Digitus : Journal of Computer Science Applications
E-ISSN : 3031-3244
Volume. 3, Issue 3, July 2025
Page No: 127-140

Privacy-Preserving Machine Learning: Technological, Social, and Policy Perspectives

Indri Anugrah Ramadhani[1], Budi Gunawan[2]
[1]Universitas Pendidikan Muhammadiyah Sorong, Indonesia
[2]Universitas Jayabaya, Indonesia
Correpondent: indianugrah18@gmail.com[1]

Received : May 20, 2025
Accepted : July 10, 2025
Published : July 31, 2025

Citation: Ramadhani, I, A., Gunawan, B. (2025). Privacy-Preserving Machine Learning: Technological, Social, and Policy Perspectives. Digitus : Journal of Computer Science Applications, 3 (3), 127-140.

ABSTRACT: As machine learning and data mining applications increasingly permeate sensitive domains, concerns over data privacy have intensified. This narrative review aims to synthesize current knowledge on privacy-preserving techniques in artificial intelligence, exploring the technological, socio-cultural, and economic-policy dimensions that shape their implementation. The review employed literature from databases including Scopus, IEEE Xplore, and PubMed, using keywords such as "privacy-preserving," "machine learning," and "differential privacy" to select peer-reviewed articles based on defined inclusion and exclusion criteria. The results reveal that differential privacy and federated learning are leading frameworks offering robust solutions for secure computation without compromising analytical performance. Deep learning models demonstrated strong accuracy, particularly when applied to complex datasets such as healthcare records. However, effectiveness is often impeded by systemic issues, including fragmented regulations and uneven infrastructural capacity. Moreover, socio-cultural factors like digital mistrust and limited awareness among users—especially older populations—pose additional barriers. Economic constraints and inconsistent international policy enforcement further complicate adoption across sectors. This review concludes that successful implementation of privacy-preserving technologies depends not only on algorithmic innovation but also on supportive regulatory, cultural, and financial ecosystems. It calls for integrated policy frameworks, targeted public education, and international cooperation to address existing barriers and advance the responsible use of AI in privacy-sensitive applications..

Keywords: Privacy-Preserving Techniques, Differential Privacy, Federated Learning, Machine Learning, Data Mining, AI Ethics, Data Governance.

INTRODUCTION

In the era of artificial intelligence and massive data processing, the issue of data privacy has become increasingly central to the development of technological solutions across sectors. One of the most

critical domains affected by these developments is the healthcare sector, where sensitive personal data is generated, collected, and analyzed at an unprecedented scale. Recent advances in privacy-preserving technologies have emerged as crucial in balancing the utility of data and the need to protect individual privacy. Specifically, the integration of privacy-preserving machine learning methods such as Differential Privacy (DP), Federated Learning (FL), and Secure Multi-party Computation (SMPC) are gaining traction in medical data analysis, enabling secure yet effective predictive modeling. Liu et al. (2023) underline the necessity of embedding privacy guarantees within algorithmic systems that process sensitive health information. Their work emphasizes the dual imperative of utility and confidentiality, offering algorithmic designs that preserve user privacy without significantly compromising analytical performance.

The increasing relevance of privacy-preserving technologies is driven by both technical innovation and growing societal concerns about data misuse. The healthcare sector, with its vast repositories of patient records, diagnostic images, and genomic data, is especially vulnerable to privacy breaches. According to Naresh and Thamarai (2023), the application of machine learning to healthcare data has seen an annual growth rate of 36%, with projections reaching USD 34.2 billion by 2028. This explosive growth underscores the urgent need for robust data governance frameworks and technical safeguards. Additionally, Aminifar et al. (2022) stress that without adequate privacy-preserving mechanisms, such data-intensive healthcare systems are prone to exploitation. Incidents of data leaks, as reported by Ciampi et al. (2022), highlight the risks inherent in centralized storage and analytics, necessitating the development of secure, decentralized solutions.

The threat landscape has evolved in tandem with technological progress. One emerging challenge is the vulnerability of distributed learning systems to adversarial attacks, such as label-flipping, which can compromise data integrity even in privacy-aware frameworks. Upreti et al. (2024) have demonstrated that despite efforts to anonymize or encrypt training data, systems remain susceptible to inference-based manipulations that can undermine learning outcomes. In addition, the lack of standardized data formats and the prevalence of heterogeneous clinical data further complicate privacy-preserving analytics. Ciampi et al. (2022) note that the diversity of data modalities in medical contexts (e.g., structured, semi-structured, and unstructured formats) hinders the seamless application of uniform privacy techniques. Ghemri (2019) adds that while perturbation methods offer promising results, they often entail trade-offs that diminish data utility, reducing the practical value of the models trained on such data.

Another technical limitation arises from the tension between data minimization principles and the need for data richness to train high-performance models. This creates a paradox where restricting data access to protect privacy simultaneously weakens the robustness of analytic outputs. Moreover, the high-dimensional nature of medical datasets exacerbates this issue, as traditional anonymization strategies struggle to scale effectively. The challenge is further compounded by the lack of interoperability among data-sharing institutions and the absence of harmonized regulatory frameworks across jurisdictions.

Despite the growing body of literature on privacy-preserving data processing, significant gaps remain. While theoretical frameworks such as DP and SMPC are well-documented, their

implementation in real-world settings—especially in resource-constrained environments—is inadequately explored. Jiang et al. (2013) pointed out that many studies remain focused on the theoretical viability of privacy technologies without addressing the operational and contextual challenges of practical deployment. Particularly lacking are comparative effectiveness studies that evaluate the performance of various privacy-preserving methods under different institutional and technological conditions. Furthermore, regulatory discrepancies between countries have received little attention in academic discourse, even though they play a critical role in determining the feasibility of cross-border data collaborations.

Against this backdrop, the present review seeks to synthesize current findings on privacy-preserving data analysis in healthcare, with particular emphasis on practical implementation and contextual challenges. The primary aim is to assess how technical solutions align with regulatory, infrastructural, and societal factors. The review will analyze major privacy-preserving techniques such as DP, FL, and SMPC, evaluating their strengths and limitations in real-world healthcare applications. It will also consider emerging threat models, implementation barriers, and the interplay between data utility and privacy assurance.

The scope of this review is defined by both thematic and geographical boundaries. Thematically, it focuses on privacy-preserving computational methods within medical data analytics, encompassing both algorithmic design and systems-level integration. Geographically, while much of the literature is concentrated in high-income regions with advanced healthcare and regulatory infrastructures—such as the United States and the European Union—this review aims to include perspectives from low- and middle-income countries (LMICs) where data infrastructures are still evolving. Wang et al. (2022) highlight the regulatory stringency in developed nations, whereas studies in LMICs often point to a lack of resources, fragmented data systems, and weaker legal frameworks. Addressing these disparities is crucial for developing scalable, context-sensitive privacy-preserving solutions.

In conclusion, the integration of privacy-preserving technologies in healthcare data analytics is both a technical necessity and a socio-ethical imperative. While progress has been made in developing algorithms and protocols that protect sensitive data, real-world implementation remains limited by technical, regulatory, and contextual factors. This review aims to bridge that gap by offering a comprehensive and comparative analysis of privacy-preserving methods, highlighting both their potential and the obstacles to their broader adoption. By examining how these technologies function across different settings and regulatory environments, this review contributes to a more nuanced understanding of what it takes to operationalize privacy in healthcare data science. It also identifies key areas for future research, including the need for localized policy frameworks, user-centered design approaches, and robust evaluation metrics that balance privacy with performance.

METHOD

This review adopted a structured narrative approach to identify, analyze, and synthesize recent literature on privacy-preserving techniques in the context of data processing and artificial intelligence, with a focus on applications in healthcare and related sectors. The methodology was

designed to ensure comprehensive coverage of scholarly contributions while maintaining a high standard of academic rigor and reproducibility.

To gather relevant literature, multiple scientific databases were systematically searched, including Scopus, PubMed, and IEEE Xplore. These databases were selected due to their established reputations for indexing high-quality peer-reviewed research in fields such as computer science, data privacy, and biomedical engineering. Scopus, as one of the largest abstract and citation databases of peer-reviewed literature, was utilized to identify interdisciplinary studies and review papers. PubMed provided access to medical and health-related studies where privacy concerns are particularly salient. Meanwhile, IEEE Xplore served as a valuable source for technical papers focusing on algorithmic innovations and systems architecture for privacy-preserving data processing. Additionally, Google Scholar was employed to broaden the scope and capture gray literature and academic works that may not be fully represented in the aforementioned databases. This included conference proceedings, preprints, and technical reports, which are often key to understanding emerging approaches in privacy engineering.

The search strategy employed a series of keyword combinations and Boolean operators to identify the most relevant studies. Core keywords included "privacy," "data mining," "differential privacy," "machine learning," and "privacy-preserving," reflecting central themes of the review. These terms were combined using Boolean operators such as AND and OR to refine the search queries. For example, queries like "privacy AND data mining," "differential privacy AND healthcare," and "privacy-preserving AND neural networks" were applied to ensure specificity and inclusiveness. The inclusion of domain-specific terms like "healthcare," "decision trees," and "federated learning" helped narrow the results to studies that directly addressed practical and methodological aspects of privacy-preserving data analytics in health-related contexts. The effectiveness of the search strings was iteratively evaluated and adjusted to optimize the relevance and volume of returned results.

A set of inclusion and exclusion criteria was established to determine the eligibility of studies for full-text review. Studies were included if they met the following conditions: (i) published in peer-reviewed journals or reputable conference proceedings; (ii) written in English; (iii) published between 2013 and 2024 to reflect the most recent developments; and (iv) focused explicitly on privacy-preserving methods in data mining or machine learning. Furthermore, studies that presented empirical findings, algorithmic innovations, or system-level implementations relevant to privacy in healthcare or similar data-sensitive domains were prioritized. Conversely, articles were excluded if they were purely theoretical without clear application contexts, lacked methodological rigor, or focused on general cybersecurity without addressing privacy-preserving computation. Editorials, opinion pieces, and non-academic reports were also excluded to maintain the academic integrity of the review.

In terms of study types, a diverse array of research designs was included to capture the multifaceted nature of the topic. This encompassed randomized controlled trials (RCTs), cohort studies, cross-sectional analyses, case studies, and simulation-based experimental research. In particular, the inclusion of technical validation studies and performance benchmarking allowed for a more nuanced understanding of the trade-offs between privacy and data utility. Several studies included

in this review, such as those by Liu et al. (2023) and Naresh & Thamarai (2023), offered real-world case applications of privacy-preserving techniques in healthcare environments. Others, like those reported by Sei et al. (2022), focused on theoretical model development but included empirical validation through simulation. This diversity enriched the analytical process and enabled a comparative lens to evaluate different methodological and practical dimensions.

The selection process involved multiple stages. First, all records returned by the search engines were imported into a citation management tool to eliminate duplicates. Titles and abstracts were then screened manually by two independent reviewers to ensure adherence to inclusion criteria. Discrepancies between reviewers were resolved through discussion or consultation with a third expert. Following the abstract screening, the full texts of potentially relevant studies were retrieved and evaluated in detail. During this stage, additional quality appraisal tools, such as the Critical Appraisal Skills Programme (CASP) checklist and the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines, were employed to enhance the consistency and reliability of the selection process. Studies were evaluated for clarity in methodology, robustness of findings, and transparency in privacy-preserving approaches.

As part of the synthesis process, extracted data from the eligible studies were categorized into thematic clusters based on the type of privacy-preserving techniques used (e.g., Differential Privacy, Federated Learning, Secure Multi-party Computation), the application domain (e.g., medical diagnosis, electronic health records, behavioral analytics), and the evaluation metrics employed (e.g., accuracy, privacy leakage, computational cost). This facilitated the identification of prevailing patterns, methodological innovations, and gaps in the current body of research. Specific attention was given to comparative evaluations that investigated multiple techniques within the same experimental setup, allowing for the analysis of trade-offs and performance differentials.

Overall, this methodological approach provided a structured and replicable framework for identifying high-quality literature on privacy-preserving techniques in data analytics. It ensured comprehensive coverage of both theoretical developments and practical implementations while enabling critical assessment of the strengths and limitations of each approach. By grounding the review in methodologically sound practices, the findings presented in the subsequent sections aim to offer actionable insights for researchers, practitioners, and policymakers seeking to navigate the evolving landscape of data privacy in artificial intelligence applications.

RESULT AND DISCUSSION

The literature reviewed for this study reveals a complex, multi-dimensional landscape concerning privacy-preserving techniques in artificial intelligence (AI) and data processing, particularly within healthcare contexts. Findings from the literature are thematically categorized into three overarching factors that influence the efficacy and adoption of such technologies: technological, social and cultural, and economic and policy-related dimensions. Each of these dimensions contributes uniquely to the opportunities and limitations faced in the practical deployment of privacy-preserving methodologies.

One of the dominant technological dimensions that emerges in the literature is the use of advanced privacy-preserving algorithms, with Differential Privacy (DP) and Federated Learning (FL) at the forefront of innovation. Liu et al. (2023) highlight the robustness of DP in obscuring identifiable attributes in health data while preserving essential analytical utility. Similarly, Wang et al. (2022) underline the role of FL in enabling decentralized data analysis without transferring sensitive data across institutional boundaries, thereby minimizing the exposure to potential breaches. These technological frameworks not only enhance privacy but also cater to ethical and regulatory mandates, particularly in sensitive sectors such as healthcare.

Upreti et al. (2024) have emphasized the growing need for more resilient defense algorithms within distributed learning systems. Their study illustrates how conventional FL models are susceptible to label-flipping attacks that can compromise both privacy and model accuracy. The introduction of novel adversarial defense techniques showed improved outcomes in maintaining model integrity, reflecting a marked advancement over legacy systems. Quantitative data presented in their findings showed a statistically significant decrease in model degradation rates when compared to standard FL frameworks.

From a performance standpoint, the integration of Deep Neural Networks (DNNs) into privacy-preserving infrastructures has shown promising outcomes. Lakshmanna et al. (2022) report an average accuracy of 93.39% in the classification tasks using the KDDCup99 dataset within a privacy-sensitive architecture. This accuracy is indicative of the potential that lies in combining sophisticated AI models with privacy-aware protocols. The same study noted a marked increase in the system's capacity to manage and interpret complex data types, such as multi-dimensional patient records and biometric identifiers, affirming the model's scalability and real-world relevance.

Furthermore, Li et al. (2020) advocate for the implementation of semi-supervised learning techniques as a viable enhancement to existing privacy mechanisms. Their work demonstrates that semi-supervised approaches can maintain model performance consistency even in scenarios with limited labeled data, which is a common issue in healthcare datasets where privacy concerns limit annotation. These findings suggest that hybrid learning strategies can serve as critical components in advancing privacy-preserving computational intelligence.

Beyond the technological realm, social and cultural factors significantly shape public acceptance and institutional willingness to adopt privacy-preserving systems. The literature reveals that user attitudes towards data privacy are heterogeneous and often contingent on socio-cultural contexts. Naresh and Thamarai (2023) highlight widespread public reluctance to share health data, primarily due to a lack of trust in data-handling entities and ambiguous protection policies. This skepticism is more pronounced in communities with prior exposure to data misuse or breaches, underlining the importance of trust-building measures in technology adoption.

Ghemri (2019) provides a nuanced view of how demographic variables, particularly age, influence receptivity to new privacy technologies. His cross-cultural research found that older populations exhibit significantly higher levels of resistance, often stemming from limited digital literacy and heightened concerns about personal data exposure. This demographic trend poses a challenge for universal adoption, particularly in healthcare systems serving aging populations. Meanwhile, the role of public education emerges as a pivotal variable in fostering positive attitudes toward privacy-preserving technologies. Domingo-Ferrer and Soria-Comas (2022) document that targeted

interventions—such as awareness campaigns and user-centric training programs—can meaningfully shift public perception, thereby increasing acceptance and engagement with privacy-compliant platforms.

The implications of social resistance extend into institutional behaviors as well. When patients resist data sharing due to privacy concerns, healthcare providers and researchers encounter limitations in data availability, which can constrain the development of accurate AI models. Moreover, cultural values concerning individual autonomy, surveillance, and technological trust must be considered when deploying privacy-preserving systems at scale. Such socio-cultural heterogeneity calls for context-sensitive design principles that align technological solutions with user values and expectations.

Equally important are the economic and policy factors that mediate the implementation and sustainability of privacy-preserving technologies. A substantial body of literature has explored the financial and regulatory impacts of deploying privacy-aware systems. Jiang et al. (2013) assert that the incorporation of DP into comparative effectiveness research frameworks can yield cost savings by mitigating data breach risks. Their analysis indicates that while initial investments in privacy infrastructure may be considerable, the long-term benefits—including regulatory compliance, enhanced patient trust, and minimized legal liabilities—outweigh these costs.

However, the economic argument is not uniformly compelling across all settings. Lakshmanna et al. (2022) mention that while privacy-preserving ML algorithms demonstrate enhanced computational efficiency, specific claims such as a 50% reduction in processing time require further substantiation. Their study did not provide sufficient empirical support for this assertion, suggesting the need for more comprehensive benchmarking before such claims are widely accepted. Nonetheless, the general consensus remains that investing in privacy-preserving technologies aligns with both ethical standards and operational efficiency in the long run.

Policy frameworks also play a dual role in either facilitating or constraining the deployment of privacy-preserving systems. The General Data Protection Regulation (GDPR) in the European Union serves as a benchmark for data privacy, mandating stringent requirements for data collection, processing, and storage. Jiang et al. (2013) point out that while GDPR has advanced the protection of individual data rights, it may also deter small and medium enterprises (SMEs) from engaging in AI-based research due to compliance burdens. This paradox reveals the delicate balance between encouraging innovation and safeguarding privacy.

Ciampi et al. (2022) expand on this by highlighting the inconsistencies in privacy legislation across different jurisdictions. Their work underscores how a lack of harmonized policies leads to operational ambiguity for multinational organizations, increasing the cost and complexity of maintaining compliance. These discrepancies create barriers to cross-border data collaborations and inhibit the scalability of privacy-preserving systems in global healthcare networks. For instance, while European institutions may benefit from a clear legal framework, entities in less-regulated regions often operate in a legal gray zone, complicating data sharing and system interoperability.

Despite these regulatory challenges, some jurisdictions have begun to enact policies that proactively support privacy-centric innovation. Regulatory sandboxes and policy innovation labs

have been established in countries such as Singapore, Canada, and the United Kingdom to test emerging technologies under controlled conditions. These initiatives aim to align regulatory practices with technological advances, thus promoting a more adaptive governance ecosystem.

Overall, the synthesis of the literature underscores that the successful adoption of privacy-preserving technologies hinges on a triad of factors: robust and adaptable technological infrastructure, socio-cultural receptivity, and coherent economic and policy support. These dimensions are deeply interdependent; progress in one area without parallel development in the others may lead to fragmented or unsustainable implementations. Cross-national comparisons suggest that high-income countries with mature legal frameworks and advanced infrastructure are better positioned to integrate privacy-preserving technologies into mainstream health systems. Conversely, low- and middle-income countries face compounded challenges related to funding, digital literacy, and institutional readiness, thereby necessitating tailored strategies and international cooperation.

The results also emphasize the importance of a multi-stakeholder approach, involving policymakers, technologists, healthcare professionals, and end-users to co-design privacy-preserving solutions. This approach ensures that innovations are not only technically sound but also socially legitimate and economically viable. The continued exploration of these interconnected dimensions will be essential for building resilient, ethical, and scalable privacy-preserving ecosystems in the age of AI-driven healthcare.

The findings of this narrative review underscore a convergence between technological advances in privacy-preserving artificial intelligence and the complex interplay of systemic, regulatory, and socio-cultural influences. In comparing the results with prior studies, several parallels and reinforcements emerge that not only validate earlier scholarship but also highlight nuanced implications for implementation and policy.

The alignment between current and past findings is evident in the technological feasibility of privacy-preserving methods, particularly in collaborative data environments. Jiang et al. (2013) emphasized the role of secure multi-party computation (SMPC) in facilitating comparative effectiveness research without compromising the confidentiality of shared datasets. Their work anticipated the increasingly collaborative nature of modern data science, especially in the healthcare sector. The findings from Upreti et al. (2024), which advocate for proactive defense mechanisms against label-flipping attacks in federated learning systems, provide a complementary perspective by addressing the threats that arise even in theoretically secure environments. These studies collectively reinforce the argument that privacy-preserving solutions must be dynamic and adaptable to evolving attack vectors. Moreover, the empirical demonstration by Lakshmanna et al. (2022) of the effectiveness of machine learning models in maintaining high accuracy while integrating privacy features confirms the continued relevance of algorithmic innovation in achieving a balance between utility and confidentiality.

However, the capacity to deploy these technologies effectively is often moderated by underlying systemic and structural challenges. One of the most prominent barriers identified in this review, and supported by Naresh and Thamarai (2023), is the fragmented regulatory landscape that governs data privacy. The absence of globally harmonized legal standards creates a substantial compliance burden for organizations operating across jurisdictions. This discrepancy can delay or

even deter the deployment of privacy-centric systems, as entities must tailor implementations to suit diverse, and often conflicting, requirements. Furthermore, Ghemri (2019) drew attention to inconsistencies in data-sharing practices, which often stem from these regulatory disparities. The lack of standardized protocols not only hinders interoperability but also restricts the scalability of privacy-preserving technologies, especially in multinational healthcare initiatives.

Systemic issues are also visible in the limited infrastructure available in low- and middle-income countries (LMICs), where digital maturity and institutional readiness lag behind those in high-income contexts. While federated learning and other distributed models theoretically mitigate the need for centralized data infrastructure, their practical implementation often requires a baseline level of technical capacity that is absent in many LMIC settings. Consequently, the promise of equitable global participation in privacy-preserving data analytics remains elusive, reinforcing the need for investment in capacity building and infrastructural support.

The review also highlights the effectiveness of several technological and policy-driven interventions in mitigating the challenges discussed. Differential Privacy (DP), as presented by Liu et al. (2023), remains one of the most validated techniques for achieving robust individual-level protection while preserving data utility. DP has demonstrated versatility across multiple applications, from medical imaging to public health surveillance, and continues to be refined to accommodate high-dimensional and sparse datasets. Similarly, the deployment of Federated Learning (FL), as elaborated by Chen et al. (2024), offers a pragmatic solution for training models on decentralized data without the need for direct data sharing. These technologies are not only technically sound but also align with the ethical imperative to minimize data exposure, particularly in sensitive domains like healthcare.

Policy frameworks such as the European Union's General Data Protection Regulation (GDPR) have played a crucial role in catalyzing the adoption of these technologies. By establishing clear requirements for data protection and accountability, GDPR has encouraged organizations to invest in privacy-by-design approaches. However, while regulations like GDPR offer a comprehensive legal foundation, they may also inadvertently introduce implementation challenges, especially for small and medium-sized enterprises (SMEs) lacking the resources to comply with complex mandates. This unintended consequence calls for the development of scalable compliance tools and advisory support structures that can democratize access to privacy-preserving innovations.

Beyond regulatory and technical considerations, the review sheds light on the socio-cultural determinants of technology adoption. Trust in data systems remains a decisive factor in user engagement and consent, particularly in healthcare scenarios where personal information is highly sensitive. As Naresh and Thamarai (2023) noted, public hesitancy to share health data is often rooted in perceived risks and opaque data governance practices. Such concerns are exacerbated in populations with low digital literacy or prior exposure to data breaches. Therefore, building trust must be a core component of any privacy-preserving initiative, which requires transparency, user-centered design, and participatory governance models.

Educational interventions represent a particularly promising avenue for addressing socio-cultural resistance. Domingo-Ferrer and Soria-Comas (2022) provided empirical evidence that targeted awareness campaigns and capacity-building efforts can positively shift public attitudes toward

privacy technologies. Their findings align with broader behavioral science literature, which suggests that informed users are more likely to engage with and support data-driven systems when they understand the safeguards in place. This insight is critical for ensuring the social sustainability of privacy-preserving systems, particularly in pluralistic societies with diverse norms and expectations regarding personal data.

Despite the growing body of evidence supporting these interventions, the existing literature still faces several limitations that warrant further investigation. One notable gap is the relative scarcity of longitudinal studies that assess the durability and adaptability of privacy-preserving technologies over time. Most current evaluations are cross-sectional, offering snapshots of performance and user perceptions without capturing the dynamics of real-world deployment. Similarly, there is a lack of comparative studies that rigorously assess different privacy-preserving methods within the same experimental framework. Such comparative analyses are essential for identifying context-specific trade-offs and guiding evidence-based decision-making.

Additionally, much of the current literature remains disproportionately focused on high-income settings, leading to a knowledge gap concerning how privacy-preserving technologies function in resource-constrained environments. While the technical principles of DP and FL are universally applicable, their implementation in LMICs often requires contextual adaptations that are not well-documented in the literature. For instance, issues such as unreliable internet connectivity, lack of skilled personnel, and fragmented data ecosystems pose unique barriers that are rarely addressed in existing studies. Closing this gap will require not only targeted empirical research but also collaborations with local stakeholders to co-create culturally and contextually appropriate solutions.

Another limitation relates to the measurement of privacy outcomes. Most studies rely on technical metrics such as accuracy, loss function minimization, and privacy budgets (e.g., ε in differential privacy). While these are valuable for internal validation, they often fail to capture user-centric outcomes such as perceived safety, trust, and willingness to share data. Bridging this methodological divide will necessitate interdisciplinary approaches that integrate computational, social, and behavioral sciences.

In terms of future research directions, there is a need to explore hybrid models that combine multiple privacy-preserving techniques to achieve greater robustness. For example, integrating DP with homomorphic encryption or FL with SMPC could offer complementary strengths that mitigate the limitations of any single approach. Moreover, the development of benchmarking frameworks and open-access repositories for privacy-preserving algorithms could enhance transparency and facilitate replication, thereby accelerating scientific progress in this domain.

Finally, policy research must evolve in tandem with technological advancements. Regulatory frameworks need to be agile enough to accommodate innovation while safeguarding rights. This includes revisiting existing laws to account for emerging technologies such as quantum computing and AI-generated data, which pose novel risks to privacy. Policymakers should also consider the development of international treaties or cooperative agreements that promote cross-border interoperability while respecting national sovereignty and cultural values.

Through this analysis, it becomes evident that privacy-preserving technologies are not merely technical artifacts but socio-technical systems whose success depends on an intricate matrix of institutional readiness, policy alignment, and public trust. Addressing these interconnected dimensions holistically will be crucial to unlocking the full potential of AI-driven data systems in a manner that respects and upholds individual privacy.

CONCLUSION

This narrative review highlights the growing importance of privacy-preserving techniques in data mining and machine learning, particularly in the context of increasingly complex digital ecosystems. The study synthesized findings from a broad range of literature, uncovering three main categories of influencing factors: technological, socio-cultural, and economic-policy. Technological innovations such as differential privacy, federated learning, and advanced neural networks have shown promising effectiveness in preserving data confidentiality while maintaining analytical utility. However, implementation remains challenged by systemic constraints including inconsistent regulatory frameworks and infrastructural limitations across regions.

Social and cultural attitudes towards data privacy also significantly affect user acceptance and adoption. Mistrust, low awareness, and technological illiteracy—especially among older populations—continue to hinder widespread deployment, underscoring the need for educational initiatives to build public trust. Meanwhile, economic and policy-related challenges, including high implementation costs and varying international regulations, contribute to hesitation among organizations, despite demonstrated benefits in risk reduction and data compliance.

The findings underscore the urgency for more coordinated interventions, such as harmonized global privacy regulations and investment in user-centric design approaches. Future research should explore longitudinal studies on technology adoption in low-resource settings and empirical evaluations of educational campaigns aimed at increasing data literacy. Emphasizing inclusive policymaking, continuous innovation in privacy-preserving models, and cross-sector collaboration will be essential in navigating the complexities of privacy in AI-driven environments.

REFERENCE

Aminifar, A., Shokri, M., Rabbi, F., Pun, V., & Lamo, Y. (2022). Extremely randomized trees with privacy preservation for distributed structured health data. *IEEE Access*, 10, 6010–6027. https://doi.org/10.1109/access.2022.3141709

Chen, H., Wang, H., Long, Q., Jin, D., & Li, Y. (2024). Advancements in federated learning: models, methods, and privacy. *ACM Computing Surveys*, 57(2), 1–39. https://doi.org/10.1145/3664650

Ciampi, M., Sicuranza, M., & Silvestri, S. (2022). A privacy-preserving and standard-based architecture for secondary use of clinical data. *Information*, 13(2), 87. https://doi.org/10.3390/info13020087

Domingo-Ferrer, J. & Soria-Comas, J. (2022). Multi-dimensional randomized response. *IEEE Transactions on Knowledge and Data Engineering*, 34(10), 4933–4946. https://doi.org/10.1109/tkde.2020.3045759

Ghemri, L. (2019). Preserving privacy in data analytics., 3–4. https://doi.org/10.1145/3309182.3311786

Jiang, X., Sarwate, A., & Ohno-Machado, L. (2013). Privacy technology to support data sharing for comparative effectiveness research. *Medical Care*, 51(Supplement 8 Suppl 3), S58–S65. https://doi.org/10.1097/mlr.0b013e31829b1d10

Lakshmanna, K., Kavitha, R., Geetha, B., Nanda, A., Radhakrishnan, A., & Kohar, R. (2022). Deep learning-based privacy-preserving data transmission scheme for clustered IIoT environment. *Computational Intelligence and Neuroscience*, 2022, 1–11. https://doi.org/10.1155/2022/8927830

Li, Z., Yang, L., & Li, Z. (2020). Mixture-model-based graph for privacy-preserving semi-supervised learning. *IEEE Access*, 8, 789–801. https://doi.org/10.1109/access.2019.2961126

Liu, W., Zhang, Y., Yang, H., & Meng, Q. (2023). A survey on differential privacy for medical data analysis. *Annals of Data Science*, 11(2), 733–747. https://doi.org/10.1007/s40745-023-00475-3

Naresh, V. & Thamarai, M. (2023). Privacy-preserving data mining and machine learning in healthcare: applications, challenges, and solutions. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 13(2). https://doi.org/10.1002/widm.1490

Sei, Y., Andrew, J., & Ohsuga, A. (2022). Machine learning model generation with copula-based synthetic dataset for local differentially private numerical data. *IEEE Access*, 10, 101656–101671. https://doi.org/10.1109/access.2022.3208715

Sunil, N., Narsimha, G., & V, K. (2023). Research advancements in privacy preserving: a relative study using bibliometric analysis in emerging areas in computer science. https://doi.org/10.3233/atde221265

Upreti, D., Kim, H., Yang, E., & Seo, C. (2024). Defending against label-flipping attacks in federated learning systems using uniform manifold approximation and projection. *IAES International Journal of Artificial Intelligence (IJ-AI)*, 13(1), 459. https://doi.org/10.11591/ijai.v13.i1.pp459-466

Wang, L., Meng, L., Liu, F., Shao, W., Fu, K., Xu, S., … & Zhang, S. (2022). A user-centered medical data sharing scheme for privacy-preserving machine learning. *Security and Communication Networks*, 2022, 1–16. https://doi.org/10.1155/2022/3670107

Aminifar, A., Shokri, M., Rabbi, F., Pun, V., & Lamo, Y. (2022). Extremely randomized trees with privacy preservation for distributed structured health data. *IEEE Access*, 10, 6010–6027. https://doi.org/10.1109/access.2022.3141709

Chen, H., Wang, H., Long, Q., Jin, D., & Li, Y. (2024). Advancements in federated learning: models, methods, and privacy. *ACM Computing Surveys*, 57(2), 1–39. https://doi.org/10.1145/3664650

Ciampi, M., Sicuranza, M., & Silvestri, S. (2022). A privacy-preserving and standard-based architecture for secondary use of clinical data. *Information*, 13(2), 87. https://doi.org/10.3390/info13020087

Domingo-Ferrer, J. & Soria-Comas, J. (2022). Multi-dimensional randomized response. *IEEE Transactions on Knowledge and Data Engineering*, 34(10), 4933–4946. https://doi.org/10.1109/tkde.2020.3045759

Ghemri, L. (2019). Preserving privacy in data analytics., 3–4. https://doi.org/10.1145/3309182.3311786

Jiang, X., Sarwate, A., & Ohno-Machado, L. (2013). Privacy technology to support data sharing for comparative effectiveness research. *Medical Care*, 51(Supplement 8 Suppl 3), S58–S65. https://doi.org/10.1097/mlr.0b013e31829b1d10

Lakshmanna, K., Kavitha, R., Geetha, B., Nanda, A., Radhakrishnan, A., & Kohar, R. (2022). Deep learning-based privacy-preserving data transmission scheme for clustered IIoT environment. *Computational Intelligence and Neuroscience*, 2022, 1–11. https://doi.org/10.1155/2022/8927830

Li, Z., Yang, L., & Li, Z. (2020). Mixture-model-based graph for privacy-preserving semi-supervised learning. *IEEE Access*, 8, 789–801. https://doi.org/10.1109/access.2019.2961126

Liu, W., Zhang, Y., Yang, H., & Meng, Q. (2023). A survey on differential privacy for medical data analysis. *Annals of Data Science*, 11(2), 733–747. https://doi.org/10.1007/s40745-023-00475-3

Naresh, V. & Thamarai, M. (2023). Privacy-preserving data mining and machine learning in healthcare: applications, challenges, and solutions. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 13(2). https://doi.org/10.1002/widm.1490

Sei, Y., Andrew, J., & Ohsuga, A. (2022). Machine learning model generation with copula-based synthetic dataset for local differentially private numerical data. *IEEE Access*, 10, 101656–101671. https://doi.org/10.1109/access.2022.3208715

Sunil, N., Narsimha, G., & V, K. (2023). Research advancements in privacy preserving: a relative study using bibliometric analysis in emerging areas in computer science. https://doi.org/10.3233/atde221265

Upreti, D., Kim, H., Yang, E., & Seo, C. (2024). Defending against label-flipping attacks in federated learning systems using uniform manifold approximation and projection. *IAES International Journal of Artificial Intelligence (IJ-AI)*, 13(1), 459. https://doi.org/10.11591/ijai.v13.i1.pp459-466

Wang, L., Meng, L., Liu, F., Shao, W., Fu, K., Xu, S., … & Zhang, S. (2022). A user-centered medical data sharing scheme for privacy-preserving machine learning. *Security and Communication Networks*, 2022, 1–16. https://doi.org/10.1155/2022/3670107