outputs.
- **Fairness and Discrimination**: AI applications can exacerbate inequalities or introduce new forms of discrimination if not properly managed.
- **Transparency and Accountability**: A lack of transparency in AI decision-making processes makes it difficult to assign responsibility in the event of errors.
- **Misinformation and Deepfakes**: AI-generated content may be used maliciously to spread disinformation, undermining public trust in media.
- **Intellectual Property Concerns**: The misuse of data or copyrighted material may give rise to intellectual property disputes.

These concerns underscore the urgent need to define a set of ethical standards to be observed during the design and development of AI applications. The evolving and increasingly complex nature of this technology calls for the establishment of systematic foundations to ensure that AI systems align with ethical principles such as transparency, fairness, and accountability.

Accordingly, this study seeks to investigate and establish a framework of ethical standards that can be adopted to support the development of AI systems that promote human values while mitigating the potential risks associated with their use.

Research Objectives

This study aims to:

1- **Identify the most critical ethical standards** relevant to the design and development of AI applications across various domains. This includes addressing issues such as privacy, fairness, bias, legal responsibility, and the broader societal implications of AI technologies.

2- **Examine the methods and strategies** that can be employed to manage these ethical concerns responsibly, ensuring a balance between rapid technological advancement and the protection of fundamental human values.

3- **Propose actionable solutions and recommendations** that guide the ethical development and deployment of AI applications in a way that supports sustainable development and reduces the risks of misuse or corruption.

4- **Analyze the current state of AI ethics**, including the scope and root causes of existing problems, their social and technological impact, previous policies, key stakeholders (both positive and negative contributors), and the prevailing indicators—both encouraging and concerning—regarding the ethical deployment of AI systems.

Research Hypotheses

This study aims to examine the following hypotheses:

1- There is a significant correlation between adherence to ethical standards in the design and development of AI applications and the reduction of challenges associated with the irresponsible use of such technologies.

2- There is a significant correlation between the existence of clear ethical standards for the design, development, and use of AI systems and the assurance of responsible and socially beneficial deployment of this technology.

Research Questions

The study seeks to answer the following questions:

1- What are the key features of the current state of AI ethics in the context of designing and developing AI applications?
2- What are the fundamental ethical principles that should be considered during the design, development, and use of AI systems?
3- What are the most prominent ethical challenges resulting from the use of AI across various domains?
4- What are the most essential ethical standards for the design and development of AI applications?

Research Methodology

This study employs a **descriptive-analytical approach**, relying on the analysis of previous research and scientific reports concerning the ethics of designing and developing AI applications. It also includes the examination of real-world case studies of AI application use in various contexts.

Definitions of Key Terms

- **Artificial Intelligence (AI):** Refers to the ability of machines to simulate human cognitive capabilities such as learning, problem-solving, decision-making, adaptation, and responding to novel situations, as well as performing other tasks that typically require human intelligence (Naqvi, 2020).
- **AI Ethics:** A set of principles and values intended to guide the responsible and ethical development and deployment of AI technologies. These principles address issues such as transparency, accountability, fairness, privacy, and security, aiming to ensure that AI serves humanity and avoids causing harm (Taddeo & Floridi, 2018).
- **Ethical Standards for AI Design and Development:** A collection of principles and values that must be considered in the design and development of AI applications to ensure their responsible and ethical use. These standards include concerns such as transparency, accountability, fairness, privacy, and security and are meant to guide the development process in a way that benefits individuals while minimizing potential harm (Jobin et al., 2015).

Main Axes of the Study

The following section outlines the key areas addressed in this research, highlighting the importance of ethical AI usage across various fields. It also explores the primary ethical challenges associated with AI deployment, examines Egypt's national strategy for AI, and considers how this strategy can contribute to the implementation of the National Anti-Corruption Strategy. Ultimately, the aim is to develop a clear and actionable list of ethical standards that should govern the design and use of AI applications across sectors.

I. The Importance of Ethical Use of AI Applications

The field of AI ethics is experiencing rapid evolution, with growing efforts to establish comprehensive and adaptable ethical frameworks. Recent trends include

a stronger emphasis on **responsible AI**, the development of **auditing and review mechanisms**, **international cooperation**, and **public engagement** (Hagendorff, 2020; UNESCO, 2021).

Undoubtedly, the increasing integration of AI technologies will significantly affect the skills required in future labor markets. Individuals will need to develop new competencies suited to the digital age, such as **critical thinking, problem-solving, collaboration, communication, creativity, innovation**, and **lifelong learning and adaptability**.

AI applications have the potential to improve living standards by enhancing productivity and creating new job opportunities. They can also help reduce social and economic disparities by offering equal access to education and training for all individuals (Harari, 2016; O'Neil, 2016).

Moreover, AI can contribute to **operational efficiency**, reducing costs and generating economic benefits that positively impact society. The European Commission anticipates that AI will **permeate various sectors and occupations**, thereby boosting productivity and supporting strong, sustainable economic growth (Craglia, 2018).

AI Ethics and Legal Frameworks

AI ethics refers to a set of guiding principles and values that inform the responsible development and deployment of artificial intelligence. These principles include **fairness**, **transparency**, **accountability**, and **safety** (Bryson & Floridi, 2019; Floridi, 2020; Hagendorff, 2020; Taddeo & Floridi, 2018; UNESCO, 2021).

In contrast, **AI law** comprises the formal rules and regulations established by governments and regulatory bodies to govern the use of artificial intelligence. These legal instruments aim to strike a balance between fostering innovation and protecting society from the potential harms posed by AI technologies (Goodman & Flaxman, 2017).

The core distinction between ethics and law lies in their **binding nature**. Ethics represents **non-binding moral guidance**, whereas laws are **mandatory legal rules** that must be followed. Ethical principles are grounded in moral values, which may vary across individuals and societies, focusing on what *ought* to be done in accordance with ideals of justice and good. Conversely, laws define what *must* be done within a specific legal jurisdiction, offering a more rigid and enforceable framework.

According to Flanagan (2024), the differences between ethics and law can be understood through the following points:

- **Basis of Decision-Making**: Ethical decisions are shaped by personal convictions and cultural norms, rendering them subjective and context-dependent. Legal decisions, however, are objective and apply uniformly to all individuals, irrespective of their beliefs or cultural backgrounds.

- **Consequences of Violations**: Ethical violations may lead to social or professional repercussions, such as loss of reputation or community rejection, but do not typically result in legal penalties. Legal violations, on the other hand, carry explicit consequences such as fines or imprisonment, making their impact more immediate and enforceable.

- **Adaptability**: Ethics is inherently dynamic and can swiftly respond to social change and technological evolution. Law, in contrast, tends to be more rigid, requiring complex legislative procedures for amendment, which may create a time lag between societal shifts and legal updates.

- **Filling Legal Gaps**: Ethics often plays a supplementary role in areas where legal regulation is ambiguous or absent—especially in emerging domains like AI ethics and genetic engineering. In such cases, ethical standards help guide human behavior and decision-making, filling the void left by existing laws.

Thus, **ethics and law together form the foundational pillars of AI governance**, jointly ensuring that the development and application of AI technologies advance justice, transparency, and accountability. While **ethics** provides the **value-oriented principles** to be upheld—such as respect for privacy, prevention of bias, and user safety—**law** enforces these values through legally binding frameworks, outlining responsibilities and sanctions for non-compliance.

Core Ethical Principles in AI Application

The ethical engagement with AI encompasses several key concepts that underpin the responsible use of these technologies (Bostrom, 2014; Zuiderveen et al., 2018; O'Neil, 2016; Taddeo & Floridi, 2018; Zuboff, 2019; Al-Garadi et al., 2020). These include:

- **Responsibility and Accountability**: Responsibility refers to identifying who is liable for the outcomes and decisions made by AI systems—be it developers, users, or the system itself. Accountability ensures that mechanisms are in place to hold relevant parties answerable for the actions of AI systems.

- **Transparency and Explainability**: AI decision-making processes should be clear, understandable, and subject to review. Users and stakeholders must be able to interpret and challenge outcomes generated by AI systems.
- **Fairness and Equity**: AI applications must treat all individuals and groups equitably, without discrimination or bias. Ensuring fairness involves recognizing and mitigating both algorithmic and data-driven inequities.
- **Privacy and Data Protection**: AI systems must respect individual privacy and protect personal data from misuse, unauthorized access, or exploitation.
- **Safety and Security**: AI systems must be designed to ensure the safety of individuals and society at large. This includes protecting systems from cyber threats and preventing malicious use or unintended harm.

The following figure illustrates these fundamental ethical dimensions of AI application and governance:

(Insert Figure: Core Ethical Principles in Artificial Intelligence Use)

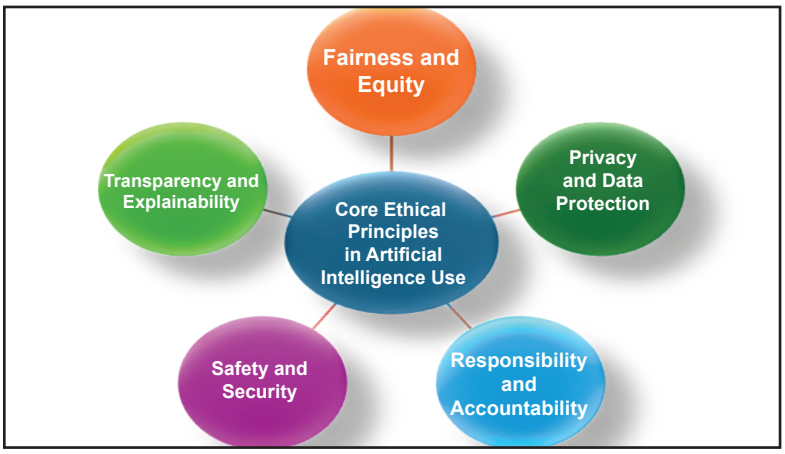

Figure 2. Ethical Dimensions of AI Applications Adapted from Taddeo & Floridi (2018)

There are numerous approaches through which ethical engagement with artificial intelligence applications can be enhanced. Among the most prominent are: establishing clear ethical standards, embedding ethics into the design and development of AI systems, training specialists in ethical best practices, raising user awareness, and promoting scientific research in the field. Considering ethical dimensions during the design and development of AI systems is essential to ensure that these technologies are used responsibly and contribute positively to society. The following figure summarizes the key strategies for promoting ethical AI implementation:

Figure (3): summarizes the key strategies for promoting ethical engagement with AI applications (Adapted from Al-Garadi et al., 2020).

There are various strategies that can significantly contribute to enhancing ethical engagement with artificial intelligence applications. These include establishing clear ethical standards, integrating ethical principles into system design and development, training professionals, raising awareness among users, and encouraging scientific research. Ethical considerations in AI design and development are essential to ensure responsible and beneficial use of the technology.

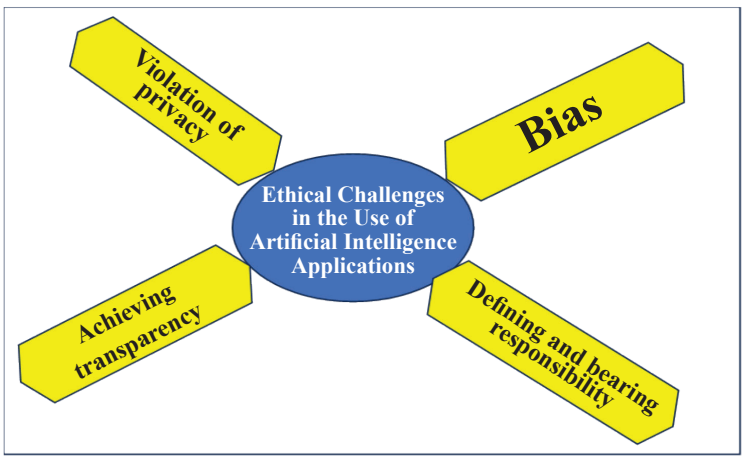

Figure 4. Ethical Challenges in the Use of Artificial Intelligence Applications. (Adapted from Brundage et al., 2020).

Artificial intelligence presents a range of ethical challenges that must be addressed to ensure its responsible and equitable use. One of the most complex issues is the **assignment of responsibility for harm** resulting from AI applications. Determining accountability can be difficult, as AI is neither a natural nor legal person, but rather a technical system based on algorithms and data (Bostrom, 2014).

Responsibility may fall upon multiple actors, including developers, manufacturers, operators, and owners. Given this complexity, there is a pressing need to establish legal frameworks that clearly define responsibilities and offer mechanisms for redress in the event of AI-related harm. These regulations must also address related issues such as **data protection**, **transparency**, and **accountability**. Some countries, including **the European Union, the United States**, and **Singapore**, have begun developing models for legal and ethical accountability in the field of AI.

Decision-making mechanisms within AI systems require a high degree of **transparency**. Both developers and users must understand how AI systems operate and how they produce their outputs. Transparency is vital for building **trust**, identifying **errors and biases**, and improving **system performance** (Doshi-Velez & Kim, 2017).

Another serious challenge is **algorithmic bias**, which can result in significant societal harm. When AI systems are trained on **imbalanced or biased datasets**, the resulting decisions may discriminate against certain populations, such as ethnic or socioeconomic minorities. This can exacerbate inequalities in critical areas such as **employment**, **healthcare**, and **criminal justice**, where biased decisions can have far-reaching and unjust consequences. Therefore, it is essential to develop **fairer algorithms** and improve **data quality** to ensure the responsible use of AI.

Equitable access to AI technologies is also a fundamental ethical concern. All individuals, regardless of background or ability, should have the opportunity to benefit from AI. This includes ensuring **universal access** and promoting **fair usage** (Goodman & Flaxman, 2017).

Furthermore, striking a **balance between security and individual freedoms** is one of the most pressing challenges of the digital age. On the one hand, societies must protect themselves from security threats. On the other, it is imperative to **safeguard individual rights and freedoms**. Achieving this balance requires the implementation of clear laws and regulations rooted in **transparency**, **accountability**, **oversight**, and **continuous improvement**.

The **ethics of data collection and use** in AI is another critical area. Data must be collected **transparently and responsibly**, and used **exclusively for its intended purposes**. Disregarding ethical data practices can result in privacy violations with direct consequences for individuals. For instance, **continuous surveillance** without informed consent may restrict personal freedom and alter behavior due to a constant sense of being watched. Additional risks include **identity theft** and **unauthorized use of private information**, potentially enabling **financial fraud** or other illicit activities (Zuboff, 2019).

The impact of AI on employment is also a subject of ongoing debate. Some argue that AI will lead to **widespread job displacement**, especially in sectors that involve **routine or repetitive tasks**. Others maintain that AI will **create new job opportunities**, particularly in fields related to **AI development, system maintenance, data analysis**, and other emerging areas of advanced technology (Agrawal et al., 2018; Brynjolfsson & McAfee, 2014; Ford, 2015).

III The Egyptian National Artificial Intelligence Strategy

The ethical decision-making capabilities of artificial intelligence (AI) remain a subject of extensive debate in academic and intellectual circles. Some argue that AI, due to its inherently technical nature, is limited by logic and data-driven analysis and lacks the ability to comprehend complex, evolving moral values. Others, however, believe that with the advancement of computational and analytical capacities, AI could—under certain circumstances—make more effective ethical decisions than humans, particularly when handling vast volumes of data and making rapid judgments.

Programming human values into AI systems remains one of the most critical challenges in the field. Ethical values are often implicit, dynamic, and influenced by a multitude of cultural, social, and personal factors. Translating such nuanced values into programmable rules risks either oversimplifying or distorting true human morality.

This issue becomes especially relevant in sensitive domains such as healthcare, law, and security, where AI-driven decisions can have significant consequences on individuals and society. For instance, in the medical field, AI might be required to make difficult decisions about allocating limited medical resources or selecting optimal treatments for complex cases. In the legal domain, AI may assist in analyzing evidence and offering recommendations to judges. In security, AI may be used to monitor borders or detect potential threats.

In response to the growing importance of AI, **Egypt launched its National Artificial Intelligence Strategy** with the aim of leveraging advanced technologies to achieve sustainable development across multiple sectors. The strategy emphasizes building **local capacity**, encouraging **research and innovation**, and developing the **digital infrastructure** necessary to support AI deployment. Additionally, it aims to enhance **international cooperation** and exchange expertise with more technologically advanced nations (Ministry of Communications and Information Technology, 2021).

The Egyptian National Artificial Intelligence Strategy encompasses several key pillars aimed at fostering a robust and sustainable AI ecosystem. These include:

- Developing AI-related skills and competencies,
- Encouraging investment in AI-driven startups, and
- Enhancing the digital infrastructure required to support the deployment and integration of AI applications.

The strategy also places strong emphasis on **building public trust** in AI technologies by ensuring **transparency** and **accountability** throughout their development and implementation (Central Agency for Public Mobilization and Statistics, 2022).

Moreover, the strategy seeks to accomplish several **strategic objectives**, including:

- **Improving the quality of public services**,
- **Fostering innovation** across various economic sectors,
- **Enhancing education and healthcare systems**, and
- **Strengthening cybersecurity frameworks**.

Through this strategy, Egypt aspires to position itself as a **regional hub for artificial intelligence** and play a proactive role in **advancing AI technologies** at the global level (National Council for Artificial Intelligence, 2023).

IV. Mechanisms for Employing AI in Anti-Corruption Efforts

Egypt has also introduced its **National Anti-Corruption Strategy**, which aims to strengthen **integrity and transparency** across state institutions and combat all forms of financial and administrative corruption. The strategy focuses on enhancing **legislative and institutional frameworks**, activating the role of **oversight bodies**, and promoting **international cooperation**. It also emphasizes raising **public awareness** and **engaging civil society** in anti-corruption efforts (Administrative Control Authority, 2023).

Key pillars of the anti-corruption strategy include:

- Promoting **transparency** and strengthening **financial audit mechanisms**.
- Empowering oversight authorities to **investigate and prosecute** corruption cases.
- Reforming the **civil service system**, particularly recruitment and promotion procedures.
- Enforcing **accountability** for misconduct and violations.

The strategy seeks to achieve several strategic goals, including:

- Reducing the prevalence of corruption across sectors.
- Enhancing public trust in government institutions.
- Improving the investment climate and attracting foreign capital.
- Advancing sustainable development and public sector efficiency.

Artificial intelligence plays an increasingly significant role in Egypt's anti-corruption efforts. AI tools can analyze massive datasets to detect suspicious patterns and predict potential violations. For example, machine learning algorithms can analyze financial and administrative data to identify signs of corruption or public fund mismanagement. AI can also be used to review **government contracts** and **monitor projects** to ensure transparency and detect fraudulent practices (Ministry of Communications and Information Technology, 2023).

The range of AI tools used in anti-corruption efforts includes:

- Big data analytics for detecting anomalies.
- Natural Language Processing (NLP) for analyzing documents and communications to flag suspicious content.
- Computer vision to inspect images and videos for signs of tampering or forgery.
- Early warning systems powered by AI to detect and respond to risks proactively.

One notable example involves the **Administrative Control Authority**, which successfully used AI-based data analysis to uncover a criminal network involved in **forging official government documents**. These forged documents falsely entitled certain individuals to benefits reserved for persons with disabilities, including exceptional pensions and customs-free vehicle imports—claims that were proven fraudulent (Administrative Control Authority, 2023).

V. Theoretical Framework of the Study

A suitable theoretical foundation for this study is the Ethical AI Governance Theory, which posits that the development and deployment of AI systems should be governed by a comprehensive framework incorporating core ethical principles such as justice, transparency, and accountability. According to this theory, institutions and regulatory bodies must adopt stringent policies to ensure that intelligent systems align with human and societal values (Floridi et al., 2018).

The theory emphasizes the importance of developing interpretable AI systems and minimizing biases that may distort outputs (Jobin, Ienca, & Vayena, 2019).

It also advocates for the principle of Ethics by Design, whereby ethical thinking is embedded in every stage of the AI development lifecycle—from data collection to real-world deployment (Mittelstadt, 2019).

Studies suggest that applying this theory contributes to public trust in AI technologies and helps mitigate social risks such as privacy violations or discriminatory decision-making (Morley et al., 2020). Furthermore, this framework aligns with international guidelines issued by organizations such as UNESCO and the European Commission, which aim to establish legal and ethical standards ensuring AI is used in ways that respect human rights and uphold community values. Thus, the Ethical AI Governance Theory offers a relevant and robust conceptual framework for analyzing ethical standards in the design and development of AI applications in the digital age (European Commission, 2020).

VI. Current Situation Analysis of AI Ethics in Egypt

Transparency is one of the most critical ethical principles in AI development and deployment. However, Egypt faces notable challenges in this regard—primarily due to limited disclosure about the algorithms used in various sectors. This lack of transparency may hinder accountability and negatively affect decision-making processes. According to Zuboff (2019) and Brey & Dainow (2023), enhancing transparency requires the enforcement of clear policies that oblige developers to disclose how intelligent systems function and their potential impacts on users. This remains a work in progress in Egypt's digital environment.

In terms of **fairness**, there is growing concern about algorithmic bias affecting outcomes in areas such as **employment**, **education**, and **public services**. Studies by Anderson (2007) and Bostrom (2014) point out that such biases often stem from datasets that fail to adequately represent all social groups, resulting in unbalanced or unjust outputs. In Egypt, there are currently limited initiatives to ensure **data fairness** in AI model training, underscoring the need for **legal and technical interventions** to establish equitable standards for AI application development and usage.

Regarding **privacy and cybersecurity**, Egypt's increasing reliance on AI raises valid concerns about protecting **personal data** and ensuring its responsible use. As highlighted by Cath (2018) and Brey & Dainow (2023), effective data protection requires both **clear legislation** and **robust enforcement mechanisms** to prevent breaches or misuse. Although Egypt has enacted personal data protection laws, **practical enforcement** remains a challenge, particularly with the rise of **big data analytics** and AI-driven platforms.

VII. Deriving a List of Ethical Standards for AI Design and Development

Establishing ethical standards for the design, development, and implementation of AI applications is of utmost importance in the modern era, as AI becomes increasingly integrated into daily life. These applications must be built in a manner that upholds **human rights**, respects **fundamental freedoms**, and avoids any form of **discrimination** or **bias**.

AI systems should be **transparent and accountable**, allowing users to understand how they function and how decisions are made (Anderson & Anderson, 2007). Ethical standards in AI development encompass a wide array of issues, including:

- **Privacy**: Ensuring that users' personal data is protected from unauthorized access or misuse.
- **Security**: Designing systems that are resilient to cyber threats and tampering.
- **Responsibility**: Clearly identifying who is accountable in cases where AI usage leads to harm.
- **Transparency**: Making AI decision-making processes interpretable and auditable.

AI systems must also be developed to prevent discrimination and promote fairness across all user groups. They should be evaluated rigorously to ensure that no hidden biases or design flaws result in unjust outcomes (Bostrom, 2014).

Additionally, AI developers must ensure that the systems they build are equitable, do not disadvantage any group, and are subject to comprehensive testing to identify and rectify potential biases (O'Neil, 2016). These measures are essential for building trustworthy, ethical, and human-centered AI systems.

Social responsibility is among the most critical ethical standards in the design of artificial intelligence applications. Designers must be fully aware of the potential social impacts of their technologies and be willing to accept responsibility for any harm that may arise from their use (Russell & Norvig, 2021).

The following is an overview of the key steps for integrating ethics into the design and development of AI applications, as proposed by Brey & Dainow (2023):

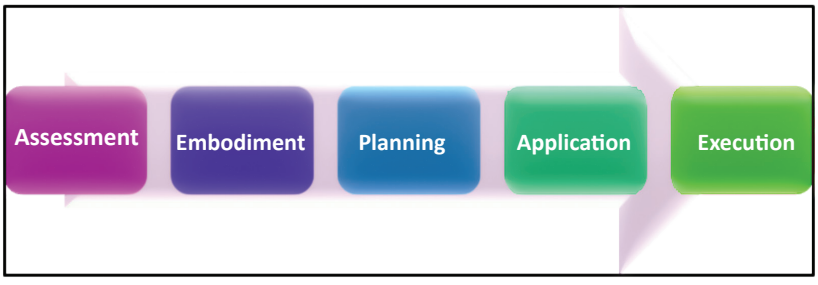

Figure 5. Steps for Applying Ethics in the Design and Development of AI Applications. (*Adapted from Brey & Dainow, 2023)*

Steps for Implementing Ethics in AI Design and Development

1. **Assessment**: Initially, the system's objectives are evaluated against core ethical values to ensure that none of these essential values are compromised.

2. **Embodiment**: Ethical values are translated into specific indicators that must be present within the AI system being used.

3. **Planning**: This phase involves converting general ethical design requirements into specific actions to be implemented during the design process. These requirements are executed through various means, including system functions, data structures, and regulatory measures.

4. **Application**: This step identifies the method by which each ethical requirement will be addressed. It helps provide a generalized development model, making it easier to integrate ethical requirements across various stages of design and implementation. In some cases, the application process may need to be revisited.

5. **Execution**: The final step in applying ethics by design during the development process. While seemingly simple, this phase constitutes the bulk of the work when developing an AI system. In some instances, ethical requirements may necessitate adding new functions, whereas in other cases, they may act as constraints, limiting allowed functionalities. Many of these requirements will also require additional regulatory processes. For example, to avoid bias in programming, a formal assessment of data bias must be conducted before data usage.

Steps to Develop a List of Ethical Standards for AI Design and Development

The following steps outline how the list of ethical standards for the design and development of AI applications was created:

1- **Defining the Purpose of the Ethical Standards List**: A clear and precise definition of the ethical standards list for AI design and development was established, with the intention of benefiting **programmers** and **developers**, as well as having a positive impact on all **end-users**.

2- **Deriving Standards and Indicators**: A number of relevant studies and literature sources were consulted to develop the initial draft of the ethical

standards list, including works such as **Jobin et al., 2019** and **Brey & Dainow, 2023**.

3. **Verification of the Authenticity of the Standards List**: After preparing the initial draft of the list (Appendix 1), the questionnaire containing the standards was presented to a group of **judges and specialists** in the field of **computer science and information technology** (Appendix 2). The feedback received addressed the following:

- The importance of the main areas covered by the standards list.
- The appropriateness of the standards to achieve the intended objectives.
- The linguistic accuracy of the list's phrasing.
- Suggestions for additions or deletions to the list.

4- Verification of the Consistency of the Standards List: Based on the feedback from the judges, the list was revised. The agreement rate on the items was calculated using the Cohen's Kappa formula:

Agreement percentage= number of agreements / number of agreements + number of disagreements ×100.

Standards with an agreement rate of **85% or higher** were retained, while those with lower agreement rates were excluded. Additionally, the majority of judges agreed on modifying the wording of certain standards, removing others, and rephrasing some indicators (Appendix 3).

5- Final Version of the Standards List: All necessary modifications were made, and the final list now includes 15 standards and 73 indicators (Appendix 4).

Research Results

The study aimed to verify a set of hypotheses related to the ethical standards for the development of AI applications. The results are as follows:

Results of Testing the First Hypothesis:

The first hypothesis of the study states: "There are key ethical standards that must be considered when developing AI applications." The results of the survey of previous studies confirmed the validity of this hypothesis. A list of **15 key ethical standards** was identified, which includes **responsibility**, **accountability**,

transparency, **fairness**, **privacy**, and **environmental sustainability**. The results also showed that adherence to these standards helps improve the reliability and efficiency of applications. Specifically, **responsibility**, **accountability**, and **transparency** enhance user trust, while focusing on **fairness** and **privacy** helps reduce ethical and legal issues.

Results of Testing the Second Hypothesis:

The second hypothesis of the study posits the need for further research to understand the impact of these ethical standards on developers and users. The theoretical analysis of relevant studies confirmed the validity of this hypothesis, emphasizing the need for in-depth exploration of how adherence to ethical standards influences user experience and how developers respond to compliance requirements.

Answers to the Research Questions:

- To answer the first question of the research, which aims to identify the most important ethical standards that should be considered when developing AI applications, the study identified a list of 15 standards, including: responsibility, accountability, transparency, fairness, privacy, and environmental sustainability.
- Regarding how to ensure that AI applications adhere to these standards, the analytical study of previous research highlights the need for clear policies governing the development process, conducting regular tests to ensure compliance with ethical standards, and involving users and experts in evaluating these applications.
- In terms of identifying the consequences of adhering to or neglecting these standards, the results showed that adherence to these standards leads to improved reliability of applications and increased user trust, while failure to comply can lead to legal, ethical, and even commercial problems, such as loss of customer trust and the emergence of biases in AI decisions.

The research results reflect the global trend toward responsible AI development, as governments, academic institutions, and technology organizations work to establish ethical standards to regulate this field. The emphasis on transparency and fairness stems from increasing concerns about bias and unfair impacts caused by some applications. Furthermore, the inclusion of environmental sustainability reflects growing awareness of the environmental impact of technology. The research

also indicates that compliance with these standards should not remain a theoretical concept but must be translated into clear policies and technical procedures during the design and development stages. This requires collaboration between developers, users, and policymakers to ensure a balance between innovation and responsibility.

Key Recommendations and Conclusions:

In light of the previous studies reviewed and the derivation of a list of ethical standards for the design and use of AI applications, the following recommendations are made:

- International organizations and governments should work on establishing clear and comprehensive ethical standards that govern the design, development, and use of AI systems. These standards should ensure that fairness and integrity are maintained in these technologies. It is essential for AI systems to be transparent and explainable, enabling users to understand how they function and make informed decisions regarding their use. Moreover, clear accountability mechanisms should be put in place to address any potential harm or negative consequences that may arise from these technologies.

- Stringent measures should be taken to protect personal data and prevent its exploitation or misuse in AI applications. This can be achieved through the enforcement of robust privacy policies and the implementation of effective regulatory mechanisms. To ensure the ethical and responsible use of these technologies, international collaboration between governments, international organizations, and civil society is critical. This collaboration will help establish unified policies that guide AI towards serving the public good.

- Public engagement should be a priority in discussions surrounding AI ethics. Open dialogues and awareness-raising events should be organized to foster public understanding of the issues associated with this technology. This will help ensure its responsible and sustainable use, benefiting all members of society.

- Further research is needed to explore ethical standards from the perspectives of both developers and users, as this will provide a deeper understanding of the challenges involved. Additionally, practical solutions should be proposed to effectively implement each ethical principle in AI systems. This will help bridge the gap between theoretical principles and practical application, leading to the development of responsible smart systems.

References

-Arabic References:

- Al-Arabi, Ahmed (2021). Artificial Intelligence: The Coming Technological Revolution. Arab Thought Foundation.
- Central Agency for Public Mobilization and Statistics (2022). The National Strategy for Artificial Intelligence in Egypt: A Future Vision. Cairo: Central Agency for Public Mobilization and Statistics.
- Central Auditing Organization (2021). Report on the Role of Artificial Intelligence in Combating Corruption. Cairo: Central Auditing Organization.
- National Council for Artificial Intelligence (2023). Report on the Progress Made in Implementing the National Artificial Intelligence Strategy. Cairo: National Council for Artificial Intelligence.
- Administrative Control Authority (2023). National Strategy for Combating Corruption. Cairo: Administrative Control Authority.
- Administrative Control Authority (2023). Administrative Control Authority Uncovers a Criminal Group Specializing in the Forgery of Government Documents. (https://aca.gov.eg/News/5217.aspx)
- Ministry of Communications and Information Technology (2021). National Strategy for Artificial Intelligence in Egypt. Cairo: Ministry of Communications and Information Technology.
- Ministry of Communications and Information Technology (2023). National Strategy for Artificial Intelligence. Cairo: Ministry of Communications and Information Technology.

- English References :

- Agrawal, A., Gans, J., & Goldfarb, A. (2018). Prediction Machines: The Simple Economics of Artificial Intelligence. Harvard Business Review Press.
- Al-Garadi, M. A., Mohamed, A., Al-Ali, A. K., Du، X., Ali, I.، & Guizani, M. (2020). A survey of machine and deep learning methods for internet of things (IoT) security. IEEE Communications Surveys & Tutorials, 22(3), 1646-1685.

- Anderson, M. (2007). Machine ethics: Creating an ethical intelligent agent. AI & Society, 22(4), 477-493.
- Anderson, M., & Anderson, S. (2007). Machine ethics: New directions. Oxford University Press.
- Berendt, B (2019) AI for the common good?! Pitfalls, challenges, and ethics pen-testing. Paladyn, Journal of Behavioral Robotics 10(1): 44–65. DOI: 10.1515/pjbr-2019-0004.
- Bostrom, N. (2014). Superintelligence: Paths, dangers, strategies. Oxford University Press.
- Brey, P., & Dainow, B. (2023). Ethics by design for artificial intelligence. AI and Ethics, 1-13.
- Brundage, L., Avin, S., Clark, J., Toner, M., Eckersley, P., Garfinkel, S., & Anderson, R. (2020). Toward trustworthy AI development: Aligning theoretical and practical perspectives. arXiv preprint arXiv:2004.07218.
- Brynjolfsson, E & McAfee, A. (2014). The Second Machine Age: Work Progress and Prosperity in a Time of Brilliant Technologies. W. W. Norton & Company.
- Bryson, J. (2018). The artificial intelligence of the ethics of artificial intelligence: An introductory overview for policymakers. Cambridge Handbook of Artificial Intelligence, 316-334.
- Bryson, J. J., Taddeo, M & Floridi, L. (2019). Ethics of artificial intelligence. The Oxford handbook of digital ethics، 109-139.
- Cath, C. (2018). Governing artificial intelligence: Ethical, legal, and societal issues. Philosophy & Technology, 31(4), 611-627.
- Cath, C; Wachter, S; Mittelstadt, B. (2018) Artificial intelligence and the 'good society': The US, EU, and UK approach. Science and Engineering Ethics 24(2): 505–528. DOI: 10.1007/s11948-017-9901-7.
- Craglia M, Annoni, A & Benczur. P. (2018). Artificial intelligence : a European perspective. Publications Office of the European Union, Luxembourg

- Doshi-Velez, F & Kim, B. (2017). Towards a rigorous science of interpretable machine learning. arXiv preprint arXiv:1702.08608.
- European Commission. (2020). White paper on artificial intelligence: A European approach to excellence and trust. Publications Office of the European Union.
- Flanagan, W. (2024). Ethics vs Laws: 5 Key Differences Experts Want You to Know in 2024. Available at: https://www.infonetica.net/articles/ethics-vs-laws
- Floridi, L. (2019). The ethics of artificial intelligence: Principles, challenges, and opportunities. Minds and Machines, 29(4), 569-589.
- Floridi, L. (2020). The ethics of artificial intelligence. Oxford University Press.
- Floridi, L; Cowls, J; Beltrametti, M. (2018). AI4People – an ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. Minds and machines 28(4): 689–707.
- Ford, M. (2015). Rise of the Robots: Technology and the Threat of a Jobless Future. Basic Books.
- Goodman, B & Flaxman, S. (2017). European Union regulations on algorithmic decision-making and a "right to explanation". AI Magazine, 38(3), 50-57.
- Hagendorff, T. (2020). The ethics of AI ethics: An evaluation of guidelines. Minds and Machines, 30، 99-120.
- Harari, Y. (2016). Homo Deus: A Brief History of Tomorrow. Harper.
- Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. Nature Machine Intelligence, 1(9), 389-399.
- Köbis, N; Starke, C & Rahwan, I. (2021). Artificial intelligence as an anti-corruption tool (AI-ACT). Potentials and pitfalls for top-down and bottom-up approaches.
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, 521(7553), 436-444.
- Machanavajjhala, a; Korolova, A & Sarma, A. (2011) Personalized social recommendations – accurate or private? Proceedings of the VLDB Endowment

(PVLDB) 4(7).

- Marwick, A. E & Lewis, R. (2017). Media Manipulation and Disinformation Online. Data & Society Research Institute.
- McKendrick, J. (2022). 7 Steps to More Ethical Artificial Intelligence. Available at: https://www.forbes.com/sites/joemckendrick/2022/06/10/7-steps-to-more-ethical-artificial-intelligence/
- Mittelstadt, B. D. (2019). Principles alone cannot guarantee ethical AI. Nature Machine Intelligence, 1(11), 501-507.
- Mittelstadt, B. D., Allo, P., Taddeo، M.& Wachter, S. (2016). The ethics of algorithms: Mapping the debate. Big Data & Society, 3(2), 2053951716679679.
- Morley, J., Floridi, L., Kinsey, L., & Elhalal, A. (2020). From what to how: An initial review of publicly available AI ethics tools, methods, and research to translate principles into practices. Science and Engineering Ethics, 26(4), 2141-2168.
- Naqvi, A. (2020). Artificial intelligence for audit، forensic accounting، and valuation: A strategic perspective. John Wiley & Sons. https://doi.org/10.1002/9781119601906
- Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. Science, 366 (6464), 447-453.
- OECD (2021). Recommendation of the council on artificial intelligence.
- O'Neil, C. (2016). Weapons of math destruction: How big data increases inequality and threatens democracy. Crown Publishing Group.
- Rahwan, I., Cebrian, M., Obradovich, N., Bongard, J., Bonnefon, J. F., Breazeal, C. & Rahwan, T. (2019). Machine behavior. Nature, 568(7753), 477-486.
- Russell, S., & Norvig, P. (2021). Artificial intelligence: A modern approach (4th ed.). Pearson.

- Schwartz, S. H. (1992). Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries. In M. Zanna (Ed.), Advances in Experimental Social Psychology (Vol. 25, 1 pp. 1-65). Academic Press.
- Shin, D. & Park, Y. (2019). Role of fairness, accountability, and transparency in algorithmic affordance. Computers in Human Behavior 98: 277–284.
- Shneiderman, B. (2020). Human-centered artificial intelligence: Reliable, safe & trustworthy. International. Journal of Human–Computer Interaction 36(6): 495–504.
- Sreenu, D. (2023). How can artificial intelligence accelerate scientific research? https://scientificnirvana.com/how-can-artificial-intelligence-accelerate-scientific-research
- Stone, P., Brooks, R., Brynjolfsson, E., Calo, R., Etzioni, O., Hager, G., & Walsh, T. (2016). Artificial intelligence and life in 2030. One Hundred Year Study on Artificial Intelligence: Report of the 2015-2016 Study Panel, Stanford University, Stanford, CA.
- Taddeo, M., & Floridi, L. (2018). How AI can be a force for good. Science, 361(6404), 751-752
- UNESCO. (2021). Recommendation on the Ethics of Artificial Intelligence.
- Wardle, C & Derakhshan, C. (2017). Information Disorder: Toward an interdisciplinary framework for research and policymaking. Council of Europe report, DGI (2017) 9.
- Zuboff, S. (2019). The age of surveillance capitalism: The fight for a human future at the new frontier of power. Public Affairs.
- Zuiderveen, F. J., Möller, J., S., Fathaigh, R., Irion, K., Dobber, T, & Hildebrandt, M. (2018). Discrimination، artificial intelligence، and algorithmic decision-making. Utrecht Law Review, 14(2), 69-201791.

Appendix (1) Initial Draft of Ethical Standards List for Designing and Developing AI Applications

Ethical Domain	Indicators	References
Transparency	– User awareness of how the application works and makes decisions, ensuring they understand the rationale behind these decisions. – Clear disclosure to all stakeholders about the system's objectives, capabilities, limitations, benefits, risks, and decisions made. – Designing AI systems so individuals can audit, inquire, challenge, or request modifications of their activities. This includes regulatory processes allowing operators to receive and assess external parties' requests. – Implementing measures to enable the tracking and monitoring of the AI system to ensure transparency and accountability. – Making decisions on ethical issues, such as how to remove biases from a dataset. – Documenting development processes and tools for ethical decisions, making it possible to understand how ethical standards were followed, enabling audits, objections, or corrections of ethical issues after deployment.	(O'Neil, 2016; Floridi, 2019; Brey & Dainow, 2023)
Clarification	– Ability to clarify the reasons for various decisions made by developers and the AI system. – AI decisions must be interpretable, and the reasons behind each decision should be clearly explained. – Providing a mechanism to explain decision-making and data interpretation for decisions made.	(Zuboff, 2019; Brey & Dainow, 2023)

Ethical Domain	Indicators	References
Accountability	– Providing accountability mechanisms in case harm results from using the application, with the possibility of obtaining compensation for damage. – Providing human oversight of operational cycles and decision–making processes in AI systems. – Including a risk assessment in the AI system deployment process, with mitigation measures in place after deployment, ready from the moment the system is operational. – Ensuring AI systems are auditable by independent parties, covering not only system decisions but also the processes and tools used during development.	(Floridi, 2019; Shneiderman, 2020; Russell & Norvig, 2021)
Accuracy	– Developers should rigorously train the AI system to operate with very high accuracy. – Minimizing errors to the lowest possible level, ensuring that they are nearly nonexistent.	(Zuiderveen et al., 2018; Naqvi, 2020)
Fairness	– Designing applications in a way that ensures no discrimination or bias against any group of people. – Avoiding algorithmic bias in AI, including bias in input data, modeling, and algorithm design. – Making AI systems globally accessible where possible and ensuring they provide the same functionalities and benefits to all users, regardless of their abilities, beliefs, preferences, or interests. – Designing AI systems to avoid negative social impacts on specific social groups, especially protected categories.	(Anderson, 2007; O'Neil, 2016; Anderson & Berendt, 2019)

Ethical Domain	Indicators	References
Privacy	– Designing applications to protect users' privacy. – Securing applications for personal data safety, preventing unauthorized access to personal data.	(Anderson, 2007; Bostrom, 2014 & Anderson)
Autonomy	– Proper training of developers for the AI system. – Ensuring the system can operate without human supervision.	(Cath, 2018; Brey & Dainow, 2023)
Human Supervision / Human Agency	– Ensuring human oversight for sensitive or important decisions. – Designing AI systems to assist humans in decision–making rather than completely replacing them. – Not designing or using AI systems in a way that deprives individuals of decisions they should be capable of making themselves. – Not designing or using AI in a way that reduces basic human freedoms, including freedom of movement, assembly, expression, and access to information. – Designing or using AI systems to create dependency on the system or its services. – AI systems should not make the final decision on important personal, ethical, or political issues but can offer recommendations.	(Obermeyer et al., 2019; Brundage et al., 2020 & Brey & Dainow, 2023)

Ethical Domain	Indicators	References
Security	– Design applications to ensure the safety of users and society, and avoid using them for harmful or unethical purposes. – Respect data owners' rights when processing personal data. – Provide mechanisms that clarify how legality, fairness, and transparency are achieved when processing personal data by AI. – Implement measures to protect data owners' rights through technical procedures, such as anonymization, and regulatory measures, such as access control systems. – Support AI systems for individuals' right to withdraw consent for the use of their personal data. – Collect, store, and process data in a way that allows for human audit.	(Bostrom, 2014; Al-Garadi et al., 2020; Russell & Norvig, 2021)
Reliability and Cybersecurity	– Protect systems from cyber-attacks and manipulation. – Ensure service continuity and prevent disruption due to threats. – Maintain the security of the service to protect users and data.	(Craglia et al., 2018; Brey & Dainow, 2023)
Social Responsibility	– Developers and producers should be aware of the social impacts of their applications. – Consider the potential impact of applications on society. – Constantly strive to design applications that serve the public good.	(Anderson & Anderson, 2007; Russell & Norvig, 2021)
Human Values	– AI systems should respect human rights. – Consider the diversity and differences between individuals and communities. – Ensure individuals' autonomy in making decisions without unjustified restrictions.	(Köbis et al., 2021; Brey & Dainow, 2023)

Ethical Domain	Indicators	References
Human, Social, and Environmental Welfare	– Ensure AI systems benefit individuals, society, and the environment. – Ensure the use of AI promotes human rights. – Contribute to achieving the general welfare through AI. – Prevent AI from being used in a way that undermines human rights and welfare.	(Brynjolfsson & McAfee, 2014; Brey & Dainow, 2023)
Sustainability and Environmental Responsibility	– Ensure that AI systems are safe for use and have no tendency to harm or significantly deteriorate the physical or mental health and welfare of any parties involved, including users, clients, data owners, and other affected parties. – Ensure the development of AI aligns with environmental sustainability principles. – Ensure AI systems do not negatively impact the quality of communication, social interaction, information flow, social relationships, or democratic processes.	(Jobin et al., 2019; Floridi, 2019; Brey & Dainow, 2023; Flanagan, 2024)

Appendix (2): List of the reviewers

No.	Name	University	Faculty	Position
1	Engy Abdel-Maboud Bayoumi	Sadat City	Computers and Artificial Intelligence	Computer Science Lecturer
2	Ahmed Abdel-Hamid Saleh	Damanhur	Computers and Information	Computer Science Lecturer
3	Rasha Mohamed Muntasir	Damanhur	Computers and Information	Information Systems Lecturer
4	Amr Fathi El-Sayed	Damanhur	Computers and Information	Computer Science Lecturer
5	Mai Helmy Shaheen	Damanhur	Computers and Information	Information Technology Lecturer
6	Ahmed Hamid Attia	Damanhur	Computers and Information	Computer Science Lecturer
7	Mohamed Abdel-Latif Said El-Ahl	Damanhur	Computers and Information	Computer Science Lecturer

Appendix (3): List of Modifications by the reviewers

Reviewer 1:

- Added a new indicator under the Transparency criterion.
- Added a new indicator under the Autonomous Independence criterion.
- Added a new indicator under the Human Oversight criterion.
- Added a new indicator under the Human Welfare criterion.
- Added a new criterion (Adaptability) and included two indicators under this criterion.

Reviewer 2:

- Revised the wording of the indicators (first, second, third, and fifth) under the Transparency criterion.
- Added a new indicator under the Accountability criterion.
- Added a supplementary part to the first indicator under the Accuracy criterion.
- Added a new indicator under the Accuracy criterion.
- Revised the wording of the fifth indicator under the Human Oversight domain.
- Added an indicator under the Social Responsibility criterion.

Reviewer 3:

- In the Accuracy criterion, the phrase “Minimizing errors to the lowest possible extent, so that they are almost nonexistent” is used. We suggest

specifying an acceptable error rate.

- There are two points that may conflict under the Independence criterion, "The system's ability to function without human supervision" and under Human Oversight, "Ensuring human oversight over critical or professional decisions." We suggest modifying the Independence criterion to: "The system's ability to function without human supervision except for critical or professional decisions."
- Under the Human Agency criterion, the phrase "Design or use of AI systems to create addiction to the system or the services it provides" is used. We suggest rephrasing it as: "Not designing or using AI systems to create addiction to the system or the services it provides."

Reviewer 4:

- Revised the wording of the first indicator under the Transparency criterion.
- Rephrased the third indicator under the Transparency criterion.
- Added a new indicator under the Clarification criterion.
- Added a new indicator under the Accuracy criterion.
- Added a new indicator under the Privacy criterion.
- Added a new indicator under the Independence criterion.
- Added a new indicator under the Security criterion.
- Added a new indicator under the Reliability criterion.
- Added a new indicator under the Social Responsibility criterion.
- Added a new indicator under the Human Values criterion.
- Added a new indicator under the Human Welfare criterion.
- Added two indicators under the Sustainability and Environmental Responsibility criterion.

Reviewer 5:

- Added a new indicator under the Privacy criterion.
- Rephrased one of the indicators under the Independence criterion.

Reviewer 6:

- Rephrased the indicators under the Accountability criterion.
- Rephrased the indicators under the Fairness criterion.

Reviewer 7:

- Rephrased the indicators under the Security criterion.
- Rephrased the indicators under the Sustainability and Environmental Responsibility criterion.

Appendix (4): Final List of Ethical Criteria for the Design and Development of Artificial Intelligence Applications

Ethical Domain	Indicators	References
Transparency	· Raising user awareness about how the intelligent application works and makes decisions, ensuring that users can understand the smart services as well as the reasons behind these decisions. · Clearly disclosing to all relevant parties the purpose of the intelligent system, its capabilities, limitations, benefits, risks, and the decisions it makes. · Ensuring system auditability and allowing for inquiries or appeals regarding its decisions, including organizational processes that enable users to receive and evaluate the services provided by the intelligent system. · Implementing measures that allow for tracking and monitoring the AI system to ensure transparency and accountability. · Making decisions on ethical issues, such as how to remove bias from a specific category of data. · Documenting the processes and tools used in ethical decision-making during development to make it possible to understand how ethical standards are followed. This will allow for audits, objections to system decisions, or correction of any ethical issues that arise after deployment. · Clearly informing users about the types of data collected and the purpose for collecting it.	(O'Neil, 2016; Floridi, 2019; Brey& Dainow, 2023)

Ethical Domain	Indicators	References
Explanation	· Clarifying the reasons behind different decisions made by both the developer and the intelligent system. · Ensuring that AI decisions are interpretable, with clear explanations of why the system made a specific decision. · Providing mechanisms for explaining decision-making and the data used, as well as ensuring the interpretability of decisions made by the system. · Developing Explainable AI (XAI) systems to ensure that AI decisions are understandable by everyone.	(Zuboff, 2019; Brey& Dainow 2023)
Accountability	· Providing accountability mechanisms in case of harm caused by the application, and the possibility of obtaining compensation in case of damage. · Ensuring human oversight over operational cycles and decision-making within AI systems. · Including a risk assessment process in the deployment of AI systems, with procedures to mitigate risks after deployment, ensuring readiness from the moment the system is operational. · Ensuring AI systems are auditable by independent parties. This audit should cover not only the system's decisions but also the processes and tools used during the development phase. · Continuous monitoring of intelligent systems to track potential unforeseen future risks.	(Floridi, 2019; Shneiderman, 2020; Russell & Norvig, 2021)

Ethical Domain	Indicators	References
Accuracy	· Ensuring developers provide intensive training for the AI system to ensure it operates with high accuracy, which requires training the system on large, multi-pattern datasets. · Using auditable machine learning techniques to ensure there are no undetected errors. · Minimizing errors to the lowest possible level, ensuring they do not exceed 1%. · Using AI model metrics to measure the accuracy of the models used by applications.	(Zuiderveen, et al., 2018; Naqvi, 2020)
Equity	· Using AI model metrics to measure the accuracy of the models used by applications. · Designing applications in a way that ensures there is no discrimination or bias against any group of people. · Avoiding algorithmic bias in AI systems, including bias in input data, modeling, and algorithm design. · Making AI systems globally accessible whenever possible, ensuring they provide the same functions and benefits to all users regardless of their abilities, beliefs, preferences, or interests. · Designing AI systems to avoid negative social impacts on social groups, especially protected groups, whenever possible.	(Anderson, 2007; O'Neil, 2016; Anderson & Berendt, 2019)

Ethical Domain	Indicators	References
Privacy	· Designing applications to protect users' privacy. · Securing applications to protect personal data and prevent unauthorized access to it. · Using techniques such as encryption and federated learning to protect data. · Right to be Forgotten: This legal principle grants individuals the right to request the deletion of their personal data from databases of companies or digital service providers when there is no legal reason to retain it. This concept is a core part of the General Data Protection Regulation (GDPR) in the European Union, enhancing individuals' privacy online.	(Anderson, 2007; Bostrom, 2014 & Anderson)
Autonomous Independence	· Properly training the developers of the intelligent system. · The system's ability to operate without human supervision except for critical or professional decisions. · Ensuring that users have full control over their personal data. · Requiring developers to provide monitoring tools that allow for human intervention in emergency situations.	(Cath, 2018; Brey& Dainow, 2023)

Ethical Domain	Indicators	References
Human Oversight / Human Agency	· Ensuring human oversight over sensitive or critical decisions. · Designing AI systems that assist humans in decision-making rather than replacing them entirely. · Establishing a list of sensitive decisions that should remain under full human oversight. · Not designing or using AI systems in a way that deprives individuals of the ability to make decisions that they should be able to make themselves. · Not designing or using AI systems in a way that diminishes fundamental human freedoms, including the freedom of movement, assembly, expression, and access to information. · Not designing or using AI systems to create addiction to the system or the services it provides. · Not allowing AI systems to make final decisions on important personal, ethical, or political matters, but they can provide recommendations. · Conducting periodic evaluations to ensure continuous compliance with ethical standards.	(Obermeyer et al., 2019; Brundage et al., 2020 & Brey& Dainow, 2023)
الأمن	· Designing applications in a way that ensures the safety of users and society, and not using them for harmful or unethical purposes. · Respecting the rights of data subjects when processing personal data. · Providing mechanisms that clarify how legitimacy, justice, and transparency are achieved when processing personal data through AI. · Implementing measures to protect data subjects' rights through technical procedures, such as anonymization, and regulatory measures, such as access control systems. · Enhancing AI systems with the right for individuals to withdraw consent for the use of their personal data. · Collecting, storing, and processing data in a way that allows it to be audited by humans. · Providing quick-response mechanisms in case of security vulnerabilities.	(Bostrom, 2014; Al-Garadi et al., 2020; Russell & Norvig, 2021)

Ethical Domain	Indicators	References
Reliability and Cybersecurity	· Protecting systems from cyberattacks and manipulation. · Ensuring service continuity and preventing disruption due to threats. · Maintaining the security of the service to protect users and data. · Requiring companies to provide emergency plans in case of service outages.	(Craglia et al., 2018; Brey& Dainow, 2023)
Social Responsibility	· Raising awareness among developers and producers about the social impacts of their applications. · Considering the potential impact of applications on society. · Constantly striving to design applications that serve the public good. · Educating about the benefits and harms of artificial intelligence. · Providing mechanisms to receive user feedback and improve systems based on it.	(Anderson, & Anderson 2007; Russell & Norvig, 2021)
Human Values	· Ensuring AI systems respect human rights. · Considering diversity and differences among individuals and communities. · Ensuring individuals' autonomy in making their decisions without unjustified constraints. · Providing training for developers on how to design systems that respect human values.	(Köbis et al., 2021; Brey& Dainow, 2023)
Human, Social and Evironmental Welfare	· Ensuring AI systems benefit individuals, society, and the environment. · Ensuring that AI usage enhances human rights. · Contributing to the achievement of public welfare through AI. · Preventing the use of AI in a way that undermines human rights and welfare. · Designing AI systems to be usable by people with disabilities or marginalized groups. · Requiring companies to conduct environmental impact assessments for their systems.	(Brynjolfsson & McAfee,2014; Brey& Dainow, 2023)

Ethical Domain	Indicators	References
Sustainability and Environmental Responsibility	· Ensuring that AI systems are safe to use and do not have the tendency to cause harm or significantly degrade the physical or mental health and well-being of any involved parties, including users, customers, data owners, and other affected parties. · Considering the development of AI in line with environmental sustainability principles. · Ensuring that AI systems do not negatively impact the quality of communication, social interaction, information flow, social relationships, or democratic processes. · Developing standards to ensure that AI systems are environmentally friendly. · Encouraging the use of renewable energy for operating data centers.	(Jobin et al., 2019; Floridi, 2019; Brey& Dainow, 2023; Flanagan, 2024)
Adaptability	• Ensuring adaptability to changes. • Ensuring the system's ability to adapt to different contexts without losing effectiveness or accuracy.	Salehie, M., & Tahvildari, L. (2009). (TAAS)