Volume 10, Issue 10, October – 2025 International Journal of Innovative Science and Research Technology
ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/25oct209

Integrating Privacy-Preserving AI Models into AI Governance Frameworks

Whenume O. Hundeyin[1]; Samson A. Adegbenro[2]; Yankat P. Rindap[3]; Chinedu Austin Adaba[4]

Publication Date: 2025/10/10

Abstract: The integration of Artificial Intelligence (AI) into Information Technology environment has transformed organizational processes, yet it has also introduced challenges around privacy, accountability, and regulatory compliance. This study explores how privacy-preserving AI (PPAI) techniques can strengthen IT governance and compliance, and identifies the governance controls internal IT auditors require in AI developments. A qualitative exploratory research design was adopted, drawing from nine peer-reviewed articles, regulatory framework (as GDPR, EU AI Act, the NIST AI RMF, and ISO/IEC standards) IT governance models (COBIT, ISACA guidelines, ISO/IEC 27001). The analytical process combined thematic analysis and comparative mapping to PPAI techniques with IT governance control and assurance checkpoints. The findings reveal that federated learning operationalizes privacy-by-design by minimizing raw data transfer, aligning with GDPR and similar principle, Secure aggregation, homomorphic encryption, and differential privacy strengthen confidentiality and safeguard model outputs against inference attacks, while immutable logging and explainability provide accountability and auditability consistent with ISO/IEC 27701 and NIST AI RMF. From an assurance perspective, auditors must expand evaluations to cover AI-specific risks, including model integrity, federated learning protocols, and privacy-preserving outputs. The study concludes that PPAI serves not only as a technical safeguard but also as a governance enabler. Recommendations include embedding PPAI in IT operations, updating governance standards, and developing dynamic audit framework tailored to AI-driven environments.

How to Cite: Whenume O. Hundeyin; Samson A. Adegbenro; Yankat P. Rindap; Chinedu Austin Adaba (2025). Integrating Privacy-Preserving AI Models into AI Governance Frameworks. *International Journal of Innovative Science and Research Technology*,10(10), 376-384. https://doi.org/10.38124/ijisrt/25oct209

I. INTRODUCTION

The integration of Artificial Intelligence into Information Technology project management and operations has revolutionized practices. AI-driven systems automate processes, optimize resources allocation, and risk management by leveraging predictive analytics and machine learning algorithms (Rusell & Norvig, 2020). In IT environments, AI is deployed in areas such as cybersecurity, predictive maintenance, and decision-support systems to improve efficiency and resilience (Gartner, 2022). However, as AI becomes more embedded within IT infrastructures, ethical, security, and legal challenges are emerging that threaten these benefits.

Among the most pressing concerns is algorithmic bias, where AI systems trained on historical or incomplete datasets can replicate and amplify discriminatory patterns (O'Neil, 2017). Regulations such as the General Data Protection Regulations (GDPR), California Consumer Privacy Act (CCPA), and the Health Insurance Portability and Accountability Act (HIPAA) reveals the importance of safeguarding data in AI-driven IT systems (Regulation (EU) 2016/679, 2016). As a result, privacy protection and ethical compliance have become central concerns in IT governance when deploying AI.

These challenges are further compounded by the rise of shadow AI, the unauthorized use of AI tools by employees outside formal IT governance structures. Shadow AI often emerges when staff adopt freely available AI solution to improve workflows or solve specific problems without IT department oversight (Yilmaz *et al.,* 2024). This introduces risks of data leakages, compliance violations, and security breaches, since such deployments fall outside established corporate governance and cybersecurity frameworks (Chi *et al.,* 2024). This highlights the need for governance mechanism that can effectively regulate both sanctioned and unsanctioned AI deployment within information Technology ecosystem.

In response, multiple governance and regulatory frameworks have been introduced to guide the responsible adoption of AI in IT environments. For example, COBIT 2019, ITIL, and ISO/IEC 27001 provide structured IT governance and security guidelines (De Haes *et al.,* 2020; Rubio & Arcilla, 2020; Hamdi *et al.,* 2019). Similarly, the EU AI Act, the NIST AI Risk Management Framework (AI RMF), and various corporate social responsibility initiative emphasize ethical AI adoption and organizational accountability (European Commission, 2021; NIST, 2023). While these framework reveals transparency, accountability, and risk management, they remain high-level, offering practical guidance that are limited on operationalising

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/25oct209

privacy-preserving AI within IT assurance and audit functions.

This limitation becomes critical given that privacy-preserving AI (PPAI) techniques including Federated Learning, Differential Privacy, and Homomorphic Encryption offer solutions for balancing AI utility with privacy and security concerns (Kairouz *et al.,* 2021). These techniques can mitigate risks associated with data breaches and unauthorized access, while also enhancing compliance with global privacy regulations. Yet, despite their technical potential, current IT governance frameworks have not incorporated these methods into assurance and audit practice. In particular, IT auditors face difficulties in verifying whether AI systems employing PPAI techniques meet compliance, fairness, and transparency standard (Rahwan *et al.,* 2019).

Existing governance model focus on policy and risk principles, while technical advances in PPAI remain disconnected from governance execution. While existing frameworks such as COBIT 2019 and ISO/IEC 27001 provide mechanism for IT governance and information security, they do not explicitly address the integration of privacy-preserving AI techniques into IT assurance and audit practices. As a result, a critical research gap emerges at the intersection of AI governance, IT assurance, and privacy-preserving AI.

➢ *Problem Statement*

The rapid integration of Artificial Intelligence into IT systems has created unprecedented opportunities for efficiency, security, and innovation. However, this transformation has also introduced challenges surrounding privacy, transparency, and accountability. Organization face difficulties in communicating AI processes to stakeholders, which undermines consent and public trust. The 'Privacy paradox" illustrates this complexity while individual express strong concern about their personal data, they continue to engage with AI-driven technologies, often feeling compelled to accept contract with little real choice (Norberg *et al.,* 2007; Peacock *et al.,* 2014; ICO, 2017). This resignation, compounded by the Opacity of AI decision-making, erodes the foundations of governance and raises ethical concerns in IT environment.

These challenges are further amplified within cloud-based ecosystems, where AI deployments are most relevant. As of 2021, nearly 80% of organizations reported cloud security incidents from systemic vulnerabilities (Edge, 2024). Although regulatory frameworks such as the EU AI Act and the NIST AI Risk Management Frameworks (AI RMF) mandate stricter controls, compliance remains low, with only 12% of AI-adopting firms implementing governance models (Writz *et al.,* 2022; McIntosh *et al.,* 2024). This gap between policy and practice not only weakens IT governance but also exposes sensitive organizational and personal data to breaches and misuse.

Privacy-Enhancing Technologies (PETs) such as differential privacy, federated learning, and homomorphic encryption offer potential solutions by embedding privacy directly into AI models (Salako *et al.,* 2024). Yet, their integration into IT governance frameworks and assurance practices remains underdeveloped, leaving auditors with limited tools to evaluate compliance, fairness, and accountability in AI deployments. As a result, this study focuses on the following objectives; to examine privacy-preserving AI techniques that strengthen IT governance and compliance, to identify IT governance controls required for internal IT auditors in privacy-preserving AI deployments.

➢ *Research Objectives*

The aim of this study is to

- examine privacy-preserving AI techniques that strengthen IT governance and compliance
- identify IT governance controls required for internal IT auditors in privacy-preserving AI deployments

➢ *Research Questions*

- How can privacy-preserving AI strengthen IT governance and regulatory compliance?
- What IT governance control does intern IT auditors require in privacy-preserving AI deployments.

II. LITERATURE REVIEW

➢ *Conceptual Review*

- *AI Governance in IT*

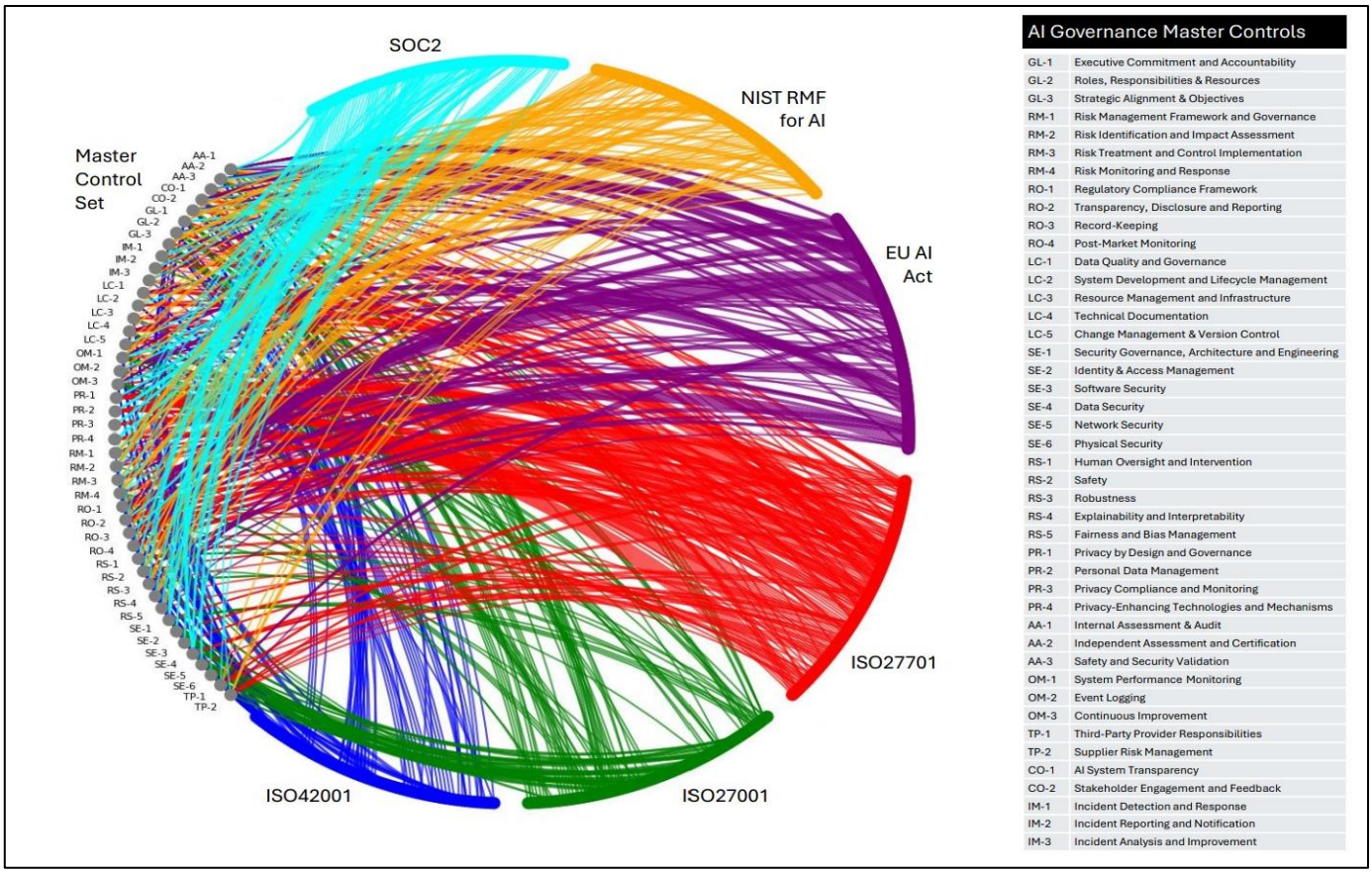
The emergence of Artificial Intelligence (AI) in IT environments has accelerated the need for governance that ensure transparency, accountability, and compliance. Traditional IT governance framework provides foundational principles for aligning IT processes with organizational objectives, but the advent of AI introduces new risk such as bias, explainability gap, and privacy concerns (Wilkin & Chenhall, 2020). Therefore, rethinking IT governance through an AII lens is imperative.

Historically, frameworks like COBIT, ITIL, and ISO/IEC 27001 have guided IT governance across strategic, operational, and security domains. COBIT, first published in 1996, evolved from an audit-focused tool into a governance framework. Its latest iteration, COBIT 2019, emphasizes value creation through benefits realization, risk optimization, and resource optimization, while differentiating governance from management (ISACA, 2019). Also, ITIL v4 advances IT service management by embedding agile and DevOps practices, ensuring responsiveness in dynamic digital ecosystems (Rubio & Arcilla, 2020). On the security front, ISO/IEC 27001 and ISO/IEC 27002 establish internationally recognized controls for confidentiality, integrity, and availability of information assets, crucial for AI systems handling sensitive data (Hamdi *et al.,* 2019).

Despite their strength, these frameworks lack mechanism for addressing AI-specific governance challenges. Recent developments in AI regulation, such as the EU AI Act and the NIST AI Risk Management Frameworks, attempts to bridge this gap by classifying AI systems based

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/25oct209

on risk and emphasizing transparency and human oversight (European Commission, ,2021; NIST, 2023). However, gaps persist in operationalizing privacy-preserving AI (PPAI) within IT assurance and audit practices. For example, while privacy-enhancing technologies like differential privacy and federated learning offer technical safeguards, governance frameworks provide limited guidance on their auditability and compliance integration (Kairouz *et al.,* 2021). The figure below maps AI governance frameworks (GDPR, NIST, ISO, EU AI Act) onto unified master controls for compliance and oversight.



AI Governance Master Controls	
GL-1	Executive Commitment and Accountability
GL-2	Roles, Responsibilities & Resources
GL-3	Strategic Alignment & Objectives
RM-1	Risk Management Framework and Governance
RM-2	Risk Identification and Impact Assessment
RM-3	Risk Treatment and Control Implementation
RM-4	Risk Monitoring and Response
RO-1	Regulatory Compliance Framework
RO-2	Transparency, Disclosure and Reporting
RO-3	Record-Keeping
RO-4	Post-Market Monitoring
LC-1	Data Quality and Governance
LC-2	System Development and Lifecycle Management
LC-3	Resource Management and Infrastructure
LC-4	Technical Documentation
LC-5	Change Management & Version Control
SE-1	Security Governance, Architecture and Engineering
SE-2	Identity & Access Management
SE-3	Software Security
SE-4	Data Security
SE-5	Network Security
SE-6	Physical Security
RS-1	Human Oversight and Intervention
RS-2	Safety
RS-3	Robustness
RS-4	Explainability and Interpretability
RS-5	Fairness and Bias Management
PR-1	Privacy by Design and Governance
PR-2	Personal Data Management
PR-3	Privacy Compliance and Monitoring
PR-4	Privacy-Enhancing Technologies and Mechanisms
AA-1	Internal Assessment & Audit
AA-2	Independent Assessment and Certification
AA-3	Safety and Security Validation
OM-1	System Performance Monitoring
OM-2	Event Logging
OM-3	Continuous Improvement
TP-1	Third-Party Provider Responsibilities
TP-2	Supplier Risk Management
CO-1	AI System Transparency
CO-2	Stakeholder Engagement and Feedback
IM-1	Incident Detection and Response
IM-2	Incident Reporting and Notification
IM-3	Incident Analysis and Improvement

Fig 1 AI Governance Master Controls
Source, Szarmach, J. (2025)

- *Privacy-Preserving AI in IT*

The adoption of AI in IT systems present governance and compliance challenges, particularly around privacy and security. Traditional centralized data collection models expose organizations to risk such as breaches, tampering, and regulatory non-compliance, especially when dealing with sensitive information like financial transactions or patient records. To address these challenges, privacy-preserving AI (PPAI) has emerged as a critical enabler of responsible AI governance.

Federated Learning (FL) represents a paradigm shift in distributed AI model training. Instead of aggregating raw data in centralized servers, FL allows model to be trained locally across multiple clients, with only model updated shared and aggregated (McMahan *et al.,* 2017; Rieke *et al.,* 2020). This approach preserves data confidentiality while enabling learning across organizations, making it relevant for critical infrastructures (Li *et al.,* 2023). In IT governance, FL reduces the privacy risk of shadow AI and support compliance by minimizing data exposure.

Complementary to federated learning, blockchain-based privacy frameworks enhance the trustworthiness of federated systems by providing audit trials, decentralized storage, and encryption-based access control (Wang *et al*., 2022; Chang *et al.,* 2021). Techniques such as homomorphic encryption enable ciphertext-level model aggregation, further securing AI deployments. Similarly, de-identification frameworks using permissioned blockchains grant individuals control over sensitive identifiers while maintaining data utility (Jennath *et al*., 2020).

Other anonymization strategies, such as K-anonymity and multi-layer distributed ledgers, provide additional safeguards by implementing ledgers, provide additional safeguards by preventing re-identification and ensuring secure distributed data sharing (Long *et al*., 2020; Tang *et al*., 2022). These approaches are increasingly relevant for IT auditors' task with verifying compliance with privacy regulations while ensuring systems reliability. Figure ii below illustrates how privacy-preserving AI techniques interconnect with encryption methods, ethical considerations, and governance challenges.

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/25oct209

Fig 2 AI Privacy and Encryption
Source, Mindmapai, (2025)

- *IT Assurance & Audit in AI*

The rapid adoption of AI within IT systems has created new complexities for assurance and audit functions. Traditional IT audit frameworks such as COBIT 2019 and ISO/IEC 27001 are designed for deterministic systems, yet AI introduces opacity, bias, and explainability challenges that undermine auditability (Rahwan *et al*., 2019). As a result, IT auditors face difficulties in verifying accountability, fairness, and compliance of AI-driven processes (Floridi and Taddeo, 2016).

Emerging thought leadership highlights the need for AI-specific assurance models. ISACA, (2020) stresses that AI assurance should extend beyond technical accuracy to include ethical accountability, transparency, and compliance alignment. Explainable AI methods are viewed as enablers of

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/25oct209

auditability, allowing internal auditors to evaluate decision path within IT systems (Arrieta *et al*., 2020). Also, privacy-preserving AI techniques such as differential privacy and federated learning can strengthen IT governance by embedding compliance into model design (Brundage *et al*., 2020).

In addition, IT assurance in AI requires a multi-layered approach that integrates governance, model transparency, and privacy safeguards. Internal auditors must adapt their evaluation criteria to cover not only data integrity and access controls but also the algorithmic accountability and social implications of AI systems (Jobin *et al*., 2019).

III. METHODOLOGY

This study adopts qualitative exploratory research design, focusing on desk research and framework development. The choice of design is justified by the emerging nature of privacy-preserving AI (PPAI) within IT governance, where little empirical evidence exists but a wealth of conceptual and practical materials can be synthesized.

This study adopts a qualitative research design to explore how privacy-preserving artificial intelligence (PPAI) can strengthen IT governance and regulatory compliance, and to identify governance control internal auditors require for PPAI deployments. The approach is grounded in an interpretivist paradigm, focusing on synthesizing insight from academic and regulatory sources to develop a conceptual framework.

The data source for this study includes nine (9) articles and three (3) reputable website sources that provide insight into privacy-preserving AI practices. In addition, key regulatory frameworks such as GDPR, EU AI Act, the NIST AI RMF, and ISO/IEC standards are analysed to align governance practices with compliance requirements. TO address audit and assurance dimensions, established IT governance standard including COBIT, ISACA guidelines, and internal audit principles are incorporated.

The analytical process involves two stages; thematic analysis and comparative mapping. Thematic analysis identifies themes across literature and frameworks, including privacy, governance, and assurance. Comparative mapping then aligns PPAI techniques with relevant IT governance control and assurance checkpoints.

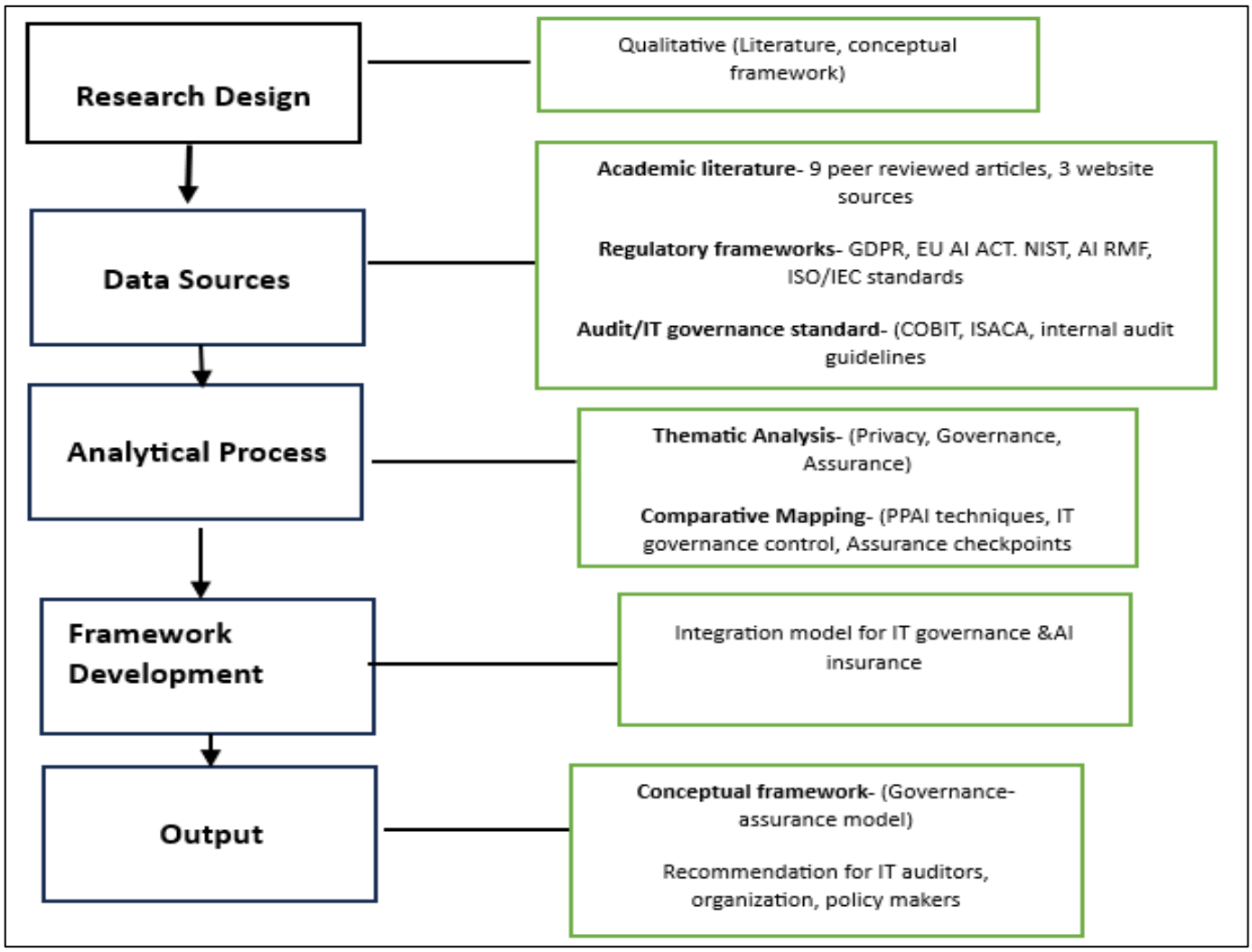


Fig 3 Methodology Framework

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/25oct209

IV. RESULTS

➢ *Hypothesis One:*

How can privacy-preserving AI strengthen IT governance and regulatory compliance?

Table 1 How Can Privacy-Preserving AI Strengthen IT Governance and Regulatory Compliance?

Theme	Description	Key author(s)
Privacy-by-Design via Federated Learning (FL)	Federated learning enables collaboration without sharing raw data, aligning with data minimization and GDPR	Zhao et al., 2025; Truong et al., 2021
Secure Aggregation & Encryption	Techniques like homomorphic encryption and SMPC ensure model updates remain confidential, even in untrusted environment	Kalodanis et al., 2025
Immutable auditing & explainability	Logging updates and decision in tamper-proof ledgers supports transparency and governance oversight	Kalodanis et al., 2025

The table highlights key themes in understanding how privacy-preserving AI can strengthen IT governance and regulatory compliance. The first theme, Privacy-by-Design via Federated Learning (FL), emphasizes that federated learning allows multiple stakeholders to collaboratively train AI models without exchanging raw data. This aligns with data minimization principles under GDPR and similar privacy raw data (Zhao *et al.,* 2025; Truong *et al.,* 2021).

The second theme, Secure Aggregation & Encryption, focuses on the technical safeguards that enhance confidentiality in distributed AI systems. Methods such as homomorphic encryption and secure multi-party computation (SMPC) allow encrypted computations and aggregation of updates, ensuring that even if communication is intercepted or environments are untrusted, data privacy remain intact (Kalodanis *et al.,* 2025).

The third theme, Immutable Auditing & Explainability, highlights the importance of transparent decision-making processes. By storing model updates and AI decisions in tamper-proof ledgers, organization can achieve traceability and accountability. This does not only satisfy audit requirements but also supports trust, as explainability mechanism make AI behaviour verifiable and defensible to regulators and stakeholders (Kalodanis *et al.,* 2025).

➢ *Hypothesis Two:*

What IT governance control does intern IT auditors require in privacy-preserving AI deployments?

Table 2 What IT Governance Control Does Intern IT Auditors Require in Privacy-Preserving AI Deployments?

Theme	Audit focus	Author(s)
Model Security & Integrity	Ensure secure model storage, access control, tamper detection, enclave usage, an adversarial resilience	Ramachandran, (2024)
Privacy in Output & Inference Controls	Validate use of differential privacy, limits on inference attack, anonymized outputs, privacy budget, and synthetic data usage	Ramachandran, (2024)
Federated Learning Controls	Review secure aggregation protocols, access validation, bias mitigation, and regulatory logs in Federated Learning deployments	Olson, (2025)
Internal Audit of AI Systems	Ensure internal audit processes include AI governance review, alignment with controls, and dynamic documentation of AI implementation	Khan, (2025)

The table outlines critical audit focus areas that internal IT auditors should verify in privacy-preserving AI deployments, highlighting how governance controls can be strengthened. The first theme, Model Security & Integrity, emphasizes the need for auditors to verify that AI models are securely stored, protected by strict access controls, and are monitored for potential tampering. In addition, safeguards such as secure enclaves and adversarial resilience strategies are essential to maintain trust in model reliability and prevent manipulations (Ramachandran, 2024).

The second theme, Privacy in Output & Inference Controls, focuses on ensuring that deployed AI models do not inadvertently leak sensitive information. Auditors should confirm the use of differential privacy techniques, limitations on inference attacks, anonymization of outputs, appropriate use of synthetic data, and management of privacy budgets (Ramachandran, 2024).

The third theme, Federated Learning Controls, highlights the governance of federated learning systems. Auditors should review secure aggregation protocols, confirm access validations, and assess whether mechanism exist to mitigate bias while maintaining regulatory logs for accountability (Olson, 2025). Finally, Internal Audit of AI Systems stresses that auditing itself must adapt. Auditors should ensure that AI governance review is integrated into standard audit cycles, with dynamic documentation and controls alignment (Khan, 2025).

V. DISCUSSION

The findings indicate that privacy-preserving AI (PPAI) can directly reinforce IT governance by operationalizing core regulatory principles through technical design. First, the emphasis on Privacy-by-Design via federated learning (FL) aligns with GDPR's data minimization and purpose-

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/25oct209

limitation mandates, as FL keeps raw personal data local while enabling collaborative model training (Kairouz *et al.,* 2021). Secure Aggregation further hardens confidentiality in distributed settings. For instance, Bonawitz *et al.,* (2017) show how updates can be aggregated without revealing individual contributions. While differential privacy (DP) offers mathematically provable guarantees against disclosure from model output (Bonawitz *et al.,* 2017; Dwork & Roth, 2014). These techniques mitigate documented risk such as membership-inference attacks (Shokri *et al.,* 2017), strengthening control effectiveness in audits of model outputs and inference channels.

Second, immutable logging and explainability map onto governance requirements for accountability and auditability. ISO/IEC 27701 extends ISO 27001 with privacy-specific controls and evidence trails supporting internal audit verification, while NIST's AI Risk Management Framework calls for traceability, transparency, and documentation across the AI lifecycles outcomes advanced by tamper-evident logging and model cards. Finally, the EU AI Act's transparency and record-keeping obligations especially for high risk and genera-purpose AI highlights the need for auditable pipelines, secure aggregation records, and clear instructions for deployers, PPAI controls directly facilitate these compliance outcomes.

VI. CONCLUSION

This study set out to explore how privacy-preserving AI (PPAI) can strengthen IT governance and regulatory compliance while also determining the governance controls internal IT auditors should verify in AI deployments. The findings reveals that PPAI operationalizes regulatory principles such as data minimization, confidentiality, and accountability through advanced technical safeguards. Federated Learning (FL) demonstrates clear alignment with GDPR and similar privacy regulations by preventing unnecessary data transfers, while technique such as homomorphic encryption, secure aggregation, and differential privacy address confidentiality and output-level risks. Moreover, immutable logging and explainability mechanisms provide auditable trails that reinforce accountability and transparency, key requirements under ISO/IEC 27701, NIST AI RMF, and the EU AI Act.

From an IT assurance perspective, internal auditors must adapt audit practices to evaluate the integrity of AI models, privacy in outputs, federated learning deployments, and overall, AI governance integration. This highlights a shift from traditional IT audits toward AI-specific assurance models, where resilience, explainability, and compliance converge. In addition, PPAI is not merely a technical safeguard but an enabler of governance and compliance, bridging the gap between high-level policy frameworks and operational IT execution.

VII. LIMITATIONS OF STUDY

- Federated learning though privacy-preserving, may not fully satisfy regulatory requirements (e.g., auditability under GDPR).
- Federated learning is still susceptible to model poisoning, inference attacks, and backdoor attack
- There are communication overhead as frequent updates between clients and central server can strain bandwidth and increase latency.

- *Suggestion for Future Research*

- Further research should examine how federated learning can align with global compliance in terms of legal accountability and organizational oversight.
- Domain such as healthcare, finance, and public governance have regulatory and ethical demands, these sectors should be explored for further research for specific framework for federated learning adoption especially in areas where trust and accountability are important
- Future research should focus on developing mechanism across different source as federal learning operate on heterogenous and often imbalanced datasets
- Regulators and internal auditors face challenges in verifying compliance when data remain decentralized. Further study could explore federated learning and blockchain-enabled logging systems that enhance transparency without compromising privacy.

RECOMMENDATION

Based on the findings, several recommendations are proposed for IT organizations, auditors, and policymakers.

- In IT organization, PPAI techniques such as federated learning, homomorphic encryption, and differential privacy should be embedded into IT operations as default safeguards. Privacy-by-design must become an integral part of AI system development and deployment to ensure resilience and compliance readiness.
- For IT auditors, assurance processes should expand to include AI-specific checkpoints. This entails verifying secure model storage, privacy-preserving outputs, federated learning protocols, and governance aligned documentation. Dynamic audit frameworks should be developed to align with evolving AI risks and regulatory requirements.
- In addition, IT governance standard such as COBIT, ISO/IEC 27001, and NIST frameworks should be updated to integrate AI-specific privacy-preserving requirements, particularly around explainability, secure aggregation, and auditability.
- Finally, for future research, empirical testing of the proposed governance-aligned frameworks in real IT environments such as cloud infrastructures, cybersecurity monitoring, and enterprise IT systems is necessary to validate effectiveness and scalability. This will advance both academic and practice knowledge in AI governance and assurance.

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/25oct209

REFERENCES

[1]. Arrieta, A.B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., García, S., Gil-López, S., Molina, D., Benjamins, R. and Chatila, R., 2020. Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information fusion*, *58*, pp.82-115.

[2]. Bonawitz, K., Ivanov, V., Kreuter, B., Marcedone, A., McMahan, H.B., Patel, S., Ramage, D., Segal, A. and Seth, K., 2017, October. Practical secure aggregation for privacy-preserving machine learning. In *proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security* (pp. 1175-1191).

[3]. Brundage, M., Avin, S., Clark, J., Toner, H., Eckersley, P., Garfinkel, B., Dafoe, A., Scharre, P., Zeitzoff, T., Filar, B. and Anderson, H., 2018. The malicious use of artificial intelligence: Forecasting, prevention, and mitigation. *arXiv preprint arXiv:1802.07228*.

[4]. Chang, Y., Fang, C. and Sun, W., 2021. A Blockchain-Based Federated Learning Method for Smart Healthcare. *Computational Intelligence and Neuroscience*, *2021*(1), p.4376418.

[5]. Chin, T., Li, Q., Mirone, F. and Papa, A., 2025. Conflicting impacts of shadow AI usage on knowledge leakage in metaverse-based business models: A Yin-Yang paradox framing. *Technology in Society*, *81*, p.102793.

[6]. De Haes, S., Van Grembergen, W., Joshi, A. and Huygh, T., 2019. COBIT as a Framework for Enterprise Governance of IT. In *Enterprise governance of information technology: Achieving alignment and value in digital organizations* (pp. 125-162). Cham: Springer International Publishing.

[7]. Dwork, C. and Roth, A., 2014. The algorithmic foundations of differential privacy. *Foundations and trends® in theoretical computer science*, *9*(3–4), pp.211-407.

[8]. Edge, (2024). Top Cloud Security Statistics in 2024. *Microsoft* (no date) *Bing*. Available at: https://www.bing.com/ck/a?\ (Accessed: August 24, 2025).

[9]. Floridi, L. and Taddeo, M., 2016. What is data ethics?. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *374*(2083), p.20160360.

[10]. Hamdi, Z., Norman, A.A., Molok, N.N.A. and Hassandoust, F., 2019, December. A comparative review of ISMS implementation based on ISO 27000 series in organizations of different business sectors. In *Journal of Physics: Conference Series* (Vol. 1339, No. 1, p. 012103). IOP Publishing.

[11]. ISACA, (2019): COBIT 2019 Framework: Introduction and Methodology. ISACA. Schaumburg, IL.

[12]. ISO/IEC, (2017). ISO/IEC 38502:2017 Information technology - Governance of IT – Framework and Model. Geneva, Switzerland: ISO/IEC.

[13]. Jobin, A., Ienca, M. and Vayena, E., 2019. The global landscape of AI ethics guidelines. *Nature machine intelligence*, *1*(9), pp.389-399.

[14]. Kairouz, P., McMahan, H.B., Avent, B., Bellet, A., Bennis, M., Bhagoji, A.N., Bonawitz, K., Charles, Z., Cormode, G., Cummings, R. and D'Oliveira, R.G., 2021. Advances and open problems in federated learning. *Foundations and trends® in machine learning*, *14*(1–2), pp.1-210.

[15]. Kalodanis, K., Feretzakis, G., Anastasiou, A., Rizomiliotis, P., Anagnostopoulos, D. and Koumpouros, Y., 2025. A Privacy-Preserving and Attack-Aware AI Approach for High-Risk Healthcare Systems Under the EU AI Act. *Electronics*, *14*(7), p.1385.

[16]. Khan, M.S. (2025). *How to audit AI and autonomous agents: A practical guide for internal auditors and GRC teams*, *Linkedin.com*. Available at: https://www.linkedin.com/pulse/how-audit-ai-autonomous-agents-practical-guide-internal-khan-av3mf (Accessed: August 23, 2025).

[17]. Long, Y., Chen, Y., Ren, W., Dou, H. and Xiong, N.N., 2020. Depet: A decentralized privacy-preserving energy trading scheme for vehicular energy network via blockchain and k-anonymity. *Ieee Access*, *8*, pp.192587-192596.

[18]. McIntosh, T.R., Susnjak, T., Liu, T., Watters, P., Xu, D., Liu, D., Nowrozy, R. and Halgamuge, M.N., 2024. From cobit to iso 42001: Evaluating cybersecurity frameworks for opportunities, risks, and regulatory compliance in commercializing large language models. *Computers & Security*, *144*, p.103964.

[19]. McMahan, B., Moore, E., Ramage, D., Hampson, S. and y Arcas, B.A., 2017, April. Communication-efficient learning of deep networks from decentralized data. In *Artificial intelligence and statistics* (pp. 1273-1282). PMLR.

[20]. Mindmapai, (2025). AI, privacy, and encryption: A comprehensive guide. *Mindmapai.app*. Available at: https://mindmapai.app/mind-mapping/ai-privacy-and-encryption (Accessed: September 12, 2025).

[21]. Norberg, P.A., Horne, D.R. and Horne, D.A., 2007. The privacy paradox: Personal information disclosure intentions versus behaviors. *Journal of consumer affairs*, *41*(1), pp.100-126.

[22]. Olson, D. (2025) *Federated learning and privacy-preserving AI*, *Linkedin.com*. Available at: https://www.linkedin.com/pulse/federated-learning-privacy-preserving-ai-douglas-olson-yyjbc (Accessed: August 23, 2025)

[23]. O'neil, C., 2017. *Weapons of math destruction: How big data increases inequality and threatens democracy*. Crown.

[24]. Peacock, S.E., 2014. How web tracking changes user agency in the age of Big Data: The used user. *Big Data & Society*, *1*(2), p.2053951714564228.

[25]. Rahwan, I., Cebrian, M., Obradovich, N., Bongard, J., Bonnefon, J.F., Breazeal, C., Crandall, J.W., Christakis, N.A., Couzin, I.D., Jackson, M.O. and Jennings, N.R., 2019. Machine behaviour. *Nature*, *568*(7753), pp.477-486.

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/25oct209

[26]. Ramachandran, A. (2024) *The transformative impact of artificial intelligence on internal controls, controls audit procedures and testing: A comprehensive analysis*, *Linkedin.com*. Available at: https://www.linkedin.com/pulse/transformative-impact-artificial-intelligence-audit-ramachandran-zbwqe (Accessed: August 23, 2025).

[27]. Rieke, N., Hancox, J., Li, W., Milletari, F., Roth, H.R., Albarqouni, S., Bakas, S., Galtier, M.N., Landman, B.A., Maier-Hein, K. and Ourselin, S., 2020. The future of digital health with federated learning. *NPJ digital medicine*, *3*(1), p.119.

[28]. Rubio, J.L. and Arcilla, M., 2019. How to optimize the implementation of itil through a process ordering algorithm. *Applied Sciences*, *10*(1), p.34.

[29]. Salako, A.O., Fabuyi, J.A., Aideyan, N.T., Selesi-Aina, O., Dapo-Oyewole, D.L. and Olaniyi, O.O., 2024. Advancing information governance in AI-driven cloud ecosystem: Strategies for enhancing data security and meeting regulatory compliance. *Asian Journal of Research in Computer Science*, *17*(12), pp.66-88.

[30]. Shokri, R., Stronati, M., Song, C. and Shmatikov, V., 2017, May. Membership inference attacks against machine learning models. In *2017 IEEE symposium on security and privacy (SP)* (pp. 3-18). IEEE.

[31]. Szarmach, J. (2025). AI governance controls mega-map, *AI Governance Library*. Available at: https://www.aigl.blog/ai-governance-controls-mega-map-feb-2025/ (Accessed: September 12, 2025).

[32]. Tang, X., Zhu, L., Shen, M., Peng, J., Kang, J., Niyato, D. and Abd El-Latif, A.A., 2022. Secure and trusted collaborative learning based on blockchain for artificial intelligence of things. *IEEE Wireless Communications*, *29*(3), pp.14-22.

[33]. Truong, N., Sun, K., Wang, S., Guitton, F. and Guo, Y., 2021. Privacy preservation in federated learning: An insightful survey from the GDPR perspective. *Computers & Security*, *110*, p.102402.

[34]. Wang, N., Yang, W., Wang, X., Wu, L., Guan, Z., Du, X. and Guizani, M., 2024. A blockchain based privacy-preserving federated learning scheme for Internet of Vehicles. *Digital Communications and Networks*, *10*(1), pp.126-134.

[35]. Wilkin, C.L. and Chenhall, R.H., 2020. Information technology governance: Reflections on the past and future directions. *Journal of Information Systems*, *34*(2), pp.257-292.

[36]. Wirtz, B.W., Weyerer, J.C. and Kehl, I., 2022. Governance of artificial intelligence: A risk and guideline-based integrative framework. *Government information quarterly*, *39*(4), p.101685.

[37]. Yilmaz, E. and Can, O., 2024. Unveiling shadows: Harnessing artificial intelligence for insider threat detection. *Engineering, Technology & Applied Science Research*, *14*(2), pp.13341-13346.

[38]. Zhao, J., Bagchi, S., Avestimehr, S., Chan, K., Chaterji, S., Dimitriadis, D., Li, J., Li, N., Nourian, A. and Roth, H., 2025. The federation strikes back: A survey of federated learning privacy attacks, defenses, applications, and policy landscape. *ACM Computing Surveys*, *57*(9), pp.1-37.