Issues in Information Systems
Volume 25, Issue 2, pp. 256-266, 2024

DOI: https://doi.org/10.48009/2_iis_2024_121

Trusting AI in cybersecurity: an empirical study of perceptions on ethical decision-making

Shaila Rana, *Purdue Global University, shaila.rana@purdueglobal.edu*
Rhonda Chicone, *Purdue Global University, rhonda.chicone@gmail.com*

Abstract

In the dynamic field of cybersecurity, this paper examines perceptions of trust and reliability in AI systems designed for ethical decision-making, employing a structured online survey based on the Technology Acceptance Model (TAM) to assess professional and public attitudes. The study aims to reveal the extent of trust in AI's ability to navigate ethical dilemmas and its perceived reliability in consistently making ethical decisions within cybersecurity contexts. The findings contribute to the discourse on AI ethics, inform policymaking, and guide the development of AI systems aligned with societal ethical values.

Keywords: cybersecurity, artificial intelligence, AI, Technology Acceptance Model (TAM), trust, reliability, ethics

Introduction

The integration of Artificial Intelligence (AI) into cybersecurity represents a paradigm shift in how digital threats are identified, analyzed, and neutralized. AI's capabilities in processing vast amounts of data at unprecedented speeds enable the development of sophisticated security mechanisms that promise enhanced protection against increasingly complex cyber threats (Sarker, Furhad, & Nowrozy, 2021). However, the delegation of decision-making processes to AI systems, especially those involving ethical considerations, raises significant concerns regarding trust and reliability. This paper explores the multifaceted perspectives on the ethical implications of AI in cybersecurity, with a specific focus on the levels of trust and reliability attributed to AI systems in making ethical decisions.

The rapid evolution of cyber threats necessitates a dynamic and adaptive security posture, where AI systems can offer significant advantages. These systems can potentially operate beyond the constraints of human bias and fatigue, offering continuous, efficient, and potentially more objective decision-making capabilities. Yet, the ethical dimension of AI decision-making—encompassing issues such as privacy, data integrity, and bias—emerges as a critical area of concern (Sanclemente, 2022). The central question revolves around the extent to which AI systems can be trusted to make ethical decisions and the reliability of these systems in consistently adhering to ethical standards in the dynamic and complex landscape of cybersecurity.

Drawing upon the Technology Acceptance Model (TAM), this study employs a quantitative research methodology to assess professional and public perceptions regarding AI's ethical implications in cybersecurity. This approach allows for the systematic exploration of the underlying factors that influence trust and reliability perceptions related to AI's ethical decision-making capacities in cybersecurity contexts. In general, this research seeks to contribute to the ongoing discourse on the ethical integration of AI in cybersecurity practices.

Background

The integration of Artificial Intelligence (AI) into cybersecurity represents a significant shift in how cyber threats are detected, analyzed, and mitigated. As cyber threats grow in complexity and sophistication, AI offers unprecedented capabilities in identifying and responding to such threats in real-time (Sontan & Samuel, 2024). AI has a pivotal role in enhancing threat intelligence, anomaly detection, and response strategies (Taddeo et al., 2019). Thus, underscoring its potential to fortify cybersecurity defenses significantly. However, this integration is not without challenges; the ethical dimensions of AI decision-making in cybersecurity, spanning privacy, data integrity, and the potential for bias emerge as critical areas of concern.

The concepts of trust and reliability are pivotal when discussing AI's role in cybersecurity. Trust in AI systems is multifaceted, encompassing the system's capability, reliability, and ethical decision-making capacity. Lee and See (2004) define trust in automation as the belief that an entity will help achieve an individual's goals in a situation characterized by uncertainty and vulnerability. Moreover, part of AI's appeal is its ability to make the cybersecurity computing process more automated and intelligence than traditional or conventional security systems (Sarker, Furhad, & Nowrozy, 2021). In general, AI is cited to have unlimited use cases (Liu & Murphy, 2020).

In the context of AI in cybersecurity, this trust extends to the system's ability to make decisions that are not only effective but also ethically sound. Reliability, on the other hand, relates to the consistency and dependability of AI systems in performing their designated tasks. In cybersecurity, this means the system's ability to detect and respond to threats accurately and efficiently. The reliability of AI systems is intricately linked to their ethical programming and the biases inherent in their data and algorithms, which can influence their decision-making processes (Sanclemente, 2022).

The ethical decision-making capacity of AI systems, especially in sensitive areas like cybersecurity, is a burgeoning field of study. Ethics in AI pertains to the system's ability to make decisions based on a set of moral principles or societal norms. Martin (2019) and Hagendorff (2020) emphasize the importance of embedding ethical considerations into AI systems from the ground up, advocating for ethical frameworks that guide AI decision-making processes. This includes considerations around privacy, data protection, and the minimization of bias. Moreover, it is noted that regulatory frameworks and standards play and will play a significant role in ensuring adherence with ethical principles, data privacy regulations, and accountability mechanisms (Konda, 2019).

For example, the EU AI Act has clear requirements and obligations regarding use cases of AI that adopts a risk-based approach to combat the growing ethical concerns of AI and its impact on various fields, including cybersecurity. The trust and reliability of AI systems in cybersecurity are deeply tied to these ethical considerations; systems perceived as ethically sound are more likely to be trusted and seen as reliable by users.

In general, studies suggest a complex interplay between trust, reliability, and ethics in the context of AI in cybersecurity. Trust in AI systems is not merely a product of their technical capabilities but also their ethical decision-making capacity. Similarly, a system's reliability is seen not just in terms of operational consistency but also in its adherence to ethical standards. This relationship underscores the need for ongoing research to better understand how these dimensions influence the acceptance and effectiveness of AI in cybersecurity.

Contribution

This paper aims to contribute to a deeper understanding of the levels of trust and reliability in attributing AI systems to make ethical decisions in cybersecurity contexts. Given that this is a multifaceted approach, understanding AI systems' ability to make ethical decisions is important to explore. Thus, by focusing on trust and reliability, it can provide empirical evidence and insights that can help refine theoretical frameworks and model related to ethical AI use. Moreover, the findings can offer valuable insights for policymakers and regulatory bodies. Understanding the levels of trust and perceptions of reliability in AI's ethical decision-making capabilities can guide the development of regulations and standards that ensure AI systems in cybersecurity are developed and used in an ethically responsible manner. This could include recommendations for transparency, accountability, and oversight mechanisms.

In regard to AI developers and cybersecurity professionals, the study highlights the importance of ethical considerations in the design, development, and deployment phases of AI systems. Insights into trust and reliability can inform strategies to enhance these aspects, potentially leading to the adoption of ethical guidelines, best practices, and design principles that prioritize ethical considerations. Finally, this study may reveal gaps in knowledge or misconceptions about AI ethics among professionals and the public, highlighting the need for educational initiatives and training programs. This can lead to the development of targeted educational materials and courses that aim to increase awareness and understanding of ethical AI use in cybersecurity.

Methodology

This study utilized a structured online survey incorporating Likert scale questions based on the Technology Acceptance Model (TAM) to quantitatively assess perceptions of AI ethics in cybersecurity. The survey consisted of 20 questions covering various aspects of ethical implications, trustworthiness, and reliability of AI in cybersecurity contexts and can be found in Appendix A.

One hundred and twenty-six participants, including professionals from cybersecurity, AI, IT, and academia, as well as interested individuals from the general public, were surveyed. One hundred and nineteen people participated in the study, resulting in a response rate of 94%. Participants rated their agreement on a five-point scale, ranging from "Strongly Disagree" (1) to "Strongly Agree" (5), to capture nuanced attitudes towards AI ethics.

Respondents were free to skip questions in the survey after the first question of whether they wanted to participate in the study. Therefore, the results varied depending on who participated in answering the questions. This study employed a structured online survey distributed through SurveyMonkey's paid response service to assess perceptions of AI ethics in cybersecurity. The respondent population was determined using a convenience sampling strategy through targeted distribution via SurveyMonkey's audience panel, ensuring representation across relevant demographics. The criteria for inclusion in the study were willingness to participate and the ability to complete the online survey through SurveyMonkey's platform.

The questions in the survey asked participants whether they believe AI can significantly improve cybersecurity, whether they are concerned about the misuse of AI in cybersecurity, if AI in cybersecurity will lead to more advanced threats, if the use of AI in cybersecurity raises significant ethical concerns, if they are worried about the potential for AI to invade personal privacy, whether the benefits of AI in cybersecurity outweigh its ethical risks, if they trust AI to make ethical decisions in cybersecurity contests,

whether AI systems are reliable enough to handle complex ethical dilemmas in cybersecurity, if human oversight is necessary to ensure the ethical use of AI in cybersecurity, if there is a stricter need for regulations governing the use of AI in cybersecurity, if participants are confident that current laws adequately address the ethical use of AI in cybersecurity, and whether organizations using AI in cybersecurity should be transparent about their ethical guidelines. The analysis included t-tests to compare mean scores across different groups, specifically between cybersecurity/IT professionals and others, for both trust and reliability ratings. Additionally, correlation analysis was conducted to explore factors influencing perceptions of trust and reliability towards AI in cybersecurity.

Research Questions

This study aims to answer the primary research question regarding to what extend do professionals and the public trust AI systems to make ethical decisions in cybersecurity contexts, and how reliable do they perceive these systems to be. Thus, to answer this primary question, we investigate the following specific research questions:

> **RQ1*:*** *Is there a difference in trust towards AI's ethical decision-making in cybersecurity between IT/cybersecurity professionals and non-professionals?*
>
> **RQ2:** *Is there a difference in perceptions of AI's reliability in handling complex ethical dilemmas in cybersecurity between IT/cybersecurity professionals and non-professionals?*
>
> **RQ3**: *Is there a relationship between trust in AI's ethical decision-making and perceptions of its reliability in cybersecurity contexts?*

Based on the above research questions, we formulate the following hypotheses:

> **$H1_o$:** *There is no significant difference in trust towards AI's ethical decision-making in cybersecurity between IT/cybersecurity professionals and non-professionals.*
>
> **$H1_a$:** *There is a significant difference in trust towards AI's ethical decision-making in cybersecurity between IT/cybersecurity professionals and non-professionals.*
>
> **$H2_o$:** *There is no significant difference in perceptions of AI's reliability in handling complex ethical dilemmas in cybersecurity between IT/cybersecurity professionals and non-professionals.*
>
> **$H2_a$:** *There is a significant difference in perceptions of AI's reliability in handling complex ethical dilemmas in cybersecurity between IT/cybersecurity professionals and non-professionals*.
>
> **$H3_o$:** *There is no significant relationship between trust in AI's ethical decision-making and perceptions of its reliability in cybersecurity contexts*.
>
> **$H3_a$:** *There is a significant relationship between trust in AI's ethical decision-making and perceptions of its reliability in cybersecurity contexts*.

These hypotheses will be tested using independent samples t-tests for H1 and H2, and correlation analysis for H3. The results will provide insights into how professional background influences trust and perceived reliability of AI in cybersecurity ethical decision-making, as well as the relationship between these two factors.

Results

Table 1 presents the descriptive statistics from all the questions in the survey which again can be found in Appendix A.

Table 1: Descriptive Statistics for All Survey Questions

	Mean	Median	Mode	SD
I believe AI technology can significantly improve cybersecurity.	3.43	3	4	1.04
I am concerned about the misuse of AI in cybersecurity.	3.93	4	4	0.974
AI in cybersecurity will lead to more advanced forms of cyber threats.	3.72	4	4	0.930
The use of AI in cybersecurity raises significant ethical concerns for me.	3.58	4	4	1.05
I am worried about the potential for AI in cybersecurity to invade personal privacy.	3.89	4	4	0.86
The benefits of AI in cybersecurity outweigh its ethical risks.	3.15	3	3	1.15
I trust AI to make ethical decisions in cybersecurity contexts.	2.86	3	3	1.24
AI systems are reliable enough to handle complex ethical dilemmas in cybersecurity.	2.85	3	3	1.16
Human oversight is necessary to ensure the ethical use of AI in cybersecurity.	3.90	4	5	1.11
There is a need for stricter regulations governing the use of AI in cybersecurity.	3.91	4	4	0.976
I am confident that current laws adequately address the ethical use of AI in cybersecurity.	2.79	3	3	1.21
Organizations using AI in cybersecurity should be transparent about their ethical guidelines.	3.95	4	5	1.09

Table 2 presents the respondents' perceptions of trust, indicating moderate trust in AI for ethical decision-making in cybersecurity contexts (Mean = 2.86, SD = 1.24). Respondents from the cybersecurity or IT field exhibited slightly higher perceptions of AI's capabilities in both trust and reliability compared to the overall sample. Table 2 also shows that respondents in the IT/Cyber field had a higher mean score (3.47) compared to the overall sample (2.86). The median for IT/Cyber professionals was also higher (4) than for the overall sample (3). This suggests a more positive perception of AI's trustworthiness in ethical decision-making within cybersecurity contexts among professionals in these fields.

Table 2: Descriptive Statistics: Perceptions of trust

Respondent Type	Mean	Median	Mode	Range	Standard deviation
All	2.86	3	3	4	1.24
In IT/Cyber Field	3.47	4	3	4	1.22

Table 3 shows the respondents' perceptions of reliability, suggesting a moderate perception of AI reliability in handling complex ethical dilemmas in cybersecurity (Mean = 2.85, SD = 1.16). Table 3 also shows that IT/Cyber professionals had a higher mean score (3.41) compared to the overall sample (2.85). However, the median values were the same (3) for both groups in this case. Consequently, these trends may reflect

the deeper understanding that IT and cybersecurity professionals have of AI technologies and their potential applications in cybersecurity. Their slightly more positive perceptions regarding AI's capabilities in both ethical decision-making and navigating complex dilemmas within cybersecurity contexts are explored further in the Discussion section.

Table 3: Descriptive Statistics: Perceptions of reliability

Respondent Type	Mean	Median	Mode	Range	Standard deviation
All	2.85	3	3	5	1.16
In IT/Cyber Field	3.41	3	3	4	1.13

Conversely, for all respondents, the means (2.86 and 2.85), medians (both 3), and modes (3) indicated a generally neutral stance on trusting AI and its reliability in ethical decision-making within cybersecurity contexts. This uniformity suggests that, on average, respondents across different backgrounds share similar perspectives on these aspects of AI.

The correlation analysis focused on two key variables extracted from our survey: trust in AI's ethical decision-making and perceptions of AI's reliability in handling ethical dilemmas. These variables were derived from specific questions in our 20-item survey. The 'trust' variable was based on responses to the question 'Do you trust AI to make ethical decisions in cybersecurity contexts?', while the 'reliability' variable was based on responses to 'Are AI systems reliable enough to handle complex ethical dilemmas in cybersecurity?'. We calculated the Pearson correlation coefficient between these two variables for different respondent groups. In summary, while overall opinions leaned towards neutrality, professionals in cybersecurity or IT expressed a slightly more favorable view of AI's capabilities in ethical decision-making compared to the broader sample, which may be attributed to their specific expertise and familiarity with AI technologies.

These coefficients indicate varying degrees of positive linear relationships between trust in AI's ethical decision-making and perceptions of AI's reliability in handling ethical dilemmas. Specifically, the correlation coefficient of 0.0638, shown in Table 4, for all respondents suggests a very weak positive relationship between trust in AI's ethical decision-making and the belief in AI systems' reliability in handling complex ethical dilemmas. Among respondents in the IT/Cyber Field, the correlation coefficient of 0.0671 indicates a slightly stronger positive relationship compared to all respondents. This suggests that IT professionals who trust AI to make ethical decisions are somewhat more likely to perceive AI systems as reliable in handling complex ethical dilemmas.

Table 4: Statistical Analysis: Correlation

Respondent Type	Pearson correlation coefficient
All	0.0638
In IT/Cyber Field	0.0671
Not in IT/Cyber Field	0.606

Conversely, respondents not in the IT/Cyber Field demonstrated a higher correlation coefficient of 0.606, indicating a moderately strong positive relationship. This correlation, while slightly weaker than that observed among all respondents and those in the IT/Cyber Field, still indicates a significant association between trust in AI's ethical decision-making and perceptions of its reliability in ethical dilemma

management. These findings highlight the nuanced perspectives across different respondent groups regarding AI's role in ethical decision-making within cybersecurity contexts.

All three groups show a moderately strong positive linear relationship between the two variables, with correlation coefficients ranging from 0.606 to 0.671. The correlation is strongest among respondents in the IT field (0.671), followed by all respondents (0.638), and weakest among those not in the IT field (0.606). The stronger correlation among IT professionals suggests that their trust in AI's ethical decision-making capabilities is more closely tied to their belief in AI systems' reliability in handling complex ethical dilemmas, compared to non-IT professionals. In summary, the correlation analysis reveals a moderately strong positive linear relationship between trust in AI's ethical decision-making and belief in AI systems' reliability in handling complex ethical dilemmas across all groups. The relationship is slightly stronger among IT professionals compared to non-IT professionals and all respondents combined.

Table 5 shows the statistical results of the t-tests that were performed. When comparing the "Trust AI" ratings between those in the Cybersecurity/IT field and those not in it, the t-test statistic is approximately -0.009 with a p-value of 0.993, indicating no statistically significant difference between the two groups. Similarly, for the "Reliable AI" ratings, the t-test statistic is 0.851 with a p-value of 0.399, also indicating no statistically significant difference between the two groups. The t-test results (t-stat=0.057, p-value=0.955) indicate no statistically significant difference between the two variables for all respondents at a 95% confidence level. Subsequently, the t-tests confirm that there is no statistically significant difference between the means of the two variables (trust in AI's ethical decision-making and belief in AI systems' reliability in handling complex ethical dilemmas) for all respondents, those in the IT field, and those not in the IT field. This suggests that the respondents' opinions on these two aspects of AI in cybersecurity are consistent across the groups. Thus, this addresses RQ1 and RQ2, and the null hypothesis cannot be rejected.

Table 5: T-tests

Comparison	Variable	t-stat	p-value	Sample Size (IT/Cyber Field)	Sample Size (Not in IT/Cyber Field)
IT/Cyber Field vs. Not in IT/Cyber Field	Trust	-0.009	0.993	36	63
IT/Cyber Field vs. Not in IT/Cyber Field	Reliability	0.851	0.399	36	63
All respondents (Trust vs. Reliability)	N/A	0.057	0.955	99	99

In summary, the data suggests that, on average, respondents have a neutral stance on trusting AI in making ethical decisions and its reliability in handling complex ethical dilemmas in cybersecurity. There is no significant difference between the two variables for all respondents. However, those in the cybersecurity or IT field have a slightly more positive view on both aspects compared to the overall sample.

Discussion

The overall neutral stance suggests that there is neither overwhelming confidence nor distrust in AI for cybersecurity across the sample. This may also indicate that there may be a lack of understanding of the implications of AI, especially within the cybersecurity domain. Consequently, the slightly more positive views of those in the IT/Cyber field could be due to greater familiarity with AI capabilities and limitations. The moderate scores for trust and reliability indicate that there's neither overwhelming confidence in nor opposition to AI's role in ethical decision-making and handling complex dilemmas in cybersecurity. This

suggests room for growth in public perception and indicates potential for increased acceptance as AI technology improves and its applications in cybersecurity become more transparent and proven.

Given the slightly higher trust and reliability scores among IT/Cyber professionals, it implies that exposure to and understanding of AI might lead to more positive perceptions. Therefore, educating the non-IT public about AI's capabilities and limitations could improve overall trust and reliability perceptions. The neutral stance also signifies a need for clear communication regarding the ethical considerations and safety measures implemented in AI systems. This could involve explaining the decision-making processes of AI and how ethical outcomes are ensured.

Moreover, the lack of a significant difference in perceptions between those in the field and those outside it could mean that professional exposure does not drastically change one's perception of AI's ethical and reliable capacities in cybersecurity. The correlation results imply that individuals' trust in AI to make ethical decisions is somewhat predictive of their belief in AI's reliability, more so for those not in the IT/Cyber field. This could suggest that non-specialists view the capability of AI in these two dimensions as more interconnected than specialists do. In terms of the implications this study has on policy and ethical framework development, the lack of significant differences in perceptions between IT/Cyber professionals and non-professionals suggests a consensus on the current state of AI in cybersecurity. This can guide policymakers to develop industry-wide ethical frameworks and standards that are generally accepted by both experts and the public. Furthermore, this study highlights the importance of demonstratable reliability. The correlation between trust in ethical decision-making and reliability perceptions, especially among non-IT professionals, indicates that demonstrating AI's reliability in handling ethical dilemmas may improve overall trust in AI. This could lead to more widespread adoption of AI in cybersecurity if the technology can be shown to be consistently reliable.

For AI developers and cybersecurity professionals, the findings point to the importance of integrating ethical considerations into the design of AI systems. Since trust and reliability are linked, AI systems that are designed to be transparent in their decision-making and consistent in their actions could be more readily accepted. With the current moderate trust and reliability scores, there is an opportunity for investment in research that aims to address ethical concerns and enhance the reliability of AI systems. Therefore, by investing in research that focuses on these areas, the cybersecurity field can develop more advanced AI solutions that are trusted and reliable. Consequently, while the current perception of AI in cybersecurity is one of cautious optimism, there is clearly potential for growth. As the technology matures and stakeholders engage in education and policy development, trust and reliability are likely to improve, leading to greater integration of AI in the field of cybersecurity.

Implications for Practice

This study aims to inform AI developers and cybersecurity professionals about the importance of integrating ethical decision-making capabilities in AI systems. Understanding which aspects of AI ethics are concerning to professionals. The public can guide the prioritization of features such as transparency, explainability, and fairness in AI system design which can lead to the development of more trustworthy and reliable AI tools for cybersecurity, where users feel confident in the system's ethical judgments and actions. Moreover, it can help with compliance with upcoming regulations, such as the EU AI Act and adherence with the Executive Order for Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence.

Public trust is crucial for the widespread acceptance and successful implementation of AI technologies in cybersecurity. Addressing the ethical concerns and emphasizing the reliability and trustworthiness of AI

systems, this study can help in crafting strategies for public engagement and communication that effectively convey the benefits and safeguards of AI in cybersecurity. This can help to demystify AI technologies, reduce public apprehension, and build a foundation of trust between technology developers, users, and the broader community.

Limitations

While this study aims to provide insights into the perceptions of trust and reliability in AI systems making ethical decisions in cybersecurity contexts, it is important to acknowledge its limitations. As with any survey-based research, there is a risk of response bias, where participants may answer in a socially desirable manner rather than truthfully. Moreover, there is a level of subjectivity in survey responses, of how the participant may be feeling in that moment of time. The voluntary nature of survey participation can lead to self-selection bias, where the respondents who choose to participate may have different opinions or characteristics from those who do not.

While efforts are made to target a diverse group, the sample may not fully represent the global demographic and professional diversity within cybersecurity, AI, and IT fields. This could limit the generalizability of the findings. Finally, the research questions, while comprehensive, might not cover all the complex ethical considerations involved in AI decision-making in cybersecurity. There may be nuanced ethical implications that are not addressed within the scope of this study.

Conclusion

The integration of AI into cybersecurity, while promising enhanced defenses against cyber threats, brings to light profound ethical considerations that must be carefully navigated to maintain societal trust and technological reliability. The quantitative analysis, grounded in the Technology Acceptance Model (TAM), indicates a moderate level of trust and reliability attributed to AI systems across the surveyed population. While professionals within the IT and cybersecurity field demonstrate a slightly more favorable perception, the overall sentiment remains cautiously optimistic, emphasizing the need for continued efforts in AI ethics and transparency. The insights garnered point to a pressing need for AI systems that not only perform with technical excellence but are also designed and perceived to be ethically sound and reliable. This dual focus on effectiveness and ethics is paramount in fostering a climate of trust, trust that is essential for the broader acceptance and successful integration of AI in cybersecurity.

As AI continues to evolve and its role in cybersecurity becomes increasingly critical, this study underscores the urgency of establishing comprehensive ethical frameworks and robust regulatory standards. The EU AI Act and the Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence are indicative of the international efforts to align technological progress with ethical governance. Thus, this study hopes to prompt further research and exploration with the ethical dimensions of AI in cybersecurity to ensure that these technologies protect against digital threats and uphold the values and trust of the society they serve.

References

EU AI Act: First Regulation on Artificial Intelligence. August 6, 2023. European Parliament. https://www.europarl.europa.eu/topics/en/article/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence

Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence. October 30, 2023. The White House. https://www.whitehouse.gov/briefing-room/presidential-actions/2023/10/30/executive-order-on-the-safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence/

Hagendorff, T. (2020). The ethics of AI ethics: An evaluation of guidelines. *Minds and machines*, *30*(1), 99-120.

Konda, S. R. (2019). Ensuring Trust and Security in AI: Challenges and Solutions for Safe Integration. *INTERNATIONAL JOURNAL OF COMPUTER SCIENCE AND TECHNOLOGY*, *3*(2), 71-86.

Lee, J. D., & See, K. A. (2004). Trust in automation: Designing for appropriate reliance. *Human factors*, *46*(1), 50-80.

Liu, X. M., & Murphy, D. (2020). A multi-faceted approach for trustworthy ai in cybersecurity. *Journal of Strategic Innovation and Sustainability*, *15*(6).

Martin, K. (2022). *Ethics of data and analytics: Concepts and cases*. Auerbach Publications.

Taddeo, M., McCutcheon, T., & Floridi, L. (2019). Trusting artificial intelligence in cybersecurity is a double-edged sword. *Nature Machine Intelligence*, *1*(12), 557-560.

Ryan, M. (2020). In AI we trust: ethics, artificial intelligence, and reliability. *Science and Engineering Ethics*, *26*(5), 2749-2767.

Sanclemente, G. L. (2022). Reliability: Understanding cognitive human bias in artificial intelligence for national security and intelligence analysis. *Security Journal*, *35*(4), 1328-1348.

Sarker, I. H., Furhad, M. H., & Nowrozy, R. (2021). Ai-driven cybersecurity: an overview, security intelligence modeling and research directions. *SN Computer Science*, *2*(3), 173.

Sontan, A. D., & Samuel, S. V. (2024). The intersection of Artificial Intelligence and cybersecurity: Challenges and opportunities. *World Journal of Advanced Research and Reviews*, *21*(2), 1720-1736.

Appendix A

Survey Questions:

1. Do you wish to participate in this survey?
    a. Yes
    b. No
2. Age Range:
    a. 18-25
    b. 26-35
    c. 36-45
    d. 46-55
    e. 56+
3. Are you in the Cybersecurity or IT field
    a. Yes
    b. No
    c. Prefer not to say
4. I believe AI technology can significantly improve cybersecurity.
5. I am concerned about the misuse of AI in cybersecurity.
6. AI in cybersecurity will lead to more advanced forms of cyber threats.
7. The use of AI in cybersecurity raises significant ethical concerns for me.
8. I am worried about the potential for AI in cybersecurity to invade personal privacy.
9. The benefits of AI in cybersecurity outweigh its ethical risks.
10. I trust AI to make ethical decisions in cybersecurity contexts.
11. AI systems are reliable enough to handle complex ethical dilemmas in cybersecurity.
12. Human oversight is necessary to ensure the ethical use of AI in cybersecurity.
13. There is a need for stricter regulations governing the use of AI in cybersecurity.
14. I am confident that current laws adequately address the ethical use of AI in cybersecurity.
15. Organizations using AI in cybersecurity should be transparent about their ethical guidelines.
16. What is your gender?
    a. Male
    b. Female
    c. Prefer not to say
17. Region
18. Device
19. Household income
20. Age range (question asked by SurveyMonkey)