Journal of Applied Artificial Intelligence
https://jaai.sabapub.com
ISSN: 2709-5908

2025 Volume 6, Issue 2: 60–70

DOI: https://doi.org/10.48185/jaai.v6i2.1918

Privacy-Preserving Artificial Intelligence: Principles, Methods, Applications, and Challenges

Olayemi Oladimeji Olasehinde[1,3*], **Boniface Kayode Alese**[2], **Ojonukpe Sylvester Eqwuche**[3]

[1] Department of Computer Science,, University of Huddersfield, Huddersfield, United Kingdom
[2] Department of Cyber Security, Federal University of Technology, Akure, Nigeria
[3] Department of Computer Science, Federal Polytechnic, Ile-Oluji Nigeria.

Received: 10.11.2025 • Accepted: 26.12.2025 • Published: 30.12.2025 • Final Version: 31.12.2025

Abstract: Artificial intelligence systems increasingly rely on large volumes of sensitive data to support decision making in domains such as healthcare, finance, education, and public administration. While these systems offer substantial benefits, their growing dependence on personal information has intensified concerns about privacy, data misuse, and loss of public trust. This study examines privacy preserving artificial intelligence as a design approach that enables meaningful data analysis while limiting exposure of sensitive information. The paper analyses key privacy preserving techniques, including federated learning, differential privacy, homomorphic encryption, and secure multi party computation, and evaluates their relevance across critical application domains. It also identifies practical challenges related to computational cost, data heterogeneity, regulatory compliance, and explainability. The study shows that privacy preserving methods can support responsible and trustworthy artificial intelligence when privacy, utility, and governance considerations are addressed together.

Keywords: Federated Learning, data privacy, secure computation, distributed learning, Homomorphic Encryption, Differential Privacy

1.0 INTRODUCTION

Artificial intelligence is now widely used in many areas of daily life, including healthcare, banking, education, and cybersecurity. These systems rely on large volumes of data collected from people, devices, and digital platforms. As AI systems become more advanced and more deeply integrated into society, important questions arise about how personal data is collected, used, stored, and protected [1].

Weak privacy protection can cause serious harm. Data breaches may lead to identity theft, financial loss, and unwanted surveillance. In some cases, AI models can even memorise sensitive data, allowing attackers to recover personal information using techniques such as membership inference or data reconstruction [2], [3]. These risks reduce public trust and limit the acceptance of AI technologies.

Privacy is closely linked to fairness, dignity, and individual autonomy. Strong privacy safeguards help people maintain control over their data and reduce the risk of discrimination or misuse. They also support compliance with laws such as the General Data Protection Regulation and the HIPAA Privacy Rule, which clearly define how data should be handled [4], [5]. When people trust that their data is treated responsibly, they are more willing to support data-driven innovation [6].

* Corresponding Author: olasehindeolayemi@yahoo.com

Traditional privacy approaches that rely only on access control or after-the-fact protections are no longer sufficient. At the same time, overly strict rules can limit collaboration and reduce system performance. This creates a need for AI systems that protect privacy by design rather than as an afterthought.

Privacy-preserving artificial intelligence addresses this need by allowing useful data analysis without exposing sensitive information. By combining techniques from machine learning, cryptography, and distributed systems, these methods help maintain privacy while still delivering useful results [7], [8]. This paper asks a central question: how can privacy-preserving AI be designed and applied in ways that protect sensitive data while remaining useful, lawful, and trustworthy?

2.0 Privacy-Preserving Framework

Figure 1 introduces a circular and iterative framework in which ethical and regulatory principles guide the development of privacy-preserving AI applications, practical and technical challenges are identified during implementation, and resulting outcomes inform continuous refinement. The framework emphasises privacy by design, showing how principles, applications, challenges, and outcomes interact to support trustworthy, compliant, and socially responsible artificial intelligence systems.

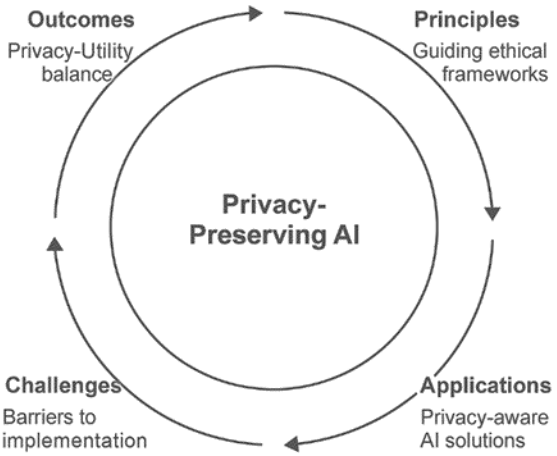

Figure 1. Privacy Preserving Artificial Intelligence Framework.

The framework illustrates how privacy-preserving AI evolves through continuous improvement as technologies, risks, and regulatory expectations change [9], [10]. At its centre is the goal of protecting sensitive information while delivering accurate and socially responsible outcomes. Surrounding this goal are four interconnected stages that operate as a feedback-driven cycle rather than a linear process.

The first stage, Principles, represents the ethical and regulatory foundations that guide responsible system development. These include fairness, transparency, accountability, data minimisation, and purpose limitation, which are widely emphasised in global regulatory frameworks and academic research on trustworthy AI [4], [11], [12]. Embedding these principles at the architectural level ensures privacy is integral to system design rather than an afterthought.

The second stage, Applications, demonstrates how privacy principles are translated into practice across critical sectors. Privacy-preserving AI methods are increasingly used in healthcare, finance,

cybersecurity, education, and public administration to analyse sensitive data without violating confidentiality constraints [13], [14]. For example, federated learning enables cross-hospital medical research without centralising patient records, while differential privacy supports safer population-level analysis in government and industry settings [6].

The third stage, Challenges, captures the technical and operational difficulties associated with deploying privacy-preserving methods at scale. Techniques such as homomorphic encryption and secure multi-party computation remain computationally expensive, limiting real-time feasibility [7], [15]. Federated learning also introduces challenges related to data heterogeneity, communication overhead, and adversarial behaviour [16]. Balancing strong privacy guarantees with acceptable model performance remains a persistent concern, particularly in high-stakes domains [17].

The fourth stage, Outcomes, reflects the benefits achieved when principles, applications, and challenges are effectively aligned. Empirical studies demonstrate that strong privacy protection can coexist with high model performance, regulatory compliance, and public trust [8], [13]. These outcomes enable collaboration across institutions that previously could not share sensitive data, supporting innovation in areas such as medical imaging and threat detection [18], [19].

The continuous interaction among these stages aligns with responsible AI governance models that emphasise ongoing evaluation rather than one-time compliance [20], [21]. This cyclical process allows privacy-preserving AI systems to adapt to evolving threats, technologies, and societal expectations.

3.0 Principles of Privacy-Preserving AI

Effective privacy-preserving artificial intelligence begins with clear principles governing data collection, processing, sharing, and protection throughout the system lifecycle. These principles ensure alignment between technical development, ethical responsibility, and legal obligations, positioning privacy at the core of system design [9], [11].

Key principles include data minimisation and purpose limitation, anonymisation and de-identification, secure computation and data sovereignty, and transparency, accountability, and fairness. Together, these principles provide a structured foundation for responsible data use and trustworthy AI deployment.

3.1 Data Minimization and Purpose Limitation

Data minimisation limits the amount of personal data collected and ensures it is used only for clearly defined purposes. This principle reduces exposure to ethical and security risks and is central to modern data protection regulation and research [22]. Collecting only essential information reduces the likelihood of misuse, unauthorised access, and unintended secondary use.

Privacy engineering research identifies minimal data exposure as a key protective strategy that reduces attack surfaces and accidental disclosure risks [6]. For example, clinical AI systems should access only data necessary for diagnosis or treatment, avoiding unrelated personal attributes unless clinically justified [22]. These practices align with regulatory expectations under GDPR and OECD guidance [4], [11].

3.2 Anonymization and De Identification

Anonymisation reduces privacy risk by removing or modifying attributes that can directly identify individuals. Common techniques include masking, tokenisation, generalisation, and aggregation,

each aiming to prevent direct linkage between records and specific persons. When applied carefully, anonymisation allows data to be reused for research, statistical analysis, and policy evaluation while lowering the likelihood of exposing personal information [23].

However, anonymisation alone does not guarantee privacy. Empirical evidence shows that individuals can often be re-identified when anonymised datasets are combined with external sources that share overlapping attributes [24]. To address this limitation, statistical privacy models such as *k-anonymity*, *l-diversity*, and *t-closeness* introduce formal safeguards that enforce similarity and diversity within data groups, thereby reducing the probability of successful re-identification [23].

3.3 Secure Computation and Data Sovereignty

Secure computation enables organisations to collaborate on data analysis without revealing raw or sensitive records. Closely linked to this is data sovereignty, which ensures that data owners retain control over where their data is stored, how it is processed, and who can access it. This control is increasingly important in cross-border settings and in sectors subject to strict regulatory oversight [14].

Techniques such as federated learning and secure multi-party computation support these goals by allowing distributed model training while keeping sensitive data at its original location [8], [25]. By embedding protection directly into the computation process, these methods follow a privacy-by-design approach and help organisations comply with data localisation requirements and sector-specific regulations, particularly in healthcare and finance [4], [13].

3.4 Transparency, Accountability, and Fairness

Transparency and accountability are essential for maintaining trust and ensuring responsible AI behaviour. Clear documentation, audit trails, and governance processes enable stakeholders to understand how data is used and how automated decisions are produced, supporting effective oversight and regulatory compliance [9].

Fairness-aware AI further strengthens privacy protection by reducing the risk that protected attributes are inferred indirectly through correlated features. Without explicit fairness controls, models may unintentionally produce biased or discriminatory outcomes even when sensitive variables are excluded [26]. Explainability techniques support accountability by helping stakeholders interpret model decisions, identify unfair behaviour, and apply corrective measures. Together, transparency, fairness, and accountability reinforce ethical deployment while complementing strong privacy safeguards.

4.0 Methods of Privacy Preserving AI

Privacy-preserving methods translate ethical and regulatory principles into concrete technical safeguards that can be applied throughout the AI lifecycle. Rather than relying on post hoc controls, these methods embed protection directly into data processing and model training workflows. Key approaches include federated learning, differential privacy, homomorphic encryption, secure multi-party computation, and synthetic data generation [8]. Each method addresses a different class of privacy risk, such as data disclosure, inference attacks, or unauthorised data transfer. When used in

combination, they provide layered protection that strengthens privacy while supporting practical deployment in regulated environments [3], [13].

4.1 Federated Learning

Federated learning enables decentralised model training by keeping data at its original location and sharing only model updates, such as gradients or parameters, with a coordinating server [8]. This design reduces the need for centralised data collection and lowers the risk of direct data exposure. Federated learning has been widely adopted in healthcare research, including collaborative medical imaging studies, and in consumer applications such as mobile text prediction systems.

Despite these advantages, federated learning introduces several challenges. Data held by participating clients is often non-identically distributed, which can affect convergence and model performance. Communication overhead may also become significant when updates are exchanged frequently. In addition, federated systems remain vulnerable to adversarial behaviours, including poisoned updates and inference attacks on shared parameters [16].

4.2 Differential Privacy

Differential privacy provides formal, mathematically grounded guarantees by adding carefully calibrated noise to data queries or model outputs. This limits the influence of any single individual record on the final result, reducing the risk of reconstruction and membership inference attacks [6]. Differential privacy has been deployed in large-scale telemetry and analytics systems and is widely regarded as one of the most mature privacy-preserving techniques.

However, achieving an effective balance between privacy and utility remains challenging. Strong privacy budgets typically require higher noise levels, which can degrade model accuracy and reduce usefulness in sensitive applications. As a result, careful parameter selection and task-specific tuning are essential to ensure that privacy protection does not undermine analytical value [27].

4.3 Homomorphic Encryption

Homomorphic encryption allows computations to be performed directly on encrypted data, ensuring that sensitive information remains protected throughout the entire processing pipeline [7], [28] This capability is particularly valuable in scenarios involving outsourced computation or cross-organisational collaboration, such as financial risk analysis and secure healthcare analytics.

Despite its strong security guarantees, homomorphic encryption remains computationally demanding. Fully homomorphic schemes, which support arbitrary computations, often incur significant performance overhead. Ongoing research therefore focuses on optimised, partially homomorphic, or application-specific schemes to improve efficiency and scalability [15].

4.4 Secure Multi Party Computation (SMPC)

Secure multi-party computation enables multiple parties to jointly compute a function without revealing their individual inputs to one another. This allows collaborative analysis and model

training under strict confidentiality constraints, making SMPC suitable for domains where direct data sharing is prohibited [25].

While SMPC provides strong privacy guarantees, it introduces additional communication and computation overhead. These costs can limit scalability, particularly for large datasets or complex machine learning models, and must be carefully managed in real-world deployments.

4.5 Synthetic Data Generation

Synthetic data generation produces artificial datasets that preserve the statistical characteristics of real data while avoiding direct exposure of original records [29]. This approach supports data sharing, testing, and benchmarking in environments where access to real data is restricted by privacy or regulatory concerns.

However, synthetic data is not inherently risk-free. If generative models overfit or replicate rare patterns too closely, privacy leakage may still occur. Consequently, synthetic datasets must be evaluated for both utility and disclosure risk, and are increasingly combined with formal privacy mechanisms to improve safety and governance [30].

5.0 Applications of Privacy Preserving AI

The practical significance of privacy-preserving artificial intelligence becomes most evident when these methods are deployed in real operational environments. Organisations that process personal, financial, or behavioural data increasingly depend on AI systems that must operate under strict privacy constraints while still delivering accurate, timely, and reliable outcomes. As data volumes grow and regulatory oversight intensifies, privacy-preserving techniques have shifted from optional safeguards to essential design components in modern data-driven systems [13].

Across domains such as healthcare, finance, cybersecurity, education, and smart city management, privacy regulations are stringent and the consequences of data misuse are substantial. Privacy-preserving AI enables institutions in these sectors to collaborate, extract insight, and make informed decisions without exposing sensitive information or violating regulatory requirements. This capability supports compliance with legal and ethical frameworks while enabling innovation in contexts where traditional data sharing is either restricted or prohibited.

Importantly, these applications demonstrate that privacy is not an impediment to technological progress. Instead, it serves as a foundation for trustworthy and sustainable AI systems. When privacy is embedded into system architecture and operational workflows, organisations are more likely to adopt advanced AI solutions with confidence, transparency, and public trust.

5.1 Healthcare

Healthcare applications of privacy-preserving AI focus on enabling collaborative analysis of sensitive clinical data without violating patient confidentiality. Federated learning has emerged as a particularly effective approach, allowing hospitals and research centres to train shared models while keeping all patient records within local infrastructure.

This distributed training paradigm improves model generalisation across diverse populations while remaining compliant with regulations such as the General Data Protection Regulation and the Health Insurance Portability and Accountability Act [13], [18]. Empirical studies in medical imaging and

disease diagnosis demonstrate that federated learning can achieve performance comparable to centralised training without exposing patient data.

5.2 Finance

In the financial sector, privacy-preserving AI supports secure collaboration on fraud detection, credit scoring, and risk assessment tasks that involve highly sensitive transactional data. Techniques such as secure multi-party computation and homomorphic encryption enable institutions to perform joint analytics without revealing individual customer information or proprietary datasets [7], [25].

These methods reduce exposure to insider threats and third-party risks while strengthening compliance with financial privacy regulations. By allowing computation over encrypted or distributed data, privacy-preserving approaches enhance trust and enable advanced analytics across organisational boundaries.

5.3 Cybersecurity

Cybersecurity applications increasingly rely on privacy-preserving AI to enable cooperative threat detection without disclosing sensitive network telemetry or operational details. Federated and differentially private intrusion detection systems allow multiple organisations to learn from shared attack patterns while keeping raw logs local [31], [32]. Such approaches are particularly valuable in distributed and encrypted network environments, where direct data sharing may introduce security or confidentiality risks. Privacy-preserving threat intelligence sharing improves collective defence capabilities while maintaining organisational autonomy.

5.4 Education

In education, privacy-preserving AI supports learning analytics and student performance modelling while protecting sensitive academic and behavioural data. Federated learning allows institutions to jointly develop predictive models for early warning systems and personalised learning while keeping student records on local systems, enabling collaboration without centralising sensitive academic data and preserving student privacy [33], [34].

This approach preserves confidentiality, supports regulatory compliance, and increases trust in educational technologies. As a result, institutions can leverage data-driven insights to improve student outcomes while maintaining ethical and legal standards for data protection.

6.0 Challenges and Open Issues

Despite notable progress, several challenges continue to constrain the effective deployment of privacy-preserving artificial intelligence in real-world systems. Strong privacy protection mechanisms can reduce model accuracy, particularly when applied to small, imbalanced, or highly complex datasets [17], [35]. This privacy–utility trade-off remains a central concern in domains such as healthcare, finance, and cybersecurity, where even modest performance degradation may have significant operational or societal consequences.

Computational complexity is another persistent limitation. Privacy-preserving techniques such as homomorphic encryption and secure multi-party computation introduce substantial processing and communication overheads, which hinder scalability and limit suitability for time-critical or large-scale

applications [15]. Although recent optimisations have improved efficiency, these methods still impose higher costs than conventional machine learning pipelines.

Additional challenges arise from data heterogeneity, limited explainability, and fragmented regulatory environments. Variations in data quality and distribution across participating entities complicate federated learning and may introduce bias or unstable convergence [16]. At the same time, privacy mechanisms can obscure internal model behaviour, reducing transparency and making auditing, accountability, and bias detection more difficult. Differences in privacy regulations across jurisdictions further slow adoption by creating uncertainty for organisations operating across borders.

Addressing these challenges requires advances in algorithmic efficiency, adaptive privacy mechanisms, and system-level design, alongside closer cooperation between researchers, practitioners, and policymakers. Such coordinated efforts are essential for developing scalable, explainable, and regulation-aware privacy-preserving AI systems that can be reliably deployed in practice.

7.0 Conclusion

Privacy-preserving artificial intelligence marks a clear shift in how intelligent systems are designed and deployed, placing privacy at the centre rather than treating it as an afterthought. Modern AI systems must balance performance with security, fairness, and regulatory compliance. This study demonstrates that it is possible to protect personal data while still enabling effective learning, which is increasingly important as AI is embedded in sensitive areas such as healthcare, finance, education, and cybersecurity. When privacy is combined with transparency and accountability, AI systems are more likely to earn public trust and social acceptance.

The analysis shows that progress in privacy-preserving AI depends on aligning ethical principles, technical solutions, and operational practices. Methods such as federated learning, differential privacy, homomorphic encryption, secure multi-party computation, and synthetic data generation address different privacy risks, but no single technique is sufficient on its own. Practical deployment therefore requires combining methods based on context, data sensitivity, and regulatory constraints, highlighting the adaptive nature of privacy-preserving AI.

Despite recent advances, important challenges remain. High computational costs, data heterogeneity, limited explainability, and fragmented regulatory frameworks continue to restrict large-scale adoption. Addressing these issues will require closer collaboration between researchers, industry practitioners, and policymakers, alongside efforts to improve efficiency, adaptability, and regulatory alignment.

Overall, privacy-preserving artificial intelligence provides a strong foundation for responsible and trustworthy AI. By embedding privacy, fairness, and accountability into system design and evaluation, future AI systems can support innovation while respecting individual rights and societal values. Continued research, standardisation, and interdisciplinary cooperation will be essential as AI adoption continues to expand.

7. 1 Future Directions

The future of privacy preserving artificial intelligence depends on finding practical ways to protect personal data while still allowing AI systems to learn and perform well. This includes improving the balance between privacy and accuracy, making privacy techniques less demanding to run, and developing clearer and more consistent rules that work across countries. It is also important that privacy protected models can still explain their decisions in ways people can understand and trust. New developments such as quantum resistant security methods, secure federated learning, and privacy aware large models are expected to help build AI systems that are reliable, transparent, and responsible in real world use.

Acknowledgements

The authors acknowledges institutional support provided in facilitating the conduct of this research. Appreciation is also extended to the developers and maintainers of the UNSW-NB15 dataset for making the dataset publicly available for academic research. No external funding was received for this study.

Author Contributions

Olayemi Oladimeji Olasehinde conceived the study and defined the research direction. Boniface Kayode Alese contributed to the theoretical framing and supervised the methodological design. Ojonukpe Sylvester Eqwuche supported the analysis of privacy preserving techniques and their application domains. All authors contributed to the reviewed of the manuscript critically for intellectual content, and approved the final version for submission.

Conflict of Interest

The author declares that there is no conflict of interest regarding the publication of this paper.

References

[1] N. Papernot, S. Song, I. Mironov, A. Raghunathan, T. Steinke, and K. Talwar, "SoK: Machine learning with membership inference," in *Proc. IEEE Symp. Security and Privacy (SP)*, San Francisco, CA, USA, 2021, pp. 1–18, doi: 10.1109/SP40001.2021.00067.

[2] R. Shokri, M. Stronati, C. Song, and V. Shmatikov, "Membership inference attacks against machine learning models," in *Proc. IEEE Symp. Security and Privacy (SP)*, San Jose, CA, USA, 2017, pp. 3–18, doi: 10.1109/SP.2017.41.

[3] N. Carlini *et al.*, "Extracting training data from large language models," in *Proc. USENIX Security Symp.*, Anaheim, CA, USA, 2023.

[4] European Union, "Regulation (EU) 2016/679 (General Data Protection Regulation)," *Official Journal of the European Union*, 2016. [Online]. Available: https://eur-lex.europa.eu/eli/reg/2016/679/oj

[5] U.S. Department of Health and Human Services, "HIPAA Privacy Rule," 2023. [Online]. Available: https://www.hhs.gov/hipaa/for-professionals/privacy/index.html

[6] C. Dwork and A. Roth, *The Algorithmic Foundations of Differential Privacy*. Hanover, MA, USA: Now Publishers, 2014, doi: 10.1561/0400000042.

[7] A. Acar, H. Aksu, A. S. Uluagac, and M. Conti, "A survey on homomorphic encryption schemes: Theory and implementation," *IEEE Commun. Surveys Tuts.*, vol. 23, no. 4, pp. 2042–2079, 2021, doi: 10.1109/COMST.2021.3061993.

[8] P. Kairouz *et al.*, "Advances and open problems in federated learning," *Found. Trends Mach. Learn.*, vol. 14, no. 1–2, pp. 1–210, 2021, doi: 10.1561/2200000083.

[9] L. Floridi and J. Cowls, "A unified framework of five principles for AI in society," *Harvard Data Science Review*, vol. 3, no. 1, 2021, doi: 10.1162/99608f92.e0c76165.

[10] L. Sweeney, "Only you, your doctor, and many others may know," *Technology Science*, 2015.

[11] OECD, "OECD Recommendation of the Council on Artificial Intelligence," OECD Publishing, Paris, France, 2019.

[12] B. Mittelstadt, "Principles alone cannot guarantee ethical AI," *Nat. Mach. Intell.*, vol. 1, pp. 501–507, 2019, doi: 10.1038/s42256-019-0114-4.

[13] N. Rieke *et al.*, "The future of digital health with federated learning," *npj Digital Medicine*, vol. 3, art. 119, 2020, doi: 10.1038/s41746-020-00323-1.

[14] Q. Yang, Y. Liu, T. Chen, and Y. Tong, "Federated machine learning: Concept and applications," *ACM Trans. Intell. Syst. Technol.*, vol. 10, no. 2, art. 12, 2019, doi: 10.1145/3298981.

[15] X. Zhou *et al.*, "Efficient homomorphic encryption for privacy-preserving machine learning: Recent advances and challenges," *IEEE Access*, 2023.

[16] L. Lyu, H. Yu, and Q. Yang, "Threats to federated learning: A survey," *ACM Comput. Surv.*, vol. 54, no. 5, art. 104, 2022, doi: 10.1145/3464421.

[17] S. Jayaraman and D. Evans, "Evaluating differentially private machine learning in practice," in *Proc. USENIX Security Symp.*, Santa Clara, CA, USA, 2019.

[18] M. J. Sheller *et al.*, "Federated learning in medicine without sharing patient data," *Sci. Rep.*, vol. 10, art. 12598, 2020, doi: 10.1038/s41598-020-69250-1.

[19] L. Mouchet *et al.*, "Privacy-preserving federated analytics for financial services using homomorphic encryption," *IEEE Security Privacy*, 2021.

[20] J. Morley, L. Floridi, L. Kinsey, and A. Elhalal, "From what to how: An overview of AI ethics tools, methods and research," *Sci. Eng. Ethics*, vol. 26, pp. 2141–2168, 2020, doi: 10.1007/s11948-019-00160-4.

[21] A. D. Selbst *et al.*, "Fairness and abstraction in sociotechnical systems," in *Proc. FAT*, Atlanta, GA, USA, 2019, pp. 59–68, doi: 10.1145/3287560.3287598.

[22] S. Gürses and J. van Hoboken, "Privacy engineering," *IEEE Security Privacy*, vol. 20, no. 2, pp. 17–27, 2022, doi: 10.1109/MSEC.2021.3136150.

[23] K. El Emam, L. Mosquera, and R. Hoptroff, *Practical Approaches to Anonymization*. Sebastopol, CA, USA: O'Reilly Media, 2020.

[24] L. Rocher, J. M. Hendrickx, and Y. A. de Montjoye, "Estimating the success of re-identifications in incomplete datasets using generative models," *Nat. Commun.*, vol. 10, art. 3069, 2019, doi: 10.1038/s41467-019-10933-3.

[25] B. Knott, A. Ciccozzi, M. Niepert, and F. Schuster, "Secure multi-party computation for machine learning," *IEEE Trans. Inf. Forensics Security*, vol. 16, pp. 389–403, 2021, doi: 10.1109/TIFS.2020.3042227.

[26] S. Barocas, M. Hardt, and A. Narayanan, *Fairness and Machine Learning*. Cambridge, MA, USA: MIT Press, 2023.

[27] J. Tang *et al.*, "Privacy loss in Apple's implementation of differential privacy," in *Proc. ACM CCS*, Dallas, TX, USA, 2017, doi: 10.1145/3133956.3134017.

[28] W. Yang *et al.*, "Homomorphic encryption for privacy-preserving healthcare applications: A review," *IEEE Rev. Biomed. Eng.*, 2023.

[29] L. Xie *et al.*, "Synthetic data generation for privacy protection: A review," *Pattern Recognit.*, vol. 138, 2023.

[30] P. Prinsloo and S. Slade, "An ethics of care approach to learning analytics," *Br. J. Educ. Technol.*, vol. 48, no. 4, pp. 1238–1250, 2017.

[31] N. Li *et al.*, "Federated learning for COVID-19 chest X-ray classification," *IEEE Trans. Med. Imaging*, 2020.

[32] Y. Zhang and Q. Zhu, "Privacy-preserving machine learning in financial services," *IEEE Security Privacy*, 2022.

[33] I. Bogdanov *et al.*, "Secure multi-party computation for privacy-preserving credit scoring," *IEEE Security Privacy*, 2014.

[34] M. Khalil, R. Shakya, and Q. Liu, “Towards privacy-preserving data-driven education: The potential of federated learning,” *Proceedings of the ACM Conference on Learning Analytics and Knowledge (LAK)*, 2024, pp. 1–11.

[35] M. Abadi *et al.*, “Deep learning with differential privacy,” in *Proc. ACM CCS*, Vienna, Austria, 2016, doi: 10.1145/2976749.2978318.