SBS Research Monograph
Chapter 12: Ethical Dilemmas in Artificial Intelligence Implementations: How to Overcome It?
DOI: https://doi.org/10.70301/SBS.MONO.2025.1.12

Chapter 12 - Ethical Dilemmas in Artificial Intelligence Implementations: How to Overcome It?

Gatis Polis 1[1]
Tatjana Volkova 2[1]

[1] BA School of Business and Finance, Riga, Latvia

Chapter Information

- **Date of Submission:** [27/02/2025]
- **Date of Acceptance:** [28/02/2025]
- **JEL Classification Codes:** M15, L86, D63;

Abstract

This research explores the ethical dilemmas of implementing artificial intelligence (AI) in businesses and discusses the ethical principles required to address them. This chapter explores the role of ethical considerations in the responsible deployment of AI. It utilizes qualitative research methods, including a systematic literature review on AI and ethics and expert interviews. The literature analysis highlights several key ethical dilemmas associated with AI, including privacy violations, biased decision-making, responsible AI usage, intellectual property rights, and challenges in monitoring AI applications, among others, that arise from AI implementation and its effects on businesses. Although there are currently no guidelines for the ethical implementation of AI, the analysis suggests the potential to create distinct guidelines that can incorporate ethical principles into AI applications, assisting organizations in addressing the identified dilemmas. The dynamic nature of AI technology and socio-economic development necessitates longitudinal research. Future studies should examine these broader societal implications to grasp AI's influence on business and society. AI applications in business processes, product design, services, and governance raise ethical dilemmas. The authors offer unique insights by exploring the complex relationship between the ethical challenges of AI implementation and the imperative for business innovation and sustainability. Understanding and applying these insights may help organizations achieve their goals by providing more trustworthy and socially acceptable AI-driven products and improving AI-led business processes. This study offers valuable insights for academics, industry leaders, and policymakers seeking to promote responsible and sustainable AI deployment by adhering to ethical principles in its application.

Keywords: *AI ethical dilemmas, AI accountability, Ethics principles, AI ethics, Sustainable AI;*

12.1 Introduction

This chapter centers on the responsible and ethical implementation of AI, highlighting the challenges and primary limitations of applying AI ethical principles in business practices to address ethical dilemmas in AI usage. The ever-increasing application of AI in business and the growing trend of its integration into various aspects of our lives—from social platforms to smart home devices—profoundly affect our daily lives by shaping our choices, often in ways we may not fully comprehend. According to the OECD, AI research publications can proxy for a country's AI development. China accounts for 22% of global AI publications, while the EU stands at 14%, followed by the US at 11%. Between 2022 and 2023, the number of AI incidents reported in the media surged by approximately 1278%, coinciding with the mainstreaming of generative AI (OECD, 2024).

Ethical issues related to AI are evolving and require close attention from different stakeholders. Overcoming ethical dilemmas in AI development and implementation requires understanding and integrating ethical principles into practice. The chapter's discussion is based on the findings of the literature review and interviews with industry experts on solutions to overcome ethical dilemmas in AI applications. The focus is on overcoming ethical dilemmas caused by AI implementation by following AI ethical principles and integrating AI ethics considerations throughout the AI system lifecycle in decision-making, governance, and management.

12.1.1 Methodology

This chapter employs a qualitative research methodology grounded in existing literature and thematic analysis. The literature review encompasses scholarly articles, industry reports, and guidelines on AI ethics. The thematic analysis identifies key themes, categories, and concepts related to the implementation of AI ethics. The expert interviews highlight the challenges and best practices in applying AI ethics.

12.2 Literature Review

A growing body of literature on ethical dilemmas highlights various aspects of AI ethics. Several authors state that AI models can perpetuate or amplify biases present in the training data, leading to discriminatory outcomes (e.g., racial bias in facial recognition or hiring algorithms (Wei & Zhou, 2022), (Olatoye et al., 2024), (Martin & Galaviz, 2022)). AI's reliance on massive datasets risks infringing on individual privacy, enabling mass surveillance or data misuse (Huang et al., 2023), (Ivchyk & Shmatko, 2024). Other scholars note that AI systems lack transparency, making it difficult to understand how decisions are made (Ivchyk & Shmatko, 2024), (Akinrinola et al., 2024). Some authors suggest that clarification is needed regarding who holds responsibility when AI causes harm: developers, users, or the AI itself (Jobin et al., 2019), (Stahl, 2023). Over-reliance on AI undermines human decision-making skills and autonomy, another ethical dilemma that requires attention (Eflova et al., 2023), (Singh & Shah, 2024), (Bubicz & Ferasso, 2024). AI-driven automation threatens jobs and exacerbates economic disparities, raising ethical concerns (Olatoye et al., 2024), (Ikkatai et al., 2022), (Farzin & Samerie, 2023). Unpredictability presents another ethical concern in the use of AI. AI systems can be hacked, misused, or behave unpredictably (Wei & Zhou, 2022),

(Huang et al., 2023), (Ayling & Chapman, 2021); thus, the consequences of using AI are often unforeseen. This complexity leads to numerous issues that scholars, professionals, and society at large must address. Many authors question AI's contribution to sustainable development, as training large AI models consumes vast amounts of energy, thereby contributing to climate change (Bolte & Wynsberghe, 2024), (Huang et al., 2023), (Shkalenko & Nazarenko, 2024). Ethical dilemmas also arise from the challenge that users often unknowingly provide data to train AI systems (Gunn & Rosas, 2024), (Ikkatai et al., 2022), (Singh & Shah, 2024). Some authors express concern that while AI is intended for societal benefits, it can also be weaponized (for example, generative AI creating deepfakes) (Hagendorff, 2020), (Ivchyk & Shmatko, 2024), (Huang et al., 2023).

12.2.1 Summary of Research and Research Gaps

A significant body of research is dedicated to formulating ethical principles and guidelines for AI development and deployment. Research shows that implementing ethical use of AI depends on organizational factors such as leadership support, resource availability, and employee training (Ali et al.,2023). There is a growing awareness of ethical AI usage, as evidenced by the increasing number of developed AI ethical guidelines (Wei, 2022). Numerous organizations from different sectors, such as Google, IBM, Microsoft, DELL, Pfizer, and Lockheed Martin, among others, have published their policies for the ethical usage of AI, typically highlighting the ethical principles to be followed in decision-making in AI applications.

Although there is no universal understanding of how many ethical principles can be applied to address ethical dilemmas, researchers agree on the following five general ethical principles: responsibility and accountability, privacy, transparency, fairness, and non-maleficence (Jobin, 2019). However, in addition to these principles, new ones have recently emerged, such as societal and environmental well-being (Inter-Parliamentary Union, 2025), (Olatoyeet et al., 2024), sustainability (Bolte & Wynsberghe, 2024), democratic control and governance (human oversight) (Huang et al., 2023), (Singh & Shah, 2024), and freedom and autonomy (Jobin et al., 2019). There are variations in the interpretation and implementation of these ethical principles, indicating that research in ethical studies is still evolving. There is a lack of consensus on the interpretation and implementation of these ethical principles in decision-making. The new field of AI ethics has recently emerged and is developing rapidly. While there exists a substantial body of research on AI ethics, significant gaps remain in translating general ethical principles into practical actions to address ethical dilemmas regarding the responsible use of AI (Adeyelu et al., 2024), (Hagendorff, 2020), (J. Ali et al., 2023).

The World Economic Forum (WEF) suggests the following ethical AI principles divided into two groups: epistemic principles and general ethical AI principles. The epistemic principles constitute the prerequisites for investigating AI ethicality and represent conditions of knowledge that enable organizations to determine whether an AI system is consistent with ethical principles. They include such principles as interpretability (explainability, transparency and provability), reliability, robustness and security. Interpretability means that an AI system should be able to explain its model decision-making overall and what drives an individual prediction to different stakeholders. Reliability, robustness and security mean that AI systems should be developed to operate reliably over long periods using the right models and datasets. According to the WEF, the general ethical principles of AI are accountability, data privacy, lawfulness and compliance, beneficial AI, human agency, safety, and fairness (World Economic Forum, 2021).

These general ethical AI principles above are coherent with ethical principles such as

responsibility, accountability, privacy, transparency, fairness and non-maleficence. They are adapted to the AI application context.

Research in AI ethics emphasizes the "black box" nature, indicating that the workings of AI are not visible or comprehensible to users (Huang et al., 2023). This represents a crucial area of investigation that should concentrate on identifying and mitigating bias. While numerous AI ethics guidelines are available, there remains a gap in understanding their effectiveness and their impact on human decision-making. There is an absence of longitudinal studies on the long-term effects of AI on society and on how AI is altering our actions over time. A research gap persists in analyses of power structures and other structural influences that give rise to ethical issues in AI development and implementation. Consequently, clear guidelines for the application of ethical principles in AI are essential.

12.2.2 Application of Principles of Ethics in AI Context

Further explanation of ethical principles is needed to understand how to overcome ethical dilemmas caused by AI implementation by applying them.

Responsibility and accountability - The significance of accountability in AI systems is emphasized in various sources (Olatoye et al., 2024), (Huang et al., 2023), (Akinrinola et al., 2024). If an AI system fails, there ought to be a mechanism to hold someone responsible, be it the designer, developer, or company (Huang et al., 2023). Measures should be established to ensure responsibility and accountability for AI systems and their outcomes, prior to and after implementation (Huang et al., 2023). AI systems must be auditable, and those in charge of AI systems should be accountable for the system's behaviors and decisions and, consequently, responsible for any harm caused (Lane, 2023).

Social responsibility in AI extends beyond technical aspects, and companies should priorities ethical considerations alongside business objectives and long-term implications (Olatoye et al., 2024), (Ivchyk & Shmatko, 2024), (Adeyelu et al., 2024). Frameworks for accountability involve clarifying the roles and responsibilities of stakeholders, including developers, organizations, and regulatory bodies (Akinrinola et al., 2024). The question of who is responsible for AI systems is complex, with possible actors including developers, users, owners, regulators, or the system itself. It is crucial to clarify legal liability and the attribution of responsibility for AI (Sthal, 2023).

AI actors must ensure they are accountable for the design and implementation of AI systems in such a way that personal information is protected throughout the life cycle of the AI system. This directly links accountability with the necessity for privacy impact assessments, emphasizing the importance of safeguarding personal information when employing AI systems (United Nations Educational, Scientific and Cultural Organization (UNESCO), 2022). The attribution of decision-making in AI systems presents challenges when trying to assign responsibility for actions or consequences. Unlike human decision-makers, AI lacks consciousness and intentionality, complicating the allocation of accountability (Akinrinola et al., 2024).

Economic incentives can easily override a commitment to ethical principles and values, suggesting that the purposes for which AI systems are developed and applied may not align with societal values or fundamental rights (Hagendorff, 2020; Bolte & Wynsberghe, 2024). Consequently, integrating ethical principles into AI systems is essential for ensuring that service enhancement aligns with ethical norms (Huang et al., 2023). To implement ethical AI principles effectively, a concerted effort is required from all AI stakeholders, including

policymakers, technologists, ethicists, and end-users (Jobin et al., 2019).

Ethical principles and values are the starting point for trust in AI (Brendel et al., 2021) and form the basis for the regulation and certification of systems based on AI (Munoko et al., 2020). That emphasizes that the development of AI should serve society and not violate ethical norms (Groşanu et al., 2024).

Responsibility and accountability in AI mean having comprehensive mechanisms to determine who is responsible for AI actions. This extends beyond merely identifying culpability. It also involves ensuring ethical development and establishing mechanisms for oversight. That includes establishing clear lines of responsibility for various stakeholders involved in the AI lifecycle, which necessitates a holistic framework that promotes responsibility and the ability to address outcomes.

Privacy - As businesses create substantial data, responsible data handling and robust data privacy and security become critical ethical considerations for AI. Protecting individual privacy rights is a central tenet of ethical AI applications (Olatoye et al., 2024), (Bubicz & Ferasso, 2024). Privacy is a value to be cherished and a right that demands protection. This concept is frequently explored in data protection and security (Jobin et al., 2019). Data protection must be integrated into every stage of an AI system's development and operation. Addressing privacy concerns requires a strong ethical foundation and clear legal guidelines for responsible AI use (United Nations Educational, Scientific and Cultural Organization (UNESCO), 2022).

Technical solutions such as differential privacy, privacy by design, and data minimization are suggested for achieving privacy (Huang et al., 2023). Transparency and access to information must be balanced with the right to privacy (Lane, 2023). However, the drive to "unbiased" AI by using ever-larger, more diverse datasets may contradict the ethical principle of giving individuals greater control over their data and its use to safeguard their privacy and autonomy (Jobin et al.,2019).

The ability of AI to learn from and combine large datasets presents significant privacy risks (Stahl, 2023, p.2). AI systems can collect, analyze, and utilize personal data in unprecedented ways, creating an increased necessity to protect individuals from unwarranted surveillance and data exploitation (Olatoye et al., 2024). Adhering to data privacy standards and obtaining informed consent from users is crucial for ethical AI development and deployment (Adeyelu et al., 2024), (Gunn-Rosas, 2024), (Singh & Shah, 2024), (Huang et al., 2023).

Privacy concerns are paramount with AI-driven social media technologies, as these systems often collect and store vast user data. Collecting and storing this data without explicit consent can violate privacy and raise concerns about unauthorized access and surveillance (Oladele et al., 2024). Many AI systems function as "black boxes," making understanding their decisions and how personal data is used is challenging. This lack of transparency can hinder efforts to identify and mitigate bias, eroding user trust (Ivchyk & Shmatko, 2024, p. 65). Users often lack control over their data once collected, rendering them vulnerable to potential misuse and privacy breaches (Oladele et al., 2024). Privacy in the context of AI is a critical ethical consideration, encompassing the protection of individual rights; businesses must priorities privacy by implementing robust safeguards that respect user control over their data. The increasing prevalence of AI applications in various domains has highlighted the imperative of data privacy and security (Khan & Mer, 2023).

Transparency - It is impossible to build stakeholder trust without transparent AI algorithms and decision-making processes. To maintain public trust and operate ethically, businesses using AI must priorities transparency and ensure that their AI decision-making processes are comprehensible.

An evolved perspective on transparency, surpassing mere technical interpretability,

requires an understanding of AI decision-making processes and a rigorous evaluation of the justifiability and rationality that underpin the entire AI lifecycle, from design and implementation to the resulting outcomes (Huang et al., 2023). This aspect of transparency insists that the design and implementation choices of AI systems are reasonable, ethically sound, aligned with societal values, and justifiably transparent, enabling stakeholder scrutiny and appraisal of the system's objectives and underlying rationale (Huang et al., 2023), (Olatoye et al., 2024), (Lane, 2023).

Transparency encompasses the visibility of the AI system's processes, alongside the justification and rationale behind its design, implementation, and resultant outcomes (Martin & Galaviz, 2022; Jobin et al., 2019). The inherent opacity in the decision-making processes of many AI systems, often referred to as their "black box" nature, presents a significant ethical challenge that undermines accountability and reduces user confidence (Akinrinola et al., 2024; Ivchyk & Shmatko, 2024).

AI systems, particularly those utilizing machine learning, often suffer from inscrutability—meaning their internal workings are difficult to comprehend even with access to the model itself—and non-intuitiveness—meaning their decision-making relies on complex, non-obvious statistical relationships (Ferrario et al., 2020).

For responsible corporate AI practices, open communication regarding AI strategies and decision-making is essential, including clear explanations of the model's reasoning, the factors it takes into account, and the relative significance assigned to each (Akinrinola et al., 2024).

Open communication plays a dual role: addressing concerns about the "black box" nature of AI algorithms that obscure decision-making and enabling public understanding of AI system implementation (Lane, 2023), (Farzin & Samerie, 2023).

Explainable AI (XAI) fosters a better understanding of AI systems, whereas open data sharing improves transparency and reduces bias (Akinrinola et al., 2024).

Increased information disclosure and transparency are essential for fostering trust in AI, a principle extensively addressed in AI ethics that greatly impacts other ethical practices (Bubicz & Ferasso, 2024).

Fairness - In AI, fairness pertains to whether an AI algorithm treats different groups of people equally; justice, on the other hand, refers to ensuring that the system for allocating goods or resources does not harmfully benefit fortunate groups (Martin & Galaviz, 2022). Given the potential to perpetuate and amplify existing societal biases, which can lead to discriminatory outcomes, fairness in AI algorithms remains a crucial concern in AI ethics, necessitating careful attention to data, algorithm design, and ongoing monitoring (Adeyelu et al., 2024).

AI systems learn from data; if that data contains biases, the AI can perpetuate those biases (Ivchyk & Shmatko, 2024). This is a significant concern because AI systems can decide or categorise individuals based on inappropriate criteria (Stahl, 2023). Moreover, AI systems can reinforce stereotypes and make biased decisions, leading to unequal access and treatment (Farzin & Samerie, 2023). This highlights that AI systems must not treat individuals as mere data subjects, but rather as people with unique circumstances (Hagendorff, 2020). For AI systems operating in critical domains such as credit scoring, criminal justice, education, and hiring, transparency and fairness are essential preconditions for ethical operation. Their absence risks engendering discriminatory outcomes, exemplified by the restriction of financial resources for marginalised communities, the perpetuation of biased policing practices, the creation of unjust student assessments, or the obstruction of diverse workforce development (Akinrinola et al., 2024 ; Ivchyk & Shmatko, 2024).

AI systems can perpetuate and amplify societal biases and discrimination, exacerbating inequality and harm to specific groups (Bolte & Wynsberghe, 2024). AI, particularly machine

learning algorithms, learns from training data. If this data reflects existing biases, the AI will replicate and worsen those biases. This can occur in various contexts, such as mortgage approval programmes or facial recognition training predominantly on white men (Martin & Galaviz, 2022). AI biases often disproportionately impact underrepresented groups, intensifying existing inequalities (Bolte & Wynsberghe, 2024). As AI systems can be designed to subtly influence user behaviour, some sources express concern that this may undermine individual worth as human beings (Eflova et al., 2023). AI systems must be developed with inclusivity and equity in mind, actively preventing the perpetuation of biases present in data. It is crucial to maintain ethical vigilance, ensuring that the pursuit of efficiency in AI does not overshadow fundamental human values such as fairness, accountability, and privacy.

Non-maleficence - In AI ethics, non-maleficence is a fundamental principle that ensures AI systems do not cause or worsen harm, implemented through safety, security, and robustness measures (Huang et al., 2023). According to the research findings, AI safety necessitates prioritizing human safety, dignity, and both physical and mental integrity by preventing harm (including discrimination, privacy violations, and bodily harm) and ensuring secure operation (Huang et al., 2023; Jobin et al., 2019). Security involves protecting AI systems from vulnerabilities and attacks to safeguard humans, the environment, and ecosystems (Huang et al., 2023; United Nations Educational, Scientific and Cultural Organization (UNESCO), 2022). Huang et al. (2023) assert that robustness is a critical component of responsible AI, alongside accountability, liability, fairness, and explainability, and is essential for avoiding harmful outcomes (Huang et al., 2023).

Guided by the principle of non-maleficence, AI ethics mandates a systematic approach to risk management throughout the entire lifecycle of AI systems. This approach includes identifying and assessing potential harms (safety, security, robustness), implementing activities for mitigation and prevention, monitoring AI operations on an ongoing basis, and establishing redress mechanisms for any harm that arises (Jobin et al., 2019; United Nations Educational Scientific and Cultural Organization (UNESCO), 2022; Huang et al., 2023; Chettiar & Prof. Jish Joy, 2024).

Human oversight is the concept and ethical principle of ensuring that humans maintain control and responsibility over AI systems and their outcomes (Huang et al., 2023; United Nations Educational, Scientific and Cultural Organization (UNESCO), 2022). AI systems gain a crucial layer of control by incorporating human oversight through Human-in-the-Loop (HITL) approaches. This allows for the identification of potential biases and consideration of ethical implications, ultimately leading to more robust, reliable, and trustworthy AI decision-making (Akinrinola et al., 2024). By extending this principle, continuous monitoring of AI systems should also be deeply integrated with HITL. By actively involving stakeholders, including end-users and experts, in the oversight process, we can ensure that AI systems remain aligned with human values, ethical considerations, and desired outcomes (Anjum et al., 2023). Additionally, the European Commission has established guidelines for trustworthy AI, mandating human oversight by emphasizing human autonomy and decision-making while ensuring that humans retain ultimate control over AI and autonomous systems (Ivchyk & Shmatko, 2024). HITL may increase users' trust in AI systems because they know humans are involved. However, if human oversight is minimal or poorly implemented, it could provide a false sense of security. From the perspective of analyzing this term throughout research, HITL should be integrated into AI, particularly in all critical decision-making processes. This leads to humans assuming unlimited control and responsibility for AI decision-making. Humans should always be responsible for AI, as they are the designers and developers (Jobin et al., 2019).

Sustainability—The concept of Sustainable AI addresses two interconnected dimensions:

using AI to advance global sustainability goals and ensuring the sustainability of AI systems themselves. Sustainable AI requires a holistic approach that considers both the benefits and risks of AI and ensures that its development and deployment are environmentally responsible.

There is growing recognition that the environmental impact of AI is significant, including carbon emissions from training AI models and the effects of AI hardware on the environment (Bolte & Wynsberghe, 2024). A structural approach is necessary to identify ethical issues related to AI on a broader scale, often involving an examination of the power structures that hinder the uncovering of these issues (Bolte & Wynsberghe, 2024).

The increasing emphasis on long-term sustainability and social responsibility, moving beyond traditional finance, necessitates an interdisciplinary approach to sustainable IT. This approach draws upon knowledge from technology, ecology, social sciences, and management to address the complex challenges of our time (Shkalenko & Nazarenko, 2024). AI applications must be developed and used responsibly to reduce energy consumption and environmental impact (Jobin et al., 2019).

Companies are being asked to create policies to ensure accountability regarding potential job losses from AI and to use sustainability challenges as opportunities for innovation (Jobin et al., 2019).

12.3 Discussion and Findings

Five experts were interviewed to gather diverse perspectives on the ethical dilemmas arising from implementing AI and the approaches for overcoming them. The selection prioritized individuals with proven experience and deep knowledge in their respective fields. Crucially, the focus was on practitioners actively engaged with AI in real-world settings, particularly in leadership roles such as department heads, leading researchers, and association presidents, aimed at capturing strategic insights. This multi-faceted approach, drawing from various sectors, including telecommunications, technology, language technology, and the non-governmental sector, sought to provide a comprehensive understanding of AI. The interview results summarized the experts' views on addressing biases in AI systems, enhancing transparency, data privacy concerns, accountability, job threats, developing resilient AI, overreliance, and AI sustainability.

12.3.1 Biases in AI

The literature analysis's findings show that addressing biases in AI systems, particularly those originating in training data, is essential for building trustworthy and equitable AI systems. A significant concern is that AI models can learn and even worsen existing biases in the data they are trained on, which poses a substantial ethical dilemma in decision-making.

Interviewed experts agree that addressing these biases will ensure that AI systems are fairer and more equitable. Experts highlighted several key approaches to identifying and reducing bias, particularly those originating from training data, emphasizing that the data used to train AI is critical. A fundamental principle is thoughtful data selection. They recommend using datasets that accurately represent real-world diversity, including demographics, geographic regions, and social situations. Furthermore, experts advise regularly auditing datasets to find and address biases within the training data itself.

Beyond data, experts stress the need for ongoing monitoring and testing of AI models. This includes implementing routine audits to detect biases in the data and in the AI model's

outputs. They suggest using established methods and metrics available to the public for bias testing. Stress-testing models with unusual or extreme scenarios are also crucial to uncover discriminatory behaviors that might not be obvious in typical use cases. To ensure a comprehensive review, experts recommend forming diverse panels of reviewers to assess AI decisions before they are put into practice, allowing for correcting biases. Furthermore, long-term evaluations involving end-users in group audits and stress tests are vital for real-world bias detection.

Experts also acknowledged the practical challenges. They pointed out that identifying and mitigating bias is an ongoing process, and even auditing can be subjective, raising the question of "the bias of the auditor?" However, in cases where unacceptable levels of bias are found, experts recommend a process of iterative refinement. This involves expanding or improving the training data and carefully adjusting the AI model to specifically reduce the identified bias. It is also advisable to check for biases early in the development process, such as when building a "Proof of Concept," potentially through questionnaires.

12.3.2 Transparency in AI

The literature review reveals that transparency in AI raises significant questions, particularly ensuring that AI systems become more understandable. Specifically, the experts were asked whether applying transparency as an ethical principle in AI decision-making can help overcome this opacity. If so, how would you implement such transparency?

Experts generally agree on the importance of enhancing AI transparency, although they recognize that achieving complete transparency may be challenging due to the inherent complexity of advanced AI models and the competitive pressures within the AI industry. Despite these challenges, experts advocate for measures to improve transparency. They emphasize that responsible AI developers play a crucial role, which includes the obligation to provide technical details about the development of their models, such as information regarding training data and the methods used to enhance performance and address safety concerns. Experts also underscore the necessity of informing users when interacting with AI systems and, where possible, providing access to the rationales behind AI-driven decisions.

Furthermore, they recommend utilizing techniques that make AI decisions more interpretable, even for non-experts. Clear documentation detailing how AI models are trained and make decisions is also essential. Providing the source of information used by AI and educating users on how answers are generated are regarded as important steps towards building trust and understanding. Some experts even suggest demonstrating the AI's "thought process" in a manner easily understandable to humans, for instance, by showing how it searches for information and constructs its outputs. Linking model outputs to the relevant source material is further recommended to enhance clarity and enable users to verify the basis and validity of the AI's conclusions. Adopting industry-wide transparency standards, such as those outlined in the EU AI Act, is also considered a valuable step towards establishing a common framework for transparency. They pointed out that the highly competitive nature of the commercial AI sector might make model developers reluctant to publicly release all detailed information about their proprietary models. Experts note that even with complete knowledge of the training data, techniques, and model parameters, the reasoning processes within large, complex AI models may remain intrinsically difficult to explain intuitively and understandably.

Given these limitations on achieving transparency, experts suggest shifting the focus to the application side of AI. This involves rigorously checking for critical risks associated with specific AI applications. To enhance understanding, experts recommend utilizing Explainable

AI (XAI) tools, carefully tracking data to identify and mitigate biases, and thoroughly documenting the decision-making process.

12.3.3 Data Privacy

Based on the analysis of the literature, opaque data collection often raises significant concerns regarding data privacy, particularly given the immense data requirements of many AI models. Experts agree on a multi-faceted approach to ensure data privacy in AI systems. A fundamental principle is data minimization: experts consistently advocate for collecting only the data that is strictly necessary for AI training. Alongside this is the crucial step of anonymization and pseudonymization, highlighting the removal of directly identifying information from datasets. Techniques such as differential privacy, which involves adding controlled noise to datasets, are also recommended to further conceal individual identities and prevent re-identification.

Legal and technical safeguards are considered essential. Experts strongly recommend adhering to established privacy frameworks such as GDPR and CCPA, as well as implementing strong encryption techniques to protect data during both storage and transmission. In addition to these fundamental measures, transparency and user consent are emphasized as critical.

To address the risks associated with centralized data collection, experts propose exploring decentralized approaches, such as federated learning, which enables AI model training without requiring central data aggregation. Furthermore, implementing technical constraints that limit applications' ability to collect excessive personal data is considered an essential protective measure.

In addition to technical and legal solutions, experts also emphasize the necessity of governance and oversight. This involves conducting regular audits and special assessments, particularly for high-risk AI applications, and establishing accountability measures such as ethics committees. The integration of privacy into AI development – a principle known as "privacy by design" – is strongly advocated.

While broadly aligned, some nuances and challenges were acknowledged. One expert expressed uncertainty regarding the "best" answer, and it was noted that data privacy concerns extend beyond AI systems to include any system accessing user data, highlighting the pervasive nature of this challenge in the digital age.

12.3.4 Accountability in AI

The question of who is truly at fault in the application of the accountability principle in AI design and application—the developers who create AI, the users who deploy it, or even the AI itself—lacks a straightforward answer. This ambiguity stems from the intricate nature of AI systems and their development processes, making the assigning and distributing of responsibility a significant ethical dilemma.

Experts have diverse opinions on this matter. One perspective regards AI as a tool, contending that human actors must bear ultimate responsibility. From this standpoint, either the developer or the user could be held accountable, depending on the context. Developers may be deemed responsible if the AI system is fundamentally flawed, providing incorrect information or taking inappropriate actions due to design errors. Conversely, users might be held accountable if they misuse the AI, applying it in ways beyond its intended purpose or in contexts for which it was not designed.

However, other experts advocate for a shared responsibility model. This perspective suggests that accountability should be distributed among various stakeholders, including developers, businesses deploying AI, regulators, and users. Within this framework, developers are primarily accountable for designing AI systems that are demonstrably fair and safe. Businesses deploying AI bear the responsibility of ensuring that their AI applications adhere to ethical standards and are used responsibly. Regulators are vital for establishing and enforcing AI accountability laws and ensuring compliance. Even users have a role, as they are expected to utilize AI responsibly and within its defined scope of intended use.

Experts have proposed several mechanisms to operationalize this shared responsibility. They suggest that legal frameworks must evolve to define liability based on factors such as the AI system's level of autonomy and the risks associated with its application. Clear governance structures are needed to explicitly outline the roles and responsibilities of developers, operators, and end-users involved with AI systems. The implementation of AI liability laws is also recommended to ensure organizations are held accountable for the AI systems they deploy, drawing parallels to the existing legal accountability of employers for their employees. Some experts advise exploring ethical and legal models similar to employer-employee accountability.

Experts advocate for practical measures such as requiring human validation for high-stakes AI decisions to ensure human oversight in critical applications. Furthermore, maintaining detailed logs of AI decisions proves valuable for tracking errors, comprehending system behavior, and assigning responsibility when harm occurs. Establishing clear channels for users to contest AI-driven decisions is also essential for ensuring fairness and providing redress. In particularly high-risk areas where AI errors could result in significant harm, some experts even propose that regulators should consider restricting or prohibiting the use of AI altogether, given the current limitations of the technology. This risk-based approach is evident in emerging AI regulations like the EU AI Act. In regions where robust AI regulation is lacking, experts call for a global or regional dialogue to establish and agree upon fundamental ethical constraints for the development and deployment of AI.

Regardless of the specific model, experts agree that assigning responsibility for AI harm is not about blaming a single entity but rather about establishing clear lines of accountability.

12.3.5 Overreliance on AI

Another ethical dilemma is a prevalent concern: becoming overly reliant on AI systems may undermine our capacity to make effective decisions and act independently.

Experts largely agree that excessive reliance on AI significantly jeopardizes human cognitive abilities. A primary concern is the decline in critical thinking and problem-solving skills. The worry is that if AI systems manage cognitive tasks for us, individuals may become mentally passive, neglecting to engage their minds. This could lead to a situation where people struggle to perform even basic tasks without AI assistance, akin to losing the ability to navigate without solely depending on navigation apps. Experts emphasize that, throughout history, technological advancements have consistently transformed human skill sets, and AI represents the latest evolution in this ongoing process. Consequently, they argue that a broader societal discussion regarding which human skills are essential to retain and actively nurture in the age of AI is crucial.

Nevertheless, some experts offer a more nuanced and potentially optimistic perspective. They argue that if AI successfully takes over numerous routine and burdensome tasks, it could free up human time and cognitive resources, providing more choices and ultimately enhancing freedom and autonomy in certain respects. These experts emphasize that the impact of AI on

human decision-making is not fundamentally different from how technological progress has reshaped human skills throughout history. Even within this more positive framing, the core risk remains: the potential for individuals to stop engaging in critical thought and to neglect the development of essential analytical, conceptual, and critical thinking skills. To navigate this complex landscape, experts advocate for a shift towards human-AI collaboration rather than pursuing complete automation whenever possible. They suggest regularly training individuals in independent decision-making and implementing AI explainability tools. They also stress that when used thoughtfully, AI can enhance decision-making—such as by suggesting dietary options based on available food and preferences—without necessarily diminishing human decision-making capacity. One expert highlights a cyclical pattern in which ease and overconfidence can lead to a decline in human capabilities if not consciously managed.

12.3.6 AI Threat to Jobs

Another ethical dilemma posed by the application of AI is the fear that increased automation, driven by AI, threatens jobs, potentially widening existing economic gaps and creating new disparities. The pressing question is how to proactively ensure that the advantages of automation are distributed more equitably throughout society rather than concentrating wealth and opportunity in the hands of a few.

Experts largely concur on the necessity for broader benefit-sharing from automation and propose various interconnected approaches. A central theme is preparing the workforce for the evolving job market. This involves fundamentally reshaping the education system to emphasize future needs and the skills required in an AI-driven era, particularly training individuals for new AI-related roles. Experts also emphasize the importance of workforce transition programmes, potentially funded by the revenue generated from automation, to support those whose jobs are displaced. Experts advocate for a strategic approach to AI applications themselves. They suggest prioritizing AI applications that complement human resources and enhance human capabilities, rather than solely focusing on complete job replacement. The overarching goal should be to harness technological change for economic growth and societal benefit. This necessitates proactively mitigating the potential distress from job losses in specific sectors and addressing the challenges faced by groups struggling to find employment in the automated economy. Experts highlight the need to educate society about the beneficial applications of AI, countering purely negative narratives. This includes increasing public awareness of AI, providing comprehensive training programmes to empower individuals, and ensuring accessibility to AI tools, particularly for those unemployed and seeking new opportunities.

The emphasis should center on leveraging AI as an enabler of growth, shifting the perception from a threat to employment towards a driver of new possibilities. Underpinning these specific recommendations is a call for a holistic and balanced approach, highlighting the importance of aligning technological advancement with practical education and robust support systems. Experts also provide a historical perspective, reminding us that significant technological changes have invariably led to the elimination of some jobs while simultaneously creating new ones. Although such transitions can be disruptive and cause temporary hardship, they do not inherently result in a decline in overall job numbers or societal wealth. The key lies in proactively managing this transition.

12.3.7 Robust and Resilient AI Systems

It is a well-founded concern that AI systems are vulnerable to hacking, misuse, and unpredictable behavior, which creates significant risks. Therefore, the crucial question is: how can we develop AI systems that are more robust and resilient, thereby making them less susceptible to these diverse threats?

Experts emphasized that AI systems pose novel risks beyond traditional cybersecurity concerns. These AI-specific threats include data poisoning (manipulating training data), prompt injection (exploiting input vulnerabilities), model distillation (unauthorized model copying), sensitive data exposure, and excessive agency (uncontrolled AI actions). Therefore, a proactive approach is essential, with both model developers and consumers actively implementing appropriate security controls and continuously monitoring for potential breaches. Transparency regarding vulnerabilities is also crucial, enabling a better understanding and quicker remediation when issues arise.

To guide these efforts, experts recommend employing established security frameworks for AI published by organizations such as NIST, ISO, OWASP, MITRE, and Google. It is hoped that these frameworks will converge over time, fostering a shared understanding of best practices for secure AI development and deployment. Clear guidelines for AI deployment are also considered essential for maintaining consistent security.

In addition to preventative measures, experts emphasize the significance of resilience and active defense. This entails ensuring human intervention remains feasible in AI decision-making processes and providing a vital override mechanism. Exposing AI systems to potential attack scenarios through stress testing is recommended as a means to actively enhance their resilience. Securing AI models and datasets from cyber threats is essential.

12.3.8 AI and Sustainability

Large AI models require immense energy, raising concerns about their contribution to climate change. Therefore, experts were asked a critical question: How can we priorities energy efficiency and sustainability in AI development?

Experts largely agree on a multi-pronged approach to tackle this challenge. One primary approach focuses on model optimization and efficiency. Experts recommend using lightweight, less energy-intensive models and employing quantization techniques to lessen the computational demands of these models. They advocate for reducing computation-heavy training cycles and consistently working to enhance the efficiency of both model training and serving (the energy consumed when the model is actively in use). According to experts, a crucial shift in focus should involve moving away from the relentless pursuit of increasingly larger models towards developing more compact and efficient ones, directly reducing energy consumption. Furthermore, they suggest leveraging advancements in GPU and supercomputer development to optimize hardware for energy-efficient AI workloads.

Beyond these technical and infrastructure solutions, experts provide essential contextual considerations and future perspectives. They recognize that improvements in efficiency are presently somewhat counterbalanced by the rising demand for increasingly sophisticated and larger AI models. Nevertheless, they foresee that a performance ceiling is likely to be reached at some stage, which will naturally shift the focus towards efficiency. Looking further ahead, some experts highlight the potential of commercially viable fusion power as a long-term solution, suggesting that AI's energy demands and capabilities may even expedite fusion

research. They assert that the pressure for energy efficiency will inherently foster innovation in model compression and reduction.

Experts have noted that while the dominant business logic in the AI field is characterized by a "race to create larger models," they anticipate a shift towards greater emphasis on efficiency and cost-effectiveness. They also provide perspective by observing that the energy consumption of cryptocurrency mining currently exceeds that of training large language models, indicating that AI, despite being energy-intensive, is not the sole or most significant contributor to digital energy demand. Consequently, experts advocate for a dual approach: enhancing the efficiency of AI training while investing in clean energy infrastructure. They frame the challenge of AI's energy use as a catalyst for developing cleaner power sources and sustainable technologies, highlighting that the pressure to power AI is already accelerating investments in areas such as improved data center cooling and more efficient computing.

12.4 Conclusions and Further Directions of Research

These ethical principles concerning AI development and implementation are interconnected, making it essential to adopt a holistic perspective when developing AI systems and to ensure that ethical considerations are integral to the entire lifecycle process. Creating ethical AI involves more than just sophisticated technology; it requires contemplating the broader societal implications of AI applications and embedding ethical principles within the organizational culture and governance framework.

Merely enhancing AI technically will not resolve the ethical issues. Ethical principles must be integrated into AI from the outset, and all parties involved should collaborate by adhering to these principles. This also entails incorporating ethical considerations into AI from the initial design and development stages, which includes thorough ethical risk assessments.

As this research emphasizes, collaboration is crucial among technologists, ethicists, policymakers, and end-users to ensure that AI development is technologically advanced, ethically sound, and safe for use. Finding the right balance between business objectives and the well-being of a broad range of stakeholders is key to ensuring that AI benefits everyone.

As the context of AI ethics continually evolves, we must keep learning and adapting our approach to mitigate the risks associated with AI applications. We must remember that AI offers numerous advantages, aiding in the achievement of sustainable development goals.

Leadership in responsible design and application of AI and oversight from managers must be ensured. Establishing committees or boards to oversee ethical aspects of AI development and ensure adherence to ethical guidelines are suggested by researchers and practitioners as a necessary condition to overcome ethical dilemmas caused by AI application (Ivchyk & Shmatko, 2024, p.65), (Akinrinola et al., 2024, p. 53), (United Nations Educational, Scientific and Cultural Organization (UNESCO), 2022, p. 28).

Using structured frameworks that emphasize transparency, fairness, and accountability are foundational pillars of responsible AI development (Olatoye et al., 2024), (Huang et al., 2023), (Lane, 2023), (Adeyeluet al., 2024), (Ivchyk & Shmatko, 2024).

Conducting ethical impact assessments to identify potential risks and benefits associated with AI systems is essential for ensuring that ethical considerations are integrated throughout the AI lifecycle (Akinrinola et al., 2024), (United Nations Educational, Scientific and Cultural Organization (UNESCO), 2022).

Engaging with diverse stakeholders to understand societal values and concerns related to AI is essential to ensure that AI development and application aligns with societal good

(Adeyelu et al., 2024), (Ivchyk & Shmatko, 2024), (Akinrinola et al., 2024), (United Nations Educational, Scientific and Cultural Organization (UNESCO), 2022).

There was broad expert consensus that effectively securing AI demands a multilayered approach. This approach combines conventional cybersecurity techniques with an innovative approach tailored to AI's unique weaknesses. A key building block is leveraging existing and reliable cybersecurity practices.

Experts stressed the importance of enforcing governance through role-based access controls and ethics committees to ensure compliance with evolving standards, such as the EU AI Act. Ultimately, building resilient AI is an ongoing process of adaptation and improvement.

Further research directions could include, but are not limited to, best practices for integrating ethical AI principles into organizational culture and decision-making, how cultural differences impact the application of ethical AI principles, the development of Ethical AI frameworks, and the key success factors facilitating the application of ethical AI principles.

12.5 References

Akinrinola, O., Okoye, C. C., Ofodile, O. C., & Ugochukwu, C. E (2024). Navigating and reviewing ethical dilemmas in AI development: Strategies for transparency, fairness, and accountability. GSC Advanced Research and Reviews, 18(3), 050–058. https://doi.org/10.30574/gscarr.2024.18.3.0088

Adeyelu, O. O., Ugochukwu, C. E., & Shonibare, M. A (2024). ETHICAL IMPLICATIONS OF AI IN FINANCIAL DECISION – MAKING: A REVIEW WITH REAL WORLD APPLICATIONS. International Journal of Applied Research in Social Sciences, 6(4), 608–630. https://doi.org/10.51594/ijarss.v6i4.1033

Ayling, J., & Chapman, A (2021). Putting AI ethics to work: are the tools fit for purpose? AI and Ethics, 2(1), 405–429. https://doi.org/10.1007/s43681-021-00084-x

Ali, S. J., Christin, A., Smart, A., & Riitta Katila (2023). Walking the Walk of AI Ethics: Organizational Challenges and the Individualization of Risk among Ethics Entrepreneurs. ArXiv (Cornell University). https://doi.org/10.1145/3593013.3593990

Anjum, P. G., Soni, P. K., Prof. Khushboo Choubey, Priyanshi Warkad, & Saniya Choubey (2023). Incorporating AI into Business Education: Examining Ethical Issues with Case Study Illustrations. International Journal of Innovative Research in Computer and Communication Engineering, 11(06), 8884–8889. https://doi.org/10.15680/ijircce.2023.1106075

Bolte, L., & Aimee van Wynsberghe (2024). Sustainable AI and the third wave of AI ethics: a structural turn. AI and Ethics. https://doi.org/10.1007/s43681-024-00522-6

Bubicz, M., & Ferasso, M (2024). Advancing Corporate Social Responsibility in AI-driven Human Resources Management: A Maturity Model Approach. International Conference on AI Research, 4(1), 82–90. https://doi.org/10.34190/icair.4.1.3016

Chettiar, F. & Joy, J (2024). Navigating Ethical Dilemmas: The Role of AI in Supply Chain Decision-Making. International Journal for Research in Applied Science and Engineering Technology, 12(10), 460–467. https://doi.org/10.22214/ijraset.2024.64552

DELL (2022) Dell Technologies Principles for Ethical Artificial Intelligence Dell Technologies Principles for Ethical Artificial Intelligence. DELL. https://www.delltechnologies.com/asset/en-us/solutions/business-solutions/briefs-summaries/principles-for-ethical-ai.pdf

Eflova, M., Yulia Vinogradova, & Aleksandr Vitushkin (2023). The impact of artificial intelligence on the development of modern society. E3S Web of Conferences, 449, 07005–07005. https://doi.org/10.1051/e3sconf/202344907005

Etzioni, A., & Etzioni, O (2017). Incorporating Ethics into Artificial Intelligence. The Journal of Ethics, 21(4), 403–418. https://doi.org/10.1007/s10892-017-9252-2

Farzin, O., & Sameie, R (2023). Ethical and Legal Challenges of Using Artificial Intelligence in Digital Businesses. Journal of Technology in Entrepreneurship and Strategic Management, 2(2), 17-27.

Ferrario, A., Loi, M., & Viganò, E (2019). In AI We Trust Incrementally: a Multi-layer Model of Trust to Analyze Human-Artificial Intelligence Interactions. Philosophy & Technology, 33(3), 523–539. https://doi.org/10.1007/s13347-019-00378-3

Google (2023). Google AI Principles. Google AI. https://ai.google/responsibility/principles/

Groșanu, A., Melinda-Timea Fülöp, & Nicolae Măgdaș (2024). Ethical Dilemmas in Digital Accounting: A Comprehensive Literature Review. Contabilitatea, Expertiza Și Auditul Afacerilor, 5(4), 56–67. https://doi.org/10.37945/cbr.2024.04.06

Gunn-Rosas, C. L (2024). Beyond the Binary: AI, Ethics, and Liability in the Legal Landscape. Texas A&M Journal of Property Law, 10(3), 389–410. https://doi.org/10.37419/jpl.v10.i3.2

Hagendorff, T (2020). The ethics of AI ethics: An evaluation of guidelines. Minds and Machines, 30(1), 99–120. https://doi.org/10.1007/s11023-020-09517-8

He, H., & Chonlavit Sutunyarak (2024). Perception of Corporate Social Responsibility, Organizational Commitment and Employee Innovation Behavior: A Survey from Chinese AI Enterprises. Journal of Risk and Financial Management, 17(6), 237–237. https://doi.org/10.3390/jrfm17060237

Huang, C., Zhang, Z., Mao, B., & Yao, X (2022). An Overview of Artificial Intelligence Ethics. IEEE Transactions on Artificial Intelligence, 4(4), 1–21. https://doi.org/10.1109/tai.2022.3194503

IBM (2024). AI Ethics | IBM. Www.ibm.com. https://www.ibm.com/impact/ai-ethics

Ifeoluwa Oladele, Adeyinka Orelaja, & Oladayo Tosin Akinwande (2024). Ethical Implications and Governance of Artificial Intelligence in Business Decisions: A Deep Dive into the Ethical Challenges and Governance Issues Surrounding the Use of Artificial Intelligence in Making Critical Business Decisions. International Journal of Latest Technology in Engineering Management & Applied Science, XIII(II), 48–56. https://doi.org/10.51583/ijltemas.2024.130207

Inter-Parliamentary Union (2025), Ethic principles: Societal and environmental well-being. https://www.ipu.org/ai-guidelines/ethical-principles-societal-and-environmental-well-being

Ikkatai, Y., Hartwig, T., Takanashi, N., & Yokoyama, H. M (2022). Octagon Measurement: Public Attitudes toward AI Ethics. International Journal of Human–Computer Interaction, 38(17), 1–18. https://doi.org/10.1080/10447318.2021.2009669

Jobin, A., Ienca, M., & Vayena, E (2019). The Global Landscape of AI Ethics Guidelines. Nature Machine Intelligence, 1(9), 389–399. https://doi.org/10.1038/s42256-019-0088-2

Lane, L (2023). Artificial Intelligence and Human Rights: Corporate Responsibility in AI Governance Initiatives. Nordic Journal of Human Rights, 1–22. https://doi.org/10.1080/18918131.2022.2137288

Lockheed Martin (2023). Artificial Intelligence and Aegis: The Future is Here. Lockheed Martin. https://www.lockheedmartin.com/en-us/news/features/2023/artificial-intelligence-and-aegis-the-future-is-here.html

Martin, K., & Villegas-Galaviz, C (2022). AI and Corporate Responsibility. Springer EBooks, 1–5. https://doi.org/10.1007/978-3-319-23514-1_1297-1

Microsoft (2019). Responsible AI Principles and Approach | Microsoft AI. Microsoft.com. https://www.microsoft.com/en-us/ai/principles-and-approach/#tabs-pill-bar-ocb9d4_tab0

Morley, J., Elhalal, A., Garcia, F., Kinsey, L., Mökander, J., & Floridi, L (2021). Ethics as a Service: a Pragmatic Operationalisation of AI Ethics. Minds and Machines, 31. https://doi.org/10.1007/s11023-021-09563-w

Morley, J., Kinsey, L., Elhalal, A., Garcia, F., Ziosi, M., & Floridi, L (2021). Operationalising AI ethics: barriers, Enablers and next Steps. AI & Society, 38. https://doi.org/10.1007/s00146-021-01308-8

OECD. (2024). Artificial intelligence. OECD. https://www.oecd.org/en/topics/policy-issues/artificial-intelligence.html

Olatoye, F. O., Awonuga, K. F., Mhlongo, N. Z., Ibeh, V., Elufioye, O. A., & Ndubuisi, L (2024). AI and ethics in business: A comprehensive review of responsible AI practices and corporate responsibility. International Journal of Science and Research Archive, 11(1), 1433–1443. https://doi.org/10.30574/ijsra.2024.11.1.0235

Olatoye, F. O., Awonuga, K. F., Mhlongo, N. Z., Ibeh, V., Elufioye, O. A., & Ndubuisi, L (2024). AI and ethics in business: A comprehensive review of responsible AI practices and corporate responsibility. International Journal of Science and Research Archive, 11(1), 1433–1443. https://doi.org/10.30574/ijsra.2024.11.1.0235

Pfizer (2023). Artificial Intelligence (AI) Responsibility in Healthcare is Critical | Pfizer. Www.pfizer.com. https://www.pfizer.com/news/articles/three_principles_of_responsibility_for_artificial_intelligence_ai_in_healthcare

Shkalenko, A. V., & Nazarenko, A. V (2024). Integration of AI and IoT into Corporate Social Responsibility Strategies for Financial Risk Management and Sustainable Development. Risks, 12(6), 87–87. https://doi.org/10.3390/risks12060087

Singh, V., & Shah, P. B (2024). AI AND ETHICS. INTERANTIONAL JOURNAL of SCIENTIFIC RESEARCH in ENGINEERING and MANAGEMENT, 08(03), 1–5. https://doi.org/10.55041/ijsrem29381

Stahl, B. C (2023). Embedding responsibility in intelligent systems: from AI ethics to responsible AI ecosystems. Scientific Reports, 13(1), 7586. https://doi.org/10.1038/s41598-023-34622-w

UNESCO (2021, November 23). Recommendation on the Ethics of Artificial Intelligence. Unesco.org. https://unesdoc.unesco.org/ark:/48223/pf0000381137

Vasyl Ivchyk, & Nataliia Shmatko (2024). EXAMINING THE ETHICAL CONSIDERATIONS OF USING ARTIFICIAL INTELLIGENCE IN BUSINESS MANAGEMENT. Scientific Notes of Taurida National v I Vernadsky University Series Economy and Management, 74(3). https://doi.org/10.32782/2523-4803/74-3-10

Waelen, R (2022). Why AI Ethics Is a Critical Theory. Philosophy & Technology, 35(1). https://doi.org/10.1007/s13347-022-00507-5

Wei, M., & Zhou, Z (2022). AI Ethics Issues in Real World: Evidence from AI Incident Database. https://doi.org/10.48550/arxiv.2206.07635

World Economic Forum (2021). 9 ethical AI principles for organizations to follow. https://www.weforum.org/stories/2021/06/ethical-principles-for-ai/?gad_source=1&gclid=CjwKCAiAiOa9BhBqEiwABCdG89MYsdVjcBjdtsZw7ItWwKOWxmsosEnS5Ry3rTN3HwH6KAU51-3WzBoCvF4QAvD_BwE