International Journal of Computer Trends and Technology
ISSN: 2231–2803 / https://doi.org/10.14445/22312803/IJCTT-V73I2P109

Volume 73 Issue 2, 71-79, February 2025
© 2025 Seventh Sense Research Group®

Original Article

Ensuring Data Confidentiality through an In-Depth Analysis of Advanced Privacy-Preserving Methodologies in Data Science

Vikas Kumar Jain[1], Jayesh Patil[2], Ashish Kumar Jain[3]

[1,2,3]*Department of Computer Engineering, Institute of Engineering and Technology Devi Ahilya Vishwavidyalaya.*

[1]*Corresponding Author : jainv6644@gmail.com*

Received: 30 December 2024 Revised: 25 January 2025 Accepted: 16 February 2025 Published: 28 February 2025

Abstract - *Privacy-preserving methods are of the greatest significance in the field of information technology as they guarantee ethical and secure data usage. The current review presents an in-depth review of state-of-the-art privacy-preserving methods such as differential privacy, Secure Multi-Party Computation (SMPC), homomorphic encryption, federated learning, and anonymization techniques. The paper presents the theoretical backgrounds, practical applications, limitations, and future advancements of the above methods, focusing on recent developments. After briefly introducing data science's privacy risk, the review presents anonymization techniques such as generalization, suppression, k-anonymity, l-diversity, and t-closeness. Homomorphic encryption, SMPC, differential privacy, and federated learning concepts and applications are also presented, citing their efficacy in protecting sensitive data while enabling data analysis collaboration. In order to emphasize the significance of privacy-preserving strategies in real applications, the study surveys real implementations in sectors such as healthcare, finance, telecommunication, social media, and government. Besides unveiling main issues such as scalability, usability, and adversarial attack resistance, the current study also presents potential future research directions for further development in this area. The current work endeavors to contribute to scholars, policymakers, and practitioners with a profound understanding of advancing ethical and sustainable data-driven decision-making by presenting an in-depth review of privacy-preserving methods and their ethical implications.*

Keywords - *Data confidentiality, Data science, Privacy algorithms, Data security, Decentralized systems.*

1. Introduction

This research aims to determine the strengths, weaknesses, and relative efficacy of different privacy-preserving techniques through a systematic review of different use cases and data types. Through critically examining emerging technologies, this research aims to guide stakeholders in selecting proper data privacy techniques for their respective domains. At the same time, as data collection technologies become more advanced and widespread, the risks of breaches, misuse, and unauthorized access to sensitive data are on the rise. It is, therefore, more important than ever that data scientists use strong privacy-preserving techniques.

Privacy-protection techniques include a wide range of techniques and tools used to protect personal data while allowing it to be used continuously for analysis and decision-making. The use of large volumes of data has become necessary for facilitating innovation in different domains and knowledge discovery in the rapidly changing field of data science. The use of large-scale data has become necessary in the rapidly changing field of data science to facilitate innovation in different domains and knowledge creation. Nevertheless, this exponential increase in data collection poses important privacy and security issues. The risks of access, misuse, and compromise of sensitive data are rising as data collection techniques become more advanced and widespread. It is, therefore, never more essential that data scientists use strong privacy-preserving techniques in data research.

2. Literature Review

The rapid expansion in data-intensive applications has generated the demand for robust methods for preserving privacy. These methods safeguard secret information without compromising beneficial analysis. Several methods have been researched, each with some merits and some demerits.

Some of them are federated learning, Secure Multi-Party Computation (SMPC), encryption, differential privacy, and anonymization methods. The paper's second half investigates previous work on these methods, their merits and demerits, and future work directions.

One of the main methods of protecting personal data in datasets is anonymization. In order to strip datasets of identifying information, techniques like k-anonymity (Sweeney 2002)[2], l-diversity (Machanavajjhala et al., 2007)[3], and t-closeness (Li et al., 2007)[4] are commonly employed. These techniques have, however, been found to be susceptible to background information and linking attacks (Narayanan & Shmatikov, 2008)[5], which is concerning regarding how effective they will continue to be in the future. Although more recent studies have sought to enhance these techniques through differential private anonymization frameworks and synthetic data (Dwork et al., 2008)[6], achieving an effective balance between data utility and privacy protection remains.

By enabling computation on encrypted data without decrypting it, encryption techniques—namely homomorphic encryption (HE) (Gentry, 2009)—have transformed how we can compute while maintaining privacy. Although it has a computationally expensive cost that makes it impractical to use in real life, Fully Homomorphic Encryption [8] (FHE) can carry out complex computations while maintaining people's anonymity (Lauter et al., 2011)[9]. Nevertheless, for secure access control in cloud computing, Searchable Encryption (SE) and Attribute-Based Encryption (ABE) (Boneh & Waters, 2007) have been explored. In spite of these developments, studies show that scaling, efficiency, and real-time processing remain significant challenges for encryption-based privacy models.

Differential privacy (DP), originally introduced by Dwork (2008), is now a mathematical framework for formal statistical analysis while maintaining individual privacy . It introduces random noise (Laplace, Gaussian, or exponential mechanisms) into query responses to avoid re-identification. In recent studies, DP has been applied to machine learning models (Abadi et al., 2016), making AI applications privacy-preserving. However, a significant limitation is still the trade-off between model accuracy and privacy budgets (ε) (Balle et al., 2020) [7]. DP also needs to be further enhanced because adversarial attacks, such as membership inference (Shokri et al., 2017), have shown weaknesses in their usability.

Several parties can cooperate to construct functions from their respective information without exposing data, thanks to Secure Multi-Party Computation (SMPC). Traditional methods like Shamir's Secret Sharing (Shamir, 1979) [10] and Yao's Garbled Circuits (Yao, 1986) [11] found secure computing with privacy guarantees. More recent innovations have demonstrated the method's usefulness in applications like healthcare and finance, for instance, privacy-preserving federated learning, by leveraging SMPC (Bonawitz et al., 2019). Research, however, suggests that security issues, communication issues, and high computing needs are ongoing impediments to the wider use of SMPC (Mohassel & Zhang, 2017).

One of the possible ways to train machine learning models on independent data sources without compromising data in privacy is called federated learning (FL) (McMahan et al., 2017). FL reduces data exposure threats while training by keeping data localized on devices [12]. FL is effective where privacy is a concern, such as edge computing, IoT, and healthcare (Kone\rný et al., 2016).

However, threats exist owing to challenges like model poisoning attacks, non-iid data distributions, and costly communications (Bagdasaryan et al., 2020) [13]. To counter such challenges, hybrid approaches combining FL with homomorphic encryption and differential privacy [14] [15] have been proposed; however, their feasibility is still under investigation.

3. Methodology

In order to select and gather appropriate studies, papers, and articles that examine privacy protection methods in data science, the research begins with a comprehensive literature review. Through conducting this first step, the analysis spans a wide range of methods and addresses the existing advancements in the field. This method provides a basis for comprehending the range of methods that have been suggested and applied in the field.

After evaluation, a list of criteria is drawn up to systematically select the most appropriate documents to be further analyzed.

To locate research that is solving privacy issues in data science, introducing novel methodologies, and making important contributions to the study of privacy protection—practically and theoretically—is the end target of the procedure. Shortlisting the most meaningful and relevant research is the target.

The privacy-preserving methods in the literature are then classified using a classification system. The framework classifies the methods into various classes, such as data perturbation methods like noise addition, data swapping, and synthetic data generation; encryption methods like homomorphic encryption, searchable encryption, and secure multi-party computation; and anonymization [16] methods like k-anonymity, l-diversity, t-closeness, and differential privacy. It also mentions access control methods like role-based access control, attribute-based encryption, blockchain-based access control [17], and privacy-preserving machine learning methods [18] like federated learning, secure aggregation, and model inversion prevention.

Finally, it considers new approaches like frameworks for privacy-preserving data exchange and privacy-preserving data mining algorithms.

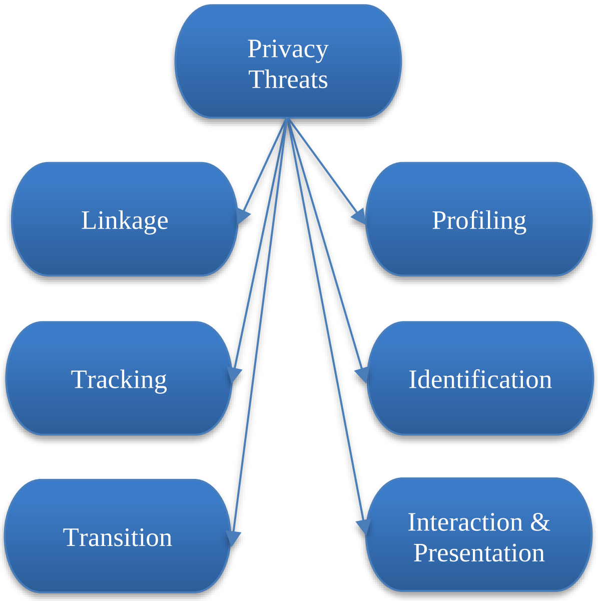

Fig. 1 Data science privacy threats

In the second phase, we will elaborate in detail on each privacy-preserving method and the corresponding algorithms and methodologies. This comparison involves describing the fundamental principles behind each method, their advantages and disadvantages as data protectors, and including case studies or examples of where they have performed best. We will also compare the approaches in each category, considering factors such as how scalable they can be, their computational capabilities, and how well they will perform with various types of data. Figure 1 illustrates all the privacy attacks and potential weaknesses in the context of data science.

In order to quantify the effectiveness of the privacy-protecting measures, the study utilizes a collection of evaluation metrics and measurements. The metrics include privacy measures like the risk of re-identification tied to anonymization methods or the epsilon value of differential privacy. The effect of these privacy-protecting measures on data usefulness—like any decrease in accuracy in machine learning models or the decrease in data available due to changes—is confirmed in addition to usability. The feasibility of using these measures in practical situations depends on technical matters, like how computationally intensive the processes are and how effective they are, which are also considered.

The method entails a detailed study of case studies and real-world data handling that utilize privacy protection measures, along with theoretical analysis. This section provides useful information on how such methods work in real life and the challenges they face. It also provides the lessons derived from their usage in different fields, such as social networks, medicine, and finance. The research concludes by synthesizing the results of the assessment metrics, algorithm analysis, classification framework, literature review, and case studies. The careful analysis determines general trends, future challenges, and research needs. Future research directions provide a blueprint for moving forward on the subject of privacy preservation in data science.

4. Data Science Privacy Risks

Although data science offers numerous solutions to making conclusions based on data, they pose serious privacy issues. Some of the privacy issues with data science are described in this section, including identity exposure, attribute disclosure, membership inference, and inference attacks [19]. These threats are known to form effective privacy-protection policies. There are potential privacy threats to individuals based on the extensive use of numerous data analysis and interpretation methods. Identity revelation occurs when an individual can identify who the people are from anonymous datasets [20]. This is one of the privacy threats due to the advanced data collection, aggregation, and analysis processes. Another issue associated with attribute disclosure is that data analysis techniques can disclose private information inadvertently. Furthermore, membership inference attacks intrude into the privacy of certain individuals by employing statistical patterns in datasets to determine whether they are in or out [21].

5. Real- World Privacy Risk Cases

5.1. De-Anonymization of Medical Information

Anonymized patient data are typically collected in research environments in an attempt to examine disease trends and patterns. However, seemingly anonymous medical records become identifiable when cross-matched with outside databases that hold demographic data such as age, gender, and ZIP code [22]. There are severe privacy risks associated with the re-identification because it can potentially disclose an individual's genetic risks or sensitive medical conditions. From the medical history, such breaches can lead to stigmatization, discrimination, or even denial of essential services.

5.2. Location Information for Mobility Research

In order to better understand traffic flows and generate more efficient services, transport firms tend to gather GPS data from individuals' mobile phones [23]. Even when anonymized, the data tend to reveal very intimate information by generating distinctive movement patterns. Frequent hospital visits, religious sites, or political demonstrations can inadvertently reveal private information about a person. This inadvertent disclosure violates individuals' privacy and exposes them to monitoring or profiling.

5.3. Behavioral Analytics and Social Media

In order to serve targeted ads and make personalized recommendations, social media platforms track user behavior and interests and post content on an ongoing basis. Yet, aggregating behavioural data across sites facilitates the creation of highly specific user profiles. Advanced algorithms can leverage seemingly harmless online behavior to infer sensitive traits such as socioeconomic status, sexual orientation, or political affiliation. There are severe privacy issues in the digital age because nefarious actors can leverage this level of data profiling for targeted harassment, disinformation operations, or unethical commerce.

6. Methods for Preserving Privacy

Scalability, accessibility, and attack resistance are principal challenges to data science privacy. These challenges must be addressed before privacy-preserving methods can be applied.

Scalability is a principal concern since methods such as homomorphic encryption and differential privacy are computationally expensive, which limits their application for real-time processing and large data.

Distributed computing, hardware acceleration through specialized hardware [25], and efficient algorithms are some means of enhancing productivity without sacrificing privacy.

Another concern is usability since privacy-preserving methods will likely require high cryptography expertise, making them difficult for data scientists and developers to apply. Accessibility and acceptance can be enhanced through training sessions, incorporation into standard frameworks such as TensorFlow, and ease-of-use tools.

Since privacy methods must be attack-resistant against attacks such as inference and adversarial attacks [24], security resistance to attack is required. Formal audits, adversarial testing, and AI-based privacy protection are required for security enhancement since these allow for real-time breach detection and threat and privacy issue mitigation.

Addressing these challenges with scalable solutions, user-friendly implementations, and strong security measures will propel the development of privacy-preserving methods in data science, ensuring effectiveness and practicality.

7. Methods of Anonymization

Since anonymization ensures privacy protection, it is a critical component of data research. Several anonymization techniques are implemented in various disciplines [26], such as generalization, suppression, k-anonymity, l-diversity, and t-closeness. Each possesses different strengths, weaknesses, and applications for data privacy preservation [27].

By replacing coarse, less identifiable attributes with exact ones, generalization reduces re-identification threats and protects identities. To further maintain privacy, suppression erases identifying information from datasets selectively. K-anonymity bars individual identification by ensuring each record is indistinguishable from at least k-1 other records based on quasi-identifiers. By ensuring that sensitive attributes have at least l distinct values in every equivalence class, L-diversity further ensures privacy protection than k-anonymity.

Organizations can maintain personal information privacy while keeping it analytically useful by implementing these anonymization practices. However, achieving a balance between data usefulness and privacy remains challenging, subject to further development in anonymization techniques.

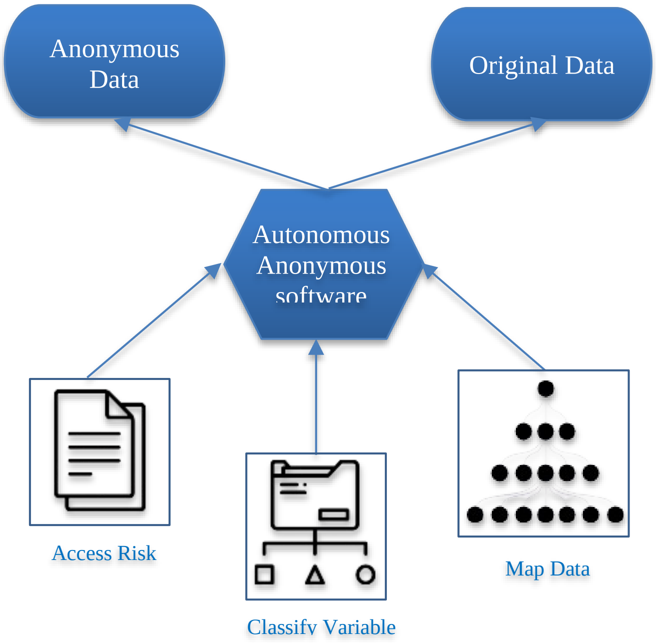

Fig. 2 Methods of anonymization

8. Differentiated Privacy

In data analysis, differential privacy is an industry standard for protecting individual privacy. In order to ensure that the presence or absence of an individual data point has minimal impact on analytical results, various mechanisms are used, such as Laplace, Gaussian, and exponential mechanisms. These mechanisms allow useful insights to be drawn from data while strong privacy guarantees are provided.

By adding random noise from a Laplace distribution to query outputs, the Laplace mechanism provides differential privacy, making it difficult to infer individual data points. By randomly selecting outputs based on their utility, the exponential mechanism balances data usability against privacy protection, enhancing privacy. However, by further obscuring individual contributions to the data set, the Gaussian approach provides a strong privacy guarantee by adding Gaussian noise to query outputs.

By integrating these differential privacy mechanisms, organizations can analyze data while protecting individual privacy. As data science develops, improvements in differential privacy methods will become essential to preserving privacy without harming analytical accuracy [28].

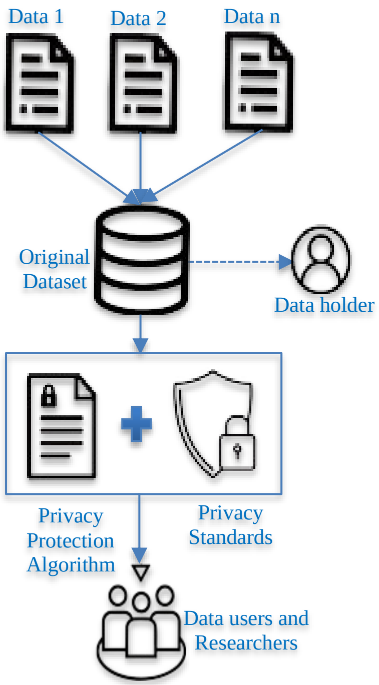

Fig. 3 Big data privacy protection model

9. The Homomorphic Encryption System

Confidentiality is ensured while performing secure data audits using privacy-preserving encryption. Fully Homomorphic Encryption (FHE), Slightly Homomorphic Encryption (SHE), and Partially Homomorphic Encryption (PHE) are the three forms of isomorphic encryption, an important technique in this area. Each form offers different security and computational convenience for secure data processing.

PHE allows only limited computations, allowing some mathematical operations, such as addition or multiplication, to be carried out on encrypted data. This capability is extended by SHE, allowing addition or multiplication under certain conditions. The most powerful form, FHE, allows infinite calculations on encrypted data without ever decrypting it, offering the highest level of privacy protection during audits and studies.

FHE is still computationally intensive despite its strong security advantages. However, processing power and encryption algorithm advances have made it more useful in safe data processing environments. Homomorphic encryption is still evolving as data privacy issues increase, offering strong solutions for private and secure data analysis.

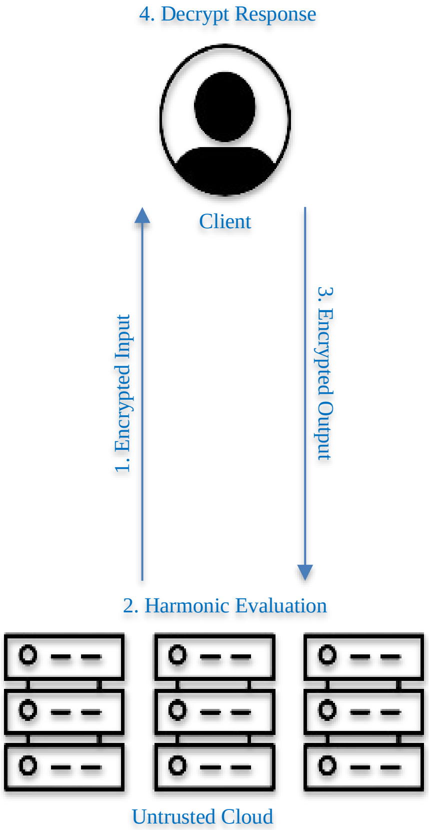

Fig. 4 Process of harmonic encryption

10. Secure Multi-Party Computation (SMPC)

Multiple parties can compute jointly with individual privacy using Secure Multi-Party Computation (SMPC). Even if multiple individuals are involved in a computation, our approach ensures that private data is kept private.

Secure function evaluation, Yao's garbled circuits, and Shamir's secret [29, 30] sharing are prominent SMPC techniques that allow parties to compute over encrypted inputs without revealing underlying data.

SMPC has widespread applications in banks, telecommunication, and medicine to enable joint data analysis, fraud detection, and privacy-preserving risk analysis.

Using SMPC, data-sharing capability is enhanced without compromising security by enabling numerous parties to securely compute joint calculations.

SMPC is limited by computationally expensive problems and large communication overhead. In order to make SMPC more practical in real-world applications, recent research attempts to optimize protocols to enhance efficiency and reduce requirements.

As improvements are made, SMPC is expected to become increasingly significant in safe and privacy-sensitive [31] data sharing.

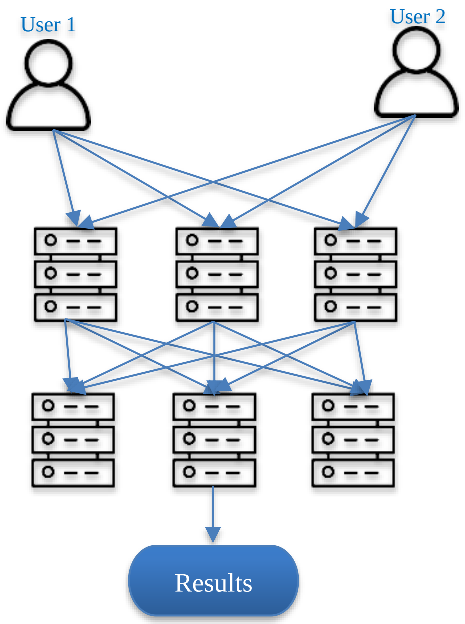

Fig. 5 The process of Secure Multi-part Computations (SMPC)

11. Federated Learning

Federated learning redefines conventional machine learning approaches such that models are trained elsewhere. It maintains privacy and security by enabling training on various devices while data remains local rather than concentrating data in one location. Some key approaches in federated learning, such as model aggregation, differential privacy, and secure aggregation protocols, enable learning to be improved and privacy to be safeguarded simultaneously.

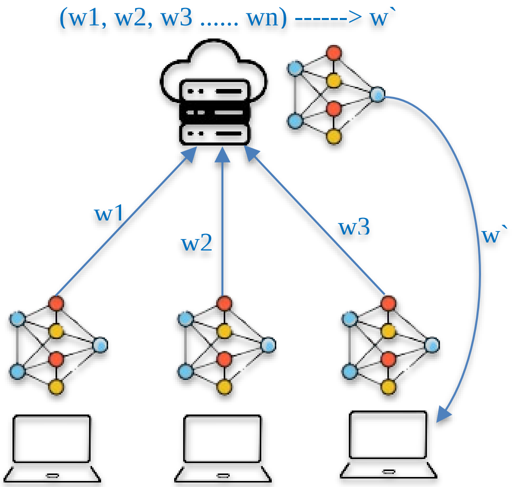

Fig. 6 Federated learning (FL)

This approach is extremely useful where it is difficult to exchange sensitive information, i.e., edge computing, the Internet of Things, and healthcare. Federated learning takes advantage of diverse computing power and maintains data privacy through joint model training [32] without exchanging raw data.

There are numerous benefits of federated learning, but its extensive deployment is hindered by issues like unstable data sources and expensive communications [33]. Nevertheless, such issues can be addressed by optimizing network design, data syncing procedures, and data privacy legislation. This makes federated learning a revolutionary method to execute secure and mass-scale machine learning.

12. Comparison of Various Privacy-Preserving Methods

Secure Multi-Party Computation (SMPC), homomorphic encryption, and federated learning are three of the strongest privacy-preserving techniques for protecting private information as it is computed and analyzed. This section compares and evaluates their strengths, weaknesses, and applicability to data science.

Secure Multi-Party Computation (SMPC), federated learning, and homomorphic encryption are significant privacy protection techniques in data analysis; each has strengths and weaknesses.

Federated learning does not move raw data but trains machine learning models on edge devices. It is scalable, heterogeneous, and maintains privacy by keeping device data. However, it has security vulnerabilities at model aggregation and requires high-frequency communication, which is costly regarding bandwidth.

Homomorphic encryption enables computation on encrypted data and maintains privacy as analysis continues. It is secure in strength and accommodates a wide range of mathematical operations. However, it is computationally expensive and inefficient for large data, and some operations are difficult to perform effectively [34]. Another significant aspect of security is key management.

SMPC maintains confidentiality by allowing different parties to compute secret data [35]. Though fault-tolerant and flexible, its implementation is complex, and its performance is hindered by communication overhead. For it to be effective, the participants must also trust each other.

All three techniques, however, maintain privacy; however, federated learning is suitable for decentralized systems, homomorphic encryption is flexible but computationally expensive, and SMPC provides secure joint computations but requires careful planning [36].

13. Uses of Privacy-Protecting Methods

To illustrate how privacy management rules work, this section provides an overview of how they are implemented across healthcare, finance, telecommunications, social media, and government industries. These practices are essential to compliance with legal rules, safeguarding personal information, and ensuring ethical use of data. Patient data is safeguarded in healthcare, and financial transactions are safeguarded from fraud in banking. Government agencies use similar practices to safeguard citizens' privacy, and telecommunications firms employ privacy steps to safeguard customers' data and regulate traffic on the network.

Also, as on social media platforms and government programs, privacy management policies are essential to safeguarding behavioral privacy. By enabling individuals to control their personal data and make informed choices, these practices empower individuals. In addition to safeguarding private data, privacy policies enable secure analysis of data and compliance with audit, security, and regulatory requirements. Industries can ensure ethical management of data while upholding individuals' right to privacy by embracing these rules.

14. Obstacles and Prospects

This study assessed secure multi-party computing (SMPC), homomorphic encryption, and federated learning as three important privacy protection strategies. We concentrated on their advantages, disadvantages, and functions in safeguarding private information. Despite their advancements, these methods still struggle with efficacy and widespread use. Usability, attack resistance, and integration with new technology are still important concerns.

Key findings show that all three methods offer strong privacy preservation. Federated learning keeps data private by storing it on local devices, homomorphic encryption allows computations on encrypted data, and SMPC enables multiple parties to collaborate while maintaining input privacy. However, each approach comes with trade-offs. Homomorphic encryption can be computationally intensive, while SMPC requires careful coordination. The choice of method depends on factors like data sharing and privacy requirements.

We suggest using improved algorithms to increase the effectiveness of these techniques in future studies. Priority should be given to integrating privacy protection strategies into current frameworks and guaranteeing strong security against intrusions. Exploring the convergence of cutting-edge technologies like blockchain and streamlining usability to make these methods more approachable may provide fresh approaches to protecting privacy while facilitating efficient data analysis.

15. Conclusion

By concentrating on these suggestions, we may promote practical and reasonably priced privacy protection measures. These initiatives are crucial for integrating validation into public sector data procedures, promoting responsible behavior, and improving individual privacy in data-driven environments. As technology develops and privacy concerns increase, ongoing innovation and cooperation will be essential to shaping a future in which data processing respects privacy.

Privacy-preserving techniques are essential for the moral and safe use of data in the era of big data and data science. This study provides a thorough review of the most recent techniques, applications, difficulties, and opportunities for privacy protection. As data increases and privacy concerns become more pressing, research and innovation in privacy-preserving strategies are essential to facilitating moral, long-term decision-making. By protecting individual privacy and encouraging trust, security, and transparency in data practices, these tactics enable people, companies, and communities to fully utilize data.

References

[1] Reza Shokri et al., "Membership Inference Attacks Against Machine Learning Models," *2017 IEEE Symposium on Security and Privacy*, San Jose, CA, USA, pp. 3-18, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[2] Latanya Sweeney, "k-Anonymity: A Model for Protecting Privacy," *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 10, no. 5, pp. 557-570, 2002. [CrossRef] [Google Scholar] [Publisher Link]

[3] Ashwin Machanavajjhala et al., "L-Diversity: Privacy Beyond k-Anonymity," *ACM Transactions on Knowledge Discovery from Data*, vol. 1, no. 1, pp. 1-52, 2007. [CrossRef] [Google Scholar] [Publisher Link]

[4] Ninghui Li, Tiancheng Li, and Suresh Venkatasubramanian, "t-Closeness: Privacy Beyond k-Anonymity and l-Diversity," *2007 IEEE 23rd International Conference on Data Engineering*, Istanbul, Turkey, pp. 106-115, 2007. [CrossRef] [Google Scholar] [Publisher Link]

[5] Arvind Narayanan, and Vitaly Shmatikov, "Robust De-Anonymization of Large Sparse Datasets," *2008 IEEE Symposium on Security and Privacy*, Oakland, CA, USA, pp. 111-125, 2008. [CrossRef] [Google Scholar] [Publisher Link]

[6] Cynthia Dwork, "Differential Privacy: A Survey of Results," *Proceedings of the 5th International Conference on Theory and Applications of Models of Computation*, Xi'an, China, pp. 1-19, 2008. [CrossRef] [Google Scholar] [Publisher Link]

[7] K. Balle et al., "Improving Differential Privacy in Machine Learning," *Journal of Privacy and Confidentiality*, 2020.

[8] Craig Gentry, "Fully Homomorphic Encryption Using Ideal Lattices," *STOC '09: Proceedings of the Forty-First Annual ACM Symposium on Theory of Computing*, MD, Bethesda, USA, pp. 169-178, 2009. [CrossRef] [Google Scholar] [Publisher Link]

[9] K. Lauter et al., "Computing on Encrypted Data," *IEEE Transactions on Information Theory*, 2011. Not Found

[10] Adi Shamir, “How to Share a Secret,” *Communications of the ACM*, vol. 22, no. 11, pp. 612-613, 1979. [CrossRef] [Google Scholar] [Publisher Link]

[11] Andrew C. Yao, “Protocols for Secure Computations,” *23rd Annual Symposium on Foundations of Computer Science*, Chicago, IL, USA, pp. 160-164, 1982. [CrossRef] [Google Scholar] [Publisher Link]

[12] Brendan McMahan et al., “Communication-Efficient Learning of Deep Networks from Decentralized Data,” *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics*, Fort Lauderdale, Florida, USA, pp. 1-10, 2017. [Google Scholar] [Publisher Link]

[13] E. Bagdasaryan et al., “Backdoor Attacks on Federated Learning,” *Advances in Neural Information Processing Systems*, 2020.

[14] Nicolas Papernot et al., “Semi-Supervised Knowledge Transfer for Deep Learning from Private Training Data,” *arXiv*, pp. 1-16, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[15] Keith Bonawitz et al., “Practical Secure Aggregation for Privacy-Preserving Machine Learning,” *CCS '17: Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security*, Dallas Texas USA, pp. 1175-1191, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[16] P. Samarati, “Protecting Respondents' Identities in Microdata Release,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 13, no. 6, pp. 1010-1027, 2001. [CrossRef] [Google Scholar] [Publisher Link]

[17] Dan Boneh, and Brent Waters, “Conjunctive, Subset, and Range Queries on Encrypted Data,” *Proceedings of the 4th Theory of Cryptography Conference*, Amsterdam, The Netherlands, pp. 535-554, 2007. [CrossRef] [Google Scholar] [Publisher Link]

[18] Y. Shokri et al., “Privacy-Preserving Deep Learning via Noisy Aggregation,” *International Conference on Learning Representations*, 2017. Not Found

[19] Daniel Kifer, and Ashwin Machanavajjhala, “No Free Lunch in Data Privacy,” *SIGMOD '11: Proceedings of the 2011 ACM SIGMOD International Conference on Management of Data*, Athens Greece, pp. 193-204, 2011. [CrossRef] [Google Scholar] [Publisher Link]

[20] Latanya Sweeney, “*Simple Demographics Often Identify People Uniquely*,” Carnegie Mellon University, Report, pp. 1-34, 2000. [Google Scholar] [Publisher Link]

[21] Benjamin C. M. Fung et al., “Privacy-preserving Data Publishing: A Survey of Recent Developments,” *ACM Computing Surveys*, vol. 42, no. 4, pp. 1-53, 2010. [CrossRef] [Google Scholar] [Publisher Link]

[22] A. El Emam, “A Globally Optimal k-Anonymity Method for the De-Identification of Health Data,” *Journal of the American Medical Informatics Association*, vol. 16, no. 5, pp. 670-682, 2009. [CrossRef] [Google Scholar] [Publisher Link]

[23] Yves-Alexandre de Montjoye et al., “Unique in the Crowd: The Privacy Bounds of Human Mobility,” *Scientific Reports*, pp. 1-5, 2013. [CrossRef] [Google Scholar] [Publisher Link]

[24] Frank McSherry, and Kunal Talwar, “Mechanism Design via Differential Privacy,” *48th Annual IEEE Symposium on Foundations of Computer Science*, Providence, RI, USA, pp. 94-103, 2007. [CrossRef] [Google Scholar] [Publisher Link]

[25] Masahiro Yagisawa, “*Fully Homomorphic Encryption without Bootstrapping*,” Cryptology ePrint Archive, Report, pp. 1-40, 2013. [Google Scholar] [Publisher Link]

[26] Josep Domingo-Ferrer, and Vicenç Torra, “A Critique of k-Anonymity and Some of Its Enhancements,” *2008 Third International Conference on Availability, Reliability and Security*, Barcelona, Spain, pp. 990-993, 2008. [CrossRef] [Google Scholar] [Publisher Link]

[27] Michael Hay et al., “Resisting Structural Re-Identification in Anonymized Social Networks,” *The VLDB Journal*, vol. 19, pp. 797-823, 2010. [CrossRef] [Google Scholar] [Publisher Link]

[28] Battista Biggio, and Fabio Roli, “Wild Patterns: Ten Years after the Rise of Adversarial Machine Learning,” *Pattern Recognition Journal*, vol. 84, pp. 317-331, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[29] Yehuda Lindell, and Benny Pinkas, “Privacy-Preserving Data Mining,” *Advances in Cryptology - CRYPTO 2000*: *Proceedings of the Institution of Mechanical Engineers 20th Annual International Cryptology Conference*, Santa Barbara, California, USA, pp. 439-450, 2000. [CrossRef] [Google Scholar] [Publisher Link]

[30] Jonathan Katz, and Yehuda Lindell, *Introduction to Modern Cryptography Principles and Protocols*, 1st ed., Chapman & Hall/CRC, pp.1-552, 2007. [CrossRef] [Google Scholar] [Publisher Link]

[31] Shafi Goldwasser, and Yehuda Lindell, “Secure Multi-Party Computation Without Agreement,” *Journal of Cryptology*, vol. 18, pp. 247-287, 1997. [CrossRef] [Google Scholar] [Publisher Link]

[32] Kallista Bonawitz et al., “Federated Learning and Privacy: Building Privacy-preserving Systems for Machine Learning and Data Science on Decentralized Data,” *Queue*, vol. 19, no. 5, pp. 87-114, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[33] Jakub Konečný et al., “Federated Optimization: Distributed Machine Learning for On-Device Intelligence,” *arXiv*, pp. 1-38, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[34] Craig Gentry, and Shai Halevi, “Implementing Gentry’s Fully Homomorphic Encryption Scheme,” *Advances in Cryptology–EUROCRYPT 2011: Proceedings of the 30th Annual International Conference on the Theory and Applications of Cryptographic Techniques*, Tallinn, Estonia, pp. 129-148, 2011. [CrossRef] [Google Scholar] [Publisher Link]

[35] Payman Mohassel, and Yupeng Zhang, “SecureML: A System for Scalable Privacy-Preserving Machine Learning,” *2017 IEEE Symposium on Security and Privacy*, San Jose, CA, USA, pp. 19-38, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[36] Eugene Bagdasaryan et al., “How to Backdoor Federated Learning,” *Proceedings of the Twenty Third International Conference on Artificial Intelligence and Statistics*, vol. 138, pp. 2938-2948, 2020. [Google Scholar] [Publisher Link]