Support the use of artificial intelligence tools and reduce its harm

Dhay Abdullateef Hammood Aljasim

Technical Institute Shatra, Southern Technical University, Basra, Iraq
Corresponding author, email: dhai.hammood@stu.edu.iq

doi: 10.17977/um063.v6.i4.2026.3

Keywords
Regulatory oversight
Education and awareness
Risk reduction
AI governance
Responsible AI use

Abstract
The goal of this study is to look into ways to support the good use of artificial intelligence (AI) tools while minimizing the potential harm they can cause to people and society. A sample of 153 people from various professional and public backgrounds participated in the study. A structured questionnaire was used to test a main hypothesis and three sub-hypotheses related to ethical guidelines, regulatory oversight, and educational awareness as ways to reduce AI-related risks using a quantitative methodology. Participants were unanimous in their belief that ethical frameworks, clear legal regulations, and increased education and training are necessary for the responsible integration of AI technologies, according to the findings. More specifically, more than 84% of respondents agreed that education and awareness play a crucial role in ensuring AI safety, while approximately 77% emphasized the significance of regulatory oversight. The main hypothesis, which is that AI harms can be significantly reduced by combining ethics, law, and education, is supported by these outcomes. The study recommends expanding educational programs, establishing enforceable regulatory frameworks, and adopting comprehensive ethical standards to improve public and professional understanding of AI. To promote public confidence in AI systems and strike a balance between innovation and safety, such an integrated approach is essential.

1. Introduction

One of the most significant technological advancements of the 21st century, artificial intelligence (AI) fundamentally reshapes how societies, businesses, and individuals interact with technology. AI is, at its core, the artificial simulation of human intelligence through learning, reasoning, and self-correction by machines. AI has become an essential component of modern innovation due to its capacity to process enormous amounts of data, recognize patterns, and make decisions at incredible speeds. Virtual assistants help us plan our days, recommendation algorithms curate our entertainment, autonomous systems are revolutionizing transportation, and predictive analytics are transforming sectors from healthcare to agriculture in both visible and invisible ways. Not only do these tools make operations run more smoothly, but they also make it possible to do things that were previously impossible, like translate languages in real time and use advanced image recognition to catch diseases early. There is a tremendous amount of potential for AI to increase productivity, stimulate economic growth, and address complex global issues. However, this era of progress also carries significant social, legal, and ethical repercussions that require careful consideration. The stakes increase and the need to strike a balance between the benefits and potential drawbacks becomes more pressing as AI systems acquire more autonomy and decision-making authority (Xu, 2021).

Despite the numerous and extensive benefits of AI, its potential to harm through misuse, unintended consequences, or inadequate oversight raises serious concerns that cannot be ignored. Algorithmic bias, in which AI systems trained on biased or incomplete datasets produce discriminatory results, is one of the most pressing issues. For instance, technologies for facial recognition have demonstrated lower accuracy rates for people of color, which has resulted in incorrect identifications and social injustice. Similarly, recruitment algorithms may unintentionally disadvantage certain groups by replicating historical gender or racial disparities. In addition, AI systems have the potential to function as "black boxes," making decisions without providing clear

explanations, which erodes trust and reduces accountability, particularly in sensitive fields like healthcare diagnostics, the criminal justice system, and financial services. The loss of privacy is another major concern because AI systems frequently rely on enormous amounts of personal data that are collected and processed without sufficient transparency or user consent. In addition, new challenges in the areas of misinformation, intellectual property, and online manipulation are brought about by the emergence of generative AI tools that can produce images, text, and videos that are exceedingly realistic. According to (Mensah, 2023), these negative effects are not just speculative; rather, they already exist and are likely to become more pronounced as AI becomes more advanced and widespread.

A comprehensive framework supporting the ethical and responsible use of AI is necessary to address these issues. This involves not only making AI system design better technically, like making it more fair, easy to understand, and strong, but also making governance mechanisms that make sure they are in line with what society values. It is necessary for community leaders, technologists, ethicists, policymakers, and others to collaborate in order to establish standards and regulatory guidelines that both promote innovation and protect the public interest. The concept of "human-centered AI," which places an emphasis on human agency, dignity, and oversight in AI decision-making processes, is essential to this endeavor. This necessitates the creation of systems that enable meaningful human control, particularly in high-risk applications, rather than replacing human capabilities. In this ecosystem, public education also plays a crucial role. As AI systems become more and more ingrained in everyday life, people must be equipped with the knowledge and ability to think critically to comprehend how these systems function, their limitations, and their implications. In addition, inclusive participation in AI development helps to mitigate narrow perspectives and ensures that AI solutions are equitable and culturally sensitive (Hanna, 2025). This is done by ensuring that diverse communities and voices are represented.

The purpose of this study is to investigate artificial intelligence's dual nature—its potential to cause harm if not properly managed and its potent tool for societal advancement. It aims to investigate the various AI applications that are currently in use, to identify patterns of risk and misuse, and to evaluate the efficacy of the mitigation strategies that are currently in place. The research will make a set of practical suggestions to encourage the safe and ethical use of AI through in-depth analysis of case studies, regulatory developments, and academic literature. Algorithm transparency, data governance, ethical auditing, and inclusive innovation will all be included in these recommendations. The study advocates for safeguards that safeguard human rights, democratic values, and social cohesion rather than adopting a pessimistic or alarmist view of AI. Instead, it emphasizes a balanced approach that acknowledges AI's transformative potential. We can ensure that this transformative technology serves the common good and contributes to a future where technological progress is aligned with the well-being of all by actively working to reduce its harms and supporting the responsible deployment of AI (Hariguna, 2024).

A worldwide change in the way tasks are carried out, decisions are made, and services are provided has been brought about by the growing integration of artificial intelligence (AI) systems into a variety of industries, including healthcare, education, security, finance, and daily life. While the adoption of AI promises enhanced efficiency, accuracy, and innovation, it simultaneously raises a host of ethical, social, economic, and legal concerns. The central problem lies in the imbalance between rapid technological advancement and the relative slowness of regulatory, ethical, and social systems to adapt. This has resulted in both overreliance on AI tools without adequate human oversight and underutilization due to fear, misunderstanding, or lack of clear policies. Moreover, unregulated AI can contribute to harmful consequences such as data bias, invasion of privacy, job displacement, and the reinforcement of social inequalities. Hence, there is a pressing need to support the responsible use of AI by maximizing its benefits while simultaneously identifying and mitigating its associated risks and harms through comprehensive policies, public education, and interdisciplinary collaboration.

The main problem addressed in this study concerns the challenges associated with supporting the responsible use of artificial intelligence while minimizing its potential harm to individuals and society. These challenges include the lack of adequate regulatory frameworks governing the ethical use of AI in sensitive sectors and the possibility that AI systems may reinforce bias and discrimination because of flawed, incomplete, or unrepresentative training data. The increasing use of AI-based

surveillance and large-scale data collection also raises serious concerns regarding privacy and data security. In addition, the growing adoption of automation has generated fears of job displacement, highlighting the urgent need for workforce reskilling and professional adaptation. Another significant issue is the limited understanding among the public and decision-makers regarding how AI systems operate, including their capabilities, limitations, and broader implications. Furthermore, the absence of clear and widely accepted international standards may hinder efforts to ensure that AI development remains aligned with human rights, ethical principles, and societal values (Mennella, 2024).

Based on these problems, this study seeks to determine how the use of artificial intelligence tools can be effectively supported while minimizing their potential harms to individuals and society. More specifically, the study examines the key benefits and opportunities offered by AI tools across various sectors and considers how their adoption can be encouraged in a responsible manner. It also investigates the most significant risks and ethical concerns associated with AI use, particularly those related to privacy, employment, bias, accountability, and automated decision-making. In addition, the study explores the strategies, policies, and governance frameworks that can be implemented to achieve a balanced approach that supports technological development and innovation while reducing the possible negative consequences of artificial intelligence.Significance of the Study:

This study is significant from a scientific standpoint because it tackles the use and governance of artificial intelligence (AI) tools, a fast developing field at the nexus of technology, ethics, and society. Scholarly research that not only demonstrates the potential of AI systems but also critically analyzes the risks and difficulties involved in their implementation is becoming more and more necessary as these systems are incorporated into vital industries like healthcare, education, finance, security, and communication. By providing an organized, analytical framework for comprehending AI's dual nature its potential to create value and advance society, as well as its potential to cause unforeseen harm if left unchecked this study advances the academic community. By highlighting important topics including algorithmic bias, data privacy, ethical responsibility, and the socioeconomic effects of automation, the study adds new perspectives and multidisciplinary relevance to the existing literature. Additionally, by relating the advancement of AI to philosophical issues of justice, openness, and the place of human agency in a digital environment, it fosters more in-depth scholarly discussion. As a result, the study is an essential point of reference for future research aiming to investigate AI in a fair, fact-based, and morally sound way (Zuiderwijk, 2021).

Practically speaking, governments, business executives, academic institutions, technological developers, and civil society organizations all find the study to be quite significant. It offers practical advice on how to encourage the use of AI tools in ways that are morally righteous, compatible with the law, and socially conscious. It is crucial to have policies in place that optimize AI's benefits while reducing any possible risks, as AI systems are increasingly affecting decisions that have an impact on people's lives, including hiring decisions, medical diagnosis, credit scoring, and law enforcement. This paper identifies practical remedies including policy creation, regulatory reform, and educational programs while highlighting real-world hazards like automation-driven job displacement, personal data exploitation, and the decline of human oversight in decision-making processes. Additionally, it highlights the significance of digital literacy and public awareness, promoting user empowerment to comprehend and challenge the technologies that influence their everyday lives. The report also urges the formation of interdisciplinary committees including members with backgrounds in technology, law, ethics, and human rights to supervise AI initiatives. The report acts as a guide for governments and groups looking to foster confidence in AI and make sure its advancement is in line with humanity's larger interests through these practical recommendations. In a world increasingly influenced by artificial intelligence, the research essentially closes the gap between theory and practice, adding to both academic knowledge and practical applications (Gupta, 2024).

This study aims to develop a comprehensive approach that supports the responsible use of artificial intelligence tools while minimizing their potential harms to individuals and society. Specifically, it seeks to identify the principal benefits and opportunities offered by artificial intelligence across various sectors and to explore appropriate ways to promote its responsible adoption. The study also aims to examine the primary ethical, social, and legal risks associated with the use of artificial intelligence technologies. In addition, it intends to formulate practical strategies,

governance frameworks, and policy recommendations that can support the safe, fair, transparent, and ethical deployment of artificial intelligence systems.

The study is based on the hypothesis that supporting the responsible use of artificial intelligence through integrated ethical, regulatory, and educational frameworks can significantly reduce its potential harms to individuals and society. It further assumes that the implementation of ethical guidelines is positively associated with the reduction of risks arising from artificial intelligence applications. Regulatory oversight and clear legal frameworks are also expected to minimize misuse, strengthen accountability, and reduce the unintended negative consequences of AI technologies. Furthermore, increasing public awareness and providing professional training on artificial intelligence are expected to enhance its safe, informed, and effective integration into various sectors.

Although many aspects of daily life have been altered by the widespread use of artificial intelligence (AI) tools, little is known about how these tools affect critical thinking. The study by (Gerlich, 2025) investigates the connection between the use of AI tools and critical thinking abilities with a focus on cognitive offloading as a mediator. Using a mixed-method approach, we interviewed 666 participants in-depth and administered questionnaires to participants of various ages and educational backgrounds. Thematic analysis of interview transcripts yielded qualitative insights, whereas ANOVA and correlation analysis were used to analyze quantitative data. Through increased cognitive offloading, the findings revealed a significant negative correlation between frequent use of AI technologies and critical thinking abilities. Younger participants had lower critical thinking scores and a greater reliance on AI technologies than older participants. In addition, there was a correlation between higher educational attainment and improved critical thinking abilities regardless of the use of AI. These findings emphasize the necessity of teaching methods that encourage critical use of AI tools and the potential cognitive costs of excessive reliance on them. By providing practical suggestions for lessening the technology's negative effects on critical thinking, this paper contributes to the expanding discussion of AI's cognitive effects. This study's findings are essential reading for developers, educators, and legislators because they demonstrate the significance of encouraging critical thinking in a world dominated by AI.

The current methods for detecting and predicting adverse drug events (ADEs), which are among the most common types of harm linked to healthcare, have a lot of room for improvement. To identify key use cases in which artificial intelligence (AI) could be used to reduce the number of ADEs, we carried out a scoping assessment. Natural language processing and cutting-edge machine learning methods were our primary subjects. The scoping review contains 78 articles. The study looked at a wide range of drugs and ADEs as well as a variety of AI methods. Through early detection to mitigate the effects and prediction to prevent ADEs, we discovered numerous significant use cases in which AI could assist in reducing the frequency and implications of ADEs. While most studies (73 [94%] of 78) assessed the effectiveness of technical algorithms, only few looked at the use of AI in therapeutic settings. The fact that 58 of the 78 papers—or 74% of the total—were published within the past five years highlights a rapidly expanding field of study. Access to unstructured clinical notes and new data types like genetic data could advance the field (Syrowatka, 2022).

Academic learning has been revolutionized by artificial intelligence (AI), which presents both problems and opportunities for students to progress. This study investigates how AI technologies impact students' learning processes and academic accomplishment, with an emphasis on students' views and the challenges of applying AI. 85 second-year students with first-hand knowledge of artificial intelligence-enhanced learning environments participated in this study, which was conducted at the National University of Science and Technology POLITEHNICA Bucharest. The subjects were chosen by deliberate sampling to ensure the study's applicability. Data were collected using a systematic questionnaire consisting of eleven items. Four open-ended questions concerning experiences, expectations, and concerns were included, along with seven closed-ended questions about attitudes, usage, and the efficacy of AI technology. While thematic analysis incorporated vertical (individual responses) and horizontal (cross-dataset) methodologies for quantitative data to find all themes, it used frequency and percentage estimates for qualitative responses. The study found that students were more engaged, performed better academically, and received more tailored training. But there are disadvantages as well, such as academic dishonesty, an over-reliance on AI, a loss of critical thinking skills, and worries about data privacy. The report highlights the necessity of a well-structured and morally sound framework in order to optimize AI's potential and minimize

hazards. In conclusion, even if artificial intelligence (AI) has a lot of promise to improve academic performance and learning effectiveness, its successful application will depend on resolving ethical concerns, accuracy, and cognitive disengagement. Equitable, effective, and responsible learning experiences in AI-enhanced learning environments require a well-rounded strategy (Vieriu, 2025).

Artificial intelligence (AI) is one useful technique that could raise the standard of care. The most frequent adverse events in the healthcare sector include falls, decompensation, venous thromboembolism, pressure ulcers, surgical complications, infections associated with healthcare, adverse pharmacological events, and diagnostic errors. This scoping review's goals were to gather pertinent research and look into the ways AI might improve patient safety in each of these eight risk areas. A thorough search turned up the pertinent MEDLINE papers. The scoping review identified studies that described the application of AI for prediction, prevention, or early detection in each harm category. Following the narrative presentation of the AI literature for each domain, occurrence, cost, and preventability were taken into consideration when making predictions about how AI can enhance safety. A total of 392 studies were included in the scoping review. Many examples of AI implementations using different approaches within each of the eight harm categories were given in the literature. The most prevalent unique data was gathered using a variety of sensing technologies, such as wearables, computer vision, pressure sensors, and vital sign monitoring. There are several chances to lower the incidence of injuries in all industries by utilizing AI and new data sources. In fields where existing approaches are inefficient and necessitate the intricate integration and analysis of new, unstructured data in order to generate precise estimates, we believe AI will have the biggest influence (Bates, 2021). Examples of these domains include adverse pharmacological events, decompensation, and diagnostic errors.

Tools based on artificial intelligence (AI) may assist hospital administrators in making better decisions. How hospital administrators currently make decisions and their perceptions of AI as a tool for supporting decisions are the focus of this study (Alves, 2024). It also examines the potential benefits of incorporating AI. In order to collect data, 15 hospital managers from various departments and organizations participated in semi-structured interviews using a qualitative descriptive and exploratory method. To identify significant themes and patterns in the responses, the interviews were transcribed, anonymized, and then thematically coded. The current inefficiencies in the decision-making processes, which frequently involve isolated decision-making, a lack of communication, and restricted data access, were brought to the attention of management at the hospital. Despite the fact that spreadsheet applications and business intelligence tools are still utilized frequently, it is evident that more sophisticated, integrated solutions are required. Despite their recognition of AI's potential to boost productivity and decision-making, managers expressed concerns regarding data privacy, moral quandaries, and the loss of human empathy. Disparities in technical skills, data fragmentation, and resistance to change were identified as the primary obstacles by the study. To guarantee a successful AI integration, managers emphasized the significance of adequate training and a robust data infrastructure. Conclusions The study revealed that, despite AI's potential advantages for hospital administration, there are also issues and concerns. With a focus on preserving human decision-making, technological, ethical, and cultural issues must be addressed in order for AI integration to be successful. Artificial intelligence is seen as a powerful tool that can complement rather than replace human judgment in hospital administration. It promises to enhance productivity, data accessibility, and analytical abilities. To get the most out of AI and reduce its risks, healthcare organizations need to have the infrastructure they need, and their management needs to receive specialized training.

The application and integration of artificial intelligence (AI) into practically every aspect of clinical treatment and public life is increasing. Care workers are increasingly impacted by AI, whether or not they accept it, because of the nature of the job. Even though AI integration and application may improve care quality and outcomes, AI use can be harmful. Like any other technology used in caregiving, AI's negative effects must be acknowledged, addressed, and, to the greatest extent possible, avoided. The use of AI tools in caregiving must be fixed whenever possible to prevent harm to caregivers and themselves. There is a social obligation on the part of nurses and other caretakers to support the recovery and well-being of everyone they serve, including themselves. Initiatives aimed at preventing, safeguarding, and recovering from the negative effects of big data and technology based on artificial intelligence are referred to as "digital defense." In this section, I propose a rethought conceptual framework for comprehending the links that exist between the

documented negative effects of AI and the general health of healthcare professionals. Following that, I provide a number of digital defensive strategies and group action plans that caregivers can use to protect one another and recover from the potential and actual negative effects of AI in healthcare (Walker, 2025).

Although the term "artificial intelligence" has been used for decades, its recent rise to prominence in public consciousness, transdisciplinary research, smart devices, and healthcare have presented nursing professionals with difficulties that necessitate new understanding. The moral implications of artificial intelligence for nursing care will be thoroughly examined in this research. When we use technology to provide nursing care, we bring an affirmative care ethic to this discussion, encouraging nurses to look beyond the obvious and truly investigate the worlds we create. We reject in this article portrayals of AI as either a threat or a blessing. Instead, we argue for a care and AI approach that is more relational and involved, recognizing that AI's potential advantages and disadvantages depend on how it is developed, applied, and oriented within healthcare systems. We address not only the most distant uses of AI/ML in the care environment by opposing reductive approaches to AI, but also issues and practices that are less obvious, such as upstream harms that are sometimes hidden from end users unless they delve into the specifics of how AI is developed. The discussion examines how redefining AI as a relational and collaborative tool could reduce damages both now and in the future, promote sustainable, patient-centered innovation, and address injustices (Dillard-Wright, 2025).

The field of artificial intelligence (AI) has seen significant advancements. The growing use of AI in healthcare has raised significant safety concerns regarding damage mitigation to prevent death, patient health information privacy, and ethical treatment that upholds patients' rights and dignity. Several safety frameworks have been agreed upon or suggested to address AI's safety. However, there has been no investigation into the extent to which the current AI tools for psychiatry adhere to these safety guidelines. This review (Jolayemi, 2025) looks at peer-reviewed test results for AI software's compliance with all or at least one AI safety framework. These test results may be important in determining how to use AI safely in psychiatry and improve clinical outcomes. A literature review was conducted in accordance with the MOOSE Meta-Analysis and Systematic Reviews for Observational Studies criteria regarding the safety and compliance of AI in psychiatry. A literature search was conducted using PubMed/MEDLINE, Google Scholar, Ovid, JoVE, and ScienceDirect with the search terms "artificial intelligence," "AI," "safety," "AI Safety," "regulations," "ethics," "regulatory compliance," "psychiatry," "healthcare," "medicine," and "harm mitigation." From February 2014 to April 2024, 72 papers were found through database scanning. After further exclusions, 24 final publications, including 51 Psychiatry AI software, were included. After two months of weekly reviews of the collected data, estimates of the frequency of peer-reviewed compliance with AI safety framework elements were made. The EU AI safety framework, the Code of Medical Ethics, the AI Bill of Rights, and the Presidential Executive Order on AI were not fully adhered to by any of the AI software (zero percent). In addition, none of the AI programs reported receiving training on safety or medical ethics during development. None of the AI programs for psychiatry (zero percent) met all of the AI safety standards when tested separately. Conclusion: Despite the growing use of AI in healthcare, safety is still a major issue that is not adequately addressed. The safety of AI software used in mental health is even more important because management involves crucial clinical judgments. This analysis revealed that none of the AI software complied completely with all AI safety guidelines. This is concerning because patients may lack the clinical judgment to recognize incorrect recommendations when using the majority of AI software. The stakes are higher for complete compliance when AI software handles life-or-death scenarios, including suicide interventions. To ensure that AI safety is prioritized, systematic procedures are required, and additional research is recommended to investigate how AI software might advance to complete compliance with safety frameworks.

2. Method

This study adopts a quantitative research methodology (Rana, 2021) to objectively analyze and measure the perceptions and attitudes of participants toward the use of artificial intelligence tools and the strategies for reducing their potential harms. The quantitative approach is suitable for this study as it enables the collection of measurable data that can be statistically analyzed to test hypotheses and identify patterns or relationships among variables. The primary data collection instrument used is a closed-ended questionnaire designed with a five-point Likert scale, ranging from

"Strongly Disagree" to "Strongly Agree." This format allows respondents to express the degree of their agreement or disagreement with a series of statements related to the benefits, risks, ethical concerns, and regulatory aspects of AI use. The questionnaire was carefully structured to ensure clarity, reliability, and alignment with the study's objectives and hypotheses. Data gathered from the respondents were analyzed using appropriate statistical tools to determine trends, correlations, and differences in opinions, thereby providing a solid empirical basis for drawing conclusions and making informed recommendations.

3. Results and Discussion

The aggregated responses across all three sub-hypotheses show a strong trend toward agreement with the main hypothesis. The majority of respondents either "Agree" or "Strongly Agree" that ethical, regulatory, and educational strategies are essential to reducing AI-related harm. This supports the main hypothesis and indicates that the proposed frameworks are considered effective by the majority of the sample.

The data strongly affirm all three sub-hypotheses, thereby validating the main hypothesis of the study. Across each domain ethical guidelines, legal regulation, and educational support respondents showed overwhelming agreement regarding their importance in reducing the potential harms of artificial intelligence. These results provide a clear mandate for policymakers, developers, and educators to collaborate in building comprehensive frameworks that both advance AI capabilities and protect against misuse. The relatively low percentage of disagreement and neutrality suggests a well-formed public perception that AI governance is not just desirable but essential. Moreover, the highest agreement was observed in the third sub-hypothesis, which emphasizes the role of training and awareness, indicating that empowering individuals with knowledge may be the most immediately impactful step toward safer AI integration. These findings can serve as a foundation for designing national or institutional strategies that balance innovation with responsibility in the field of artificial intelligence.

3.1. Sub-Hypothesis 1

There is a positive relationship between the implementation of ethical guidelines and the reduction of risks associated with artificial intelligence applications.

Table 1. Frequency Distribution of Responses for Sub-Hypothesis 1

Response	Frequency
Strongly Disagree	5
Disagree	10
Neutral	18
Agree	70
Strongly Agree	50

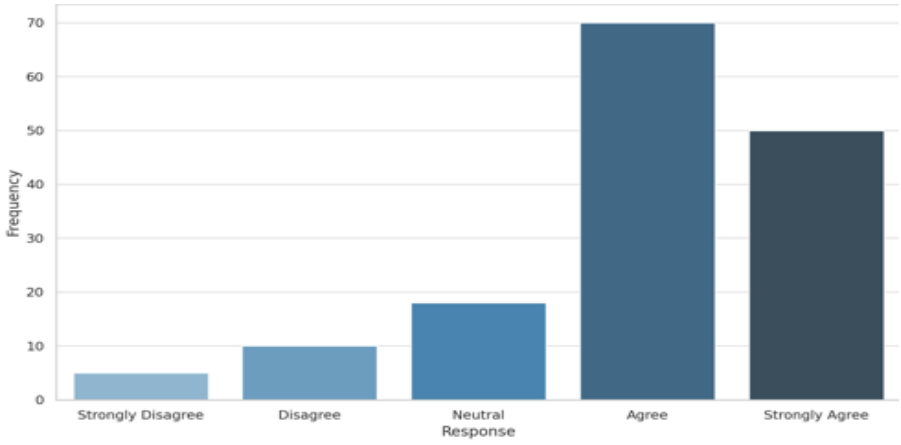

Figure 1. Hypothesis 1: Ethical Guidelines and Risk Reduction

3.2. Result Interpretation

A combined total of 120 out of 153 respondents (approximately 78.4%) either Agree or Strongly Agree that ethical guidelines play a crucial role in reducing AI-related risks. Only 15 respondents (9.8%) disagreed to any extent. This indicates a clear consensus that ethics-driven development and deployment are foundational to mitigating AI-related harms.

3.3. Sub-Hypothesis 2

Regulatory oversight and clear legal frameworks contribute to minimizing misuse and unintended negative consequences of AI technologies.

Table 2. Frequency Distribution of Responses for Sub-Hypothesis 2

Response	Frequency
Strongly Disagree	4
Disagree	9
Neutral	22
Agree	68
Strongly Agree	50

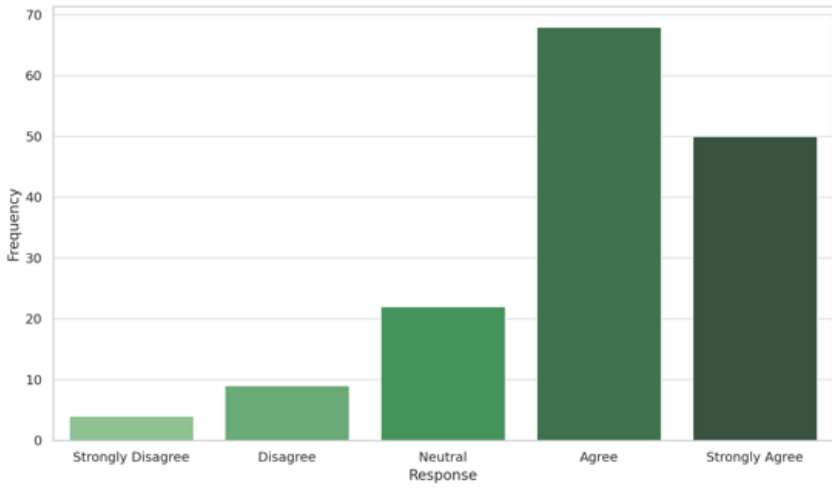

Figure 2. Hypothesis 2: Regulatory Frameworks and Minimizing Misuse

3.4. Result Interpretation

The responses indicate that 118 participants (approximately 77.1%) either Agree or Strongly Agree with the importance of regulation in AI safety. With only 13 responses (8.5%) expressing disagreement and 22 (14.4%) remaining neutral, the findings support the sub-hypothesis. This underlines the perceived necessity of formal, enforceable legal structures to ensure accountability and control in AI applications.

3.5. Sub-Hypothesis 3

Increasing public awareness and professional training about artificial intelligence enhances its safe and effective integration into various sectors.

Table 3. Frequency Distribution of Responses for Sub-Hypothesis 3

Response	Frequency
Strongly Disagree	3
Disagree	6
Neutral	15
Agree	60
Strongly Agree	69

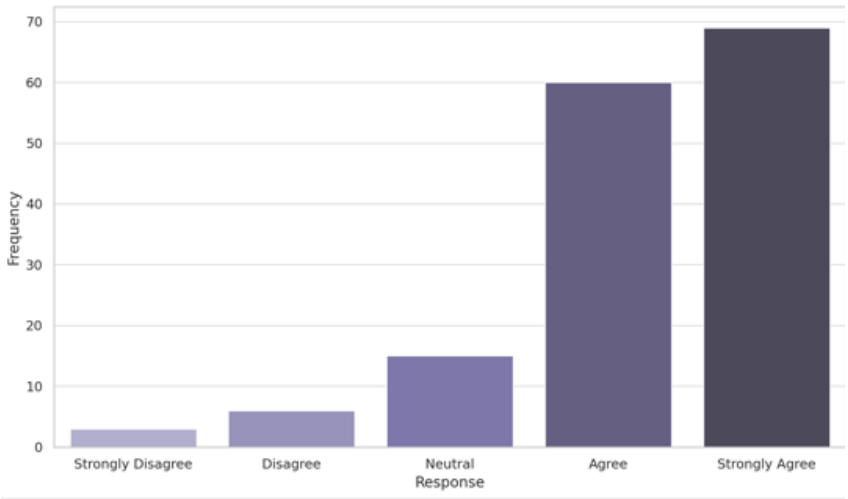

Figure 3. Hypothesis 3: Awareness, Training and Safe AI Integration

3.6. Result Interpretation

This hypothesis received the strongest support. A total of 129 respondents (84.3%) expressed agreement, with 69 choosing Strongly Agree. The low level of disagreement (only 5.9%) further confirms the vital role of education and capacity-building in ensuring that AI is implemented safely and with public trust.

The findings from this study offer significant insight into the critical role that ethical, regulatory, and educational frameworks play in mitigating the potential harms associated with artificial intelligence (AI) technologies. Based on the responses of 153 participants, there is a strong, consistent consensus supporting the main hypothesis: that the responsible and well-structured use of AI tools, guided by clear principles and practices, substantially reduces risks to individuals and society at large.

The first sub-hypothesis, which examined the relationship between ethical guidelines and risk reduction, received robust support from the participants. Approximately 78.4% of respondents affirmed that implementing ethical standards is fundamental to managing the risks posed by AI applications. This consensus reflects a growing recognition that AI development cannot proceed effectively without embedding ethical considerations throughout its lifecycle from design and deployment to monitoring and ongoing adjustment. Ethical frameworks help prevent misuse, protect privacy, ensure fairness, and uphold human dignity, all of which are crucial to fostering trust in AI systems. The relatively small minority expressing disagreement suggests that while ethical adherence is largely viewed as necessary, there may be differing opinions on the specific guidelines or how they should be enforced, indicating an area for further nuanced dialogue among stakeholders.

The second sub-hypothesis highlighted the importance of regulatory oversight and legal frameworks in curbing AI misuse and unintended consequences. With around 77.1% agreement among participants, the data indicates a widespread belief in the necessity of formal regulations to complement ethical efforts. Unlike voluntary ethical codes, regulatory structures possess the authority to enforce compliance and impose accountability, which are indispensable in complex AI ecosystems that often transcend borders and sectors. Participants' support reflects awareness that AI's transformative power brings both tremendous opportunity and significant risk, necessitating clear laws and policies to govern its applications. The presence of some neutral and dissenting voices may reflect concerns over regulatory overreach, stifling innovation, or the challenge of keeping laws abreast of rapid technological changes. Nonetheless, the results advocate for balanced yet firm governance frameworks that protect public interest without hindering progress.

The strongest agreement emerged for the third sub-hypothesis, which focused on the role of education, training, and awareness in the safe integration of AI. An overwhelming 84.3% of respondents agreed that increasing both public knowledge and professional expertise is essential to realizing AI's benefits responsibly. This underscores that beyond top-down controls, empowering

individuals whether they are developers, users, or policy-makers with comprehensive understanding is a powerful tool for harm reduction. Education builds capacity to identify risks, implement best practices, and foster a culture of ethical AI use. It also promotes transparency and trust, which are vital for social acceptance of AI innovations. The high level of consensus here suggests that educational initiatives may be the most direct and scalable intervention to achieve safer AI adoption in diverse sectors such as healthcare, finance, education, and public services.

The convergence of strong support across all three pillars ethics, regulation, and education forms a compelling narrative that responsible AI governance requires a holistic approach. The findings suggest that no single strategy is sufficient; rather, their integration creates a robust safety net. Ethical guidelines set the foundational values and norms; regulatory frameworks provide enforceable boundaries and accountability; and education ensures widespread capability and vigilance. Together, these components empower society to harness AI's potential while minimizing unintended harms such as bias, privacy invasion, misinformation, and autonomous decision errors.

The relatively low levels of disagreement and neutrality throughout the survey affirm a growing public and professional awareness of AI's dual-edged nature and the need for proactive governance. This consensus can inform policy formulation and strategic planning at national and organizational levels. Moreover, the prominence of education as a key factor signals that investments in AI literacy, professional development, and public engagement campaigns may yield substantial dividends in risk reduction.

For policymakers, these results reinforce the urgency to enact clear, adaptive legal frameworks that safeguard rights without impeding innovation. For developers and industry leaders, the findings call for embedding ethics systematically into AI design processes and actively participating in shaping standards and regulations. For educators and institutions, the study highlights the critical need to expand AI-related curricula and training programs across disciplines and sectors. Lastly, for the broader society, it signals that informed public discourse and participation are essential to shaping an AI-powered future aligned with shared values and interests.

This study validates the hypothesis that supporting the responsible use of AI tools through combined ethical, regulatory, and educational frameworks substantially reduces their potential harm. As AI technologies continue to evolve and permeate every facet of life, the lessons drawn from this research advocate for coordinated, multi-dimensional approaches that promote innovation while safeguarding human and societal well-being. Embracing these frameworks not only mitigates risks but also enhances the trust and acceptance necessary for AI to fulfill its transformative promise in a sustainable and equitable manner.

4. Conclusion

This study concludes that the responsible adoption of artificial intelligence requires an integrated framework combining ethical principles, legal regulation, education, stakeholder collaboration, and transparency. The findings from 153 respondents strongly support the main hypothesis, with 78.4% agreeing that ethical guidelines reduce AI-related risks, 77.1% affirming the importance of regulatory and legal oversight, and 84.3% emphasizing that education, professional training, and public awareness are essential for the safe and effective integration of AI. These results demonstrate that no single strategy is sufficient to address the complex risks associated with AI, including algorithmic bias, privacy violations, misinformation, discrimination, overreliance on automated systems, and unclear accountability. Therefore, policymakers, technology developers, educational institutions, private organizations, civil society, and affected communities should collaboratively establish adaptable ethical standards that uphold fairness, privacy, transparency, accountability, human rights, and meaningful human oversight throughout the AI lifecycle. These ethical commitments should be reinforced through enforceable laws, certification requirements, independent audits, clear liability provisions, and proportionate penalties for violations, while ensuring that regulation does not unnecessarily obstruct beneficial innovation. Given that education received the strongest support, substantial investment should also be directed toward AI literacy, ethics, safety, and responsible-use training in schools, universities, workplaces, and public awareness programs so that developers, professionals, decision-makers, and users understand AI's capabilities, limitations, and risks. In addition, explainable decision-making, accessible reporting of AI impacts, open evaluation processes, and continuous monitoring should be promoted to strengthen public

trust and enable external scrutiny. Through this coordinated and human-centered approach, society can maximize AI's contributions to productivity, healthcare, education, decision-making, and economic development while minimizing potential harm and ensuring that technological progress remains sustainable, equitable, socially acceptable, and aligned with the common good.

Data Availability

The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

Conflicts of Interest

Author in this publication declare no conflict of interest regarding the title, data, location, and results of the research.

Funding Statement

This research was conducted independently by the researcher without any financial support or funding from external institutions or organizations.

Acknowledgments

The author would like to thank all those who have helped in the preparation of this article.

Supplementary Materials

This study does not include any supplementary materials.

Declaration on AI Use

The author declare that no artificial intelligence (AI) or AI-assisted tools were used in the preparation of this manuscript. AI were used only to improve readability and language under strict human oversight; no content, ideas, analyses, or conclusions were generated by AI.

References

Alves, M. (2024). *Use of Artificial Intelligence tools in supporting decision-making in hospital management*. Retrieved from https://www.researchgate.net/publication/385252045_Use_of_Artificial_Intelligence_tools_in_supporting_decision-making_in_hospital_management

Bates, D. W. (2021). *The potential of artificial intelligence to improve patient safety: A scoping review*. Retrieved from https://pmc.ncbi.nlm.nih.gov/articles/PMC7979747/

Dillard-Wright, J. (2025). *An ethics of artificial intelligence for nursing*. Retrieved from https://www.researchgate.net/-publication/392383691_An_Ethics_of_Artificial_Intelligence_for_Nursing

Gerlich, M. (2025). *AI Tools in Society: Impacts on Cognitive Offloading and the Future of Critical Thinking*. Retrieved from https://www.mdpi.com/2075-4698/15/1/6

Gupta, A. (2024). *Ethical Considerations in the Deployment of AI*. Retrieved from https://www.researchgate.net/publication/-380518796_Ethical_Considerations_in_the_Deployment_of_AI

Hanna, M. G. (2025). *Ethical and Bias Considerations in Artificial Intelligence/Machine Learning*. Retrieved from https://www.sciencedirect.com/science/article/pii/S0893395224002667

Hariguna, T. (2024). *Assessing the impact of artificial intelligence on customer performance: a quantitative study using partial least squares methodology*. Retrieved from https://www.sciencedirect.com/science/article/pii/S2666764924000018

Jolayemi, A. (2025). *Safety Exploration of Artificial Intelligence in Psychiatry: A Systematic Review*. Retrieved from https://www.researchgate.net/publication/392399087_Safety_Exploration_of_Artificial_Intelligence_in_Psychiatry_A_Systematic_Review

Mennella, C. (2024). *Ethical and regulatory challenges of AI technologies in healthcare: A narrative review*. Retrieved from https://pmc.ncbi.nlm.nih.gov/articles/PMC10879008/

Mensah, G. B. (2023). *Artificial Intelligence and Ethics: A Comprehensive Review of Bias Mitigation, Transparency, and Accountability in AI Systems*. Retrieved from https://www.researchgate.net/publication/375744287_Artificial_Intelligence_and_Ethics_A_Comprehensive_Review_of_Bias_Mitigation_Transparency_and_Accountability_in_AI_Systems

Rana, J. (2021). *Quantitative methods.* Retrieved from https://www.researchgate.net/publication/-352356475_Quantitative_Methods

Syrowatka, A. (2022). *Key use cases for artificial intelligence to reduce the frequency of adverse drug events: A scoping review*. Retrieved from https://www.sciencedirect.com/science/article/pii/S2589750021002296

Vieriu, A. M. (2025). *The Impact of Artificial Intelligence (AI) on Students' Academic Development*. Retrieved from https://www.mdpi.com/2227-7102/15/3/343

Walker, R. (2025). *Digital Defense Toolkit: Protecting Ourselves from Artificial Intelligence-Related Harms.* Retrieved from https://www.researchgate.net/publication/392386126_Digital_Defense_Toolkit_Protecting_Ourselves_from_Artificial_Intelligence-Related_Harms

Xu, Y. (2021). *Artificial intelligence: A powerful paradigm for scientific research*. Retrieved from https://www.sciencedirect.com/science/article/pii/S2666675821001041

Zuiderwijk, A. (2021). *Implications of the use of artificial intelligence in public governance: A systematic literature review and a research agenda*. Retrieved from https://www.sciencedirect.com/science/article/pii/S0740624X21000137