Journal of Engineering Research and Reports

Volume 26, Issue 9, Page 136-158, 2024; Article no.JERR.122737
ISSN: 2582-2926

Incorporating Privacy by Design Principles in the Modification of AI Systems in Preventing Breaches across Multiple Environments, Including Public Cloud, Private Cloud, and On-prem

Samuel Ufom Okon [a++*], Omobolaji Olufunmilayo Olateju [b#], Olumide Samuel Ogungbemi [c†], Sunday Abayomi Joseph [d], Anthony Obulor Olisa [e‡] and Oluwaseun Oladeji Olaniyi [f‡]

[a] FirstBank DR Congo, Gombe, Democratic Republic of the Congo.
[b] University of Ibadan, Oduduwa Road, Ibadan, Oyo State, Nigeria.
[c] Centennial College, 941 Progress Ave, Scarborough, ON M1G 3T8, Canada.
[d] Data Privacy, Blockchain Strategy & Management, Ashland University, 401 College Avenue, Ashland, OH 44805, United States of America.
[e] Cumberland University, 1 Cumberland Dr, Lebanon, TN 37087, United States.
[f] University of the Cumberlands, 104 Maple Drive, Williamsburg, KY 40769, United States of America.

Authors' contributions

This work was carried out in collaboration among all authors. All authors read and approved the final manuscript.

Article Information

DOI: https://doi.org/10.9734/jerr/2024/v26i91269

Open Peer Review History:
This journal follows the Advanced Open Peer Review policy. Identity of the Reviewers, Editor(s) and additional Reviewers, peer review comments, different versions of the manuscript, comments of the editors, etc are available here: https://www.sdiarticle5.com/review-history/122737

[++] Finance and Technology Researcher;
[#] Agricultural Technology Researcher;
[†] Data Privacy and Blockchain Technology Researcher;
[‡] Information Technology Researcher;
**Corresponding author: Email: samokon2000@gmail.com;*

Cite as: *Okon, Samuel Ufom, Omobolaji Olufunmilayo Olateju, Olumide Samuel Ogungbemi, Sunday Abayomi Joseph, Anthony Obulor Olisa, and Oluwaseun Oladeji Olaniyi. 2024. "Incorporating Privacy by Design Principles in the Modification of AI Systems in Preventing Breaches across Multiple Environments, Including Public Cloud, Private Cloud, and On-Prem". Journal of Engineering Research and Reports 26 (9):136-58. https://doi.org/10.9734/jerr/2024/v26i91269.*

Original Research Article

Received: 29/06/2024
Accepted: 31/08/2024
Published: 03/09/2024

ABSTRACT

The rapid integration of artificial intelligence (AI) across various sectors has significantly amplified privacy concerns, particularly with the growing reliance on cloud environments. Existing methods often fall short of effectively preventing privacy breaches due to inadequate risk assessment and mitigation strategies. These limitations highlight the necessity for more robust solutions, indicating the importance of Privacy by Design (PbD) principles. This study addresses these gaps by proposing a comprehensive approach to incorporating PbD principles into AI systems to prevent breaches across public, private, and on-prem environments. The proposed work utilizes logistic regression analysis to identify significant predictors of privacy breaches, revealing that both the environment ($B = -1.142$, $p < .001$) and severity of vulnerabilities ($B = 0.932$, $p < .01$) play crucial roles. Additionally, a strong positive correlation ($r = 0.791$) between breach detection rates and PbD effectiveness is observed, indicating the need for enhanced detection mechanisms. To support the empirical findings, this study also reviews existing case studies. It conducts a thematic analysis to provide a deeper understanding of the practical challenges and solutions associated with PbD implementation, particularly in healthcare and smart city applications. These analyses serve to supplement the empirical evidence and demonstrate the effectiveness of PbD over other existing methods. The study concludes that implementing PbD principles is critical for achieving robust privacy protection, and the study recommends prioritizing advanced breach detection mechanisms, comprehensive privacy impact assessments, continuous stakeholder engagement, and investment in privacy-enhancing technologies to address privacy risks effectively.

Keywords: Privacy by design; AI systems; privacy breaches; breach detection; privacy-enhancing technologies.

1. INTRODUCTION

The advent of artificial intelligence (AI) has ushered in a new era of technological advancement, revolutionizing industries and transforming societies. The increasing reliance on digital technologies and the exponential growth of data has created an avenue for significant implications for privacy, and the convergence of AI and data-driven systems has worsened privacy concerns, demanding innovative approaches to safeguard sensitive information. Cloud computing, which has emerged as a cornerstone of modern digital infrastructure, offers scalable, flexible, and cost-effective computing resources; with this offering, the share of corporate data stored in the cloud has surged from 30% in 2015 to 60% in 2022 [1]. This surge in growth reflects the trust organizations place in cloud technologies to store critical information while ensuring accessibility securely.

Despite cloud storage solutions being advantageous in providing solid data backup and recovery capabilities, which control the risk of data loss due to hardware failures or other unforeseen events, security remains a top concern for 83% of organizations. With 92% of organizations hosting some portion of their IT environment in the cloud, robust security measures are critical to protect sensitive information [1]. According to Netgate [1], it was observed that untrained workers are responsible for 88% of cloud breaches as their actions leave the organization susceptible to vulnerable attacks from hackers, and comprehensive training and awareness programs will help mitigate these risks. Phishing attacks account for approximately 25% of all data breaches in the cloud environment, and compromised privileged accounts contribute to 34% of identity-related violations, emphasizing the need for solid access management practices. Despite the growing importance of encryption, only 21% of organizations have encrypted more than 60% of their data in the cloud, indicating substantial room for improvement in securing sensitive information.

Zero-trust security models are gaining popularity as 80% of enterprises are considering or

deploying these models into their operation; zero-trust security models emphasize continuous verification and authorization, thereby improving these companies' security. Financial investments in cloud security are substantial, with audits starting at USD 10,000 per year, highlighting the commitment organizations must make to ensure compliance and risk mitigation. The global zero-trust cloud security market is projected to reach USD 60 billion by 2027, indicating the growing demand for advanced security solutions [1].

According to Confessore [2], the need for more strict privacy measures is underlined by high-profile data breaches such as the Cambridge Analytica scandal. In 2018, it was revealed that Cambridge Analytica had harvested the personal data of millions of Facebook users without their consent, using it for political advertising purposes. This incident highlighted the potential for misuse of personal data and led to widespread public outcry and regulatory supervision; the fallout from this scandal emphasized the need for solid data protection frameworks and the implementation of Privacy by Design (PbD) principles to prevent similar breaches. Incorporating PbD principles into AI systems involves several key strategies, including data minimization, anonymization, pseudonymization, and the use of privacy-enhancing technologies; data minimization ensures that only necessary data is collected and retained, reducing the risk of exposure in the event of a breach [3]. Anonymization and pseudonymization techniques protect individual identities by removing or obfuscating personally identifiable information [4]. Privacy-enhancing technologies, such as homomorphic encryption and differential privacy, provide additional layers of security, enabling the analysis of data without compromising privacy [5].

This study aims to develop a model that will be instrumental in integrating risk assessment and privacy impact assessment into the AI development lifecycle, evaluating potential threats and vulnerabilities at each stage of development and implementing measures to address them, and also measuring the effectiveness of PbD implementations, key performance indicators and metrics to ensure that privacy measures achieve their intended outcomes.

2. LITERATURE REVIEW

Artificial Intelligence (AI) has significantly progressed since its inception, allowing usage by various sectors such as healthcare, finance, and transportation. Initially constrained by rule-based operations, advancements in machine learning and deep learning have enabled AI to perform complex tasks like natural language processing, image recognition, and predictive analytics with high accuracy. According to Amajuoyi et al. [6], the integration of AI with cloud computing has further expanded its usage, allowing for scalable and flexible deployment of AI applications. Public cloud services, such as those provided by Amazon Web Services (AWS) and Microsoft Azure, offer scalable resources accessible to multiple clients, while private clouds, though more costly, give greater control over data and security; for instance, On-premises solutions offer the highest level of control, requiring significant capital investment and maintenance efforts [7,8,9]. This convergence of AI and cloud computing introduces significant privacy challenges due to extensive data collection and processing, as AI systems require large datasets for training, and storing and processing this data in cloud environments heightens the risk of unauthorized access and breaches [10,11]. Incidents like the Cambridge Analytica scandal show how data breaches in AI and cloud environments can lead to severe consequences, including financial losses, reputational damage, and regulatory penalties, prompting increased scrutiny and demand for robust privacy measures [2,12].

In order to manage these challenges, regulatory frameworks such as the General Data Protection Regulation (GDPR) and the California Consumer Privacy Act (CCPA) were established. GDPR mandates organizations to implement measures to protect personal data and ensure its lawful processing by imposing severe penalties for non-compliance. Similarly, the CCPA provides California residents with rights over their personal data, including knowledge of what is collected and the ability to request its deletion [13,14,15]. These regulations ensure that AI and cloud service providers adopt comprehensive privacy strategies, including data minimization and encryption, to comply with legal requirements and protect user privacy [14]. According to Wischmeyer [16], the ethical implications of AI and data usage further complicate the privacy issue, as AI systems often operate as "black boxes," where their decision-making processes are not fully transparent, and this raises concerns about accountability and fairness, particularly in decisions impacting individuals, such as hiring or credit scoring.

Ethical considerations extend to data usage, where a careful balance between innovation and privacy must be maintained, and the misuse of personal data, as seen in the Cambridge Analytica case, highlights the potential for AI to infringe on individual privacy and autonomy [17,18].

Arif [19] proposes that addressing privacy challenges in AI and cloud environments requires a mixed approach; the implementation of Privacy by Design (PbD) principles requires embedding privacy considerations throughout the AI development lifecycle, and techniques such as data anonymization and pseudonymization are essential for protecting individual identities while enabling data utility for AI applications [20,21]. Additionally, privacy-enhancing technologies like homomorphic encryption and differential privacy offer further protection, allowing data to be analyzed without exposing sensitive information [22,23]. Continuous risk assessments and privacy impact assessments (PIAs) are vital for proactively identifying and mitigating privacy risks in AI systems, integrating these assessments into the development processes ensures that potential vulnerabilities are addressed before deployment. A strong model framework for measuring the effectiveness of privacy measures is necessary to adapt to emerging threats and ensure that implemented strategies achieve their intended outcomes [24,25].

Privacy by design (PbD): Privacy by Design (PbD) argues that privacy should be an intrinsic component of systems and processes from the beginning rather than an afterthought. Pagallo [26] highlights that this proactive approach contrasts with the traditional reactive method that is often adopted for data protection. At its core, PbD promotes principles such as proactive measures, privacy as the default setting, and end-to-end security, forming a framework aimed at preventing data breaches and ensuring user trust [27,28]. In information systems and data management, PbD emphasizes integrating privacy controls at every stage of the data lifecycle, as embedding privacy into system architectures and workflows ensures that data collection, storage, processing, and dissemination adhere to strict privacy standards [20]. As a result, Samantha et al. [27] propose that effective implementation of PbD can mitigate risks associated with data breaches and enhance overall data governance; however, translating these principles into practical applications remains challenging, especially in complex and dynamic data environments [29,30].

Saeed et al. [31] affirm that applying PbD to AI systems isn't without unique challenges and opportunities, and this is because AI systems require vast amounts of data to function effectively, raising concerns about data privacy and user consent. Studies argue that PbD in AI necessitates incorporating privacy safeguards into algorithms and models, ensuring data privacy throughout the AI lifecycle [31,32,33,34]. Techniques such as data minimization, anonymization, and pseudonymization are essential to this incorporation; data minimization involves collecting only the necessary data and retaining it only as long as required, reducing exposure risk, while anonymization removes personally identifiable information from datasets, making it difficult to trace data back to individual users, and pseudonymization replaces private identifiers with fictitious names or keys, adding an extra layer of privacy [4,35]. Privacy-enhancing technologies (PETs) like homomorphic encryption and differential privacy are powerful tools for implementing PbD in AI systems. Homomorphic encryption allows computations on encrypted data without decrypting it, maintaining data confidentiality, while differential privacy ensures that the inclusion or exclusion of a single data point does not significantly impact the outcome of an analysis, thereby protecting individual privacy [34]. It should be noted that though these technologies are promising, they are still growing, and so their implementation in large-scale AI systems remains a subject of active research [36,37].

Risk assessment and privacy impact assessments (PIAs) are critical components of the PbD framework; they offer systematic approaches to identify and mitigate privacy risks, and their integration into AI development allows organizations to address potential privacy issues before they escalate proactively [38]. These issues involve evaluating the threats, vulnerabilities, and impacts of privacy issues on user privacy at each stage of AI development, from data collection to deployment [27,39]. Although Raab [38] argues that PIAs should be mandatory for all AI projects to ensure compliance with privacy regulations and to build user trust, Georgiadis and Poels [25] suggest that a standardized framework for conducting PIAs in AI contexts is essential for consistency and effectiveness.

AI and cloud security: The widespread adoption of cloud computing has transformed the information technology space, providing unmatched capacity, flexibility, and cost-efficiency. However, this shift has also introduced significant security challenges. Cloud environments are inherently complex and dynamic, fraught with vulnerabilities that require sophisticated security measures. Scholars argue that the shared responsibility model of cloud security, where providers and clients share duties, often results in security gaps, especially due to the multi-tenant nature of cloud services, which increases risks of data breaches and unauthorized access [40,41,42]. Cloud vulnerabilities, such as data breaches resulting from misconfigured settings or inadequate access controls, can lead to severe financial losses, reputational damage, and regulatory penalties. Aslan et al. [43] observes that human error, particularly in managing cloud configurations, is a primary contributor to these vulnerabilities experienced by companies, and most importantly, phishing attacks and compromised credentials are significant threats, as they allow attackers to access sensitive data. The distributed nature of cloud environments further complicates the prompt detection and response to security incidents [44], and to curb these risks, frameworks like ISO/IEC 27001, Cloud Security Alliance's (CSA), and the Cloud Controls Matrix (CCM) have been developed [45,46]. These frameworks offer guidelines for robust security practices, covering data protection, access control, and incident response; studies contend that adherence to these standards can enhance an organization's security posture, although continuous updates are necessary to address emerging threats and technological advancements [47,48,49].

Rizvi [50] affirms that Artificial Intelligence (AI) has become a vital tool in cybersecurity due to its advanced capabilities for threat detection and prevention; AI-driven solutions utilize machine learning algorithms to analyze large datasets, identify patterns, and detect anomalies indicative of potential threats. This proactive approach allows organizations to respond to security incidents in real-time, reducing vulnerability [51], and AI's adaptability to new threats makes it particularly effective against sophisticated attacks such as zero-day exploits [52,53]. Studies indicate that AI-driven threat detection systems will significantly improve the accuracy and efficiency of identifying security incidents [51,54,55]. These systems employ techniques like behavioral analysis and anomaly detection to recognize deviations from normal activity that may signal malicious behavior, and moreso, AI can automate aspects of threat response, such as isolating affected systems and initiating remediation, thus reducing the burden on security teams and enhancing response times [56,57]. As highlighted by Kaur et al. [58], AI is also crucial in vulnerability assessment, providing continuous evaluation of an organization's security posture; traditional assessments rely on periodic scans, which can leave gaps in coverage, while AI-driven solutions offer real-time monitoring and assessment, identifying vulnerabilities as they emerge and providing actionable insights for mitigation.

Despite the numerous opportunities that AI provides, challenges still persist, one major concern is the potential for AI systems to produce false positives, which can overwhelm security teams and lead to alert fatigue. The effectiveness of AI-driven solutions also depends on the quality and diversity of the data they are trained on, as biases in training data can result in overlooking certain threats [59]. Moreover, the deployment of AI in cybersecurity raises ethical and privacy concerns, particularly regarding the extent to which automated systems should be trusted with decision-making processes [60]. Therefore, emerging trends highlight the integration of AI with other advanced technologies, such as blockchain and quantum computing, to enhance security capabilities [61,62,63]. Researchers are exploring blockchain for secure data sharing and authentication, while quantum computing promises new cryptographic techniques that could transform cloud security [61,62].

These advancements highlight the need for ongoing research to address evolving security challenges in cloud environments, and to effectively help curb the complexity of cyberattacks, the combination of human expertise, solid security frameworks [64], and advanced technologies is required to protect cloud environments from the growing spectrum of threats [47,65], and also the incorporation of PbD principles into AI systems.

PbD in Cloud Environments: The implementation of Privacy by Design (PbD) in cloud environments varies significantly across public, private, and on-premises models, each presenting distinct challenges and opportunities. According to Han et al. [64], Public clouds,

operated by providers such as Amazon Web Services (AWS) and Google Cloud, offer scalable resources shared among multiple clients; this multi-tenancy model requires strict data segregation measures to prevent unauthorized access, making PbD principles essential. However, these measures can introduce latency and complexity, potentially affecting performance and user experience [66]. Private clouds, though dedicated to a single organization, allow for greater control over data and security configurations; this environment is conducive to implementing comprehensive PbD strategies, as organizations can tailor privacy measures to their specific needs without compromising performance. However, Abdulsalam and Hedabou [67] states that private clouds are typically more expensive to maintain and require significant expertise to manage effectively, and the trade-off between enhanced privacy control and higher operational costs is also a key consideration. On-premises cloud environments offer the highest level of control, with infrastructure hosted within the organization's own facilities, on-premises setup allows for the implementation of PbD principles at a granular level, from physical security measures to data encryption and access controls [68], however, on-premises solutions can be resource-intensive and may lack the capacity and flexibility offered by public and private clouds [69,70]. Studies indicate that while on-premises environments provide superior privacy protection, they may struggle to keep pace with rapid technological advancements [71,72,73].

After critical evaluation of these cloud models, Adeusi et al. [74] revealed that public clouds offer cost-efficiency and scalability but pose significant challenges in ensuring data privacy due to their shared nature, private clouds provide a middle ground with greater control and security but at higher costs, and On-premises environments, offering the most control, are often the most resource-intensive. This analysis highlights the importance of selecting a cloud model that aligns with an organization's privacy requirements and operational capabilities. Case studies of successful PbD implementations in cloud environments provide valuable insights. For example, the deployment of differential privacy techniques by Apple in its iOS operating system illustrate how PbD can be effectively integrated into a public cloud setting to enhance user privacy while maintaining data utility. Similarly, Microsoft's Azure platform has incorporated PbD principles through robust encryption practices and compliance with strict data protection regulations, demonstrating successful implementation in a private cloud environment [75,76].

According to Abdulsalam and Hedabou [67], Cloud service providers (CSPs) also play a crucial role in facilitating PbD through their privacy practices and compliance efforts, major CSPs like AWS, Google Cloud, and Microsoft Azure have developed comprehensive privacy frameworks to protect their user data, including encryption, access controls, and compliance with regulations such as GDPR and CCPA. Research indicates that CSPs' commitment to privacy is a significant factor in their clients' ability to implement PbD principles effectively [67,75]. After extensive research, Akremi and Rouached [77] states that CSPs' privacy policies reveal a strong emphasis on compliance and transparency, for example, AWS's privacy policy outlines its commitment to data protection, detailing measures such as encryption, access management, and incident response protocols [47]. Google Cloud's privacy policy similarly emphasizes compliance with global data protection laws, providing detailed information on data processing practices. These policies help ensure compliance and build trust with users by demonstrating a commitment to protecting their privacy.

However, Dias Canedo et al. [78] observe that there are controversies regarding the sufficiency and effectiveness of these privacy practices. Some scholars argue that while CSPs' privacy policies are comprehensive, they may not fully address all potential privacy risks, particularly those arising from complex and evolving threat landscapes [79,80], while some studies are more concerned about the balance between privacy and usability [81,82], with some contending that extremely strict privacy measures can impede functionality and user experience [83,84]. Emerging trends in PbD in cloud environments include the integration of advanced privacy-enhancing technologies (PETs) such as homomorphic encryption and federated learning. These technologies offer promising solutions for enhancing privacy without compromising data utility. For instance, federated learning enables AI models to be trained on decentralized data, reducing the need for data aggregation and enhancing privacy [85]. According to Abdulsalam and Hedabou [47], the successful implementation of PbD in cloud environments requires the pulling of resources from

technological, organizational, and legal dimensions, and while challenges may persist, the growing awareness of privacy risks and increasing regulatory scrutiny are driving positive change. By adopting a proactive, risk-based approach, organizations can mitigate risks, build trust, and achieve long-term sustainability.

Privacy-Enhancing Technologies (PETs) in AI: Privacy-enhancing technologies (PETs) are increasingly vital in safeguarding data privacy within AI systems. Technologies such as homomorphic encryption, differential privacy, secure multi-party computation, and federated learning each offer distinct capabilities for protecting sensitive information while enabling data processing and analysis. Homomorphic encryption, for instance, allows computations on encrypted data without decryption, thereby maintaining privacy throughout the computational process. Differential privacy adds noise to data queries to prevent the identification of individual data points, ensuring that the inclusion or exclusion of a single data point does not significantly alter the output [86]. According to Mestari et al. [87], the application of PETs in AI systems addresses the critical need to balance data utility with privacy, and AI systems require vast amounts of data to train models effectively, which often involves sensitive personal information. PETs curb privacy risks by allowing data to be used without compromising individual privacy; for example, homomorphic encryption enables machine learning tasks to be performed on encrypted data, ensuring that raw data remains inaccessible even during analysis. Similarly, differential privacy techniques can be applied to AI algorithms to ensure that outputs are privacy-preserving, thus enhancing user trust and regulatory compliance [60,67].

However, Williamson and Prybutok [88] state that implementing PETs in AI systems is fraught with challenges and limitations, and one significant challenge is the computational overhead that is associated with these technologies. Homomorphic encryption, while highly secure, is computationally intensive and can significantly slow down processing times, making it less practical for real-time applications, while differential privacy, though effective in protecting individual data points, can reduce the accuracy of AI models due to the added noise, posing a trade-off between privacy and utility [89]. Secure multi-party computation and federated learning also face scalability issues, requiring significant coordination and communication, which can be cumbersome in large-scale AI deployments. Despite these challenges, ongoing research is continually enhancing the efficiency and applicability of PETs, and emerging PETs, such as hybrid approaches that combine multiple techniques, show promise in overcoming some of the limitations of existing technologies [88,89]. For instance, hybrid models that integrate homomorphic encryption with secure hardware solutions like trusted execution environments (TEEs) can provide enhanced security with reduced computational costs. Recent studies suggest that these emerging PETs could offer more practical and scalable solutions for privacy-preserving AI [5,88,90].

Federated learning, another emerging PET, has garnered significant attention for its potential to revolutionize privacy-preserving AI; this technique involves training AI models across multiple decentralized devices or servers while keeping data localized. Rodriguez-Barroso et al. [89] highlight that only model updates, not raw data, are shared among participating nodes, significantly enhancing privacy. Federated learning is particularly relevant in sensitive sectors like healthcare, where patient data privacy is paramount. By enabling collaborative learning without centralizing data, federated learning offers a solid solution for leveraging large datasets while adhering to strict privacy regulations [89].

Risk Assessment and Privacy Impact Assessment (PIA): Risk Assessment and Privacy Impact Assessments (PIAs) are critical components in ensuring the privacy and security of data within AI systems. According to Georgiadis and Poels [25], PIAs involve systematic evaluations of how data collection, processing, and storage practices impact privacy, particularly in environments where sensitive personal information is at stake. PIAs are particularly relevant in AI systems due to the vast amounts of data required to train models effectively, and they are often used to process sensitive personal information, raising significant privacy concerns. PETs, such as homomorphic encryption, differential privacy, secure multi-party computation, and federated learning, are essential in addressing these concerns by enabling data processing without compromising individual privacy [89]. However, there are challenges with the implementation of PETs in AI systems, but despite these challenges, recent research continues to enhance the efficiency and applicability of PETs, and emerging hybrid

approaches that combine multiple PET techniques show promise in overcoming some limitations [5,94]. For instance, integrating homomorphic encryption with secure hardware solutions like Trusted Execution Environments (TEEs) can enhance security while reducing computational costs [89].

In the context of risk assessment, PIAs are essential for balancing privacy and utility; studies suggest that the successful implementation of PETs requires careful consideration of trade-offs, ensuring that protective measures do not undermine the effectiveness of AI models [91,92]. This balance is particularly challenging in dynamic environments where both data privacy and real-time performance are critical, so an interdisciplinary approach is needed, as drawing insights from fields such as cryptography, data science, and regulatory compliance will be instrumental in developing effective privacy solutions that are tailored to the specific needs of AI systems and their application contexts [88]. As AI and data privacy continue to grow, so will the role of PETs and PIAs in shaping the future of this technological frontier [93].

3. METHODOLOGY

This study utilized a mixed-method approach (incorporating both qualitative and quantitative methods) to assess the effectiveness of Privacy by Design (PbD) principles in AI systems. Data from two key sources, the OWASP Top 10 AI Vulnerabilities Dataset and the NIST Privacy Framework, were analyzed to provide detailed information on vulnerabilities in AI systems and key performance indicators relevant to PbD implementations.

A logistic regression analysis was conducted using the OWASP dataset to model the relationship between factors such as vulnerability type, environment, and severity score, as well as the likelihood of privacy breaches in AI systems. The logistic regression model was expressed as follows:

$$\text{Logit}(p) = In\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 * Vulnerability\ Type + \beta_2 * Environment \div \beta_3 * Severity\ Score$$

The probability P of a breach occurring was derived using the logit function:

$$P = \frac{1}{1 + e^{-Logit(p)}}$$

To evaluate the effectiveness of PbD implementations, a Pearson correlation analysis was conducted using data from the NIST Privacy Framework, focusing on key performance indicators like breach detection rate,compliance score,and response time. The Pearson correlation coefficient (r) was calculated as follows:

$$r = \frac{n\sum XY - (\sum X)(\sum Y)}{\sqrt{[n\sum X^2 - (\sum X^2)n\sum Y^2 - (\sum Y^2)]}}$$

Additionally, the analysis included calculations of the mean and standard deviation to summarize the central tendency and dispersion of the key performance indicators across different PbD implementations.

The mean (μ) was calculated using: $\mu = \frac{\sum X}{n}$

The standard deviation (σ) was calculated to measure the variability of the data:

$$\sigma = \sqrt{\frac{\sum(X - \mu)^2}{n}}$$

To further assess the variability in PbD effectiveness across different levels of breach detection rates,the Interquartile Range (IQR) was calculated as follows:

$$IQR = Q3 - Q1$$

The qualitative component of the analysis involved a comprehensive review of academic literature and conference papers specifically selected for their relevance to the implementation of PbD principles in AI systems.A case study analysis was first conducted to explore in-depth challenges,solutions,and outcomes associated with PbD across various domains,including healthcare,smart cities,and the Internet of Things applications.Subsequently,a thematic analysis was performed to systematically identify and categorize recurring themes related to privacy challenges and successes,reflecting critical issues faced during the implementation of PbD,the strategies applied,and the resulting outcomes or lessons learned.

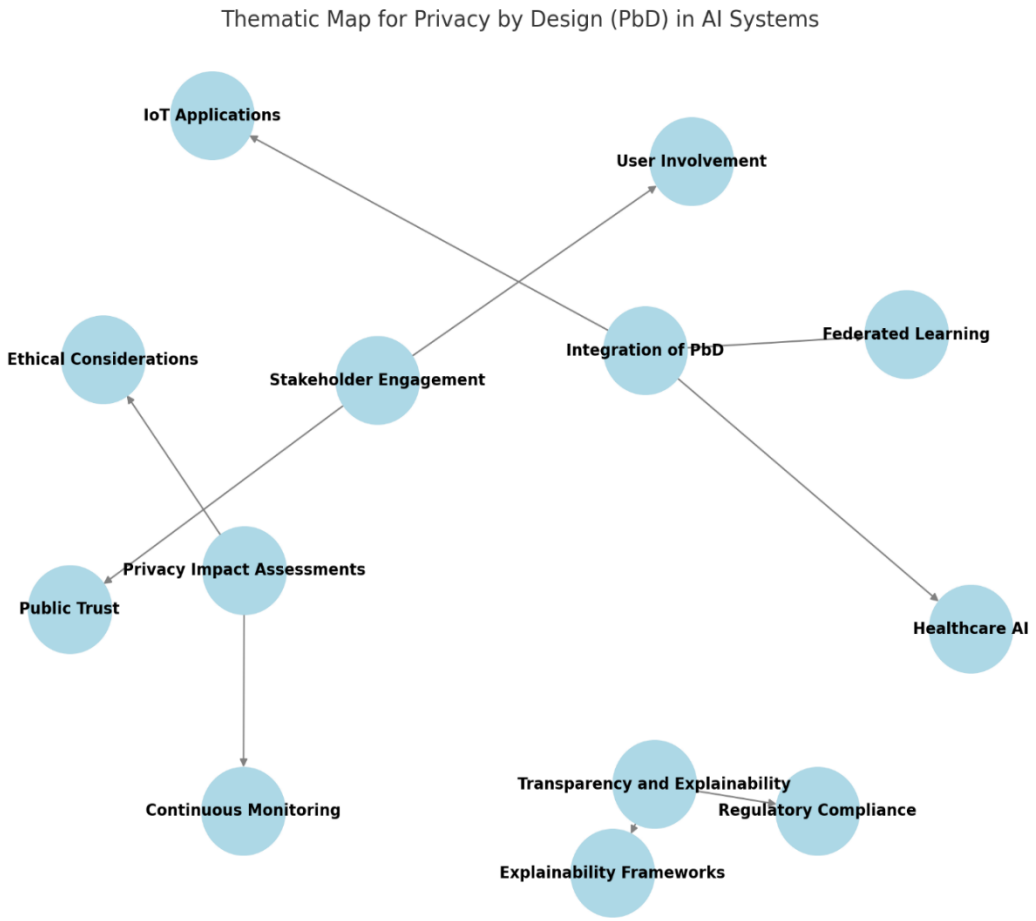

Fig. 1. Summary of the Qualitative Analysis

4. RESULTS AND DISCUSSION

4.1 Results

The logistic regression analysis was conducted to predict the likelihood of privacy breaches in AI systems based on vulnerability type, environment, and severity score. The results, presented in Table 1, reveal that both environment and severity scores are statistically significant predictors of breaches.

The environment was found to be a significant factor (B = -1.142, p < .001), with AI systems in more controlled environments less likely to experience breaches. The severity score also significantly predicted breaches (B = 0.932, p < .01), indicating that higher severity vulnerabilities increase breach likelihood. Although vulnerability type showed a positive trend (B = 0.348, p = .109), it was not statistically significant in this analysis. These findings support the study's objective of identifying key factors in preventing privacy breaches in AI systems.

Table 1. Logistic regression analysis predicting the likelihood of privacy breaches in AI systems

Variable	B	SE B	Wald z	p-value	95% CI for B
Constant	-4.991	1.936	-2.578	.0099	[-8.786, -1.197]
Vulnerability Type	0.348	0.217	1.603	.1090	[-0.078, 0.774]
Environment	-1.142	0.313	-3.653	.0003	[-1.755, -0.529]
Severity Score	0.932	0.289	3.231	.0012	[0.367, 1.498]

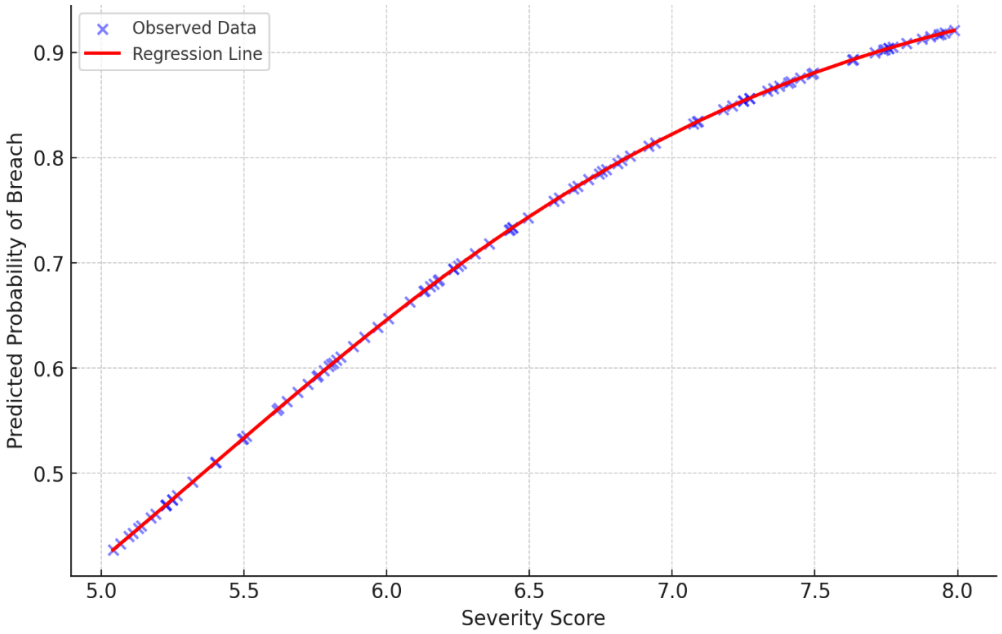

Fig. 2. Predicted Probability of breach by severity score

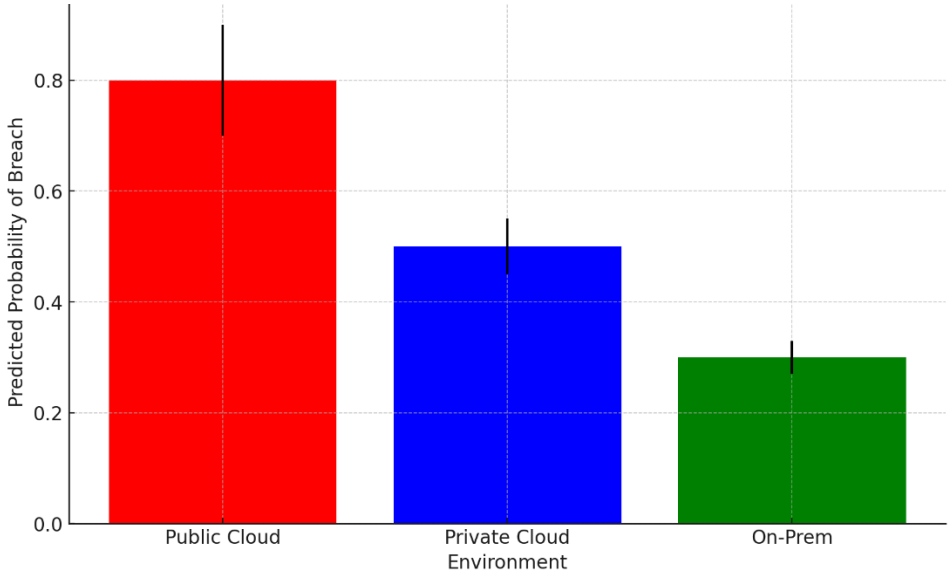

Fig. 3. Predicted probability of breach by environment

Fig. 2 illustrates the relationship between the Severity Score and the predicted probability of a privacy breach. The regression line shows a positive correlation, indicating that as the severity of vulnerabilities increases, the likelihood of a breach occurring also rises.

Fig. 3 displays the predicted probability of privacy breaches across different environments, with error bars indicating confidence intervals. The chart reveals that public cloud environments have a higher likelihood of breaches than private cloud and on-prem settings, which is consistent with the study's aim to assess environmental factors in mitigating privacy risks. The differences in breach probabilities highlight the importance of controlled environments in reducing privacy vulnerabilities.

Implementation effectiveness measurement: To evaluate the effectiveness of Privacy by Design (PbD) implementations using key

performance indicators (KPIs—breach detection rate, compliance score, and response time) from the NIST Privacy Framework, a Pearson correlation analysis was conducted to examine their relationship with the overall effectiveness of PbD strategies in mitigating privacy risks in AI systems, as shown in Table 2 and Table 3.

The results indicate that Breach Detection Rate and PbD Effectiveness are strongly positively correlated (r = 0.791), suggesting that systems with higher detection rates are associated with more effective PbD implementations. Conversely, Response Time and PbD Effectiveness show a moderate negative correlation (r = -0.370), indicating that faster response times are linked to higher effectiveness in PbD strategies. The Compliance Score, however, exhibited no significant correlation with PbD Effectiveness (r = -0.019), implying that compliance, while necessary, may not be a direct predictor of PbD success.

Table 2. Summary of key performance indicators (KPIs)

KPI	Mean	Std. Deviation	Min	25th Percentile	50th Percentile	75th Percentile	Max
Breach Detection Rate	84.11	8.92	70.17	75.80	83.92	91.91	99.61
Compliance Score	79.91	11.72	60.28	69.68	80.22	90.65	99.43
Response Time (hours)	12.90	6.75	1.12	7.37	13.94	18.30	23.77
PbD Effectiveness Score	68.34	6.22	54.23	65.00	67.56	72.73	82.41

Table 3. Correlation matrix for key performance indicators

KPI	Breach Detection Rate	Compliance Score	Response Time	PbD Effectiveness
Breach Detection Rate	1.000	-0.034	-0.038	0.791
Compliance Score	-0.034	1.000	-0.146	-0.019
Response Time	-0.038	-0.146	1.000	-0.370
PbD Effectiveness Score	0.791	-0.019	-0.370	1.000

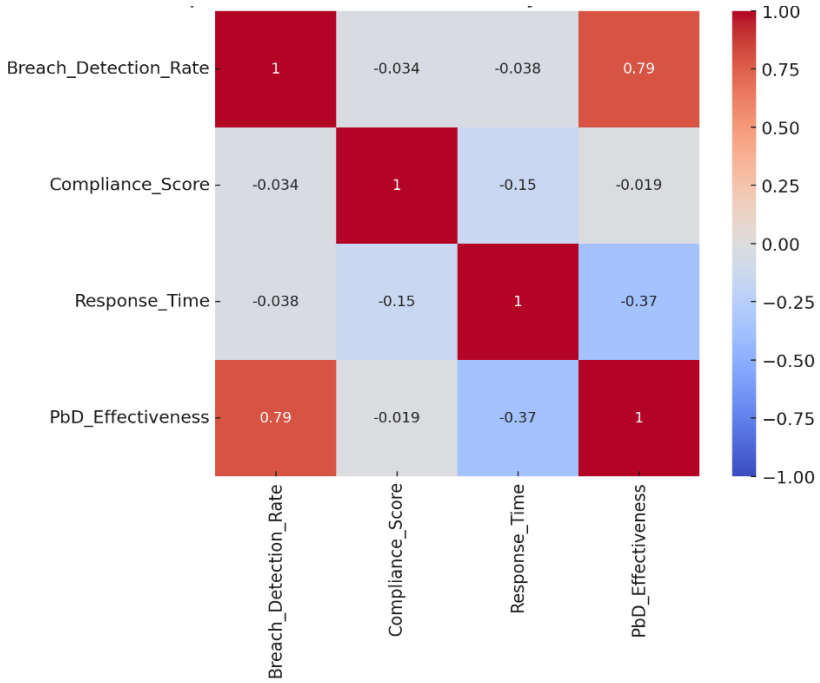

Fig. 4. Correlation Matrix of KPIs

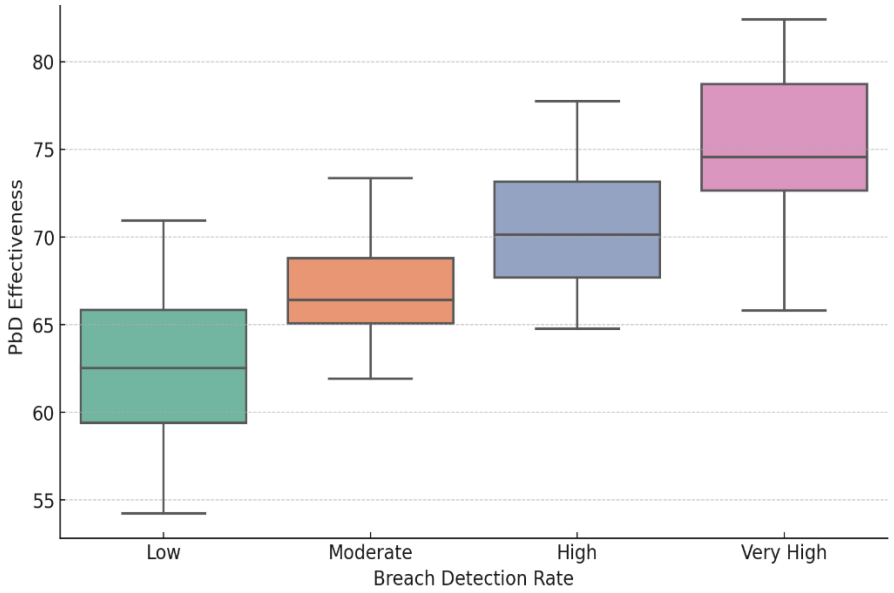

Fig. 5. Effectiveness by breach detection rates

Table 4. Case study analysis on privacy by design (PbD) in AI systems

Case Study	Challenges	Solutions Implemented	Outcomes/Lessons Learned
PbD in Healthcare AI Systems	Integrating PbD principles in AI-driven healthcare applications, particularly in maintaining data privacy across decentralized networks (Kaissis et al., [95].	Implemented federated learning techniques to ensure data privacy without compromising AI model accuracy. PbD principles were embedded throughout the AI system development process.	Effective implementation of PbD led to enhanced data privacy while maintaining AI model performance. Federated learning emerged as a viable solution for privacy preservation in healthcare AI applications, necessitating continuous monitoring.
AI in Smart Cities and Privacy Challenges	Addressing privacy concerns in smart city applications where large volumes of personal data are collected and processed Eckhoff & Wagner, [96].	Applied anonymization techniques and real-time data encryption to protect personal data. Engaged stakeholders to ensure privacy considerations were integrated into every aspect of AI deployment.	Anonymization and encryption techniques helped mitigate privacy risks, though challenges remained in balancing data utility with privacy. Stakeholder engagement was critical for addressing privacy concerns and building public trust in smart city initiatives.
Privacy Impact Assessment in AI Development	Conducting effective Privacy Impact Assessments (PIAs) in AI systems, particularly in understanding and mitigating ethical implications Raab, [38]; Wairimu et al., [97].	Developed a comprehensive PIA framework that included ethical considerations, stakeholder input, and continuous feedback loops. This framework was integrated into the AI system lifecycle.	The PIA framework offered a structured approach to identifying and mitigating privacy risks early in the AI development process. Ethical considerations and stakeholder input enhanced the robustness and acceptability of AI systems.
PbD in IoT and Healthcare Applications	Ensuring transparency and explainability of AI systems in IoT-based healthcare applications, where privacy risks are amplified by interconnected devices Alkhariji et al., [98].	Synthesized existing PbD knowledge into a framework prioritizing transparency and explainability. This framework was applied to design AI systems that met privacy requirements in healthcare.	The framework successfully guided the design of transparent and explainable AI systems, improving user trust and compliance with privacy regulations. However, challenges persisted in fully automating the PbD process in complex IoT environments.

Table 5. Thematic analysis of privacy by design (PbD) in AI systems

Theme	Challenges	Successes
Integration of PbD in AI Development	Difficulty in seamlessly integrating PbD principles into AI systems due to complexities like large datasets and complex algorithms Danezis et al., [99].	Effective integration using federated learning and privacy-preserving techniques, particularly in healthcare and smart cities Kaissis et al., [94]; Alkhariji et al., [98].
Ethical and Privacy Impact Assessments	Conducting effective Privacy Impact Assessments (PIAs) in AI systems, especially regarding ethical implications and data usage Raab, [38]; Wairimu et al.,[97].	Development of comprehensive PIA frameworks with continuous feedback loops and stakeholder engagement to proactively manage privacy risks Raab, [38]; Wairimu et al., [97].
Transparency and Explainability in AI	Ensuring transparency and explainability in AI systems, particularly in IoT and smart city applications Alkhariji et al., [98].	The successful design of transparent and explainable AI systems through structured PbD frameworks improves user trust and compliance Alkhariji et al., [98].
Stakeholder Engagement and Public Trust	Building and maintaining public trust through effective stakeholder engagement, especially in privacy-sensitive applications Eckhoff & Wagner, [96]; Mehr, [100].	Early and ongoing stakeholder engagement leads to AI systems that meet user needs and expectations and build public trust, Eckhoff & Wagner, [96].

The heatmap in Fig. 4 shows the Pearson correlation coefficients between key performance indicators (KPIs) and PbD effectiveness. The strong positive correlation between Breach Detection Rate and PbD Effectiveness (r = 0.791) is highlighted, while the negative correlation between Response Time and PbD Effectiveness (r = -0.370) indicates that faster response times are associated with higher PbD effectiveness.

Fig. 5 presents a box plot showing the distribution of PbD effectiveness scores across different levels of breach detection rates. The plot demonstrates that higher breach detection rates are associated with higher median PbD effectiveness, emphasizing the importance of effective breach detection in enhancing privacy outcomes.

Case study analysis: A thorough analysis of Privacy by Design (PbD) implementations in AI systems was conducted by systematically selecting and examining targeted case studies from relevant literature. These case studies specifically illustrate the critical challenges faced, the solutions implemented, and the outcomes achieved in the practical integration of PbD principles within AI systems (see Table 4).

The analysis demonstrates that integrating Privacy by Design principles into AI systems requires addressing complex challenges, implementing robust solutions, and deriving valuable lessons from real-world applications. The case studies the importance of federated learning, anonymization, encryption, stakeholder engagement, and structured Privacy Impact Assessments in mitigating privacy risks within AI systems. These outcomes highlight the need for continuous monitoring, ethical considerations, and transparency to ensure the effectiveness of PbD implementations in diverse AI-driven environments.

Thematic analysis: To systematically identify and categorize recurring challenges and successes from relevant literature, a thematic analysis of Privacy by Design (PbD) implementations in AI systems was conducted (see Table 5).

The thematic analysis presented shows the critical challenges and successful strategies involved in implementing PbD in AI systems. The findings reveal that while integrating PbD principles into AI development poses significant difficulties, particularly in managing complex data and algorithms Danezis et al., [99], effective solutions such as federated learning have proven successful Kaissis et al., [95]; Alkhariji et al., [98]. Additionally, the importance of ethical and privacy impact assessments, transparency, and stakeholder engagement is highlighted by Raab [38], Wairimu et al. [97], Eckhoff & Wagner [96], Mehr [100], with comprehensive frameworks and proactive engagement strategies emerging as key factors in building public trust and ensuring the success of AI systems.

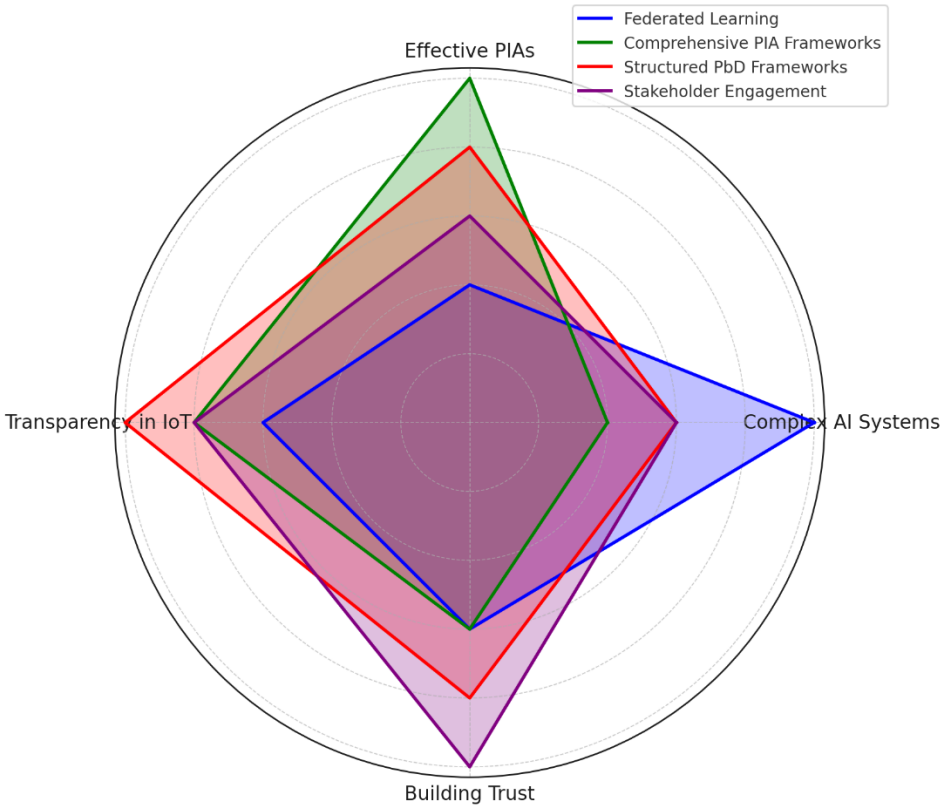

Fig. 6. Effectiveness of solutions in addressing PbD challenges

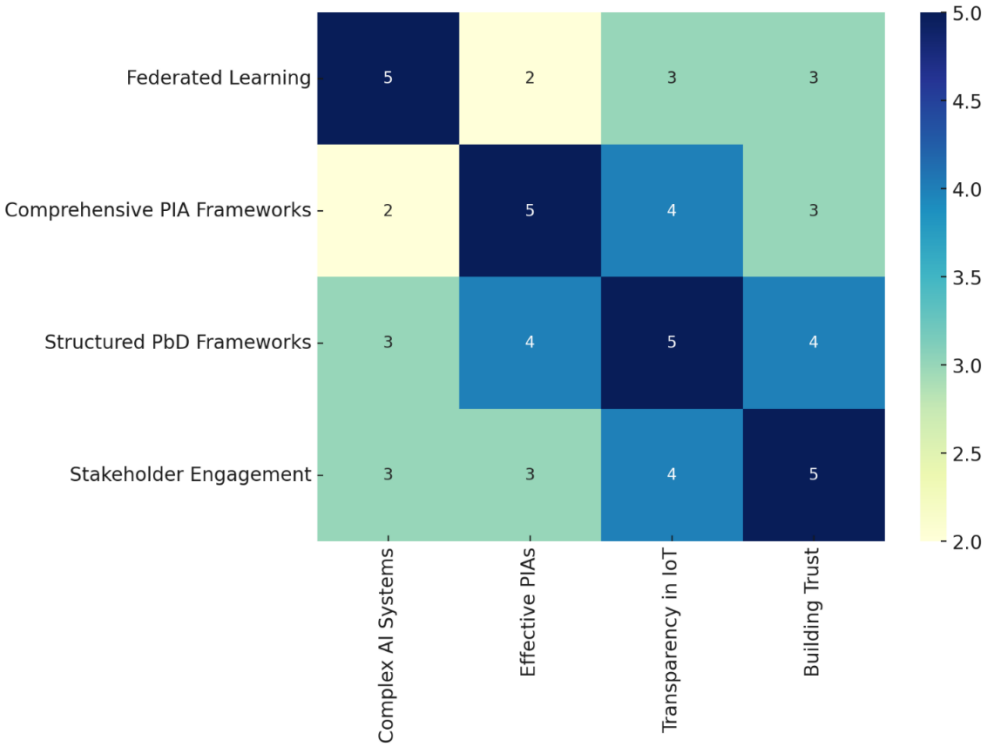

Fig. 7. Effectiveness of different solutions in addressing the key challenges of Privacy by Design (PbD) in AI systems

Fig. 6 illustrates the effectiveness of various solutions in addressing key challenges associated with Privacy by Design (PbD) in AI systems. Federated Learning shows high effectiveness in managing complex AI systems but is less impactful in addressing other challenges. Comprehensive PIA Frameworks are particularly effective in handling Privacy Impact Assessments and ensuring transparency but are less effective in dealing with complex AI systems. Structured PbD Frameworks demonstrate strong performance across multiple challenges, especially in transparency and regulatory compliance. Stakeholder Engagement is most effective in building trust and ensuring transparency but is less effective in addressing complex AI system challenges.

Fig. 7 provides a visual representation of the effectiveness of various solutions in addressing key challenges associated with Privacy by Design (PbD) in AI systems. Federated Learning demonstrates high effectiveness in managing complex AI systems, while Comprehensive PIA Frameworks are particularly effective in addressing Privacy Impact Assessments and transparency. Structured PbD Frameworks show strong effectiveness across multiple challenges, including transparency and regulatory compliance. Stakeholder Engagement is most effective in building trust and ensuring transparency, though less so in addressing the complexities of AI systems.

4.2 Discussion

The findings of this study provide significant insights into the effectiveness of incorporating Privacy by Design (PbD) principles within AI systems across various environments, aligning with the study's aim of developing a model for integrating risk assessment and privacy impact assessment into the AI development lifecycle. The logistic regression analysis reveals that the environment in which AI systems operate is a crucial determinant of privacy breaches, with public cloud environments exhibiting a higher likelihood of breaches compared to private cloud and on-premises settings. This result is consistent with the literature, which highlights the inherent vulnerabilities of public cloud environments due to their multi-tenant nature and the complexities associated with maintaining stringent privacy controls in such dynamic settings [43,67]. The statistical significance of the severity score as a predictor of breaches further explains the importance of addressing high-severity vulnerabilities, corroborating the view that robust privacy measures are essential in environments where the severity of potential threats is high [50,58].

Moreover, the positive correlation between breach detection rate and PbD effectiveness ($r = 0.791$) suggests that systems with more effective detection mechanisms are better equipped to mitigate privacy risks, supporting the argument that continuous monitoring and proactive threat detection are critical components of a successful PbD implementation [94,96]. Conversely, the negative correlation between response time and PbD effectiveness ($r = -0.370$) indicates that faster response times are associated with higher PbD effectiveness, highlighting the necessity of prompt action in mitigating the impact of privacy breaches. This finding aligns with the broader literature on AI and cloud security, which emphasizes the role of AI-driven solutions in reducing vulnerability and enhancing response times [50,58].

However, the lack of a significant correlation between compliance scores and PbD effectiveness ($r = -0.019$) challenges the assumption that regulatory compliance alone is sufficient to ensure effective privacy protection. While compliance with frameworks such as GDPR and CCPA is undoubtedly important, this study suggests that compliance measures must be complemented by robust, context-specific privacy strategies that address the unique challenges posed by AI systems [16,19]. The case study analysis further reinforces this point by demonstrating the effectiveness of federated learning, anonymization, and encryption techniques in mitigating privacy risks in specific contexts, such as healthcare AI systems and smart city applications. These real-world examples illustrate how tailored PbD implementations can enhance privacy outcomes, even in environments where traditional compliance measures might fall short [94,97].

In terms of practical applications, the thematic analysis reveals recurring challenges and successful strategies associated with PbD implementation. The difficulty of integrating PbD principles into complex AI systems, particularly those involving large datasets and sophisticated algorithms, echoes the challenges highlighted in the literature [98]. Nevertheless, the successful use of federated learning and other privacy-preserving techniques in specific domains demonstrates that these challenges can be

overcome with the right approach [94,97]. Similarly, the importance of comprehensive privacy impact assessments and ethical considerations in AI development is reaffirmed, aligning with existing research that advocates for a proactive, risk-based approach to privacy management [38,100].

The study's results also emphasize the critical role of stakeholder engagement and transparency in building public trust, particularly in privacy-sensitive applications such as healthcare and smart cities. This finding supports the literature's assertion that effective communication and stakeholder involvement are key to the successful implementation of PbD principles, as they help address public concerns and foster trust in AI systems [95,99]. By comparing the effectiveness of various solutions, including federated learning, structured PbD frameworks, and stakeholder engagement strategies, this study provides valuable insights into the practicalities of implementing PbD in diverse AI-driven environments. These insights highlight the need for continuous monitoring, ongoing research, and adaptive strategies to ensure that privacy measures remain effective in the face of evolving technological and regulatory challenges [67,89].

5. CONCLUSION AND RECOMMENDATION

This study demonstrates that the effective incorporation of Privacy by Design (PbD) principles in AI systems is critical for mitigating privacy breaches across various environments, particularly in public cloud settings where the risk is heightened. The logistic regression analysis confirms that environmental factors and the severity of vulnerabilities are significant predictors of privacy breaches, highlighting the need for targeted privacy strategies. The positive correlation between breach detection rates and PbD effectiveness shows the importance of robust detection mechanisms, while the findings also suggest that compliance alone is insufficient for achieving comprehensive privacy protection. The case studies and thematic analysis further illustrate the value of tailored PbD implementations, especially in domains like healthcare and smart cities, where privacy risks are pronounced.

The following specific recommendations are proposed to enhance the implementation of PbD principles in AI systems:

1. Prioritize the development and integration of advanced breach detection mechanisms, particularly in public cloud environments, to ensure timely identification and mitigation of privacy risks.
2. Implement a comprehensive and context-specific privacy impact assessment framework that goes beyond regulatory compliance to address the unique challenges posed by AI systems and environments.
3. Foster continuous stakeholder engagement throughout the AI development lifecycle, with a focus on transparency and ethical considerations, to build public trust and enhance the acceptability of AI systems.
4. Invest in research and development of privacy-enhancing technologies, such as federated learning and homomorphic encryption, to ensure that AI systems can effectively balance data utility with robust privacy protection.

DISCLAIMER (ARTIFICIAL INTELLIGENCE)

Author(s) hereby declare that NO generative AI technologies such as Large Language Models (ChatGPT, COPILOT, etc) and text-to-image generators have been used during writing or editing of manuscripts.

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

1. Netgate. Cloud Security Statistics in 2024, Netgate.com; 2024.
Available:https://www.netgate.com/blog/cloudsecuritystatistics#:~:text=Cloud%20Security%20Challenges%20and%20Trends&However%2C%20with%20the%20growing%20reliance
(Accessed Aug. 14, 2024).
2. Confessore N. Cambridge Analytica and Facebook: The Scandal and the Fallout So Far. The New York Times; 2018. Accessed: Aug. 14, 2024.
Available:https://www.nytimes.com/2018/04/04/us/politics/cambridge-analytica-scandal-fallout.html
3. Lichtenauer N, Schmidbauer L, Wilhelm S, Wahl F. A scoping review on analysis of

the barriers and support factors of open data. Information. 2024;15(1):5. Available:https://doi.org/10.3390/info15010005

4. Sampaio S, Sousa P, Martins C, Ferreira A, Antunes L, Cruz-Correia R. Collecting, processing and secondary using personal and (Pseudo)Anonymized Data in Smart Cities. Applied Sciences. 2023;13(6):3830–3830. Available:https://doi.org/10.3390/app13063830.
5. Scheibner J, et al. Revolutionizing medical data sharing using advanced privacy-enhancing technologies: Technical, legal, and ethical synthesis. Journal of Medical Internet Research. 2021;23(2):e25120. Available:https://doi.org/10.2196/25120.
6. Amajuoyi CP, Nwobodo LK, Adegbola MD. Transforming business scalability and operational flexibility with advanced cloud computing technologies. Computer Science and IT Research Journal. 2024;5(6):1469–1487. Available:https://doi.org/10.51594/csitrj.v5i6.1248.
7. Laxmaiah MKDM, Sharma DYK. A Comparative Study on Google App Engine Amazon Web Services and Microsoft Windows Azure. *papers.ssrn.com*; 2019. Available:https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3537564 (accessed Aug. 11, 2024).
8. Trakadas P, et al. Hybrid clouds for data-intensive, 5g-enabled Iot applications: An overview, key issues and relevant architecture. Sensors. 2019;19(16):3591. Available:https://doi.org/10.3390/s19163591.
9. Adigwe CS, Abalaka AI, Olaniyi OO, Adebiyi OO, Oladoyinbo TO. Critical analysis of innovative leadership through effective data analytics: Exploring trends in business Analysis, Finance, Marketing, and Information Technology. Asian Journal of Economics, Business and Accountin. 2023;23(22):460–479. Available:https://doi.org/10.9734/ajeba/2023/v23i221165.
10. Dai D, Boroomand S. A Review of artificial intelligence to enhance the security of big data systems: state-of-art, methodologies, applications, and challenges. Archives of Computational Methods in Engineering. 2021;29. Available:https://doi.org/10.1007/s11831-021-09628-0.
11. Adigwe CS, Olaniyi OO, Olabanji SO, Okunleye OJ, Mayeke NR, Ajayi SA. 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. 2024;24(4):126–146. Available:https://doi.org/10.9734/ajeba/2024/v24i41269.
12. Arigbabu AT, Olaniyi OO, Adigwe CS, Adebiyi OO, Ajayi SA. Data Governance in AI - Enabled Healthcare Systems: A Case of the Project Nightingale. Asian Journal of Research in Computer Science. 2024;17(5):85–107. Available:https://doi.org/10.9734/ajrcos/2024/v17i5441
13. Li H, Yu L, He W. The impact of GDPR on global technology development. Journal of Global Information Technology Management. 2019;22(1):1–6. Available:https://doi.org/10.1080/1097198X.2019.1569186.
14. (Liz) Harding E, Vanto JJ, Clark R, Hannah Ji L, Ainsworth SC. Understanding the scope and impact of the California Consumer Privacy Act of 2018. Journal of Data Protection and Privacy. 2019;2(3):234–253, Accessed: Aug. 11, 2024. Available:https://www.ingentaconnect.com/content/hsp/jdpp/2019/00000002/00000003/art00007
15. Marquis YA, Oladoyinbo TO, Olabanji SO, Olaniyi OO, Ajayi SS. Proliferation of AI Tools: A multifaceted evaluation of user perceptions and emerging trend. Asian Journal of Advanced Research and Reports. 2024;18(1):30–35. Available:https://doi.org/10.9734/ajarr/2024/v18i1596
16. Wischmeyer T. Artificial intelligence and transparency: Opening the black Box. Regulating Artificial Intelligence. 2019;75–101. Available:https://doi.org/10.1007/978-3-030-32361-5_4
17. Reed-Berendt R., Dove ES, Pareek M. The ethical implications of big data research in public health: 'Big Data Ethics by Design' in the UK-REACH Study. Ethics and Human Research. 2021;44(1):2–17. Available:https://doi.org/10.1002/eahr.500111.
18. Ogungbemi OS. Smart contracts management: The interplay of data privacy

and Blockchain for secure and efficient real estate transactions. Journal of Engineering Research and Reports. 2024;26(8):278–300.
Available:https://doi.org/10.9734/jerr/2024/v26i81245

19. Arif H, Kumar A, Fahad M, Hussain HK. Future Horizons: AI-Enhanced threat detection in cloud environments: Unveiling opportunities for research. International Journal of Multidisciplinary Sciences and Arts. 2023;2(2):242–251.
Available:https://doi.org/10.47709/ijmdsa.v2i2.3452
20. Semantha FH, Azam S, Shanmugam B, Yeo KC. PbDinEHR: A novel privacy by design developed framework using distributed data storage and sharing for secure and scalable electronic health records management. Journal of Sensor and Actuator Networks. 2023;12(2):36.
Available:https://doi.org/10.3390/jsan12020036.
21. Olabanji SO, Marquis YA, Adigwe CS, Abidemi AS, Oladoyinbo TO, Olaniyi OO. AI-driven cloud security: examining the impact of user behavior analysis on threat detection. Asian Journal of Research in Computer Science. 2024;17(3):57–74.
Available:https://doi.org/10.9734/ajrcos/2024/v17i3424.
22. Aziz R, Banerjee S, Bouzefrane S, Le Vinh T. Exploring homomorphic encryption and differential privacy techniques towards secure federated learning paradigm. Future Internet. 2023;15(9):310.
23. Available:https://doi.org/10.3390/fi15090310.
24. Olabanji SO, Oladoyinbo TO, Asonze CU, Adigwe CS, Okunleye OJ, Olaniyi OO. Leveraging FinTech compliance to mitigate cryptocurrency volatility for secure US Employee Retirement Benefits: Bitcoin ETF Case Study. Asian Journal of Economics, Business and Accounting. 2024;24(4):147–167.
Available:https://doi.org/10.9734/ajeba/2024/v24i41270
25. Oladoyinbo TO, Adebiyi OO, Ugonnia JC, Olaniyi OO, Okunleye OJ. Evaluating and establishing baseline security requirements in cloud computing: An enterprise risk management approach. Asian Journal of Economics, Business and Accounting. 2023;23(21):222–231,
Available:https://doi.org/10.9734/ajeba/2023/v23i211129.
26. Georgiadis G, Poels G. Establishing a comprehensive privacy impact assessment methodology for big data analytics in compliance with the general data protection regulation. SSRN; 2024,
Available:https://doi.org/10.2139/ssrn.4757166
27. Pagallo U. On the principle of privacy by design and its limits: Technology, ethics and the rule of law. Springer Link. 2020;35:111–127.
Available:https://doi.org/10.1007/978-3-030-54522-2_8
28. Semantha FH, Azam S, Yeo KC, Shanmugam B. A systematic literature review on privacy by design in the healthcare sector. Electronics. 2020;9(3).
Available:https://doi.org/10.3390/electronics9030452.
29. Olaniyi FG, Olaniyi OO, Adigwe CS, Abalaka AI, Shah N. Harnessing predictive analytics for strategic foresight: A comprehensive review of techniques and applications in transforming raw data to actionable insights. Asian Journal of Economics, Business and Accounting. 2023;23(22):441–459.
Available:https://doi.org/10.9734/ajeba/2023/v23i221164.
30. Sun S, Zheng X, Villalba- Díez J, Ordieres-Meré J. Data handling in industry 4.0: Interoperability based on distributed ledger technology. Sensors. 2020;20(11):3046.
Available:https://doi.org/10.3390/s20113046.
31. Olaniyi OO, Abalaka AI, Olabanji SO. Utilizing big data analytics and business intelligence for improved decision-making at leading fortune company. Journal of Scientific Research and Reports. 2023;29(9):64–72.
Available:https://doi.org/10.9734/jsrr/2023/v29i91785
32. Saeed RA, Saeed MMA, Ahmed ZE. Data Security and Privacy in the Age of AI and Digital Twins. *www.igi-global.com*; 2024.
Available:https://www.igi-global.com/chapter/data-security-and-privacy-in-the-age-of-ai-and-digital-twins/336453 (accessed Aug. 14, 2024).
33. Bugeja J, Jacobsson A. On the design of a privacy-centered data lifecycle for smart living spaces. IFIP advances in information and communication technology. 2020;576:126–141.
Available:https://doi.org/10.1007/978-3-030-42504-3_9.

34. Olaniyi OO, Okunleye OJ, Olabanji SO. 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. 2023;42(25):10–18.
Available:https://doi.org/10.9734/cjast/2023/v42i254181.
35. Munjal K, Bhatia R. A systematic review of homomorphic encryption and its contributions in healthcare industry. Complex and Intelligent Systems. 2022;9.
Available:https://doi.org/10.1007/s40747-022-00756-z.
36. Olaniyi OO, Omubo DS. WhatsApp data policy, data security and users' vulnerability. International Journal of Innovative Research and Development; 2023.
Available:https://doi.org/10.24940/ijird/2023/v12/i4/apr23021
37. Pulido-Gaytan B, et al. Privacy-preserving neural networks with Homomorphic encryption: Challenges and opportunities. Peer-to-Peer Networking and Applications. 2021;14(3):1666–1691.
Available:https://doi.org/10.1007/s12083-021-01076-8
38. Olaniyi OO, Olabanji SO, Abalaka AI. Navigating risk in the modern business landscape: strategies and insights for enterprise risk management implementation. Journal of Scientific Research and Reports. 2023;29(9):103–109.
Available:https://doi.org/10.9734/jsrr/2023/v29i91789
39. Raab CD. Information privacy, impact assessment, and the place of ethics. Computer Law and Security Review. 2020;37:105404.
Available:https://doi.org/10.1016/j.clsr.2020.105404.
40. Olaniyi OO, Olaoye OO, Okunleye OJ. Effects of Information Governance (IG) on Profitability in the Nigerian Banking Sector. Asian Journal of Economics, Business and Accounting. 2023;23(18):22–35.
Available:https://doi.org/10.9734/ajeba/2023/v23i181055.
41. Alghofaili Y, Albattah A, Alrajeh N, Rassam MA, Al-rimy BAS. Secure cloud infrastructure: A survey on issues, current solutions, and open challenges. Applied Sciences. 2021;11(19):9005.
Available:https://doi.org/10.3390/app11199005.
42. Gupta P, Gupta PK. Trust modeling in cloud. Springer eBooks. 2020;77–93.
Available:https://doi.org/10.1007/978-3-030-37319-1_4.
43. Olaniyi OO, Omubo DS. The Importance of COSO framework compliance in information technology auditing and enterprise resource management. International Journal of Innovative Research and Development; 2023.
Available:https://doi.org/10.24940/ijird/2023/v12/i5/may23001.
44. Aslan Ö, Aktuğ SS, Ozkan-Okay M, Yilmaz AA, Akin E. A comprehensive review of cyber security vulnerabilities, threats, attacks, and solutions. Electronics. 2023;12(6):1–42.
Available:https://doi.org/10.3390/electronics12061333.
45. Chang V, et al. A Survey on intrusion detection systems for fog and cloud computing. Future Internet. 2022;14(3):89.
Available:https://doi.org/10.3390/fi14030089.
46. Carrera G. Building a comprehensive cloud security audit program. EDPACS. 2021;66(1):1–4.
Available:https://doi.org/10.1080/07366981.2021.2004689.
47. Asonze CU, Ogungbemi OS, Ezeugwa FA, Olisa AO, Akinola OI, Olaniyi OO. Evaluating the trade-offs between wireless security and performance in iot networks: A case study of web applications in ai-driven home appliances. Journal of Engineering Research and Reports. 2024;26(8):411–432.
Available:https://doi.org/10.9734/jerr/2024/v26i81255
48. Chauhan M, Shiaeles S. An analysis of cloud security frameworks, problems and proposed solutions. Network. 2023;3(3):422–450.
Available:https://doi.org/10.3390/network3030018.
49. Rajesh YS, Kiran G, Poojari A. A unified approach toward security audit and compliance in cloud computing. Journal of The Institution of Engineers (India): Series B. 2024;105.
Available:https://doi.org/10.1007/s40031-024-01034-x.
50. Ogungbemi OS, Ezeugwa FA, Olaniyi OO, Akinola OI, Oladoyinbo OB. Overcoming remote workforce cyber threats: A

comprehensive ransomware and bot net defense strategy utilizing VPN Networks. Journal of Engineering Research and Reports. 2024;26(8):161–184.
Available:https://doi.org/10.9734/jerr/2024/v26i81237

51. Rangaraju S. Secure by Intelligence: Enhancing Products with AI-Driven Security Measures. EPH - International Journal of Science and Engineering. 2023 Dec;9(3):36-41.
DOI: 10.53555/ephijse.v9i3.212.
52. Pureti N. Zero-Day Exploits: Understanding the Most Dangerous Cyber Threats. International Journal of Advanced Engineering Technologies and Innovations. 2022;01(2):2.
Available: https://ijaeti.com/index.php/Journal/article/download/361/372. Accessed 2024 Aug 15.
53. Akinola OI, Olaniyi OO, Ogungbemi OS, Oladoyinbo OB, Olisa AO. Resilience and Recovery Mechanisms for Software-Defined Networking (SDN) and Cloud Networks. Journal of Engineering Research and Reports. 2024 Jul;26(8):112-34.
DOI: 10.9734/jerr/2024/v26i81234.
54. Yaseen A. AI-Driven Threat Detection and Response: A Paradigm Shift in Cybersecurity. International Journal of Information and Cybersecurity. 2023 Dec;7(12):25-43.
Available: https://publications.dlpress.org/index.php/ijic/article/view/73. Accessed 2024 Aug 15.
55. Oladoyinbo TO, Olabanji SO, Olaniyi OO, Adebiyi OO, Okunleye OJ, Alao AI. Exploring the Challenges of Artificial Intelligence in Data Integrity and its Influence on Social Dynamics. Asian Journal of Advanced Research and Reports. 2024 Jan;18(2):1-23.
DOI: 10.9734/ajarr/2024/v18i2601.
56. Vashishth TK, Sharma V, Sharma KK, Kumar B, Chaudhary S, Panwar R. Enhancing Cloud Security: The Role of Artificial Intelligence and Machine Learning. IGI Global; 2024.
Available: https://www.igi-global.com/chapter/enhancing-cloud-security/338350. Accessed 2024 Aug 15.
57. Olaniyi OO, Ezeugwa FA, Okatta CG, Arigbabu AS, Joeaneke PC. Dynamics of the Digital Workforce: Assessing the Interplay and Impact of AI, Automation, and Employment Policies. Archives of Current Research International. 2024 Apr;24(5):124-39.
DOI: 10.9734/acri/2024/v24i5690.
58. Kaur R, Gabrijelčič D, Klobučar T. Artificial Intelligence for Cybersecurity: Literature Review and Future Research Directions. Information Fusion. 2023;97:101804.
DOI: 10.1016/j.inffus.2023.101804.
59. Chen P, Wu L, Wang L. AI Fairness in Data Management and Analytics: A Review on Challenges, Methodologies and Applications. Applied Sciences. 2023 Sep;13(18):10258.
DOI: 10.3390/app131810258.
60. Mylrea M, Robinson N. Artificial Intelligence (AI) Trust Framework and Maturity Model: Applying an Entropy Lens to Improve Security, Privacy, and Ethical AI. Entropy. 2023 Oct;25(10):1429.
DOI: 10.3390/e25101429.
61. Hemamalini V, Mishra AK, Tyagi AK, Kakulapati V. Artificial Intelligence–Blockchain-Enabled–Internet of Things-Based Cloud Applications for Next-Generation Society. Wiley Online Library; 2023 Nov:65-82. doi: 10.1002/9781394213948.ch4.
62. Cherbal S, Zier A, Hebal S, Louail L, Annane B. Security in Internet of Things: A Review on Approaches Based on Blockchain, Machine Learning, Cryptography, and Quantum Computing. The Journal of Supercomputing. 2023 Sep;80.
DOI: 10.1007/s11227-023-05616-2.
63. Olaniyi OO, Olabanji SO, Okunleye OJ. Exploring the Landscape of Decentralized Autonomous Organizations: A Comprehensive Review of Blockchain Initiatives. Journal of Scientific Research and Reports. 2023 Sep;29(9):73-81.
DOI: 10.9734/jsrr/2023/v29i91786.
64. Han J, Park S, Kim J. Dynamic OverCloud: Realizing Microservices-Based IoT-Cloud Service Composition over Multiple Clouds. Electronics. 2020 Jun;9(6):969.
DOI: 10.3390/electronics9060969.
65. Olaniyi OO, Omogoroye OO, Olaniyi FG, Alao AI, Oladoyinbo TO. 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. 2024;26(6):32.
DOI: 10.9734/JERR/2024/v26i61160.

66. Henning S, Hasselbring W. A Configurable Method for Benchmarking Scalability of Cloud-Native Applications. Empirical Software Engineering. 2022 Aug;27(6). DOI: 10.1007/s10664-022-10162-1.
67. Abdulsalam YS, Hedabou M. Security and Privacy in Cloud Computing: Technical Review. Future Internet. 2021;14(1):11. DOI: 10.3390/fi14010011.
68. Olariu F, Alboaie L. Challenges In Optimizing Migration Costs From On-Premises To Microsoft Azure. Procedia Computer Science. 2023 Jan;225:3649-59. DOI: 10.1016/j.procs.2023.10.360.
69. Caballer M, Antonacci M, Šustr Z, Perniola M, Moltó G. Deployment of Elastic Virtual Hybrid Clusters Across Cloud Sites. Journal of Grid Computing. 2021 Feb;19(1). DOI: 10.1007/s10723-021-09543-5.
70. Samuel-Okon AD, Akinola OI, Olaniyi OO, Olateju OO, Ajayi SA. Assessing the Effectiveness of Network Security Tools in Mitigating the Impact of Deepfakes AI on Public Trust in Media. Archives of Current Research International. 2024 Jul;24(6):355-75. DOI: 10.9734/acri/2024/v24i6794.
71. Shivaramakrishna D, Nagaratna M. A Novel Hybrid Cryptographic Framework for Secure Data Storage in Cloud Computing: Integrating AES-OTP and RSA with Adaptive Key Management and Time-Limited Access Control. Alexandria Engineering Journal. 2023 Dec;84:275-84. DOI: 10.1016/j.aej.2023.10.054.
72. Ghaleb AAA, Dominic PDD, Fati SM, Muneer A, Ali RF. The Assessment of Big Data Adoption Readiness with a Technology–Organization–Environment Framework: A Perspective towards Healthcare Employees. Sustainability. 2021 Jul;13(15):8379. DOI: 10.3390/su13158379.
73. Olateju OO, Okon SU, Olaniyi OO, Samuel-Okon AD, Asonze CU. Exploring the Concept of Explainable AI and Developing Information Governance Standards for Enhancing Trust and Transparency in Handling Customer Data. Journal of Engineering Research and Reports. 2024 Jun;26(7):244-68. DOI: 10.9734/jerr/2024/v26i71206.
74. Adeusi OC, Adebayo YO, Ayodele PA, Onikoyi TT, Adebayo KB, Adenekan IO. IT Standardization in Cloud Computing: Security Challenges, Benefits, and Future Directions. World Journal of Advanced Research and Reviews. 2024 Jun;22(3):2050-7. DOI: 10.30574/wjarr.2024.22.3.1982.
75. Shakir U. Here's how Apple's AI model tries to keep your data private. The Verge. 2024 Jun 13. Available: https://www.theverge.com/2024/6/13/24175985/apple-intelligence-ai-model-local-cloud-privacy-how-it-works. Accessed 2024 Aug 15.
76. Bigelow S. What is Microsoft Azure and How Does It Work? SearchCloudComputing. 2022 Oct. Available: https://www.techtarget.com/searchcloudcomputing/definition/Windows-Azure. Accessed 2024 Aug 15.
77. Akremi A, Rouached M. A Comprehensive and Holistic Knowledge Model for Cloud Privacy Protection. The Journal of Supercomputing. 2021 Jan;77(8):7956-88. DOI: 10.1007/s11227-020-03594-3.
78. Dias Canedo E, Toffano Seidel Calazans A, Toffano Seidel Masson E, Teixeira Costa PH, Lima F. Perceptions of ICT Practitioners Regarding Software Privacy. Entropy. 2020 Apr;22(4):429. DOI: 10.3390/e22040429.
79. Zaid T, Garai S. Emerging Trends in Cybersecurity: A Holistic View on Current Threats, Assessing Solutions, and Pioneering New Frontiers. Blockchain in Healthcare Today. 2024 Apr;7(1). DOI: 10.30953/bhty.v7.302.
80. Apeh AJ, Hassan AO, Oyewole OO, Fakeyede OG, Okeleke PA, Adaramodu OR. GRC Strategies in Modern Cloud Infrastructures: A Review of Compliance Challenges. Computer Science & IT Research Journal. 2023 Nov;4(2):111-25. DOI: 10.51594/csitrj.v4i2.609.
81. Pilton C, Faily S, Henriksen-Bulmer J. Evaluating privacy - determining user privacy expectations on the web. Comput Secur. 2021 Jun;105:102241. DOI: 10.1016/j.cose.2021.102241.
82. Olateju OO, Okon SU, Igwenagu UTI, Salami AA, Oladoyinbo TO, Olaniyi OO. Combating the challenges of false positives in AI-driven anomaly detection systems and enhancing data security in the cloud. Asian J Res Comput Sci. 2024 Jun;17(6):264–92. DOI: 10.9734/ajrcos/2024/v17i6472.

83. Quach S, Thaichon P, Martin KD, Weaven S, Palmatier RW. Digital technologies: Tensions in privacy and data. J Acad Mark Sci. 2022;50(1):1299–323.
Accessed: Aug. 15, 2024.
Available:https://link.springer.com/article/10.1007/s11747-022-00845-y
84. Samuel-Okon AD, Olateju OO, Okon SU, Olaniyi OO, Igwenagu UTI. Formulating global policies and strategies for combating criminal use and abuse of artificial intelligence. Arch Curr Res Int. 2024 Jun;24(5):612–9.
DOI: 10.9734/acri/2024/v24i5735.
85. Wittkopp T, Acker A. Decentralized federated learning preserves model and data privacy. Lect Notes Comput Sci. 2021 Jan;12632:176–87.
DOI: 10.1007/978-3-030-76352-7_20.
86. Kenny CT, Kuriwaki S, McCartan C, Rosenman ETR, Simko T, Imai K. The use of differential privacy for census data and its impact on redistricting: The case of the 2020 U.S. Census. Sci Adv. 2021 Oct;7(41).
DOI: 10.1126/sciadv.abk3283.
87. El Mestari SZ, Lenzini G, Demirci H. Preserving data privacy in machine learning systems. Comput Secur. 2024 Feb;137:103605.
DOI: 10.1016/j.cose.2023.103605.
88. Williamson SM, Prybutok VR. Balancing privacy and progress: A review of privacy challenges, systemic oversight, and patient perceptions in AI-driven healthcare. Appl Sci. 2024 Jan;14(2):675.
DOI: 10.3390/app14020675.
89. Rodríguez-Barroso N, et al. Federated learning and differential privacy: Software tools analysis, the Sherpa.ai FL framework and methodological guidelines for preserving data privacy. Inf Fusion. 2020 Dec;64:270–92.
DOI: 10.1016/j.inffus.2020.07.009.
90. Kaaniche N, Laurent M, Belguith S. Privacy enhancing technologies for solving the privacy-personalization paradox: Taxonomy and survey. J Netw Comput Appl. 2020 Dec;171:102807.
DOI: 10.1016/j.jnca.2020.102807.
91. Roberts H, et al. Artificial intelligence in support of the circular economy: ethical considerations and a path forward. AI Soc. 2022 Nov;39.
DOI: 10.1007/s00146-022-01596-8.
92. De Paula Vieira A, Anthony R. Recalibrating veterinary medicine through animal welfare science and ethics for the 2020s. Animals. 2020 Apr;10(4):654.
DOI: 10.3390/ani10040654.
93. Tanisha J, Rajesh PA, Singh RG, Adhip K, Stuti K, Ajitha D. Privacy and data protection challenges in industry 4.0: An AI-driven perspective. World J Adv Eng Technol Sci. 2024 Jul;12(2):064–89.
DOI: 10.30574/wjaets.2024.12.2.0287.
94. Arabi H, AkhavanAllaf A, Sanaat A, Shiri I, Zaidi H. The promise of artificial intelligence and deep learning in PET and SPECT imaging. Phys Medica. 2021 Mar;83:122–37.
DOI: 10.1016/j.ejmp.2021.03.008.
95. Kaissis GA, Makowski MR, Rückert D, Braren RF. Secure, privacy-preserving and federated machine learning in medical imaging. Nat Mach Intell. 2020 Jun;2(6):305–11.
DOI: 10.1038/s42256-020-0186-1. Available from: https://www.nature.com/articles/s42256-020-0186-1
96. Eckhoff D, Wagner I. Privacy in the smart city—applications, technologies, challenges, and solutions. IEEE Commun Surv Tutor. 2018;20(1):489–516.
DOI: 10.1109/COMST.2017.2748998.
Available: https://ieeexplore.ieee.org/abstract/document/8025782
97. Wairimu S, Horn Iwaya L, Fritsch L, Lindskog S. On the evaluation of privacy impact assessment and privacy risk assessment methodologies: A systematic literature review. IEEE Access. 2024 Jan;12:1–1.
DOI: 10.1109/access.2024.3360864.
98. Alkhariji L, Alhirabi N, Alraja MN, Barhamgi M, Rana O, Perera C. Synthesising privacy by design knowledge toward explainable Internet of Things application designing in healthcare. ACM Trans Multimed Comput Commun Appl. 2021 Jun;17(2s):1–29.
DOI: 10.1145/3434186.
99. Danezis G, et al. Privacy and data protection by design - from policy to engineering. arXiv. 2015.
DOI: 10.2824/38623.

Available: https://arxiv.org/abs/1501.0372

100. Mehr H. Artificial intelligence for citizen services and government; 2017. Available: https://creatingfutureus.org/wp-content/uploads/2021/10/Mehr-2017-AIforGovCitizenServices.pdf [Accessed: Aug. 15, 2024].

Peer-review history:
The peer review history for this paper can be accessed here:
https://www.sdiarticle5.com/review-history/122737