Exploring the Concept of Explainable AI and Developing Information Governance Standards for Enhancing Trust and Transparency in Handling Customer Data

Abstract

The increasing integration of Artificial Intelligence (AI) systems in diverse sectors has raised concerns regarding transparency, trust, and ethical data handling. This study investigates the impact of Explainable AI (XAI) models and robust information governance standards on enhancing trust, transparency, and ethical use of customer data. A mixed-methods approach was employed, combining a comprehensive literature review with a survey of 342 respondents across various industries. The findings reveal that the implementation of XAI significantly increases user trust in AI systems compared to black-box models. Additionally, a strong positive correlation was found between XAI adoption and the ethical use of customer data, highlighting the importance of transparency frameworks and governance mechanisms. Furthermore, the study underscores the critical role of user education in fostering trust and facilitating informed decision-making regarding AI interactions. The results emphasize the need for organizations to prioritize the integration of XAI techniques, establish robust information governance frameworks, invest in user education, and foster a culture of transparency and ethical data use. These recommendations provide a roadmap for organizations to harness the benefits of AI while mitigating potential risks and ensuring responsible and trustworthy AI practices.

Keywords: Explainable AI, Information Governance, Trust, Transparency, Ethical Data Use

1. Introduction

Artificial intelligence (AI) has become a transformative force across various industries, including healthcare, finance, retail, and customer service, as the technology's ability to process vast amounts of data, identify patterns, and make decisions autonomously has significantly improved operational efficiency and innovation [1]. However, as AI systems become more integrated into organizational operations, concerns about transparency, trust, and ethical data handling have emerged as critical issues that need to be addressed [2,3]. Traditional AI models, often referred to as black-box models, are characterized by their lack of transparency on how decisions are made, leading to significant mistrust among users and stakeholders, who may be wary of systems they do not fully understand [4]. The conventional response to this critical issue of transparency in AI's modus operandi is the rise of Explainable AI (XAI), which aims to make AI decision-making processes more transparent and understandable, thereby enhancing user trust and ensuring that AI systems are used more responsibly [4,5].

Another critical issue necessitating the exploration of XAI is the ethical use of customer data, considering that AI systems often rely on large datasets to function effectively. According to Dhirani et al. [6], organizations that fail to manage customer data ethically risk violating user privacy, which can lead to severe legal and reputational consequences. Therefore, the adoption of XAI could potentially promote more ethical data practices by making the decision-making processes of AI systems more transparent. In addition, robust information governance standards are essential for ensuring data security and minimizing the risk of data breaches [7]. Olaniyi [8] further assert that effective information governance can help organizations comply with regulatory requirements, protect user data, and maintain trust.

The lack of transparency in AI systems can lead to mistrust among users, which can hinder the adoption and effectiveness of AI technologies, since users are more likely to reject or question the outputs of AI systems if they do not understand how decisions are made [6,7]. This mistrust can reduce the effectiveness of AI systems and limit their potential benefits. Moreso, AI systems that handle customer data without transparent and ethical practices can lead to privacy violations and misuse of personal information

[9,10]. These issues can result in legal penalties, loss of customer trust, and damage to an organization's reputation [11]. Thus, there is a necessity for comprehensive strategies that integrate explainable AI models and robust information governance standards to enhance trust, transparency, and ethical use of customer data [12].

Robust information governance standards are essential for preventing data breaches and ensuring that customer data is handled securely. In the abscence of transparency in AI models and systems, users will continue to mistrust AI systems, limiting their adoption and effectiveness [12]. Without ethical data practices, organizations risk privacy violations and misuse of personal information, leading to legal penalties and loss of customer trust [11,12]. In addition, the risk of data breaches will remain high, resulting in severe financial and reputational consequences. Therefore, this study aims to investigate the impact of Explainable AI models and robust information governance standards on enhancing trust, transparency, and ethical use of customer data in organizations, and to formulate a strategy for improving user trust through effective education on AI transparency and data governance. The study objectives are:

1. To evaluate the effect of Explainable AI models on user trust in AI systems and analyze the relationship between their adoption and the ethical use of customer data in organizations.

2. To assess the impact of robust information governance standards on the incidence of data breaches in organizations utilizing AI systems.

3. To examine the role of user education on AI transparency and data governance in increasing the perceived trustworthiness of AI systems.

4. To formulate a comprehensive strategy for organizations to implement Explainable AI and robust information governance standards, thereby enhancing user trust and ensuring ethical data practices.

2. Literature Review

The concept of Explainable AI (XAI)

Explainable AI (XAI) has emerged as a pivotal field within artificial intelligence, addressing the critical need for transparency and understandability in AI decision-making processes [4]. Traditionally, AI models, especially deep learning and other complex algorithms, have been characterized as black-box models due to their opaque nature, as users and stakeholders often find it challenging to comprehend how these models arrive at specific decisions, leading to mistrust and reluctance to adopt AI technologies [7,13]. Thus, explainable AI seeks to mitigate these issues by providing clear, interpretable, and actionable explanations for AI-generated outcomes [7].

de Bruijn [4] contends that although early AI systems were relatively simple and their decision-making processes were more transparent; the complexity and opacity of models increased as AI technologies advanced, particularly with the advent of deep learning, resulting to a pressing demand for methods that could demystify these sophisticated models. In response, various explainable AI (XAI) techniques have emerged, aimed at making AI more interpretable, such as model-agnostic methods, including LIME (Local Interpretable Model-agnostic Explanations) and SHAP (SHapley Additive exPlanations), which provide post-hoc explanations by approximating the behavior of complex models with simpler, interpretable ones [14,15,16]. Other approaches, such as attention mechanisms in neural networks and inherently interpretable models like decision trees and linear models, aim to build transparency directly into the model architecture [17,18].

Despite its potential, XAI faces several challenges including the trade-off between model accuracy and interpretability, considering that studies have shown that complex models often achieve higher accuracy but are less interpretable, whereas simpler models are more transparent but may sacrifice performance [4,19]. Another challenge is the subjective nature of explanations, as different users might require different types of explanations depending on their expertise and needs, making it difficult to design universally satisfactory explainability techniques [19,20,21]. Furthermore, there is a lack of standardized metrics to evaluate the effectiveness of XAI methods, complicating efforts to assess and compare different approaches [22].

The benefits of XAI are however significant, particularly in enhancing user trust in AI systems. By providing understandable explanations, XAI helps users feel more confident in the decisions made by AI, thereby fostering greater acceptance and trust, which is particularly important in sensitive fields such as healthcare, finance, and criminal justice, where the consequences of AI decisions can be profound [19,21,23]. Transparent AI systems can also facilitate better human-AI collaboration, as users are more likely to trust and effectively interact with systems they understand [24]. Moreover, XAI can aid in identifying and mitigating biases in AI models, promoting universality, accuracy, fairness and ethical use of AI [24].

Despite these benefits, the implementation of XAI remains controversial, as some studies argue that post-hoc explanations might not fully capture the decision-making processes of complex models, potentially leading to misleading interpretations [422]. Additionally, the effectiveness of XAI in actually improving user trust is still an area of active research, with some studies suggesting that the mere presence of explanations is not always sufficient to enhance trust, especially if the explanations are not perceived as credible or useful [19,22].

Information Governance Standards

The importance of information governance has grown significantly in recent years due to the exponential increase in data generation and the corresponding rise in data breaches and privacy concerns [25]. Information governance have proven to be essential in management and organizational strategy, considering its role in ensuring that data is handled in a way that meets regulatory requirements, protects privacy, and maintains the integrity and availability of information especially in industries where data sensitivity is paramount, such as healthcare, finance, and retail [25,26].

Key standards and frameworks in information governance include the General Data Protection Regulation (GDPR), the Health Insurance Portability and Accountability Act (HIPAA), and the ISO/IEC 27001 standard for information security management [18,27]. GDPR, implemented in the European Union, has set a high benchmark for data protection and privacy, influencing global practices and legislation [28]. HIPAA focuses on the protection of health information in the United States, while ISO/IEC 27001 provides a systematic approach to managing sensitive company information so that it remains secure [29]. These frameworks establish clear guidelines and best practices for organizations to follow, ensuring that data governance is not only comprehensive but also standardized across different sectors and regions [30,31].

The role of information governance in reducing data breaches cannot be overstated as it establishes stringent controls over how data is accessed, used, and stored, so that organizations can significantly mitigate the risk of unauthorized access and data leaks [31,32]. Studies have shown that organizations with robust information governance frameworks experience fewer data breaches and are better equipped to respond effectively when breaches occur [33,34]. This is particularly relevant in the context of AI systems, which often handle vast amounts of sensitive data. Effective governance ensures that AI systems are designed and operated in a manner that prioritizes data security, thereby reducing vulnerabilities and enhancing trust among users and stakeholders [6].

However, the implementation of information governance standards encounters several challenges such as the dynamic nature of regulatory requirements, which can vary significantly across different jurisdictions[8]. Organizations operating globally are obligated to navigate a complex landscape of regulations, making it difficult to develop a cohesive governance strategy [1]. Moreover, the rapid pace of technological advancement means that governance frameworks must continuously evolve to address new types of data and emerging threats, thereby necessitating ongoing investment in training, technology, and compliance monitoring, which can be resource-intensive [8].

By establishing clear guidelines for data handling, governance frameworks promote accountability and transparency, ensuring that organizations use data responsibly [13]. This is particularly important in the context of AI, where concerns about bias, discrimination, and unethical data usage are prevalent. Effective governance can help organizations identify and mitigate biases in their AI models, ensuring that decisions made by XAI systems are fair and equitable [35][36]. Additionally, governance frameworks that emphasize transparency can enhance user trust by providing clear information about how data is collected, used, and protected.

Trust and Transparency in AI Systems

Trust and transparency are critical factors in artificial intelligence (AI) systems, influencing their acceptance and integration across various domains. Trust in AI refers to the confidence users have in the system's ability to perform tasks reliably, ethically, and without causing harm [37]. Transparency, on the other hand, involves the clarity and openness with which AI systems operate, including the understandability of their decision-making processes and the accessibility of information regarding their functioning [38].

For XAI technologies to be widely accepted and integrated into critical areas such as healthcare, finance, and law enforcement, users and stakeholders must believe that these systems are reliable, fair, and secure [39][40]. Trust mitigates fears of unpredictability and potential misuse, while transparency helps users understand and verify the actions of AI systems, fostering a sense of control and accountability. Without these elements, the adoption of XAI is likely to face significant resistance, limiting its potential benefits [40].

Asserting the crucial role of AI in fostering trust and transparency, Schmidt et al [41] alludes that users are more likely to trust AI systems when they can understand how decisions are made. By providing clear and interpretable explanations for AI decisions, XAI helps demystify complex models and makes them more accessible to non-expert users. For instance, Elon Musk has expressed concerns about AI bias on several occasions. One notable instance occurred in June 2020, when Musk tweeted about his concern regarding OpenAI's GPT-3, the language model developed by the AI research lab he co-founded. In his tweet, Musk suggested that the AI had a political bias [42]. This reflects his broader apprehension about AI systems reflecting the biases of their creators and the potential implications of these biases on society [43]. Another factor is the perceived fairness of AI systems. If users believe that an AI system is biased or discriminatory, their trust in the system will diminish [41]. Ensuring that AI models are developed and trained with fairness and inclusivity in mind is essential for building trust. Additionally, the robustness and security of AI systems are critical. Users need to be confident that AI technologies can withstand attacks and operate reliably under various conditions [43]. AI directly addresses the transparency issue by making the decision-making processes of AI systems more interpretable. Techniques such as LIME (Local Interpretable Model-agnostic Explanations) and SHAP (SHapley Additive exPlanations) allow users to see which features influenced a particular decision, thereby providing insights into the model's functioning [44][45]. This level of transparency not only helps users understand AI decisions but also enables them to identify and address potential biases and errors. By fostering greater transparency, XAI enhances trust in AI systems, making users more willing to rely on these technologies [45].

Ethical Use of Customer Data

As AI systems increasingly leverage vast amounts of personal data to drive decision-making processes, ethical considerations have become crucial to ensure that data practices align with societal norms and legal standards. Cite avers that when companies misuse customer data or fail to protect it adequately, they risk violating privacy rights, leading to significant legal repercussions and financial penalties [46]. High-profile data breaches, such as those experienced by Equifax and Facebook, have not only resulted in substantial fines but have also caused long-lasting damage to the organizations' reputations, and also users trust and usage of their services. Users affected by such breaches face risks such as identity theft, financial loss, and emotional distress [47][48]. Moreover, unethical data practices can erode trust in digital services, reducing user engagement and willingness to share data, which is detrimental to businesses and systems which rely on data-driven insights and processes [48]. Therefore, ensuring the transparency of AI decision-making processes ensure that data is used responsibly. When users can see how their data influences AI outcomes, they are more likely to trust that their data is being handled ethically [38]. Furthermore, XAI enables organizations to audit and validate their AI systems, ensuring compliance with ethical standards and identifying potential biases that could lead to unfair or discriminatory outcomes [36]. This transparency is particularly important in sectors like finance and healthcare, where decisions can have significant personal impacts.

Several case studies highlight the necessity of transparency and trust in systems and ethical data practices. For instance, the misuse of data by Cambridge Analytica in the 2016 U.S. presidential election is a prime example of unethical data practices [49]. The company exploited personal data from millions of Facebook users without their consent to influence voter behavior, leading to widespread condemnation and regulatory scrutiny. Such case as the Cambridge Analytica suggests an insight to the endless possibility of users' data at the disposal of organizations, and even worse when such an organization operates a black box model, in which users are unsure how their information is being utilized. In contrast, IBM's deployment of AI in their Watson Health project demonstrates ethical data use as the company employs rigorous data governance frameworks and XAI techniques to ensure that patient data is used transparently and ethically, enhancing patient outcomes while maintaining trust.

However, Kitchin [50] argues that while detailed data can provide valuable insights and drive innovation, it also increases the risk of privacy breaches; hence, striking the right balance requires continuous dialogue between policymakers, technologists, and the public [51]. Additionally, the fast-paced evolution of AI technologies often outstrips the development of regulatory frameworks, leading to gaps in governance and oversight.

User Education on AI Transparency and Data Governance

User education on AI transparency and data governance is increasingly recognized as a crucial component in the adoption and ethical use of AI systems [52]. As AI technologies permeate various aspects of daily life, it becomes essential for users to understand how these systems work, how their data is being used, and what measures are in place to protect their privacy [53]. Effective education initiatives can demystify AI, foster trust, and empower users to make informed decisions about their interactions with AI systems. The importance of user education in AI lies in its potential to bridge the knowledge gap between AI developers and end-users [54]. Educated users are more likely to trust AI systems if they understand the underlying mechanisms and the safeguards protecting their data. This trust is crucial for the widespread adoption of AI technologies, particularly in sectors where data sensitivity is paramount, such as healthcare and finance [45]. Moreover, informed users can engage more effectively with AI systems, providing valuable feedback that can guide the development of more user-centric AI solutions.

Several case studies highlight best practices in user education. For example, Google's "AI for Everyone" initiative offers free online courses designed to educate the public about AI's basics, its applications, and ethical considerations [55]. This initiative has been praised for its accessible content and broad reach, helping to demystify AI for a global audience. Similarly, IBM's Watson Health program includes educational components that explain how AI assists in medical decision-making, enhancing transparency and trust among patients and healthcare professionals [56]. Emerging trends in user education emphasize interactive and experiential learning. AI-driven platforms that offer personalized learning experiences based on user interactions are gaining traction [57]. These platforms can adapt to individual learning paces and preferences, providing a more effective and engaging educational experience. Additionally, collaboration between tech companies, educational institutions, and policymakers is increasingly seen as essential for developing comprehensive AI education strategies [58][59].

Existing Strategies and Best Practices

Existing strategies for implementing Explainable AI (XAI) and robust information governance have shown varied levels of effectiveness, with some organizations setting benchmarks in enhancing trust and transparency. One prominent strategy for XAI implementation is the adoption of model-agnostic explainability techniques, such as LIME (Local Interpretable Model-agnostic Explanations) and SHAP (SHapley Additive exPlanations) [45]. These methods have gained popularity due to their flexibility in providing post-hoc explanations for a wide range of AI models. Organizations like Microsoft and Google have integrated these techniques into their AI frameworks, allowing users to understand the factors influencing AI decisions [46]. This transparency fosters greater trust among users, as they can see and verify how their data is being utilized.

Robust information governance strategies often involve comprehensive frameworks like the General Data Protection Regulation (GDPR) and the ISO/IEC 27001 standard for information security management [61][62]. These frameworks mandate strict data handling practices and regular audits to ensure compliance. For instance, GDPR's emphasis on user consent and data minimization has led to more transparent data practices across organizations operating in the European Union. Similarly, ISO/IEC 27001 certification has become a hallmark of trust, signaling to stakeholders that an organization takes data security seriously [62][63].

Best practices from organizations that have successfully enhanced trust and transparency offer valuable insights. IBM's Watson Health, for example, employs rigorous data governance practices combined with custom designed XAI models to ensure that AI-driven medical insights are both reliable and understandable to healthcare professionals [64]. This dual focus on explainability and robust governance not only enhances trust but also ensures compliance with regulatory standards in healthcare, where data sensitivity is paramount [65,66,67].

Methodology

The study's proposed hypotheses are:

H_1: The implementation of Explainable AI models significantly increases user trust in AI systems compared to black-box models.

H_2: There is a positive correlation between the adoption of Explainable AI and the ethical use of customer data in organizations.

H_3: Robust information governance standards reduce the incidence of data breaches in organizations utilizing AI systems.

H_4: User education on AI transparency and data governance increases the perceived trustworthiness of AI systems.

To achieve the research objectives and evaluate the proposed hypotheses, the study employed a mixed-methods approach, integrating both qualitative and quantitative analyses. The qualitative analysis involved an extensive review of academic literature and industry reports from reputable databases including Google Scholar, IEEE Xplore, and the ACM Digital Library. Stringent inclusion and exclusion criteria were applied to select 21 papers for in-depth analysis, focusing on those that specifically addressed XAI, information governance, trust, transparency, and ethical data use in the context of AI systems. The qualitative data extracted from these sources were subjected to thematic analysis, identifying key themes and patterns related to the research objectives.

The quantitative analysis involved the development and administration of a survey questionnaire to a diverse sample of 342 respondents representing various roles and industries. The questionnaire comprised both Likert-scale and open-ended questions designed to assess the perceived impact of XAI and information governance on trust, transparency, and ethical data use. The quantitative data collected were analyzed using statistical techniques, including descriptive statistics, correlation analysis, and logistic regression.

Table 1: Qualitative result from the Literature and Industry

Hypothesis	Literature Source	Industry Source	Qualitative Data	Content Analysis	Insights
H1	Adadi & Berrada [74], Alufaisan et al. [70], Bach et al. [76], Markus et al. [15], Shin [75], Shin [71]	Google, IBM, Microsoft Research, The Alan Turing Institute	XAI methods enhance transparency leading to increased user trust.	Clear explanations and user-centric design improve trust.	Technical complexity is a challenge, but design and education are key.
H2	Arrieta et al. [82], Felzmann [38], Georgieva et al. [80], Schmidt et al. [41] Palladino [79]	Google, IBM, OpenAI, Microsoft Research, The Alan Turing Institute	Ethical AI practices linked to XAI adoption.	Transparency frameworks and robust governance support ethical use.	Ethical guidelines and strong governance frameworks enhance trust.
H3	Saeed & Omlin [81], Fejza et al. [73], Bell et al. [17], Vereschak et al. [72]	Google, IBM, Microsoft Research	Best practices in data governance ensure data integrity.	Governance and compliance strategies reduce breaches.	Effective governance reduces breaches, but scalability can be a challenge.
H4	Liaison [69], Shin	Google, IBM,	Education on AI	Continuous education	Educational programs

	[71], Akinrinola et al. [78], Singhal et al. [77], Larsson and Heintz [83], Mohseni et al. [68]	Microsoft Research, The Alan Turing Institute	transparency builds trust.	efforts and diverse design approaches enhance trust.	and public engagement are crucial for trust and transparency.

The study, as detailed in Table 1, shows that adopting Explainable AI (XAI) significantly enhances user trust through transparency, supported by sources such as Adadi & Berrada [74] and industry leaders like Google and IBM. Clear explanations and user-centric designs are crucial, though technical complexity remains a challenge. There is a positive correlation between XAI adoption and ethical use of customer data, with robust transparency frameworks and governance enhancing trust, as supported by Arrieta et al. [82] and organizations like OpenAI. Robust information governance standards are essential in reducing data breaches, though scalability is a challenge, as indicated by Markus et al. [15]. Continuous user education on AI transparency and data governance is vital for maintaining trust, supported by Liaison [69] and industry practices from The Alan Turing Institute.

Hypothesis 1: The implementation of Explainable AI models significantly increases user trust in AI systems compared to black-box models.

Table 2: Convergent Validity Assessment

Construct	Indicator	Item Loading	Item Communality
Implementation of Explainable AI (XAI)	XAI1: Fully implemented	0.84	0.71
	XAI2: Partially implemented	0.81	0.66
	XAI3: Considering implementation	0.78	0.61
	XAI4: Not implemented	0.75	0.57
Perceived Trust in AI Systems	TRUST1: Significantly increased	0.85	0.72
	TRUST2: Somewhat increased	0.83	0.69
	TRUST3: No change	0.79	0.62

	TRUST4: Somewhat decreased	0.77	0.59
	TRUST5: Significantly decreased	0.75	0.57
Transparency of AI Decisions	TRANS1: Extremely transparent	0.87	0.76
	TRANS2: Moderately transparent	0.85	0.72
	TRANS3: Neutral	0.82	0.67
	TRANS4: Slightly transparent	0.79	0.62
	TRANS5: Not transparent	0.76	0.58
Ethical Use of Customer Data	ETHICS1: Strongly agree	0.86	0.74
	ETHICS2: Agree	0.84	0.71
	ETHICS3: Neutral	0.81	0.66
	ETHICS4: Disagree	0.78	0.61
	ETHICS5: Strongly disagree	0.76	0.58

Table 3: Reliability and Validity Metrics

Construct	Cronbach's Alpha	Composite Reliability	AVE
Implementation of Explainable AI (XAI)	0.87	0.89	0.67
Perceived Trust in AI Systems	0.88	0.91	0.68
Transparency of AI Decisions	0.89	0.92	0.71
Ethical Use of Customer Data	0.91	0.93	0.7

Table 4: Discriminant Validity (Fornell-Larcker Criterion & HTMT Ratio)

Construct	Indicator	Item Loading	Item Communality

Implementation of Explainable AI (XAI)	XAI1: Fully implemented	0.84	0.71
	XAI2: Partially implemented	0.81	0.66
	XAI3: Considering implementation	0.78	0.61
	XAI4: Not implemented	0.75	0.57
Perceived Trust in AI Systems	TRUST1: Significantly increased	0.85	0.72
	TRUST2: Somewhat increased	0.83	0.69
	TRUST3: No change	0.79	0.62
	TRUST4: Somewhat decreased	0.77	0.59
	TRUST5: Significantly decreased	0.75	0.57
Transparency of AI Decisions	TRANS1: Extremely transparent	0.87	0.76
	TRANS2: Moderately transparent	0.85	0.72
	TRANS3: Neutral	0.82	0.67
	TRANS4: Slightly transparent	0.79	0.62
	TRANS5: Not transparent	0.76	0.58
Ethical Use of Customer Data	ETHICS1: Strongly agree	0.86	0.74
	ETHICS2: Agree	0.84	0.71
	ETHICS3: Neutral	0.81	0.66
	ETHICS4: Disagree	0.78	0.61
	ETHICS5: Strongly disagree	0.76	0.58

Table 5: Structural Model Analysis Results (Bootstrapping)

Construct	XAI	Trust	Transparency	Ethics
Implementation of Explainable AI (XAI)	0.82			
Perceived Trust in AI Systems	0.55	0.83		
Transparency of AI Decisions	0.5	0.6	0.85	

Ethical Use of Customer Data	0.45	0.5		0.55	0.84

Table 6: Bootstrapping Result

Path	Path Coefficient (β)	t-test	p-Value	95% Confidence Interval
XAI -> Trust	0.5	6.75	<0.001	0.35 - 0.65
Trust -> Transparency	0.6	8.1	<0.001	0.50 - 0.70
Transparency -> Ethics	0.45	5.85	<0.001	0.35 - 0.55
XAI -> Trust (indirect via Transparency)	0.3	4.1	<0.001	0.20 - 0.40
Trust -> Ethics (indirect via Transparency)	0.35	5	<0.001	0.25 - 0.45

The study results presented from table 2 to table 6 above supports Hypothesis 1, demonstrating that the implementation of Explainable AI (XAI) models significantly increases user trust in AI systems compared to black-box models. Table 1 shows high item loadings and communalities for constructs related to XAI implementation, perceived trust, transparency, and ethical use of customer data, indicating strong convergent validity. Reliability and validity metrics in Table 2 confirm the constructs' robustness, with Cronbach's Alpha and Composite Reliability values all exceeding 0.87, and Average Variance Extracted (AVE) values above 0.67. Discriminant validity metrics in Table 3, based on the Fornell-Larcker criterion and HTMT ratio, further validate these constructs.

The Structural model analysis in Table 4 shows significant path coefficients, with XAI directly impacting trust ($\beta = 0.5$, $p < 0.001$) and trust influencing transparency ($\beta = 0.6$, $p < 0.001$). Transparency, in turn, enhances the ethical use of customer data ($\beta = 0.45$, $p < 0.001$). Indirect effects also support the model, with XAI impacting trust through transparency ($\beta = 0.3$, $p < 0.001$) and trust influencing ethics via transparency ($\beta = 0.35$, $p < 0.001$). These results underscore the critical role of XAI in fostering trust, transparency, and ethical practices in AI systems.

Hypothesis 2: There is a positive correlation between the adoption of Explainable AI and the ethical use of customer data in organizations.

Table 7: Correlations between Information Governance, Data Breaches, Effectiveness, and Compliance

	Adoption Scores	Breaches Scores	Effectiveness Scores	Compliance Scores
Adoption Scores	1.00			

Breaches Scores	-0.90	1.00		
Effectiveness Scores	0.85	-0.75	1.00	
Compliance Scores	0.92	-0.82	0.87	1.00

The study's findings provide substantial support for Hypothesis 2 (H2), which posits a positive correlation between the adoption of Explainable AI (XAI) and the ethical use of customer data in organizations. The data presented in Table 5 demonstrates a strong positive correlation between adoption scores and compliance scores ($r = .92$), indicating that higher levels of XAI adoption are closely associated with improved compliance and ethical data practices. Additionally, the results depicted in chart H2Q14 reveal that a significant majority of respondents perceive that the implementation of XAI has enhanced the ethical use of customer data. These converging lines of evidence provide robust support for accepting Hypothesis 2.

Hypothesis 3: Robust information governance standards reduce the incidence of data breaches in organizations utilizing AI systems.

Table 8: Logistic Regression Analysis Results Predicting Adoption of Information Governance Standards

Variable	Coefficient (β)	Standard Error	z-value	p-value	95% Confidence Interval
Intercept	-1.24	0.35	-3.54	<0.001	-1.92 to -0.56
Availability of Education Programs (ref: No programs)					
Extensive programs available	1.15	0.42	2.74	0.006	0.33 to 1.97
Some programs available	0.75	0.37	2.03	0.042	0.03 to 1.47
Considering developing programs	0.45	0.30	1.50	0.134	-0.13 to 1.03
Perceived Trustworthiness (ref: No change)					
Significantly increased	1.30	0.38	3.42	<0.001	0.56 to 2.04
Somewhat increased	0.90	0.35	2.57	0.010	0.21 to 1.59
Somewhat decreased	-0.75	0.40	-1.88	0.061	-1.54 to 0.04
Significantly decreased	-1.20	0.42	-2.86	0.004	-2.02 to -0.38

Understanding of AI Transparency (ref: No understanding)					
Excellent understanding	1.45	0.43	3.37	<0.001	0.61 to 2.29
Good understanding	1.00	0.39	2.56	0.011	0.23 to 1.77
Moderate understanding	0.70	0.35	2.00	0.046	0.01 to 1.39
Basic understanding	0.35	0.31	1.13	0.258	-0.26 to 0.96
Effectiveness of Education (ref: Not effective)					
Highly effective	1.50	0.44	3.41	<0.001	0.64 to 2.36
Moderately effective	1.05	0.41	2.56	0.011	0.24 to 1.86
Slightly effective	0.60	0.37	1.62	0.105	-0.12 to 1.32

The logistic regression analysis results provide strong support for the acceptance of Hypothesis 3, which posits that the availability of education programs on AI transparency and data governance significantly increases the perceived trustworthiness of AI systems. The significant positive coefficients for the availability of extensive education programs ($\beta = 1.15$, $p = .006$) and some programs ($\beta = 0.75$, $p = .042$) indicate that such programs are associated with higher adoption of information governance standards. Additionally, the significant positive effects of perceived trustworthiness and understanding of AI transparency further reinforce the hypothesis. Therefore, based on these results, Hypothesis 3 can be accepted.

Hypothesis 4: User education on AI transparency and data governance increases the perceived trustworthiness of AI systems.

Table 9: ANOVA Test Results for Hypothesis 4

Source	SS	df	MS	F	p-value
Between Groups	45.634	4	11.41	8.75	< .001
Within Groups	442.27	337	1.31		
Total	487.90	381			

The ANOVA test results in Table 9 further validate Hypothesis 4. The analysis shows a significant effect of best practices in XAI and information governance on trust levels in AI systems, with an F-value of 8.75 and a p-value of less than .001. The between-groups sum of squares (SS = 45.634) indicates substantial variation due to the implementation of these best practices, while the within-groups sum of squares (SS = 442.27) reflects individual differences among respondents.

Discussion

The study affirms that the implementation of Explainable AI (XAI) models significantly increases user trust in AI systems compared to black-box models. This finding aligns with previous research, which has consistently emphasized the importance of transparency and interpretability in fostering trust in AI [4,7,13]. As Adadi & Berrada [74] and Alufaisan et al. [70] noted, XAI methods enhance transparency by providing clear explanations for AI-generated outcomes, thereby mitigating the opacity associated with black-box models. The study's quantitative analysis further solidifies this notion, revealing a strong positive correlation between XAI implementation and perceived trust in AI systems. This correlation is consistent with the findings of Bach et al. [76] and Markus et al. [15], who highlighted the critical role of transparency and user interface design in building trust.

In addition, the study shows that there is a positive correlation between the adoption of Explainable AI (XAI) and the ethical use of customer data in organizations. This result is consistent with the literature, which emphasizes the role of XAI in promoting ethical AI practices [18,27]. For instane, Arrieta et al. [82] proposed a taxonomy of XAI methods that facilitate ethical AI practices, while Bell et al. [17] highlighted the importance of transparency frameworks in supporting ethical data use. The study's quantitative analysis further strengthens this link, demonstrating a strong positive correlation between XAI adoption scores and compliance scores. This suggests that organizations that have embraced XAI are more likely to adhere to ethical data practices, aligning with the findings of Georgieva et al. [80] and Palladino [79], who emphasized the importance of bridging the gap between ethics principles and practice.

Considering the role of user education, the study affirms the positive effect of user education on AI transparency and data governance, which increases the perceived trustworthiness of AI systems. This result aligns with the literature, which consistently emphasizes the importance of user education in fostering trust in AI [52,53]. As Liaison [69] noted, education on AI explainability and transparency can significantly improve user trust and acceptance. The study's quantitative analysis further reinforces this notion, revealing a significant positive effect of education programs on perceived trustworthiness. This finding is consistent with the work of Akinrinola et al. [78] and Singhal et al. [77], who highlighted the importance of educating users on AI fairness, accountability, and transparency.

Finally, the study's findings strongly validate that a synergistic approach of explainable AI (XAI) and information governance enhances overall trust in AI systems. This result is consistent with literature, which asserts the importance of both technical and governance aspects in building trust in AI [61,62]. The adoption of model-agnostic explainability techniques, as advocated by Microsoft and Google, and the implementation of robust information governance frameworks like GDPR and ISO/IEC 27001, have been shown to significantly enhance transparency and trust [45,46]. The study's quantitative analysis, particularly the ANOVA test results, further solidifies this notion, revealing a significant effect of best practices on trust levels. This finding aligns with the case studies of IBM's Watson Health and JP Morgan's AI implementation in finance, which demonstrate the effectiveness of combining XAI with robust governance to enhance trust and transparency [64,66].

Conclusion and Recommendation

The findings of this study highlights the significant impact of XAI in mitigating the opacity of black-box models and enhancing user trust. The positive correlation between XAI adoption and ethical data use further emphasizes the importance of transparency frameworks and governance mechanisms in ensuring responsible AI practices. Moreover, the study's findings underscore the critical role of user education in bridging the knowledge gap between AI developers and end-users, thereby fostering trust and facilitating informed decision-making regarding AI interactions.

Based on these findings, this study suggests the following recommendations:

1. Organizations should actively invest in the development and implementation of XAI methods to enhance the transparency and interpretability of AI decision-making processes. This can be achieved through the adoption of model-agnostic explainability techniques, the development of inherently interpretable models, or a combination of both approaches.

2. Organizations should establish comprehensive information governance frameworks that encompass data handling practices, privacy protection measures, and regular audits to ensure compliance with regulatory standards and ethical guidelines. This includes the adoption of industry-standard frameworks such as GDPR and ISO/IEC 27001, as well as the development of internal policies and procedures that prioritize data security and ethical use.

3. Organizations should prioritize user education initiatives that focus on AI transparency, data governance, and ethical considerations. These programs should be designed to demystify AI, empower users to understand how their data is being used, and foster trust in AI systems. This can be achieved through various channels, including online courses, workshops, and interactive platforms that offer personalized learning experiences.

4. Organizations should cultivate a culture that values transparency and ethical data practices at all levels. This involves promoting open communication about AI systems, encouraging employees to raise concerns about potential biases or ethical issues, and ensuring that decision-makers are held accountable for the responsible use of AI and customer data.

COMPETING INTERESTS DISCLAIMER:

Authors have declared that they have no known competing financial interests OR non-financial interests OR personal relationships that could have appeared to influence the work reported in this paper.

References

[1] R. Girasa, *Artificial Intelligence as a Disruptive Technology: Economic Transformation and Government Regulation*. Springer Nature, 2020. Accessed: Jun. 01, 2024. [Online]. Available: https://books.google.com/books?hl=en&lr=&id=QHDJDwAAQBAJ&oi=fnd&pg=PR8&dq=Artificial+intelligence+(AI)+has+become+a+transformative+force+across+various+industries

[2] H. Felzmann, E. Fosch-Villaronga, C. Lutz, and A. Tamò-Larrieux, “Towards Transparency by Design for Artificial Intelligence,” *Science and Engineering Ethics*, vol. 26, no. 6, pp. 3333–3361, Nov. 2020, doi: https://doi.org/10.1007/s11948-020-00276-4

[3] N. Omrani, G. Rivieccio, U. Fiore, F. Schiavone, and S. G. Agreda, “To trust or not to trust? An assessment of trust in AI-based systems: Concerns, ethics and contexts,” *Technological Forecasting and Social Change*, vol. 181, p. 121763, Aug. 2022, doi: https://doi.org/10.1016/j.techfore.2022.121763

[4] H. de Bruijn, M. Warnier, and M. Janssen, “The perils and pitfalls of explainable AI: Strategies for explaining algorithmic decision-making,” *Government Information Quarterly*, vol. 39, no. 2, p. 101666, Dec. 2021, doi: https://doi.org/10.1016/j.giq.2021.101666

[5] M. B. Fazi, “Beyond Human: Deep Learning, Explainability and Representation,” *Theory, Culture & Society*, p. 026327642096638, Nov. 2020, doi: https://doi.org/10.1177/0263276420966386

[6] L. L. Dhirani, N. Mukhtiar, B. S. Chowdhry, and T. Newe, “Ethical Dilemmas and Privacy Issues in Emerging Technologies: A Review,” *Sensors*, vol. 23, no. 3, p. 1151, Jan. 2023, doi: https://doi.org/10.3390/s23031151

[7] Apostolos Vorras and L. Mitrou, “Unboxing the Black Box of Artificial Intelligence: Algorithmic Transparency and/or a Right to Functional Explainability,” *Springer eBooks*, pp. 247–264, Jan. 2021, doi: https://doi.org/10.1007/978-3-030-69583-5_10

[8] O. O. Olaniyi, “Ballots and Padlocks: Building Digital Trust and Security in Democracy through Information Governance Strategies and Blockchain Technologies,”

Asian Journal of Research in Computer Science, vol. 17, no. 5, pp. 172–189, Mar. 2024, doi: https://doi.org/10.9734/ajrcos/2024/v17i5447

[9] H. H. H. Aldboush and M. Ferdous, “Building Trust in Fintech: An Analysis of Ethical and Privacy Considerations in the Intersection of Big Data, AI, and Customer Trust,” *International Journal of Financial Studies*, vol. 11, no. 3, p. 90, Sep. 2023, doi: https://doi.org/10.3390/ijfs11030090

[10] J. Carmody, S. Shringarpure, and G. Van de Venter, “AI and privacy concerns: a smart meter case study,” *Journal of Information, Communication and Ethics in Society*, vol. ahead-of-print, no. ahead-of-print, Jul. 2021, doi: https://doi.org/10.1108/jices-04-2021-0042

[11] I. Varošanec, “On the path to the future: mapping the notion of transparency in the EU regulatory framework for AI,” *International Review of Law, Computers & Technology*, vol. 36, no. 2, pp. 95–117, Apr. 2022, doi: https://doi.org/10.1080/13600869.2022.2060471

[12] M. Mylrea and N. Robinson, “Artificial Intelligence (AI) Trust Framework and Maturity Model: Applying an Entropy Lens to Improve Security, Privacy, and Ethical AI,” *Entropy*, vol. 25, no. 10, p. 1429, Oct. 2023, doi: https://doi.org/10.3390/e25101429

[13] T. O. Oladoyinbo, S. O. Olabanji, O. O. Olaniyi, O. O. Adebiyi, O. J. Okunleye, and A. I. Alao, “Exploring the Challenges of Artificial Intelligence in Data Integrity and its Influence on Social Dynamics,” *Asian Journal of Advanced Research and Reports*, vol. 18, no. 2, pp. 1–23, Jan. 2024, doi: https://doi.org/10.9734/ajarr/2024/v18i2601

[14] S. Bramhall, H. Horn, M. Tieu, and N. Lohia, “QLIME-A Quadratic Local Interpretable Model-Agnostic Explanation Approach,” *SMU Data Science Review*, vol. 3, no. 1, Apr. 2020, Available: https://scholar.smu.edu/datasciencereview/vol3/iss1/4/

[15] A. F. Markus, J. A. Kors, and P. R. Rijnbeek, “The role of explainability in creating trustworthy artificial intelligence for health care: A comprehensive survey of the terminology, design choices, and evaluation strategies,” *Journal of Biomedical Informatics*, vol. 113, Jan. 2021, doi: https://doi.org/10.1016/j.jbi.2020.103655

[16] C. S. Adigwe, O. O. Olaniyi, S. O. Olabanji, O. J. Okunleye, N. R. Mayeke, and S. A. Ajayi, “Forecasting the Future: The Interplay of Artificial Intelligence, Innovation, and Competitiveness and its Effect on the Global Economy,” *Asian journal of economics, business and accounting*, vol. 24, no. 4, pp. 126–146, Feb. 2024, doi: https://doi.org/10.9734/ajeba/2024/v24i41269

[17] A. Bell, O. Nov, and J. Stoyanovich, “Think about the stakeholders first! Toward an algorithmic transparency playbook for regulatory compliance,” *Data & Policy*, vol. 5, p. e12, Jan. 2023, doi: https://doi.org/10.1017/dap.2023.8

[18] A. T. Arigbabu, O. O. Olaniyi, C. S. Adigwe, O. O. Adebiyi, and S. A. Ajayi, “Data Governance in AI - Enabled Healthcare Systems: A Case of the Project Nightingale,” *Asian Journal of Research in Computer Science*, vol. 17, no. 5, pp. 85–107, Mar. 2024, doi: https://doi.org/10.9734/ajrcos/2024/v17i5441

[19] S. Chesterman, “Through a Glass, Darkly: Artificial Intelligence and the Problem of Opacity,” *The American Journal of Comparative Law*, vol. 69, no. 2, pp. 271–294, Apr. 2021, doi: https://doi.org/10.1093/ajcl/avab012

[20] S. O. Olabanji, Y. A. Marquis, C. S. Adigwe, A. S. Abidemi, T. O. Oladoyinbo, and O. O. Olaniyi, “AI-Driven Cloud Security: Examining the Impact of User Behavior Analysis on Threat Detection,” *Asian Journal of Research in Computer Science*, vol. 17, no. 3, pp. 57–74, Jan. 2024, doi: https://doi.org/10.9734/ajrcos/2024/v17i3424

[21] K. Sokol and P. A. Flach, “Explainability Is in the Mind of the Beholder: Establishing the Foundations of Explainable Artificial Intelligence,” Dec. 2021, doi: https://doi.org/10.48550/arxiv.2112.14466

[22] W. Ding, M. Abdel-Basset, H. Hawash, and A. M. Ali, “Explainability of Artificial Intelligence Methods, Applications and Challenges: A Comprehensive Survey,” *Information Sciences*, Oct. 2022, doi: https://doi.org/10.1016/j.ins.2022.10.013

[23] Y. A. Marquis, T. O. Oladoyinbo, S. O. Olabanji, O. O. Olaniyi, and S. S. Ajayi, “Proliferation of AI Tools: A Multifaceted Evaluation of User Perceptions and Emerging Trend,” *Asian Journal of Advanced Research and Reports*, vol. 18, no. 1, pp. 30–35, Jan. 2024, doi: https://doi.org/10.9734/ajarr/2024/v18i1596

[24] S. Caldwell *et al.*, “An Agile New Research Framework for Hybrid Human-AI Teaming: Trust, Transparency, and Transferability,” *ACM Transactions on Interactive Intelligent Systems*, vol. 12, no. 3, pp. 1–36, Sep. 2022, doi: https://doi.org/10.1145/3514257

[25] V. L. Lemieux, C. Rowell, M.-D. L. Seidel, and C. C. Woo, “Caught in the middle?,” *Records Management Journal*, vol. 30, no. 3, pp. 301–324, Apr. 2020, doi: https://doi.org/10.1108/rmj-09-2019-0048

[26] O. O. Olaniyi and D. S. Omubo, “The Importance of COSO Framework Compliance in Information Technology Auditing and Enterprise Resource Management,”

International journal of innovative research and development, Jun. 2023, doi: https://doi.org/10.24940/ijird/2023/v12/i5/may23001

[27] B. Alamri, K. Crowley, and I. Richardson, "Cybersecurity Risk Management Framework for Blockchain Identity Management Systems in Health IoT," *Sensors*, vol. 23, no. 1, p. 218, Dec. 2022, doi: https://doi.org/10.3390/s23010218

[28] S. Bakare, N. Adekunle, C. U. Akpuokwe, and N. E. Eneh, "DATA PRIVACY LAWS AND COMPLIANCE: A COMPARATIVE REVIEW OF THE EU GDPR AND USA REGULATIONS," *Computer science & IT research journal*, vol. 5, no. 3, pp. 528–543, Mar. 2024, doi: https://doi.org/10.51594/csitrj.v5i3.859

[29] M. Parker, "Chapter 7 - Managing threats to health data and information: toward security," *ScienceDirect*, Jan. 01, 2023. https://www.sciencedirect.com/science/article/pii/B9780323908023000162

[30] F. A. Ezeugwa, O. O. Olaniyi, J. C. Ugonnia, A. S. Arigbabu, and P. C. Joeaneke, "Artificial Intelligence, Big Data, and Cloud Infrastructures: Policy Recommendations for Enhancing Women's Participation in the Tech-Driven Economy," *Journal of Engineering Research and Reports*, vol. 26, no. 6, pp. 1–16, May 2024, doi: https://doi.org/10.9734/jerr/2024/v26i61158

[31] O. O. Olaniyi, O. O. Omogoroye, F. G. Olaniyi, A. I. Alao, and T. O. Oladoyinbo, "CyberFusion Protocols: Strategic Integration of Enterprise Risk Management, ISO 27001, and Mobile Forensics for Advanced Digital Security in the Modern Business Ecosystem," *Journal of Engineering Research and Reports*, vol. 26, no. 6, p. 32, 2024, doi: https://doi.org/10.9734/JERR/2024/v26i61160

[32] V. Bandari, "Enterprise Data Security Measures: A Comparative Review of Effectiveness and Risks Across Different Industries and Organization Types," *International Journal of Business Intelligence and Big Data Analytics*, vol. 6, no. 1, pp. 1–11, Jan. 2023, Available: https://research.tensorgate.org/index.php/IJBIBDA/article/view/3

[33] L. Thomas, I. Gondal, T. Oseni, and S. Sally) Firmin, "A framework for data privacy and security accountability in data breach communications," *Computers & Security*, vol. 116, p. 102657, Feb. 2022, doi: https://doi.org/10.1016/j.cose.2022.102657

[34] T. R. Reshmi, "Information security breaches due to ransomware attacks - a systematic literature review," *International Journal of Information Management Data Insights*, vol. 1, no. 2, p. 100013, Nov. 2021, doi: https://doi.org/10.1016/j.jjimei.2021.100013

[35] U. T. I. Igwenagu, A. A. Salami, A. S. Arigbabu, C. E. Mesode, T. O. Oladoyinbo, and O. O. Olaniyi, "Securing the Digital Frontier: Strategies for Cloud Computing Security, Database Protection, and Comprehensive Penetration Testing," *Journal of Engineering Research and Reports*, vol. 26, no. 6, pp. 60–75, May 2024, doi: https://doi.org/10.9734/jerr/2024/v26i61162

[36] R. Isley, "Algorithmic Bias and Its Implications: How to Maintain Ethics through AI Governance," *Volume 2, Online*, vol. 2, no. 1, Oct. 2022, doi: https://doi.org/10.21428/4b58ebd1.0e834dbb

[37] M. Ryan, "In AI We Trust: Ethics, Artificial Intelligence, and Reliability," *Science and Engineering Ethics*, vol. 26, no. 5, Jun. 2020, doi: https://doi.org/10.1007/s11948-020-00228-y

[38] H. Felzmann, E. F. Villaronga, C. Lutz, and A. Tamò-Larrieux, "Transparency you can trust: Transparency requirements for artificial intelligence between legal norms and contextual concerns," *Big Data & Society*, vol. 6, no. 1, pp. 1–14, Jan. 2019, doi: https://doi.org/10.1177/2053951719860542

[39] A. A. Salami, U. T. I. Igwenagu, C. E. Mesode, O. O. Olaniyi, and O. B. Oladoyinbo, "Beyond Conventional Threat Defense: Implementing Advanced Threat Modeling Techniques, Risk Modeling Frameworks and Contingency Planning in the Healthcare Sector for Enhanced Data Security," *Journal of Engineering Research and Reports*, vol. 26, no. 5, pp. 304–323, Apr. 2024, doi: https://doi.org/10.9734/jerr/2024/v26i51156

[40] S. R. Konda, "Ensuring Trust and Security in AI: Challenges and Solutions for Safe Integration," *INTERNATIONAL JOURNAL OF COMPUTER SCIENCE AND TECHNOLOGY*, vol. 3, no. 2, pp. 71–86, Jun. 2019, Available: http://ijcst.com.pk/index.php/IJCST/article/view/388

[41] P. Schmidt, F. Biessmann, and T. Teubner, "Transparency and trust in artificial intelligence systems," *Journal of Decision Systems*, vol. 29, no. 4, pp. 260–278, Sep. 2020, doi: https://doi.org/10.1080/12460125.2020.1819094

[42] B. Nolan, "Elon Musk says OpenAI's GPT-4o reveal 'made me cringe,'" *Business Insider*, May 14, 2024. https://www.businessinsider.com/elon-musk-openai-chatgpt-flirty-cringe-update-gpt4-2024-5 (accessed Jun. 02, 2024)

[43] J. C. Ugonnia, O. O. Olaniyi, F. G. Olaniyi, A. A. Arigbabu, and T. O. Oladoyinbo, "Towards Sustainable IT Infrastructure: Integrating Green Computing with Data Warehouse and Big Data Technologies to Enhance Efficiency and Environmental

Responsibility," *Journal of Engineering Research and Reports*, vol. 26, no. 5, pp. 247–261, Apr. 2024, doi: https://doi.org/10.9734/jerr/2024/v26i51151

[44] S. O. Olabanji, "AI for Identity and Access Management (IAM) in the Cloud: Exploring the Potential of Artificial Intelligence to Improve User Authentication, Authorization, and Access Control within Cloud-Based Systems," *Asian Journal of Research in Computer Science*, vol. 17, no. 3, pp. 38–56, 2024, doi: https://doi.org/10.9734/ajrcos/2024/v17i3423

[45] Viswan Vimbi, Noushath Shaffi, and M. Mahmud, "Interpreting artificial intelligence models: a systematic review on the application of LIME and SHAP in Alzheimer's disease detection," *Brain informatics*, vol. 11, no. 1, Apr. 2024, doi: https://doi.org/10.1186/s40708-024-00222-1

[46] N. J. King and V. T. Raja, "Protecting the privacy and security of sensitive customer data in the cloud," *Computer Law & Security Review*, vol. 28, no. 3, pp. 308–319, Jun. 2012, doi: https://doi.org/10.1016/j.clsr.2012.03.003

[47] O. O. Olaniyi, O. J. Okunleye, S. O. Olabanji, C. U. Asonze, and S. A. Ajayi, "IoT Security in the Era of Ubiquitous Computing: A Multidisciplinary Approach to Addressing Vulnerabilities and Promoting Resilience," *Asian Journal of Research in Computer Science*, vol. 16, no. 4, pp. 354–371, Dec. 2023, doi: https://doi.org/10.9734/ajrcos/2023/v16i4397

[48] Y. Li, A. Yazdanmehr, J. Wang, and H. R. Rao, "Responding to identity theft: A victimization perspective," *Decision Support Systems*, vol. 121, pp. 13–24, Jun. 2019, doi: https://doi.org/10.1016/j.dss.2019.04.002

[49] O. O. Olaniyi, O. J. Okunleye, and S. O. Olabanji, "Advancing Data-Driven Decision-Making in Smart Cities through Big Data Analytics: A Comprehensive Review of Existing Literature," *Current Journal of Applied Science and Technology*, vol. 42, no. 25, pp. 10–18, Aug. 2023, doi: https://doi.org/10.9734/cjast/2023/v42i254181

[50] B. Méndez Egea, "Data's use and abuse in political communication. Cambridge analytica and the brexit campaign, a case study," *repositorio.comillas.edu*, 2021, Available: https://repositorio.comillas.edu/xmlui/handle/11531/46838

[51] R. Kitchin, "Getting smarter about smart cities: Improving data privacy and data security," Jan. 2016. Available: https://mural.maynoothuniversity.ie/7242/1/Smart

[52] O.L. Olorunfemi, O.O. Amoo, A. Atadoga, O.A Fayayola, O. Abrahams, and P. Olaseni, "TOWARDS A CONCEPTUAL FRAMEWORK FOR ETHICAL AI

DEVELOPMENT IN IT SYSTEMS," *Computer science & IT research journal*, vol. 5, no. 3, pp. 616–627, Mar. 2024, doi: https://doi.org/10.51594/csitrj.v5i3.910

[53] N. R. Mayeke, A. T. Arigbabu, O. O. Olaniyi, O. J. Okunleye, and C. S. Adigwe, "Evolving Access Control Paradigms: A Comprehensive Multi-Dimensional Analysis of Security Risks and System Assurance in Cyber Engineering. ," vol. 17, no. 5, pp. 108–124, 2024, doi: https://doi.org/10.9734/ajrcos/2024/v17i5442

[54] F. Cabitza, A. Campagner, and C. Balsano, "Bridging the 'last mile' gap between AI implementation and operation: 'data awareness' that matters," *Annals of Translational Medicine*, vol. 8, no. 7, Apr. 2020, doi: https://doi.org/10.21037/atm.2020.03.63

[55] S.-C. Kong, W. M.-Y. Cheung, and G. Zhang, "Evaluating artificial intelligence literacy courses for fostering conceptual learning, literacy and empowerment in university students: Refocusing to conceptual building," *Computers in Human Behavior Reports*, vol. 7, p. 100223, Aug. 2022, doi: https://doi.org/10.1016/j.chbr.2022.100223

[56] G. Edison, "The Integration of AI in the Doctor's Toolkit: Enhancing Medical Decision-making," *BULLET : Jurnal Multidisiplin Ilmu*, vol. 2, no. 3, pp. 604–613, Jun. 2023, Available: https://www.journal.mediapublikasi.id/index.php/bullet/article/view/2898

[57] V. Willis, "The Role of Artificial Intelligence (AI) in Personalizing Online Learning," *Journal of Online and Distance Learning*, vol. 3, no. 1, pp. 1–13, 2023, doi: https://doi.org/10.47941/jodl.1689

[58] A. D. Samuel-Okon and O. O. Abejide, "Bridging the Digital Divide: Exploring the Role of Artificial Intelligence and Automation in Enhancing Connectivity in Developing Nations," *Journal of Engineering Research and Reports*, vol. 26, no. 6, pp. 165–177, May 2024, doi: https://doi.org/10.9734/jerr/2024/v26i61170

[59] O. O. Olaniyi, C. U. Asonze, S. A. Ajayi, S. O. Olabanji, and C. S. Adigwe, "A Regressional Study on the Impact of Organizational Security Culture and Transformational Leadership on Social Engineering Awareness among Bank Employees: The Interplay of Security Education and Behavioral Change," *Asian Journal of Economics, Business and Accounting*, vol. 23, no. 23, pp. 128–143, Dec. 2023, doi: https://doi.org/10.9734/ajeba/2023/v23i231176

[60] F. Pedro, M. Subosa, A. Rivas, and P. Valverde, "Artificial Intelligence in Education : Challenges and Opportunities for Sustainable Development," *MINISTERIO DE EDUCACIÓN*, 2019

[61] A. Bhattacharya, *Applied Machine Learning Explainability Techniques: Make ML models explainable and trustworthy for practical applications using LIME, SHAP, and*

more. Packt Publishing Ltd, 2022. Accessed: Jun. 02, 2024. [Online]. Available: https://books.google.com/books?hl=en&lr=&id=KaI3EAAAQBAJ&oi=fnd&pg=PP1&dq=Microsoft+and+Google+have+integrated++LIME+and+SHAP+into+their+AI+frameworks

[62] Y. Maleh, A. Sahid, M. Alazab, and M. Belaissaoui, "IT Governance and Information Security," Nov. 2021, doi: https://doi.org/10.1201/9781003161998

[63] O. O. Olaniyi, S. O. Olabanji, and O. J. Okunleye, "Exploring the Landscape of Decentralized Autonomous Organizations: A Comprehensive Review of Blockchain Initiatives," *Journal of Scientific Research and Reports*, vol. 29, no. 9, pp. 73–81, Sep. 2023, doi: https://doi.org/10.9734/jsrr/2023/v29i91786

[64] R. F. Smallwood, *Information Governance: Concepts, Strategies and Best Practices*. John Wiley & Sons, 2019. Accessed: Jun. 02, 2024. [Online]. Available: https://books.google.com/books?hl=en&lr=&id=z3nADwAAQBAJ&oi=fnd&pg=PT11&dq=ISO/IEC+27001+certification+has+become+a+hallmark+of+trust

[65] C. S. Adigwe, N. R. Mayeke, S. O. Olabanji, O. J. Okunleye, P. C. Joeaneke, and O. O. Olaniyi, "The Evolution of Terrorism in the Digital Age: Investigating the Adaptation of Terrorist Groups to Cyber Technologies for Recruitment, Propaganda, and Cyberattacks," *Asian journal of economics, business and accounting*, vol. 24, no. 3, pp. 289–306, Feb. 2024, doi: https://doi.org/10.9734/ajeba/2024/v24i31287

[66] O. O. Olaniyi, F. A. Ezeugwa, C. G. Okatta, A. S. Arigbabu, and P. C. Joeaneke, "Dynamics of the Digital Workforce: Assessing the Interplay and Impact of AI, Automation, and Employment Policies," *Archives of current research international*, vol. 24, no. 5, pp. 124–139, Apr. 2024, doi: https://doi.org/10.9734/acri/2024/v24i5690

[67] P. John, "Exploratory Analysis of Artificial Intelligence (AI) Impact and Opportunities for Financial Services Compliance - ProQuest," *www.proquest.com*, 2022. https://search.proquest.com/openview/732d40ec2eaf18a056582dfc46c9efab/1?pq-origsite=gscholar&cbl=18750&diss=y

[68] S. Mohseni, N. Zarei, and E. D. Ragan, "A Multidisciplinary Survey and Framework for Design and Evaluation of Explainable AI Systems," *ACM Transactions on Interactive Intelligent Systems*, vol. 11, no. 3–4, pp. 1–45, Dec. 2021, doi: https://doi.org/10.1145/3387166

[69] Liaison, "The Importance of Transparency in Education When Adopting AI," *Liaison*, Feb. 29, 2024. https://www.liaisonedu.com/the-importance-of-transparency-in-education-when-adopting-ai/

[70] Y. Alufaisan, L. R. Marusich, J. Z. Bakdash, Y. Zhou, and M. Kantarcioglu, “Does Explainable Artificial Intelligence Improve Human Decision-Making?,” *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 35, no. 8, pp. 6618–6626, May 2021, doi: https://doi.org/10.1609/aaai.v35i8.16819

[71] D. Shin, “The effects of explainability and causability on perception, trust, and acceptance: Implications for explainable AI,” *International Journal of Human-Computer Studies*, vol. 146, p. 102551, Feb. 2021, doi: https://doi.org/10.1016/j.ijhcs.2020.102551

[72] O. Vereschak, G. Bailly, and B. Caramiaux, “How to Evaluate Trust in AI-Assisted Decision Making? A Survey of Empirical Methodologies,” *Proceedings of the ACM on Human-Computer Interaction*, vol. 5, no. CSCW2, pp. 1–39, Oct. 2021, doi: https://doi.org/10.1145/3476068

[73] A. Fejza, P. Genevès, N. Layaïda, and J.-L. Bosson, “Scalable and interpretable predictive models for electronic health records,” 2018, pp. 341–350. doi: https://doi.org/10.1109/DSAA.2018.00045

[74] A. Adadi and M. Berrada, “Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI),” *IEEE Access*, vol. 6, pp. 52138–52160, 2018, doi: https://doi.org/10.1109/access.2018.2870052

[75] D. Shin, “User Perceptions of Algorithmic Decisions in the Personalized AI System:Perceptual Evaluation of Fairness, Accountability, Transparency, and Explainability,” *Journal of Broadcasting & Electronic Media*, vol. 64, no. 4, pp. 1–25, Dec. 2020, doi: https://doi.org/10.1080/08838151.2020.1843357

[76] T. A. Bach, A. Khan, H. Hallock, G. Beltrão, and S. Sousa, “A Systematic Literature Review of User Trust in AI-Enabled Systems: An HCI Perspective,” *International Journal of Human–Computer Interaction*, vol. 40, no. 5, pp. 1–16, Nov. 2022, doi: https://doi.org/10.1080/10447318.2022.2138826

[77] A. Singhal, N. Neveditsin, H. Tanveer, and V. Mago, “Toward Fairness, Accountability, Transparency, and Ethics in AI for Social Media and Health Care: Scoping Review,” *JMIR Medical Informatics*, vol. 12, no. 1, p. e50048, Apr. 2024, doi: https://doi.org/10.2196/50048

[78] O. Akinrinola, C. C. Okoye, O. C. Ofodile, and C. E. Ugochukwu, “Navigating and reviewing ethical dilemmas in AI development: Strategies for transparency, fairness, and accountability,” *GSC Advanced Research and Reviews*, vol. 18, no. 3, pp. 050–058, Mar. 2024, doi: https://doi.org/10.30574/gscarr.2024.18.3.0088

[79] N. Palladino, “A ‘biased’ emerging governance regime for artificial intelligence? How AI ethics get skewed moving from principles to practices,” *Telecommunications Policy*, vol. 47, no. 5, p. 102479, Dec. 2022, doi: https://doi.org/10.1016/j.telpol.2022.102479

[80] I. Georgieva, C. Lazo, T. Timan, and A. F. van Veenstra, “From AI ethics principles to data science practice: a reflection and a gap analysis based on recent frameworks and practical experience,” *AI and Ethics*, Jan. 2022, doi: https://doi.org/10.1007/s43681-021-00127-3

[81] W. Saeed and C. Omlin, “Explainable AI (XAI): A systematic meta-survey of current challenges and future opportunities,” *Knowledge-Based Systems*, vol. 263, p. 110273, Jan. 2023, doi: https://doi.org/10.1016/j.knosys.2023.110273

[82] A. B. Arrieta *et al.*, “Explainable Artificial Intelligence (XAI): Concepts, Taxonomies, Opportunities and Challenges toward Responsible AI,” *ArXiv*, 2019, Available: https://www.semanticscholar.org/paper/Explainable-Artificial-Intelligence-(XAI)%3A-and-AI-Arrieta-Rodr%C3%ADguez/530a059cb48477ad1e3d4f8f4b153274c8997332

[83] S. Larsson and F. Heintz, “Transparency in artificial intelligence,” *Internet Policy Review*, vol. 9, no. 2, May 2020, Available: https://policyreview.info/concepts/transparency-artificial-intelligence

[84] Google, “Google AI,” *Google AI*, 2024. https://ai.google/

[85] Microsoft, “Microsoft Research,” *Microsoft Research*, Dec. 05, 2018. https://www.microsoft.com/en-us/research

[86] “Home | The Alan Turing Institute,” *Turing.ac.uk*, 2020. https://www.turing.ac.uk/

[87] “OpenAI Research,” *OpenAI*, Jul. 30, 2021. https://openai.com/research/

[88]“IBM Research Publications | IBM Research,” *IBM Research Publications*, Feb. 09, 2021. https://research.ibm.com/publications

Appendices

Demographic Category	**N**	**Percentage (%)**
Role in Organization		
Executive	41	12
Data Governance Officer	72	21.1
AI Developer	96	28.1
IT/Security Personnel	53	15.5
Compliance Officer	63	18.4
Other	17	5
Organization's Industry		
Financial Services	102	29.8
Healthcare	61	17.8
Technology	88	25.7
Manufacturing	39	11.4
Government	28	8.2
Other	24	7
Organization Size		
Small (1-100 employees)	132	38.6
Medium (101-500 employees)	126	36.8
Large (>500 employees)	84	24.6
Years of Experience		
Less than 1 year	29	8.5
1-5 years	98	28.7
6-10 years	107	31.3
More than 10 years	108	31.6
Age Group		
Under 25	44	12.9
25-34	106	31
35-44	89	26
45-54	66	19.3
55-64	25	7.3
65 or older	12	3.5
Gender		
Male	172	50.3
Female	160	46.8

Non-binary/Third gender	5	1.5
Prefer not to say	5	1.5
Other	0	0

Table 1: Literature review extraction of data

Hypothesis	**Literature Source**	**Theme**	**Qualitative Data Extracted**	**Content Analysis**	**Lessons Learned**
H1: XAI Increases User Trust (vs. Black-Box)	Adadi & Berrada (2018)	Transparency & Trust	XAI enhances transparency, leading to increased user trust.	Various XAI methods and their benefits are highlighted.	Explainability improves transparency, building user trust. Common challenge: technical complexity.
	Alufaisan et al. (2021)	Decision-Making & Trust	Clear AI explanations improve decision accuracy and user trust.	Empirical data showing improved decision-making accuracy with XAI.	Clear explanations correlate with higher trust and better decisions. Pattern: Need for user-centric explanations.
	Bach et al. (2022)	Factors Influencing Trust	Transparency and user interface design are critical for building trust.	Review of factors affecting trust in AI systems, emphasizing transparency.	Design and transparency are key to trust. Lesson: Prioritize user interface design.
	Markus et al. (2021)	Explainability in Healthcare	Explainability in healthcare AI enhances trustworthiness.	Detailed survey of explainability in healthcare AI and its	Trust in healthcare AI grows with explainability. Lesson: Invest in

				impact on trust.	explainability features.
	Shin (2020)	Perception & Trust	User perceptions of transparency and explainability affect trust.	Analysis of user surveys showing perception's impact on trust.	Positive perceptions of transparency boost trust. Pattern: Perception management is crucial.
	Shin (2021)	Explainability & Trust	Education on AI explainability improves trust and acceptance.	Survey results on the effects of explainability on trust.	Explainability education enhances trust. Lesson: Implement educational programs.
H2: XAI and Ethical Data Use	Arrieta et al. (2019)	Ethical AI Practices	Taxonomy of XAI methods facilitates ethical AI practices.	Overview of ethical considerations in XAI adoption.	Ethical use of data linked to XAI adoption. Lesson: Develop ethical guidelines for XAI.
	Bell et al. (2023)	Algorithmic Transparency	Framework for transparency aids ethical data use.	Analysis of regulatory frameworks supporting ethical data use.	Transparency frameworks support ethical data practices. Pattern: Regulatory compliance enhances trust.
	Georgieva et al. (2022)	Ethics & Practice Gap	Bridging the gap between ethics principles	Gap analysis between ethical AI	Closing the ethics-practice gap is essential.

			and practice enhances ethical data use.	principles and practice.	Lesson: Implement practical ethics training.
	Palladino (2022)	Governance & Ethics	Governance regimes help align AI practices with ethical standards.	Review of governance practices and their impact on ethical AI use.	Strong governance aligns practices with ethics. Pattern: Governance frameworks are effective.
H3: Data Governance and Breaches	Chon & Alexander (2023)	Data Governance	Best practices in data governance ensure data integrity and compliance.	Case studies on data governance in clinical trials.	Effective governance reduces breaches. Lesson: Adopt best practices in data governance.
	Fejza et al. (2018)	Interpretable Models	Interpretable models improve transparency and data security.	Discussion on scalable and interpretable predictive models.	Interpretability aids transparency and security. Pattern: Scalability is a common challenge.
	Bell et al. (2023)	Regulatory Compliance	Compliance strategies for transparency reduce data breaches.	Analysis of transparency frameworks and compliance strategies.	Compliance strategies are key to reducing breaches. Lesson: Develop robust

					compliance protocols.
	Vereschak et al. (2021)	Trust Evaluation	Methodologies for evaluating trust in AI-assisted decision making.	Survey of empirical methodologies.	Evaluation methodologies help understand trust. Lesson: Implement diverse evaluation methods.
H4: User Education and Trust	Liaison (2024)	Education & Transparency	Transparency in educational AI systems builds trust.	Reports on the role of transparency in educational AI.	Education on transparency builds trust. Pattern: Need for continuous education efforts.
	Shin (2021)	Explainability & Trust	Education on AI explainability improves trust and acceptance.	Survey results on the effects of explainability on trust.	Explainability education enhances trust. Lesson: Implement educational programs.
	Olatunji Akinrinola et al. (2024)	Ethical Dilemmas & Education	Strategies for transparency and fairness in AI development enhance trust.	Case studies on ethical dilemmas in AI and educational strategies.	Transparency and fairness strategies boost trust. Pattern: Ethical dilemmas need proactive education.
	Singhal et al. (2024)	AI in Social Media & Healthcare	Education on AI fairness, accountabilit	Comprehensive review of AI	Education on fairness and transparency

			y, and transparency in social media and healthcare builds trust.	transparency and educational initiatives.	is crucial. Lesson: Promote educational initiatives across sectors.
	Ehsan et al. (2019)	Human-AI Interaction	Strategies for user education on AI transparency.	Analysis of educational programs and their effectiveness.	Human-AI interaction improves with education. Lesson: Develop comprehensive education programs.
	Mohseni et al. (2021)	Design & Evaluation	Survey of XAI design and evaluation methods.	Multidisciplinary perspectives on XAI.	Diverse design and evaluation methods enhance trust. Lesson: Adopt multidisciplinary approaches.
General: Implementation and Impact	Saeed & Omlin (2023)	XAI Challenges	Current challenges and future opportunities in XAI.	Systematic meta-survey of methodologies and applications.	Addressing XAI challenges improves implementation. Lesson: Focus on overcoming technical barriers.
	Teixeira et al. (2022)	AI Governance	AI risks and governance strategies.	Exploratory insights from international	Responsible governance mitigates AI

				governance conferences.	risks. Lesson: Implement comprehensive governance strategies.
	Umbrello & Yampolskiy (2021)	Explainability & Verifiability	Designing AI for explainability and verifiability.	Value-sensitive design approach.	Value-sensitive design enhances trust. Lesson: Incorporate ethical values in design.
	Olatunji Akinrinola et al. (2024)	Ethical Dilemmas & Education	Strategies for transparency and fairness in AI development enhance trust.	Case studies on ethical dilemmas in AI and educational strategies.	Transparency and fairness strategies boost trust. Pattern: Ethical dilemmas need proactive education.
	Rubén González-Sendino et al. (2024)	Fair Data Generation	Methods for fair data generation and bias mitigation.	Impact on transparency and explainability in AI decision-making.	Fair data generation reduces bias. Lesson: Use causal models for fairness.
	Larsson & Heintz (2020)	AI Transparency	Aspects of AI transparency impacting trust and ethical use.	Review of transparency practices.	Transparency practices improve ethical use. Lesson: Promote

					transparent AI practices.
	Floridi et al. (2019)	AI Ethics	Ethical implications of AI: accountability, responsibility, and transparency.	Analysis of AI ethics frameworks.	Ethical AI frameworks enhance trust. Lesson: Implement robust ethics frameworks.
	Zwitter & Gstrein (2020)	Data Governance	Principles, policies, and practices for data governance.	Comparative analysis of governance frameworks.	Effective data governance principles are critical. Lesson: Standardize governance practices.

Table 2: Industrial Based Data

Organization	**XAI Methods & Tools**	**Trust Impact Metrics**	**Ethical Data Use & Governance**	**Data Breach Incidents (Publicly Reported)**	**User Education & Resources**	**Lessons Learned**
Google	Integrated Gradients, What-If Tool, TensorFlow Explainability	User surveys, A/B testing	AI Principles, internal data governance policies	Yes, but with significant security investments	Courses, publications, Explainable AI resources	Importance of robust security and user-centric design in enhancing trust.
IBM	AI Explainability 360 toolkit	User feedback, case studies	Trust and Transparency Principles for AI, robust governance practices	Yes, but with ongoing security improvements	Training, workshops, AI Explainability 360 resources	Continuous improvement in security and transparency boosts user trust and ethical use.
OpenAI	Research on XAI for NLP models	Qualitative feedback, user studies	Guidelines for ethical AI research, internal data policies	None publicly reported	Research papers, blog posts, educational materials	Ethical guidelines and transparent research practices are crucial for maintaining trust.
Microsoft Research	InterpretML, Fairlearn	User surveys, feedback on model	Principles for ethical AI, strong governance practices	Yes, but with enhanced security	Training, workshops, resources on AI fairness and interpretability	User feedback and strong governance practices are essential for

		explanations		measures		maintaining trust and security.
The Alan Turing Institute	Research on XAI for various domains	Case studies, impact assessments of XAI	Guidelines for ethical AI research, internal data policies	None publicly reported	Courses, workshops, public engagement events	Cross-sector impact assessments and public engagement are vital for effective XAI implementation.

Table 3: Rating of the Data using Coding

Organization	XAI Methods & Tools	Trust Impact Metrics	Ethical Data Use & Governance	Data Breach Incidents (Publicly Reported)	User Education & Resources
Google	5	4	5	2	5
IBM	5	4	5	2	4
OpenAI	4	3	4	1	3
Microsoft Research	4	4	4	2	4
The Alan Turing Institute	4	3	4	1	4