International Journal of Scientific Research and Engineering Development-– Volume 9 Issue 1, Jan-Feb 2026
Available at ***wwwijsred.com***

RESEARCH ARTICLE OPEN ACCESS

Privacy-Preserving Healthcare Fraud Detection: A Federated and Data Mining–Centric Survey

Kalaiyarasi .D * Dr.John Paul.C **
*Research Scholar, St.Joseph University Tindivanam,TN,India
E-Mail: kalai2334@gmail.com
**Research Supervisor, *Associate Professor,*
Department of Computer Science St. Joseph University Tindivanam,TN, India ,
E-Mail: johnpaul.sju@gmail.com

Abstract—

Healthcare fraud poses a significant financial and operational burden on healthcare systems, while the sensitive nature of medical data imposes strict privacy requirements on analytical solutions. Traditional fraud detection approaches often rely on centralized data collection, which conflicts with regulatory and ethical constraints related to data sharing. To address this challenge, privacy-preserving data mining and machine learning techniques have gained increasing attention. This paper presents a concise survey of privacy-preserving approaches for healthcare fraud detection, with a particular focus on federated learning, differential privacy, cryptographic methods, and privacy-aware data mining techniques. We introduce a taxonomy that categorizes existing methods based on their underlying privacy mechanisms and analyze representative approaches through a structured comparative study. Furthermore, we discuss practical deployment scenarios and identify key open challenges that hinder real-world adoption, including privacy–utility trade-offs, scalability, and system heterogeneity. By synthesizing recent advances and highlighting unresolved research gaps, this survey aims to provide researchers and practitioners with a clear understanding of the current landscape and future directions of privacy-preserving healthcare fraud detection.

Keywords— Healthcare fraud detection, Privacy-preserving data mining, Federated learning, Differential privacy, Secure computation, medical data analytics.

I. INTRODUCTION

Healthcare fraud has emerged as a critical challenge for modern healthcare systems, leading to substantial financial losses, reduced service quality, and erosion of public trust. Fraudulent activities such as false insurance claims, billing manipulation, and provider abuse have become increasingly sophisticated with the digitization of healthcare services and records. Recent studies highlight that the growing scale and complexity of healthcare data have amplified both the opportunities for fraud and the consequences of undetected misuse [1], [2].

To address these challenges, data mining and machine learning techniques have been widely adopted for healthcare fraud detection due to their ability to identify complex and hidden patterns within large-scale datasets. Supervised, unsupervised, and hybrid learning models have demonstrated promising performance in detecting anomalous behaviour that are difficult to capture using rule-based systems [3], [4]. However, the effectiveness of such approaches heavily depends on access to large volumes of high-quality data, which is often restricted in healthcare environments.

Healthcare data are inherently sensitive, containing personal, clinical, and financial information that must be protected under strict ethical and regulatory frameworks. Concerns related to patient privacy, data misuse, and informed consent, along with compliance requirements imposed by regulations such as HIPAA and GDPR, significantly limit centralized data collection and sharing [5]–[7]. These constraints introduce a fundamental tension between effective fraud detection and privacy preservation. Traditional centralized fraud detection frameworks typically require aggregating data from multiple healthcare entities, including hospitals, insurers, and regulatory bodies. Such centralized architectures not only increase the risk of data breaches but also raise trust and governance issues across institutions, making large-scale collaboration difficult to achieve [8], [9]. As a result, many existing solutions struggle to balance analytical performance with privacy and security requirements.

In response, privacy-preserving data mining and machine learning techniques have gained increasing attention. Approaches such as federated learning, differential privacy, and cryptographic computation enable collaborative model training and analysis without exposing raw data, making them particularly attractive for healthcare fraud detection scenarios [10]–[13]. These methods aim to mitigate privacy risks while maintaining acceptable detection accuracy. Although several surveys have explored privacy-preserving machine learning, secure data analytics, and ethical considerations in healthcare, most existing works focus on general-purpose frameworks or domain-independent applications [14]–[16]. A focused and structured survey that specifically examines privacy-preserving techniques through the lens of healthcare fraud detection remains lacking.

Contributions: This paper presents (i) a taxonomy of privacy-preserving approaches for healthcare fraud detection based on underlying privacy mechanisms, (ii) a comparative analysis of representative methods, (iii) an overview of practical deployment scenarios, and (iv) a discussion of open challenges and future research directions.

II. TAXONOMY OF PRIVACY-PRESERVING HEALTHCARE FRAUD DETECTION

To systematically analyze existing research on privacy-preserving healthcare fraud detection, this section presents a taxonomy that categorizes approaches based on the primary privacy-preservation mechanism employed. This taxonomy serves as the structural foundation of the survey and enables a coherent comparison of diverse techniques that address privacy constraints while supporting fraud detection tasks.

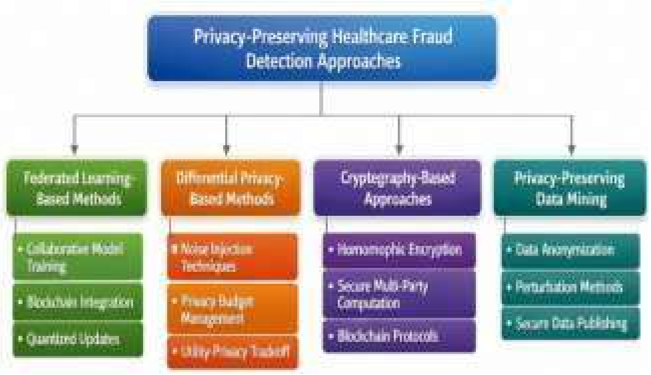
Privacy-Preserving Healthcare Fraud Detection Approaches
Federated Learning-Based Methods
Differential Privacy-Based Methods
Cryptography-Based Approaches
Privacy-Preserving Data Mining
• Collaborative Model Training
• Blockchain Integration
• Quantized Updates
• Noise Injection Techniques
• Privacy Budget Management
• Utility-Privacy Tradeoff
• Homomophic Encryption
• Secure Multi-Party Computation
• Blockchain Protocols
• Data Anonymization
• Perturbation Methods
• Secure Data Publishing

Fig. 1. Privacy-Preserving Healthcare Fraud Detection Approaches.

Figure 1 illustrates the proposed taxonomy, which organizes existing approaches into four major categories: (A) federated learning–based methods, (B) differential privacy–based methods, (C) cryptography-based approaches, and (D) privacy-preserving data mining techniques.

A. Federated Learning–Based Approaches

Federated learning (FL) enables multiple healthcare entities to collaboratively train fraud detection models without sharing raw data. Instead, local model updates are exchanged and aggregated, preserving data locality and reducing exposure risks. Recent studies demonstrate the effectiveness of FL in distributed and privacy-sensitive environments, including healthcare and IoT ecosystems, where data heterogeneity and institutional boundaries are prevalent [17]–[20]. Enhancements such as perturbation mechanisms, blockchain integration, and quantized model updates have been proposed to strengthen robustness and trust in federated settings [21], [22]. Due to its scalability and regulatory compatibility, FL has emerged as one of the most promising paradigms for multi-institutional healthcare fraud detection.

B. Differential Privacy–Based Methods

Differential privacy (DP) provides formal and quantifiable privacy guarantees by introducing controlled noise into data or model parameters. In healthcare fraud detection, DP-based methods aim to prevent sensitive information leakage while enabling statistical analysis and learning. Recent works explore noise-injection strategies, privacy-aware model training, and system-level optimizations to balance detection accuracy and privacy budgets [23] – [25]. Although DP offers strong theoretical guarantees, its application in fraud detection often faces challenges related to performance degradation and parameter tuning in highly imbalanced datasets.

C. Cryptography-Based Privacy Preservation

Cryptographic techniques protect data confidentiality by enabling computation over encrypted data or through secure multi-party protocols. Approaches based on homomorphic encryption, secure multi-party computation, and blockchain technologies allow healthcare stakeholders to jointly perform fraud analytics without revealing sensitive inputs [26]–[29]. While these methods provide strong security assurances, their high computational and communication overhead can limit practical deployment in large-scale, real-time fraud detection systems.

D. Privacy-Preserving Data Mining Techniques

Privacy-preserving data mining (PPDM) focuses on protecting sensitive information at the data level prior to analysis. Techniques such as anonymization, perturbation, and secure data publishing have been applied to healthcare datasets to support fraud detection while minimizing privacy risks [30] – [32]. Although PPDM methods are relatively simple to implement, they often offer weaker privacy guarantees compared to FL or cryptographic approaches.

III. COMPARATIVE SURVEY OF EXISTING APPROACHES

This section presents a comparative analysis of representative privacy-preserving approaches relevant to healthcare fraud detection. Rather than exhaustively reviewing all existing studies, the comparison focuses on methodologically representative works that reflect diverse privacy mechanisms, learning models, and deployment settings. The goal is to highlight key trade-offs affecting accuracy, scalability, and privacy assurance.

Table 1 summarizes the selected studies across multiple dimensions, including application domain, privacy mechanism, learning model, dataset type, evaluation metrics, and key limitations.

A. Federated Learning–Based Approaches

Federated learning–based approaches have gained prominence due to their ability to enable collaborative model training without centralized data sharing. Several recent studies demonstrate the effectiveness of FL frameworks in privacy-sensitive and distributed environments by incorporating mechanisms such as secure aggregation, perturbation, and quantized updates [23] – [25]. Blockchain-enabled FL architectures further improve trust and accountability among participating entities, which is particularly relevant for multi-institutional healthcare ecosystems [26]. However, FL-based methods often face challenges related to communication overhead, data heterogeneity, and vulnerability to poisoning attacks.

B. Differential Privacy–Based Methods

Differential privacy–based approaches aim to provide formal privacy guarantees through controlled noise injection during data analysis or model training. Recent works explore privacy-aware learning pipelines that balance detection accuracy and privacy budgets in sensitive domains [27], [28]. While these methods offer strong theoretical guarantees, selecting appropriate noise levels remains challenging, especially in highly imbalanced fraud detection datasets where excessive noise can degrade performance.

C. Privacy-Preserving Data Mining Approaches

Privacy-preserving data mining techniques protect sensitive information at the data level prior to analysis. Techniques such as anonymization and perturbation have been applied to healthcare datasets to support fraud detection while reducing disclosure risks [30]. Although simple to deploy, these methods generally offer weaker privacy guarantees and may be vulnerable to inference attacks.

TABLE I. COMPARATIVE ANALYSIS OF PRIVACY-PRESERVING APPROACHES

Ref.	Application Domain	Privacy Mechanism	Learning / DM Model	Dataset Type	Evaluation Metrics	Key Strength	Key Limitation
[23]	IoT / Distributed Systems	Federated Learning	Deep Neural Network	Real	Accuracy, F1-score	Preserves data locality	Sensitive to data heterogeneity
[24]	Distributed ML	Federated Learning + Perturbation	Neural Network	Synthetic	Accuracy, Robustness	Improved resistance to attacks	Accuracy loss due to noise
[25]	Smart Infrastructure	Federated Learning	Transformer Network	Real	Precision, Recall	High detection accuracy	High communication cost
[26]	Healthcare IoT	Blockchain-enabled FL	Deep Learning	Real	Accuracy, Latency	Trustworthy collaboration	System complexity
[27]	Privacy-Aware ML	Differential Privacy	Statistical ML	Synthetic	Privacy budget (ε), Accuracy	Formal privacy guarantees	Utility degradation
[28]	Healthcare Analytics	Differential Privacy	Machine Learning	Real	AUC, ε	Regulatory compliance	Difficult privacy tuning
[29]	Secure Analytics	Homomorphic Encryption / MPC	Encrypted ML	Synthetic	Runtime, Security	Strong confidentiality	High computational overhead
[30]	Healthcare Fraud	Privacy-Preserving Data Mining	Classification Models	Real	Precision, Recall	Simple deployment	Weak privacy guarantees
[21]	Big Data Environments	PP Data Mining	Clustering / Classification	Real	Accuracy	Scalable data handling	Limited adversary resistance
[22]	Healthcare Informatics	PPDM	Data Mining Models	Real	Accuracy, Sensitivity	Domain-specific applicability	Susceptible to inference attacks
[18]	General PPML	Hybrid Privacy Models	ML / DL	Mixed	Accuracy, Privacy	Comprehensive framework	Not fraud-specific
[20]	Privacy-Preserving ML	System-level PPML	ML / DL	Mixed	Efficiency, Accuracy	End-to-end perspective	High system complexity

IV. PRIVACY-PRESERVING HEALTHCARE FRAUD DETECTION FRAMEWORK

This section presents a conceptual end-to-end framework for privacy-preserving healthcare fraud detection that synthesizes the techniques discussed in earlier sections. Rather than introducing a new algorithmic contribution, the framework provides a system-level abstraction that reflects how existing privacy-preserving methods can be orchestrated in practical healthcare fraud detection pipelines, as motivated by recent studies on privacy-aware analytics in sensitive domains [18], [20].

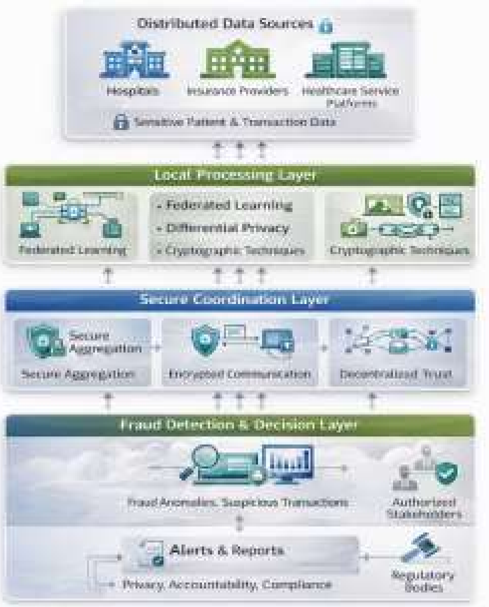

Fig. 2. Privacy-Preserving Healthcare Fraud Detection Approaches.

Figure 2 illustrates the overall workflow of the proposed framework. The process begins with distributed data sources, such as hospitals, insurance providers, and healthcare service platforms, where sensitive patient and transactional data are locally stored. Centralized data pooling is avoided due to regulatory and trust constraints, a limitation widely recognized in healthcare analytics and fraud detection systems [25], [27].

In the local processing layer, data preprocessing and feature extraction are performed within each organization. Privacy-preserving mechanisms are applied depending on system requirements. For example, federated learning enables collaborative model training through the exchange of model updates rather than raw data, while perturbation or noise-based mechanisms can be used to further limit information leakage [23], [27]. Such strategies have been shown to improve regulatory compliance while maintaining acceptable analytical performance.

The secure coordination layer aggregates privacy-protected information from participating entities. This layer may employ secure aggregation, encrypted communication, or decentralized trust mechanisms to ensure confidentiality and integrity during collaboration [26], [29]. By abstracting institutional heterogeneity, the coordination layer supports scalable and cross-organizational fraud detection without compromising data ownership.

Finally, the fraud detection and decision layer utilizes aggregated or globally trained models to identify anomalous behaviours, suspicious transactions, or high-risk entities. Detection outcomes can be shared with authorized stakeholders in a controlled manner, enabling accountability and auditability while preserving privacy [30]. This layered framework demonstrates how privacy-preserving techniques can be systematically integrated to balance analytical effectiveness with ethical, legal, and operational constraints.

V. PREFERRED DEPLOYMENT PLACES AND PRACTICAL SCENARIOS

The effectiveness of privacy-preserving healthcare fraud detection solutions is highly influenced by the deployment environment, regulatory constraints, and operational characteristics of different stakeholders. Healthcare ecosystems involve heterogeneous entities such as hospitals, insurers, regulatory agencies, and cloud service providers, each imposing distinct requirements on data governance, trust, and computational resources. Consequently, the choice of privacy-preserving technique must be aligned with the intended deployment context.

Hospitals and healthcare providers are primary generators of sensitive patient and transactional data. In such environments, federated learning–based approaches are particularly suitable, as they enable collaborative fraud detection across multiple institutions without centralized data aggregation. This paradigm preserves data locality, supports regulatory compliance, and has been shown to be effective in privacy-sensitive healthcare and distributed analytics scenarios [23], [26]. Local training combined with privacy-aware model updates allows institutions to retain control over their data while benefiting from shared intelligence.

Insurance companies and claim processing agencies manage large volumes of billing and reimbursement records and are frequent targets of fraudulent activities. For these stakeholders, differential privacy–based techniques are well suited, as they provide formal privacy guarantees while enabling large-scale statistical analysis and reporting. Prior studies highlight the suitability of differential privacy for regulated environments where aggregate-level insights are sufficient and individual-level disclosure risks must be minimized [27], [28].

Government agencies and regulatory bodies often require cross-organizational analysis to detect systemic fraud patterns and ensure compliance. In such high-stakes environments, cryptography-based approaches, including secure multi-party computation and encrypted analytics, enable collaborative fraud detection without exposing sensitive data across institutions. These techniques offer strong confidentiality guarantees and are commonly adopted in scenarios involving inter-organizational trust limitations [29].

Cloud-based healthcare platforms and multi-institutional networks benefit from hybrid deployment models that combine federated learning with secure coordination mechanisms. Such architectures support scalability, interoperability, and controlled information sharing, making them suitable for regional or nationwide fraud detection initiatives [30].

Table 2 summarizes the preferred deployment places and their corresponding privacy-preserving techniques, along with key rationales and operational considerations.

TABLE II. SUITABLE PRIVACY-PRESERVING TECHNIQUES FOR HEALTHCARE FRAUD DETECTION

Deployment Place	Primary Stakeholders	Suitable Privacy-Preserving Technique	Operational Considerations
Hospitals and Clinics	Healthcare providers, patients	Federated Learning	Data heterogeneity and communication efficiency
Insurance Companies	Insurers, claim processors	Differential Privacy	Careful tuning of privacy budgets
Government and Regulatory Bodies	Regulators, auditors	Secure MPC / Cryptographic Analytics	High computational overhead
Cloud-Based Healthcare Platforms	Platform operators, service providers	Hybrid FL with Secure Coordination	Requires secure aggregation and access control
Multi-Institutional Healthcare Networks	Hospitals, insurers, regulators	Federated Learning with Blockchain	Increased system complexity
Research and Analytics Organizations	Analysts, policy makers	Privacy-Preserving Data Mining	Weaker privacy guarantees than cryptographic methods

VI. OPEN CHALLENGES AND FUTURE RESEARCH DIRECTIONS

A. Challenges

Despite significant progress, several open challenges continue to limit the widespread adoption of privacy-preserving healthcare fraud detection systems. One major challenge is the privacy–utility trade-off, where stronger privacy guarantees often lead to degraded fraud detection performance. This issue is particularly pronounced in highly imbalanced healthcare fraud datasets, where noise injection or constrained model updates can obscure rare but critical fraud patterns [27], [28].

Another challenge lies in data heterogeneity and system scalability. Healthcare institutions vary widely in terms of data distributions, infrastructure, and participation levels, which can negatively impact collaborative learning frameworks. Communication overhead and synchronization requirements further complicate deployment in large-scale, multi- institutional environments [23], [26]. Addressing these issues remains critical for real-world adoption.

Security and robustness also present unresolved concerns. Privacy-preserving frameworks, especially federated learning systems, may still be vulnerable to adversarial behaviors such as model poisoning or inference attacks. While cryptographic techniques offer stronger confidentiality guarantees, their computational complexity and latency limit applicability in time-sensitive fraud detection scenarios [29]. Balancing robustness, efficiency, and privacy remains an open problem.

B. Future Research Directions

Looking forward, several research directions offer promising opportunities to advance privacy-preserving healthcare fraud detection. Developing adaptive privacy mechanisms that dynamically adjust privacy parameters based on data sensitivity and fraud risk could help mitigate the privacy–utility trade-off. Similarly, hybrid frameworks that combine federated learning, differential privacy, and cryptographic primitives in a modular manner may offer better flexibility and performance.

Another important direction is the integration of explainable and interpretable models within privacy-preserving frameworks. Explainability is essential for regulatory compliance, trust, and decision support in healthcare settings, yet remains underexplored under strict privacy constraints.

Finally, future research should focus on real-world validation, including deployment on large-scale, cross-institutional healthcare datasets and alignment with evolving regulatory requirements. Establishing standardized benchmarks and evaluation protocols would further accelerate progress in this domain.

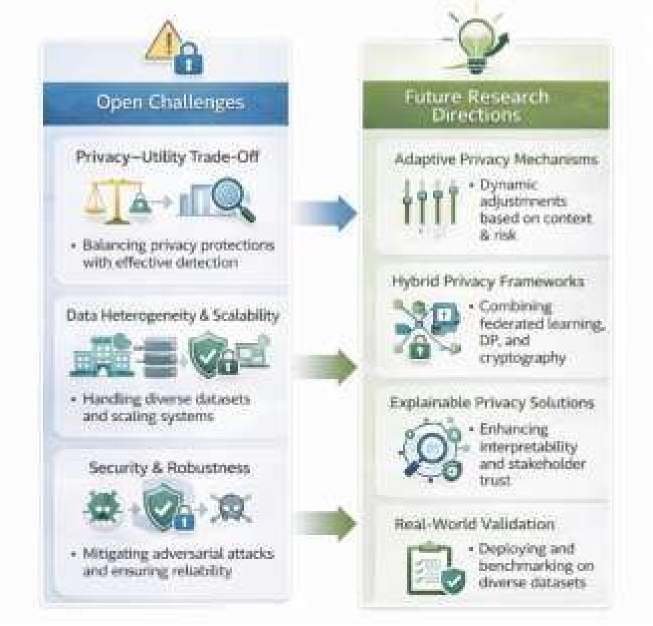
Open Challenges
Privacy–Utility Trade-Off
• Balancing privacy protections with effective detection
Data Heterogeneity & Scalability
• Handling diverse datasets and scaling systems
Security & Robustness
• Mitigating adversarial attacks and ensuring reliability
Future Research Directions
Adaptive Privacy Mechanisms
• Dynamic adjustments based on context & risk
Hybrid Privacy Frameworks
• Combining federated learning, DP, and cryptography
Explainable Privacy Solutions
• Enhancing interpretability and stakeholder trust
Real-World Validation
• Deploying and benchmarking on diverse datasets

Fig. 3. Open Challenges & Future Directions.

Figure 3. Open challenges and future research directions in privacy-preserving healthcare fraud detection. The figure highlights key challenges such as privacy–utility trade-offs, data heterogeneity, and system robustness. It also outlines promising future directions, including adaptive privacy mechanisms, hybrid frameworks, explainable models, and real-world validation.

VII. CONCLUSION

This paper presented a concise survey of privacy-preserving approaches for healthcare fraud detection, addressing the growing need to balance effective analytics with stringent privacy requirements. A structured taxonomy was introduced to categorize existing methods based on their underlying privacy mechanisms, followed by a comparative analysis of representative approaches. The survey further discussed practical deployment scenarios and highlighted key challenges that hinder real-world adoption, including privacy–utility trade-offs, scalability, and system robustness. By synthesizing recent advances and identifying open research directions, this work provides a clear and unified view of the current landscape of privacy-preserving healthcare fraud detection. The insights presented in this survey are intended to support researchers, practitioners, and policymakers in designing and deploying privacy-aware fraud detection systems that are both effective and compliant with ethical and regulatory constraints.

REFERENCES

[1] Ali, Danish, Sundas Iqbal, Sumaira Rafique, Merzougui Hanane, and Irshad Khalil. "Ethical implications and data privacy concerns in AI and ML applications within food science." In Artificial Intelligence in Food Science, pp. 763-780. Academic Press, 2026.

[2] Gautam, Priyanka, Shabir Ali, Dhirendra Kumar Shukla, and Arvind Dagur. "Security and privacy solutions for EHR systems: A comparative study of modern frameworks." In Artificial Intelligence and Sustainable Innovation, pp. 799-808. CRC Press, 2026.

[3] Puviarasu, A., and V. K. Sudha. "Enhanced IoT security: privacy-preserving federated learning model for accurate, real-time intrusion detection across devices." Ain Shams Engineering Journal 17, no. 1 (2026): 103866.

[4] Peng, Chuanyu, and Hequn Xian. "Lightweight orthogonal perturbation for privacy-preserving federated learning against poisoning attacks." Journal of Information Security and Applications 97 (2026): 104345.

[5] Alotaibi, Jamal. "A Multi-Scale Fusion Transformer Network for Federated Privacy - Preserving Smart Parking Slot Detection." Concurrency and Computation: Practice and Experience 38, no. 1 (2026): e70506.

[6] Reddy, Ganta Raghotham, Venkatagurunatham Naidu Kollu, Hashim Elshafie, and Shamimul Qamar. "A novel blockchain-federated learning framework with quantum neural networks and wavelet transforms for secure IoT healthcare monitoring." Biomedical Signal Processing and Control 113 (2026): 108759.

[7] Bhutta, Muhammad Nasir Mumtaz, Ghulam Irtaza, Abid Mehmood, Rabeya Hamood, Imran Makhdoom, Mourad Elhadef, and Muhammad Habib Ur Rehman. "A systematic review of secure federated learning based on blockchain and Multi-Party computation." Peer-to-Peer Networking and Applications 19, no. 1 (2026): 7.

[8] Innan, Nouhaila, Alberto Marchisio, Mohamed Bennai, and Muhammad Shafique. "Exploring quantum federated learning: Algorithms, applications, and challenges." Quantum Computational AI (2026): 25-50.

[9] Kong, Minxue, Feifei Shen, Zhi Li, Xin Peng, and Weimin Zhong. "Finite-Time Performance Mask Function-Based Distributed Privacy-Preserving Consensus: Case Study on Optimal Dispatch of Energy System." IEEE Transactions on Signal and Information Processing over Networks 10 (2024): 776-787.

[10] Alomari, Saleh Ali. "ChainGuard 6G+: A Secure and Private Architecture for Wireless Communication Using Federated Learning and Blockchain in IoT Networks." Journal of Intelligent Systems & Internet of Things 18, no. 1 (2026).

[11] Kasiri, Nasrin, Mohammadreza Mollahoseini-Ardakani, and Mostafa Ghobaei-Arani. "Improving prediction accuracy in serverless edge computing using a federated learning." Cluster Computing 29, no. 1 (2026): 69.

[12] Sun, Longlong, Hui Li, Qingcai Luo, Yanguo Peng, and Jiangtao Cui. "Ophiuchus: Privacy-preserving training service with user-controlled pseudo-noise information generation." Information Processing & Management 63, no. 2 (2026): 104443.

[13] Balaji, Sangeetha, Naveen Jeyaraman, Swaminathan Ramasubramanian, and Madhan Jeyaraman. "Cryptography-based fog computing and IOT-based technology in smart healthcare systems." In Fundamentals of Fog Computing and the Internet of Things for Smart Healthcare, pp. 253-272. Elsevier, 2026.

[14] Kumar, Mandeep, and Bhaskar Mondal. "Quantum secure multiparty secret sharing using quantum non-locality." Quantum Studies: Mathematics and Foundations 13, no. 1 (2026): 3.

[15] Gurunath, R., Debabrata Samanta, and Yashas G. Goutham. "Progressions and unfilled gaps in homomorphic encryption for emerging application areas: A comprehensive literature review and preface." IoT Security (2026): 333-357.

[16] Pacharla, Nagaraju, and Srinivasa Reddy Konda. "A Review on Secure

Outsourcing and Privacy-Preserving Traffic Monitoring in Fog-Enabled VANETs." Recent Advances in Computer Science and Communications 19, no. 1 (2026): E26662558339759.

[17] Guru, R. Pavithra, Thomas M. Chen, and Mithileysh Sathiyanarayanan. "ABO optimized hybrid Trans-CNN-Bi-GRU approach for intrusion detection in IoT networks: a privacy-preserving solution." Cluster Computing 29, no. 1 (2026): 50.

[18] Fatima, Ruksar, Ayesha Siddiqua, M. tech Student, Aliza Mahvash, and Syeda Sheeba. "Privacy-Preserving Machine Learning: Techniques, Frameworks, and Future Directions." Journal of Systems Engineering and Electronics (ISSN NO: 1671-1793) 35, no. 12 (2025).

[19] Khan, Firoz, C. Sai Varun, Balamurugan Balusamy, and Jeevanandam Jotheeswarn. "Quantized Hybrid Privacy Preserving Approach for Federated Learning with Flower Framework." In 2025 13th International Symposium on Digital Forensics and Security (ISDFS), pp. 1-7. IEEE, 2025.

[20] Zeng, Wenxuan, Tianshi Xu, Yi Chen, Yifan Zhou, Mingzhe Zhang, Jin Tan, Cheng Hong, and Meng Li. "Towards efficient privacy-preserving machine learning: A systematic review from protocol, model, and system perspectives." arXiv preprint arXiv:2507.14519 (2025).

[21] Anoop, M., G. Michael, and Jeni Gracia RC. "Privacy-Preserving Data Mining Techniques in Big Data Environments." In 2025 3rd International Conference on Sustainable Computing and Data Communication Systems (ICSCDS), pp. 696-702. IEEE, 2025.

[22] Shukla, Abhay, Shubham Chaurasia, Gaurav Pandey, Sanjeev Kumar Shukla, Subhash Singh Parihar, and Edwin Prabhakar PB. "Privacy-Preserving Data Mining Methods Metrics and Applications in Healthcare Informatics." In ITM Web of Conferences, vol. 76, p. 04002. EDP Sciences, 2025.

[23] Bhavani, B., and Haritha Donavalli. "Privacy-Aware ML Framework for Dynamic Query Formation in Multi-Dimensional Data." International Journal of Advanced Computer Science & Applications 16, no. 9 (2025).

[24] Hossain, Ismail, Sai Puppala, Md Jahangir Alam, Samrendra Roy, Kazuma Kobayashi, Syed Bahauddin Alam, and Sajedul Talukder. "A Privacy-Aware Cyber Attack Detection Framework for Advanced Reactors Using Data Fusion and Quantum Deep Learning." In 2025 Nuclear Plant Instrumentation and Control and Human-Machine Interface Technology, NPIC and HMIT 2025, pp. 1377-1386. American Nuclear Society, 2025.

[25] Al Kinoon, Mohammed. "A comprehensive and comparative examination of healthcare data breaches: assessing security, privacy, and performance." (2024).

[26] Bonett, Stephen, Willey Lin, Patrina Sexton Topper, James Wolfe, Jesse Golinkoff, Aayushi Deshpande, Antonia Villarruel, and José Bauermeister. "Assessing and improving data integrity in web-based surveys: Comparison of fraud detection systems in a COVID-19 study." JMIR formative research 8 (2024): e47091.

[27] Najar, Ali Vafaee, Leili Alizamani, Marziye Zarqi, and Elaheh Hooshmand. "A global scoping review on the patterns of medical fraud and abuse: integrating data-driven detection, prevention, and legal responses." Archives of Public Health 83, no. 1 (2025): 43.

[28] Ahmed, Mohamed Mustaf, Olalekan John Okesanya, Majd Oweidat, Zhinya Kawa Othman, Shuaibu Saidu Musa, and Don Eliseo Lucero-Prisno III. "The ethics of data mining in healthcare: challenges, frameworks, and future directions." BioData mining 18, no. 1 (2025): 47.

[29] Hamid, Zain, Fatima Khalique, Saba Mahmood, Ali Daud, Amal Bukhari, and Bader Alshemaimri. "Healthcare insurance fraud detection using data mining." BMC Medical Informatics and Decision Making 24, no. 1 (2024): 112.

[30] AlMarshoud, Mishri, Mehmet Sabir Kiraz, and Ali H. Al-Bayatti. "Security, privacy, and decentralized trust management in VANETs: A review of current research and future directions." ACM Computing Surveys 56, no. 10 (2024): 1-39.