CENTRAL ASIAN JOURNAL OF MATHEMATICAL THEORY AND COMPUTER SCIENCES

https://cajmtcs.casjournal.org/index.php/CAJMTCS

Volume: 07 Issue: 02 | April 2026 ISSN: 2660-5309

Article

A Comprehensive Review of Privacy-Preserving Techniques in Artificial Intelligence and Machine Learning: Challenges, Solutions, and Future Directions

Hayder Majid Sachit

1. University of Wasit, College of Science, Iraq
* Correspondence: haider.majid.s@uowasit.edu.iq

Abstract: Fast progress of new technologies like Artificial Intelligence (AI) and Machine Learning (ML) is bringing forth important issues when it comes to data privacy in different fields. AI and ML solutions depend on great volumes of data that frequently include personal or organizational data, some of which is sensitive. Preserving privacy with no degradation of model performance presents major technical, ethical and legal issues. This survey presents an extensive overview of the most widely adopted privacy-preserving methodologies in AI and ML, including encryption based techniques, federated learning, differential privacy, data anonymization, blockchain-inspired approaches as well as private AI APIs and synthetic data. 875 Each approach is evaluated by its respective advantage, disadvantage and applicable conditions. Moreover, a comparison shows trade-offs between security level, performance and scalability and hybrid solutions are found most promising for practical use cases. Finally, the paper concludes by providing a few research directions, in particular highlighting the mandatory support for adaptive privacy preservation measures, scalable solutions and consistent global legislations. This is a thorough reference text for those who wish to design secure, performant and privacy-focused AI and ML systems.

Keywords: Privacy-Preserving, Artificial Intelligence (AI), Machine Learning (ML), Federated Learning, Differential Privacy

Citation: Sachit, H. M A Comprehensive Review of Privacy-Preserving Techniques in Artificial Intelligence and Machine Learning: Challenges, Solutions, and Future Directions. Central Asian Journal of Mathematical Theory and Computer Sciences 2026, 7(2), 106-112

Received: 09th Nov 2025
Revised: 17th Dec 2025
Accepted: 04th Jan 2026
Published: 28th Feb 2026

1. Introduction

Background and Significance

Artificial Intelligence (AI) and Machine Learning (ML) were brought to the forefront of multiple modern applications such as health care, finance, smart cities and personalized services. These technology-based techniques depend on significant data gathering and processing, with sensitive private or organizational resources frequently involved. AI and ML are attractive due to their efficiency, accuracy, and automation - at the same time, they bring significant privacy concerns. Unauthorized access, data breaches, and unintended information exposure all pose direct threats to privacy and expose the enterprise to regulatory violations. As AI continues to adopt, privacy preservation has become a crucial concern for research developers and policy makers.

Problem Statement

Though the significance of privacy is increasing day by day, AI/ML based systems are susceptible to a wide range of privacy threats. Legacy approaches to data security may be inadequate for the nature of modern AI models, which are complex and spread across multiple locations. For example, model inversion attacks, membership inference

attacks and HIV propagation data reconstruction from aggregated mobility data have shown that even when a dataset is seemingly anonymized it may be exploited to infer sensitive information. Therefore, how to protect privacy with minimal sacrifice of model performance has become a research hotspot in both academic field and industry.

Objectives of the Review

This survey covers privacy-preserving approaches in AI/ML in the realm of AI and ML. The specific objectives are:

1. To discover and classify the existing methods of indirect personalization in AI and ML systems.
2. To assess the pros and cons, and to compare use cases between these techniques.
3. To outline trends and to recommend future research avenues for improved privacy in AI and ML.

Privacy Challenges in AI and ML

AI/ML systems depend on large set of training and decision-making data. This data based method for achieving high accuracy, however, comes with privacy let down opening up the door to tremendous risks. Data privacy is an issue for AI and ML, which factored into maintaining information about the time of the exposure, model vulnerability scanning, and regulatory requirements. This subsection describes key privacy challenges of AI and ML systems. [1]

Data-Related Challenges

The twin has always been the data in AI and ML, which tends to be sensitive, distributed, heterogeneous and so on. Some of the leading issues are [2].

Data Gathering and Storage: Bulk data collection usually means that you are being subjected to personal or confidential data collection. Poor storage procedures, or weak security measures, can produce unauthorized access of information, often called data breaches.

Data Sharing and Transfer: Collaborative AI projects often involve data sharing between institutions. The data when shared without strong privacy protection becomes liable to potential abuse.

Anonymization and Data Quality: Merely anonymizing datasets may not be effective, as adversaries can re-identify individuals based on auxiliary data. "And it's also about striking a balance between preserving utility and privacy."

Model-Related Challenges

Farther out, AI and ML models can also inadvertently leak sensitive data of their own accord: [3].

1. Model Inference Attacks — Attackers infer sensitive attributes based on the trained model structure, without having access to data.
2. Membership Inference Attacks: whether an individual record was part of the training data or not, which is a significant privacy concern .
3. Weaknesses of Deep Learning Models: The complexity of deep neural networks is the potential source for leakage. Small modifications or queries **can leak sensitive training data**

Regulatory and Ethical Challenges

AI and ML systems need to comply with judicial and ethical systems in order to protect privacy: [4].

Data Privacy Regulations: Rules like the EU's General Data Protection Regulation (GDPR) and the California Consumer Privacy Act (CCPA) enforce stringent privacy guidelines. Adhering to these laws and regulations can be difficult for AI systems, especially those that operate globally.

Ethics: Apart from compliance to the law, she adds considerations such as Informed Consent, Fairness and Transparency. It is incumbent on companies to actually ensure their AI systems adhere as much as possible to the privacy rights of individuals while limiting harm.

2. Materials and Methods

The current study employs a systematic literature review process to provide a comprehensive coverage of privacy-preserving methods in AI and ML systems. The review process entails systematic identification, selection and analysis of articles from peer-reviewed journals, international conferences, and authoritative reports focused on encryption-oriented techniques, federated learning, differential privacy, data anonymization, blockchain-inspired solutions, and private AI APIs that can generate synthetic data. We conducted searches of the academic literature using pre-specified keywords in electronic databases such as IEEE Xplore, SpringerLink, ScienceDirect, and Google Scholar, with keywords including "privacy-preserving AI", "federated learning", "differential privacy", "homomorphic encryption", and "secure multi-party computation". We included studies that discussed privacy problems with AI/ML but included defined technical methodologies to solve or evaluate performance metrics to demonstrate how they were evaluated, regulatory implications, etc., and we excluded non-peer-reviewed, old or untechnical discussions. We classified selected studies by type of technique applied and performed a comparative analysis based on the criteria of security level, computational cost, scalability, performance degradation, and normative compliance. This work took a qualitative synthesis approach to identify cross-study emergent recurring patterns, trade-offs, and hybrid solutions trends. In addition, the review included analytical comparison tables to demonstrate the relative strengths and weaknesses of techniques in different application contexts based on healthcare, finance, and distributed systems. Such a systematic approach allowed not only to have a fair and evidence-driven evaluation of existing privacy-preserving approaches, but also to point out the research voids and the emerging paths towards adaptive and scalable privacy-aware AI systems.

3. Results and Discussion

Privacy-Preserving Techniques

To overcome the privacyrelated issues in AI and ML systems, a number of privacy preserving techniques have been proposed. These techniques seek to preserve sensitive information and yet provide models with high performance and utility. The details of the most popular methods are covered in the following sub-sections. [5].

Encryption-Based Techniques

The encryption-based techniques are aimed to protect data at rest, in motion and in use. One of the most prominent methods is based on Homomorphic Encryption (HE) that allows computation to be carried out over encrypted data without decryption. This is designed to keep all sensitive data private for the period of computation. Regarding its security, although HE provides strong security guarantees it is computationally expensive and can generate high latency in large dataset processing. Another promising tool is SMPC (Secure Multi-Party Computation), which allows several parties to jointly compute a function over their respective inputs without revealing the raw input data. Collaborative AI offers an especially strong use case for SMPC: in collaborative environments, entities may want to provide model updates without revealing sensitive proprietary data. [6].

Federated Learning

Federated Learning (FL) is a distributed model training method that enables multiple clients to jointly build a common AI model without sharing their private data with the central server. In this setting, each client trains the model on its local data and only communicates with a central party in an update or gradient pusing fashion. These updates are subsequently aggregated at the aggregator for global model further refinement. In this sense a system trained with FL has less access to sensitive information and becomes more adequate for data protection legislation. However, there are also concerns regarding FL, i.e., communication overhead, susceptibility to model poisoning attacks as well as potential of privacy leakage via gradients. [7]

Differential Privacy

Differential Privacy (DP) is known for its principled, provable privacy guarantees by carefully adding noise to data or outputs of models. The central DP principle is that if one point of in a dataset becomes present or absent, this should not substantially change the output of any computation. That way it prevents that detailed information can be extracted from the added up outputs. DP has been used in different AI and ML applications such as model training, query-answering systems and large scale analytics. Federated learning In conjunction with federated learning, differential privacy can even be used to improve the privacy of distributed AI systems by covering the updates that clients share [8].

Data Anonymization and Pseudonymization

Data anonymization and pseudonymization are the standard methods of lowering privacy threats. Data anonymization is a known mechanism to remove PII in datasets to prevent attackers from re-identifying the individuals. Pseudonymization, meanwhile, substitutes PII with dummy or arbitrary alphanumeric strings to lower the risk of being directly identified. Although both approaches achieve some level of privacy protection, they are not perfect; there are attacks that can be powerful enough to in many cases reconstruct information considered as sensitive. Hence, they are typically integrated with other cryptosystem or modern PPMs to achieve better security predictions [9].

Blockchain-Based Privacy Solutions

Blockchain has been proposed as a promising candidate to facilitate PP-FD. Blockchain offers security and transparency for data sharing in the absence of a trusted central institution, thanks to its decentralized tamper-evident architecture. It provides oversight and audit/logging of data access, which is critical in regulated domains like healthcare and finance. However, due to scalability and computa- tional overhead issues it is still a research challenge for blockchain-driven solutions to become applicable in large-scale AI systems [10].

Private AI APIs and Synthetic Data

Other methods for privacy-side techniques include private AI APIs and synthetic data creation. Private AI APIs enable cloud AI to operate on encrypted or obfuscated data and offer end users privacy assurances. Synthetic data production generates synthetic datasets matching the statistical properties of real-world data without disclosing sensitive information. An advantage of this methodology is that it can be applies to train AI models when access to real, sensitive data may be limited or controlled [11].

Comparative Analysis of Techniques

Full-ranging comprehension on privacy-preserving techniques for AI and ML In addition to providing with details of above methods, overview should be given as well how good or bad each method is and when you can use what. This subsection compares the methods in agreement with above section, w.r.t security, performance, scalability and practicality [10].

Comparative Overview

Different privacy-preserving methods differ in complexity as well the computational runtime and strength of privacy guaranty. Encryption based methods (e.g., Homomorphic Encryption and SMPC) provide strong security guarantees, as the data are kept encrypted during computing. But the computation is heavy and can be very slow, it is less applicable for large scale data or real time application. In comparison, federated learning (FL) securely minimizes data exposure with raw data remaining on local devices; however, traditional FL might be susceptible to gradient-based privacy attacks and is subject to high communication overhead in large distributed networks [12].

Differential Privacy (DP) [13] affords formal privacy guarantees with added noise, and is appropriate for statistical analysis and learning settings. However, too much noise can decrease the accuracy of models and hence a trade-off must be chosen between

privacy and utility. Anonymization and pseudonymization data approaches can easily be applied, leading to a decrease in direct identifiability, but are still prone to re-identification attacks so long as additional information exists. Blockchain solutions provide transparent, immutable and decentralized trust, that are highly suitable for privacy-sensitive (multi-party) data sharing, however the scalability is still an issue. Finally, private AI APIs and synthetic data are flexible solutions that could be pragmatic in reality in certain regulated scenarios but are only as safe as the synthetic data quality and encryption implementation.

Strengths and Weaknesses

Each of which enjoys the pros and cons: [14][15]

Encryption approach: Very secures but needs huge computational.

Federated learning: Reduces data exposure, compliant-aware, but vulnerable in communications and attacks.

Differential Privacy: Provable privacy guarantee, may compromise accuracy (if the noise is high).

Data anonymization/pseudonymization: Basic but economical, insufficient against advanced threats.

Decentralized solutions: Transparency and trust, but can be constrained by scalability and computational cost.

Private AI APIs and synthetic data: Practical and privacy-friendly — Provided the implementation is of high-standards.

Future Directions

With the ongoing efforts in AI and ML, further solid PP solutions will be sought. Although notable progress has been achieved in the past years, there are still a number of research gaps and technical challenges to be addressed. In this section, we explore possible future directions to improve privacy in AI and ML systems.

Research and Technical Advances

Further studies will need to continue to work on the effciency and practicality of current PP techniques. For example, federated learning could employ novel aggregation mechanisms (to reduce communication) and noise mechanisms for more robust privacy protection against model gradients. In the same line, combining differential privacy with deep learningbased approaches presents an interesting prospect to keep data confidential when it is distributed over a large-scale network, without compromising too much on performance. Hybrid methods for example the combination of homomorphic encryption with secure multi-party computation could be investigated, that could provide better privacy but with computational efficiency. Furthermore, the design of novel algorithms for adaptive privacy preserving mechanism can facilitate a dynamic trade-off between privacy preservation and model accuracy according to application scenarios.

Policy and Legal Frameworks

As well as technological advances, the conservation of privacy in AI relies on strong legal and regulatory standards. There is an obvious minimum standard in international standards (like GDPR and CCPA) to protect your data, but local variations make it difficult to deploy global AI. Recommendations for future direction should focus on international uniform regulations; frameworks regarding cross-border data sharing and come to compliance with AI models. Ethics and best practices related to privacy preserving AI, too, shall become part of the deal to make sure technologies respect individuals' rights while fostering innovation.

Open Challenges

There are many open problems in privacy preserving AI. Trade-off between privacy, utility and efficiency is a challenging problem in deep learning based real-time AI systems. New attacks, aiming at sophisticated model inversion and membership inference attacks, need to be followed up with ongoing research in predicting and pre-empting future privacy violations. Furthermore, the scalability of privacy-preserving

techniques (in particular within massive distributed network and IoT ecosystems) is also a technical challenge. Solving these challenges will be a matter of interdisciplinary research, involving AI and cryptography, data science and legal studies.

4. Conclusion

"Preserving privacy is important during the design phase and when deploying AI/ML systems, especially since (the ML portion) often works with sensitive personal and corporate data," it was stressed. This survey has outlined the major challenges of preserving privacy, such as data-related vulnerabilities, model-based attacks, and regulatory and ethical issues. Several privacy preserving methods have been investigated such as encryption-based, federated learning, differential privacy, data anonymization, blockchain based solutions, private AI APIs and synthetic data. Each of these has its own pros and cons, which is potentially suited for one data type as well as an application domain due to factors like the data type, availability of computational resources, regulatory requirements etc.

Comparative examinations show that there is no singular approach that can completely cover privacy issues. Hybrid solutions, which exploit more than one techniques, usually allow the system to have better privacy guarantees as well as to support performance and scalability. Future work Future research should streamline the methods and make them more practical, look into privacy adaptive mechanisms, and investigate scalable solutions for large distributed/real-time AIsystems. What we need As well, harmonized global standards and ethical guidelines and regulatory frameworks that support compliance in respect of AI technologies are critical to uphold privacy rights and innovation.

To conclude, progressing privacy-preserving AI necessitates a combined technical, regulatory and ethical approach. Addressing these issues, researchers and practitioners can design AI/ML technologies which not only work effectively but are also trustworthy; thus guaranteeing that the benefits of these technologies will be captured, and at the same time avoiding violations on individual or organizational privacy.

REFERENCES

[1] E. N. Kucur, T. Buyuktanir, M. Ugurelli, and K. Yildiz, "Privacy-Preserving Machine Learning Techniques: Cryptographic Approaches, Challenges, and Future Directions," Appl. Sci., vol. 16, no. 1, p. 277, 2025.

[2] R. N. SABER and Y. H. MASEEH, "Privacy-preserving machine learning: a review of federated learning techniques and applications," Int. J., vol. 11, no. 1, pp. 30–39, 2025, doi: https://doi.org/10.14419/af03y111.

[3] J. Lu, "Survey on Privacy-Preserving Techniques for Federated Learning".

[4] H. Schwarz, "Comprehensive review on privacy-preserving machine learning techniques for exploring federated learning," Eur. J. Intell. Autom. Res., vol. 3, no. 2, 2024.

[5] E. V. Svanovich, A. T., & Petrova, "Privacy Preserving AI: Federated Learning and Differential Privacy," AI Sci. Acad. J..

[6] A. M. Akinsiku, "A comprehensive survey of federated learning approaches for privacy-preserving machine learning," Tech-sph. J. Pure Appl. Sci., vol. 2, no. 1, 2025.

[7] E. Shalabi, W. Khedr, E. Rushdy, and A. Salah, "A comparative study of privacy-preserving techniques in federated learning: A performance and security analysis," Information, vol. 16, no. 3, p. 244, 2025.

[8] E.D.Kanmani Ruby, "Advanced Privacy-Preserving Federated Learning in 6G Networks Using Differential Privacy and Homomorphic Encryption," Int. J. Intell. Syst. Appl. Eng., vol. 12, no. 23s SE-Research Article, pp. 1–7, Aug. 2024, [Online]. Available: https://ijisae.org/index.php/IJISAE/article/view/6427

[9] S. Barański, "A Survey on Privacy-Preserving Machine Learning Inference," TASK Q., vol. 28, no. 2, 2024.

[10] W. Jin et al., "FedML-HE: An efficient homomorphic-encryption-based privacy-preserving federated learning system," arXiv Prepr. arXiv2303.10837, 2023.

[11] T. H. Rafi, F. A. Noor, T. Hussain, and D.-K. Chae, "Fairness and privacy preserving in federated learning: A survey," Inf. Fusion, vol. 105, p. 102198, 2024.

[12] V. S. Naresh, A. Venkata Raju, and O. Srinivasa Rao, “Secure Multiparty Computation for Privacy-Preserving Machine Learning in Healthcare: A Comprehensive Survey,” Wiley Interdiscip. Rev. Comput. Stat., vol. 17, no. 3, p. e70046, 2025.
[13] C. Dwork, “Differential privacy: A survey of results,” in International conference on theory and applications of models of computation, Springer, 2008, pp. 1–19.
[14] F. Liu, Z. Zheng, Y. Shi, Y. Tong, and Y. Zhang, “A survey on federated learning: a perspective from multi-party computation,” Front. Comput. Sci., vol. 18, no. 1, p. 181336, 2024.
[15] C. Dwork, “Differential privacy: A survey of results,” in Proc. Int. Conf. Theory and Applications of Models of Computation (TAMC). Berlin, Germany: Springer, 2008, pp. 1–19.