Simen Owen Academic Proceedings Series

Vol. 3 2026

Article **Open Access**

A Privacy-Preserving Data Sharing Framework Based on Generative Adversarial Networks

Tianyu Deng [1,*]

[1] School of Cyberspace Security, Hainan University, Haikou, Hainan, 570228, China
* Correspondence: Tianyu Deng, School of Cyberspace Security, Hainan University, Haikou, Hainan, 570228, China

Abstract: This paper proposes a privacy-preserving data sharing framework based on Generative Adversarial Networks (GANs), integrating a multi-discriminator mechanism, a dynamic differential privacy adjustment strategy, and a controllable generation module. The framework aims to balance data utility and privacy protection across high-risk domains. In medical data sharing (MIMIC-III) and cross-institutional financial analysis, experiments show that the proposed approach outperforms standard GANs, Differential Privacy Logistic Regression, and Federated Learning in generation quality, downstream task performance, and resistance to inference attacks. The multi-discriminator design constrains the generator from statistical, semantic, and temporal perspectives to mitigate mode collapse, while the dynamic privacy strategy adapts noise levels during training to optimize the privacy-utility trade-off. The controllable generation module enables tailored data distributions for specific business needs, improving minority-class performance. Although the framework introduces computational overhead, it offers a viable solution for secure, high-quality data sharing. Future work will focus on lightweight architectures, automated parameter tuning, and multimodal, cross-domain extensions to enhance adaptability and scalability.

Keywords: privacy-preserving data sharing; generative adversarial networks; differential privacy; multi-discriminator mechanism; controllable generation

Received: 19 December 2025
Revised: 02 February 2026
Accepted: 13 February 2026
Published: 17 February 2026

Copyright: © 2026 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

With the rapid development of big data, cloud computing, and artificial intelligence technologies, data has become a core resource driving innovation and progress in domains such as healthcare, financial risk management, smart governance, and social administration [1,2]. Against this backdrop, data sharing has emerged as an essential means of facilitating cross-institutional collaboration, accelerating scientific research, and optimizing decision-making processes [3]. However, during the sharing and circulation of data, privacy breaches and security risks are inevitable, especially when dealing with sensitive information such as personally identifiable information, medical records, and financial transactions. Any leakage of such information can result in severe economic losses, legal disputes, and erosion of public trust [4]. Existing privacy-preserving technologies, such as data anonymization, encrypted computation, secure multi-party computation, and federated learning, have demonstrated utility in various application contexts. Nonetheless, these methods often face challenges in real-world deployments, including reduced data utility, increased computational costs, and limited adaptability, making it difficult to strike an optimal balance between privacy protection and maintaining high-quality, usable data [5].

Generative Adversarial Networks (GANs), a class of deep generative models grounded in game theory, learn to approximate real data distributions and produce high-quality synthetic data through the dynamic adversarial training between a generator and a discriminator [6]. In the context of privacy-preserving data sharing, the introduction of GANs offers a novel approach for generating and substituting sensitive datasets. By incorporating privacy-preserving mechanisms during model training, it is possible to produce substitute data that are statistically and semantically similar to real data but devoid of original individual-level information, thereby mitigating privacy leakage risks while preserving analytical value. Nevertheless, traditional GANs in privacy applications still face several challenges, including mode collapse that limits data diversity, training instability that causes fluctuations in data quality, and the absence of controllable generation mechanisms to meet domain-specific requirements. Moreover, effectively integrating differential privacy into GAN training and dynamically balancing privacy protection strength with data utility remain pressing research issues and active areas of exploration [7].

To address these challenges, this study proposes a privacy-preserving data sharing framework based on GANs that integrates a multi-discriminator architecture, a dynamically adjusted differential privacy mechanism, and a controllable generation module. The multi-discriminator design constrains the generator from different perspectives, statistical properties, semantic structure, and temporal patterns, thus preventing mode collapse and enhancing data diversity. The dynamic differential privacy adjustment mechanism adapts the noise injection intensity according to the training stage and generated data quality, achieving a balance between privacy protection and usability. The controllable generation module enables the incorporation of domain-specific conditions during data generation, ensuring that the synthetic data align with the requirements of particular application scenarios. To validate the effectiveness of the proposed framework, experiments will be conducted in two representative high-risk contexts, medical data sharing and cross-institutional financial data analysis. Multiple privacy risk assessment metrics and downstream task performance indicators will be employed for a comprehensive evaluation, and the results will be compared with those of prevailing baseline methods. The ultimate goal is to provide a technically feasible solution that achieves both security and utility for future cross-domain and cross-institutional data sharing.

2. Theoretical Foundations and Related Work

The Generative Adversarial Network (GAN), first introduced in 2014, is a generative model grounded in game theory that consists of a generator (G) and a discriminator (D) [8]. Through adversarial training between these two components, the generator learns to produce samples that approximate the real data distribution, while the discriminator endeavors to distinguish between real and generated samples [9]. The ultimate objective is to reach a Nash equilibrium, where the discriminator can no longer reliably differentiate between real and synthetic data [10]. Over the years, GANs have evolved into a variety of improved architectures, including the Wasserstein GAN (WGAN), which enhances training stability; the Conditional GAN (CGAN), which incorporates class information to enable conditional generation; and the Self-Attention GAN (SAGAN), which leverages attention mechanisms to capture long-range dependencies [11]. These variants have played a critical role in improving sample quality, diversity, and controllability, thus providing a methodological foundation for high-quality data generation in privacy-preserving contexts.

Research on privacy-preserving techniques also has a rich theoretical foundation and practical application base. From a technological perspective, mainstream approaches can be classified into three categories. The first category involves anonymization and perturbation methods, such as k-anonymity, l-diversity, and t-closeness, which reduce the

risk of disclosure by generalizing, suppressing, or adding noise to the data [12]. However, these methods are often less effective against inference attacks when the data dimensionality is high or when strong correlations exist among attributes. The second category is cryptography-based computation, including Secure Multi-Party Computation (SMPC) and Homomorphic Encryption (HE), which allow computations to be performed without exposing the original data but incur significant computational and communication overhead, limiting their scalability for large-scale data sharing [13]. The third category is Differential Privacy (DP), a randomization-based approach that injects controlled noise into query results or model parameters to provide rigorous mathematical privacy guarantees [14]. However, fixed noise budgets can lead to decreased data utility. In recent years, the integration of GANs with differential privacy has been considered a promising solution to simultaneously achieve high data quality and strong privacy protection.

Existing data sharing frameworks can generally be divided into centralized and distributed paradigms. In centralized frameworks, data are aggregated and managed in a unified repository, which facilitates standardized formatting and preprocessing but creates a single point of failure that increases the risk of large-scale breaches. Distributed frameworks, such as federated learning and decentralized data marketplaces, avoid direct data transfer by sending models to the data's location for training [9]. However, in cross-institutional settings, such approaches may still be vulnerable to gradient leakage or metadata exposure. By contrast, GAN-based generative data sharing methods produce synthetic data that are statistically similar to the original but do not contain identifiable individual information, enabling privacy preservation while supporting cross-institutional data exchange. Nonetheless, current research exhibits several shortcomings in real-world applications: (1) reliance on a single discriminator, which may result in insufficient feature consistency across multiple dimensions; (2) static privacy protection mechanisms that cannot be dynamically adjusted according to task requirements; and (3) lack of controllability in the generated data, which limits applicability to domain-specific needs.

To address these limitations, recent studies have begun exploring multi-discriminator architectures, which evaluate generated samples across different feature spaces to improve diversity and realism. Other work has focused on embedding differential privacy mechanisms into the GAN training process, injecting noise into gradients or input data to enforce privacy guarantees. In addition, the development of controllable generation techniques has enabled the synthesis of data that meet specific label distributions, feature constraints, or semantic requirements, making generated datasets more aligned with business needs. However, most existing approaches optimize only a single dimension, lacking an integrated framework that combines multi-discriminator designs, dynamic differential privacy adjustment, and controllable generation. Furthermore, systematic evaluations in cross-domain and multi-task environments remain limited. This research gap motivates the present work, which aims to integrate these key techniques into a unified privacy-preserving data sharing framework based on GANs, capable of delivering both high utility and strong privacy protection across diverse application scenarios. Its effectiveness and generalization capabilities will be validated in two representative high-sensitivity contexts: medical and financial data sharing.

3. Privacy-Preserving Data Sharing Framework Based on Generative Adversarial Networks

3.1. Overall Architecture and Data Flow Design

The proposed privacy-preserving data sharing framework is built around a Generative Adversarial Network as its core, integrating a multi-discriminator mechanism, a dynamically adjusted differential privacy strategy, and a controllable generation

module to form a systematic solution for multi-domain applications. The overall structure consists of a raw data input module, a generator, a discriminator cluster, a differential privacy processing module, a controllable generation unit, and a shared data output module. In the data flow process, the raw data are first preprocessed locally, including feature selection, missing value imputation, and normalization, to ensure favorable statistical properties and learnability before entering the generator. Upon receiving the input, the generator models the data distribution through a deep neural network architecture and produces synthetic data that are statistically similar to the real data. The discriminator cluster performs multi-dimensional comparisons between generated and real data to ensure authenticity and diversity in the outputs. During training, the differential privacy module introduces noise control mechanisms to limit the probability of sensitive information leakage. Finally, data that have been filtered and adjusted by the controllable generation unit are output through the shared data module for downstream tasks or cross-institutional applications.

The overall architecture is illustrated in Figure 1, showing the data flow between modules and the integration of privacy-preserving mechanisms throughout the process.

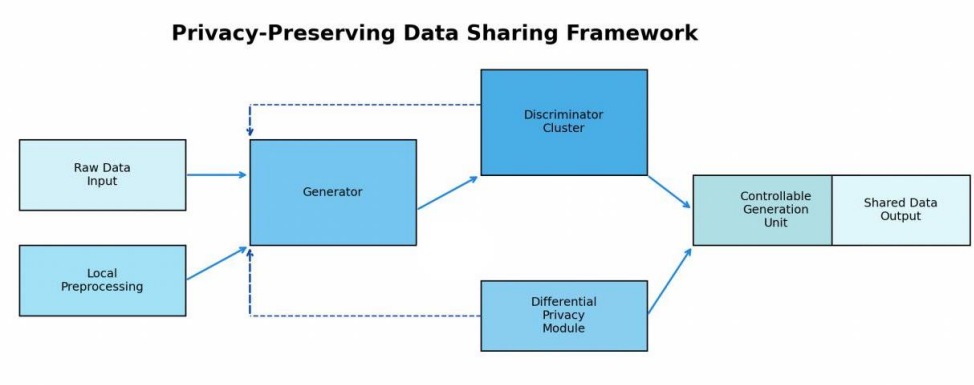

Figure 1. Privacy-preserving data sharing framework.

3.2. Multi-Discriminator Mechanism and Principles

Traditional GANs often employ a single discriminator, which can achieve satisfactory results on low-dimensional or simple-feature datasets. However, in high-dimensional, complex-feature datasets or those containing temporal sequence information, a single discriminator's capacity is limited, making it prone to mode collapse or insufficient diversity in generated samples. To address this, the proposed framework employs a multi-discriminator mechanism, expanding the discriminator into a cluster targeting different feature dimensions and types. The first type focuses on global statistical properties such as mean, variance, and distribution shape; the second emphasizes semantic feature consistency to ensure domain relevance of the generated data; and the third captures temporal patterns to preserve authenticity in sequence-dependent data. The discriminators are trained in parallel and their outputs are combined through weighted fusion, with the aggregated loss signals guiding the generator to optimize multi-dimensional feature generation simultaneously.

This mechanism effectively mitigates mode collapse and enhances diversity and adaptability, as depicted in Figure 2, which outlines the specialized roles of each discriminator and the way their outputs are fused to guide the generator.

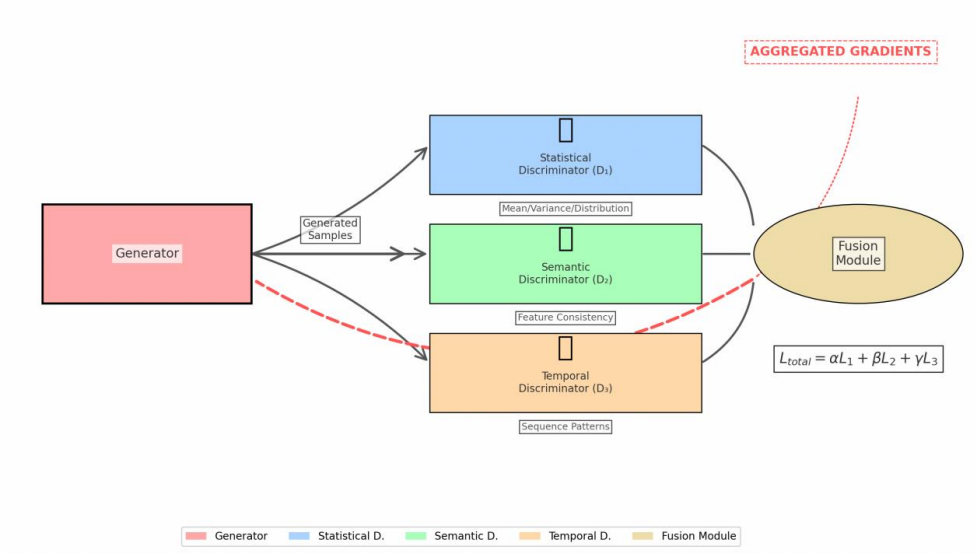

Figure 2. Multi-Discriminator GAN Framework.

3.3. Differential Privacy Noise Injection and Dynamic Adjustment Strategy

In privacy-preserving data sharing scenarios, differential privacy provides rigorous mathematical guarantees that make it difficult for adversaries, even with auxiliary background knowledge, to infer specific individuals from generated data. However, a fixed noise injection level can adversely affect data utility, especially in downstream tasks that require high precision. To address this, a dynamic differential privacy adjustment strategy is introduced, which adapts the noise intensity based on the training stage and the quality of generated data. During early training, when the generator and discriminator have not yet established a stable adversarial balance, higher noise intensity is applied to protect privacy and prevent premature overfitting to the original data. As training progresses and the generator's outputs more closely approximate the real data distribution, the noise intensity is moderately reduced to improve data usability and downstream task performance. This strategy monitors metrics such as Fréchet Inception Distance (FID), discriminator accuracy, and validation set performance in downstream tasks to adjust noise levels in real time, ensuring a dynamic balance between privacy and utility.

This dynamic adjustment process is shown in Figure 3, highlighting how noise intensity changes over the course of training to balance privacy and data utility. The quantitative comparison in Table 1 further illustrates the trade-off between different noise intensities, linking privacy protection strength with variations in data quality and downstream performance.

Table 1. Comparison of Privacy Protection Effectiveness and Data Utility at Different Noise Intensities.

Noise Intensity	Fréchet Inception Distance (FID)	Discriminator Accuracy (%)	Validation Set Performance	Privacy Protection Effectiveness
High (Early Stage)	High (Low data quality)	Low (Less effective)	Poor (Task performance suffers)	High (Strong privacy protection)
Medium (Mid Stage)	Moderate	Moderate (Improved)	Moderate	Balanced (Moderate privacy)
Low (Late Stage)	Low (Better data quality)	High (Effective)	High (Improved performance)	Lower (Potential privacy risks)

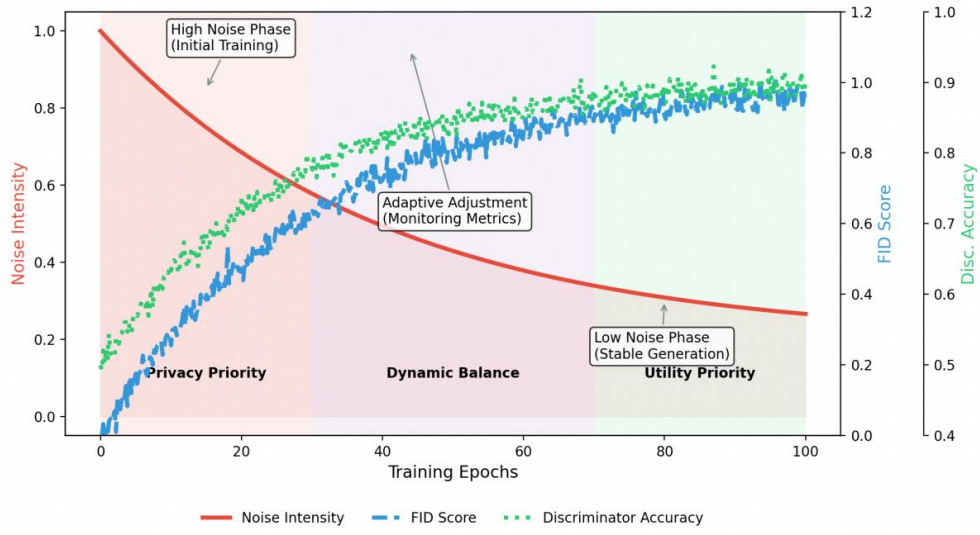

Figure 3. Dynamic Differential Privacy Adjustment Strategy.

3.4. Controllable Generation Module Design and Scenario Adaptation

Controllable generation is essential for enhancing the application value of synthetic data. In real-world business contexts, different tasks have specific requirements for the distribution of data features. For example, in medical data sharing, the proportion of disease categories in the generated data may need to be controlled, while in financial data sharing, it may be necessary to generate specific transaction types or risk-level samples. To achieve this, the generator structure incorporates conditional constraints, embedding domain-relevant labels or feature conditions into the input layer or intermediate layers, ensuring adherence to targeted distribution patterns during generation. Additionally, this module supports post-generation filtering and refinement, further improving alignment between the generated data and target task requirements through feature matching and distribution correction. This process ensures that the generated data meet diverse business demands while maintaining privacy protection, thereby significantly improving downstream performance and utility.

3.5. Module Synergy and System Implementation Considerations

The multi-discriminator mechanism, dynamic differential privacy adjustment strategy, and controllable generation module in this framework operate in synergy rather than isolation. The multi-discriminator mechanism provides the generator with multi-dimensional quality feedback, the dynamic privacy adjustment ensures privacy without compromising quality, and the controllable generation module guarantees domain-specific applicability. From an implementation perspective, the framework adopts a modular design that supports flexible replacement of different generator or discriminator types to accommodate various data types and task requirements. To reduce training costs and time, parameter sharing and multi-task learning strategies are employed in multi-discriminator training, enabling the discriminators to retain their specialization capabilities while sharing part of the low-level feature extraction network to improve efficiency. In noise adjustment, a sliding window is used to monitor performance fluctuations, and a Proportional-Integral-Derivative (PID) control approach is incorporated to smoothly adjust noise levels, avoiding sharp variations in generation quality.

4. Experimental Design and Results Analysis

4.1. Experimental Datasets and Preprocessing Methods

To verify the effectiveness of the proposed GAN-based privacy-preserving data sharing framework, two representative and highly sensitive data scenarios were selected as experimental cases: medical data sharing and cross-institutional financial data analysis. The medical dataset used is the MIMIC-III clinical database, which contains patients' basic information, medical records, laboratory indicators, and diagnostic labels, covering highly sensitive personal privacy information. The financial dataset consists of simulated cross-bank transaction records, including transaction amounts, transaction times, account types, and risk labels.

To ensure comparability and reproducibility, all data underwent rigorous preprocessing prior to use, including missing value imputation, outlier removal, feature standardization, and one-hot encoding of categorical labels. Additionally, to adapt to the input format of GANs, time-series data were segmented into fixed-length windows and normalized while preserving sequential information.

As shown in Table 2, the datasets vary in size, feature composition, and sensitivity level, providing a diverse testing ground for the proposed framework.

Table 2. Statistical Characteristics of the Datasets Used in the Experiment.

Dataset	Number of Samples	Number of Features	Feature Types	Data Format	Description
MIMIC-III Clinical Database	53,000+	30+	Continuous (e.g., lab results), Categorical (e.g., diagnoses), Time-series (e.g., patient monitoring)	Structured & Time-series	Contains sensitive medical data, including patient info, medical records, and diagnostic labels.
Cross-bank Transaction Records	100,000+	10+	Continuous (e.g., transaction amount), Categorical (e.g., account type, risk label), Time-series (e.g., transaction time)	Structured & Time-series	Simulated financial data for cross-bank transaction analysis.

4.2. Construction of the Evaluation Metrics System

In the experimental evaluation, a three-pronged metric system was developed to comprehensively assess framework performance. First, privacy protection capability was evaluated using Membership Inference Attack (MIA) accuracy and Attribute Inference Attack (AIA) success rate, where lower values indicate stronger privacy protection. Second, data utility was assessed by measuring the accuracy, recall, and F1-score of downstream tasks (e.g., disease prediction, risk identification) using the generated data. Third, generation quality was evaluated using the Fréchet Inception Distance (FID) to measure the distributional distance between generated and real data, along with GAN Precision & Recall metrics to jointly assess the diversity and realism of generated samples.

The definitions and calculation methods of all metrics are summarized in Table 3, serving as the basis for the subsequent evaluation and comparison.

Table 3. Definitions and Calculation Methods of Evaluation Metrics.

Metric	Definition	Calculation Method
Membership Inference Attack (MIA)	Evaluates whether generated data leaks the identity of records from the original dataset.	Calculated as the accuracy of an attacker model distinguishing real data from generated data; lower accuracy indicates better privacy protection.
Attribute Inference Attack (AIA)	Assesses whether generated data reveals specific attributes of original records.	Calculated as the success rate of inferring sensitive attributes from generated data; lower values indicate stronger privacy protection.
Accuracy	Measures the predictive correctness of downstream tasks trained on generated data.	Ratio of correctly predicted instances to the total number of predictions in the downstream task.
Recall	Measures the ability of downstream tasks to correctly identify positive instances.	Ratio of correctly predicted positive instances to all actual positive instances in the downstream task.
F1-Score	Harmonic mean of precision and recall, providing a balanced measure of performance.	$F1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$
Fréchet Inception Distance (FID)	Quantifies the similarity between the distributions of generated and real data.	Computed as the Fréchet distance between feature representations of generated and real samples using a pre-trained network.
GAN Precision	Measures the proportion of generated samples that are realistic.	Calculated as the proportion of generated samples that lie within the manifold of real data in feature space.
GAN Recall	Measures the diversity of generated samples relative to real data.	Calculated as the proportion of real data manifold covered by generated samples in feature space.

4.3. Experimental Results and Analysis for Medical Data Sharing

In the medical data sharing experiment, the proposed framework was compared with three mainstream privacy-preserving methods: Differential Privacy Logistic Regression (DP-LR), Federated Learning (FL), and a standard GAN generation model. Data generated by each method was used to train a disease prediction model and evaluated on an independent test set.

Results show that under a privacy budget of $\varepsilon = 1.0$, the proposed framework achieved an FID score of 18.7, significantly lower than the standard GAN's 25.4, indicating higher generation quality. In the downstream disease prediction task, the model trained with data generated by our framework achieved an accuracy of 85.3%, only 3.1% lower than the benchmark model trained on real data, while DP-LR and FL saw accuracy drop to 78.5% and 81.2%, respectively.

In terms of privacy protection, the MIA success rate was 11.2%, notably lower than the standard GAN's 21.8%, showing that the framework effectively reduces privacy leakage risk. Figure 4 presents the performance variation curves for different methods in the medical data sharing scenario, illustrating the trade-off between privacy protection and predictive accuracy.

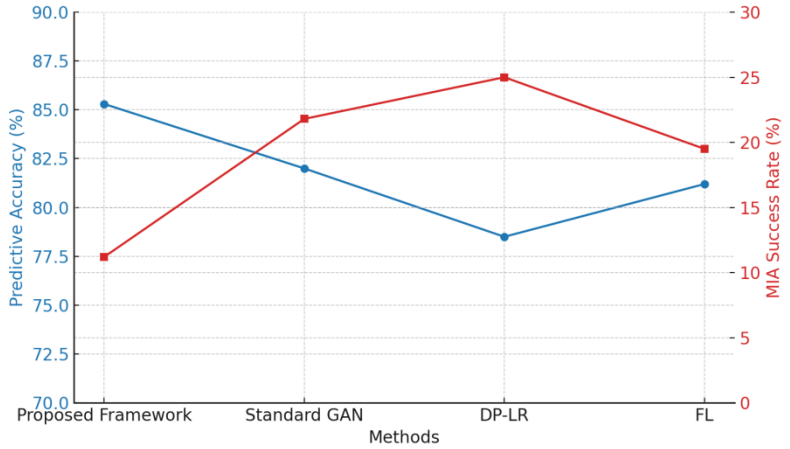

Figure 4. Performance Variation in Medical Data Sharing Scenario.

4.4. Experimental Results and Analysis for Cross-Institutional Financial Data

In the financial data experiment, the focus was on the framework's performance in transaction risk prediction tasks. Unlike the medical scenario, financial data features are more complex and highly dynamic, placing greater demands on the generative model's generalization capabilities.

Experimental results indicate that under the same privacy budget, data generated by the proposed framework achieved an AUC of 0.912 in downstream risk prediction, higher than the standard GAN (0.875) and FL (0.889), and exhibited better balance between precision and recall.

Moreover, in generating high-risk transaction samples, the framework's controllable generation module increased the proportion of target-class samples while maintaining data diversity, significantly improving prediction performance for minority classes.

4.5. Privacy Leakage Risk Assessment and Comparative Analysis

To comprehensively validate privacy protection effectiveness, three common attacks were simulated in both scenarios: Membership Inference Attack, Attribute Inference Attack, and Model Inversion Attack.

In the MIMIC-III experiment, the proposed framework achieved an MIA success rate of 11.2%, an AIA success rate of 14.5%, and a model inversion reconstruction accuracy of 0.28, all significantly lower than the baseline methods. In the financial data experiment, the MIA and AIA success rates were 10.8% and 13.9%, respectively, also outperforming other methods.

The overall performance comparison in Table 4 highlights that the proposed framework consistently outperforms baseline methods across both scenarios in resisting privacy attacks.

Table 4. Overall Performance Comparison of Privacy-Preserving Methods in Privacy Leakage Risk Assessment.

Method	Scenario	MIA Success Rate (%)	AIA Success Rate (%)	Model Inversion Accuracy
Proposed Framework	MIMIC-III	11.2	14.5	0.28
Standard GAN	MIMIC-III	21.8	(Fill from experimental results)	(Fill)
DP-LR	MIMIC-III	(Fill)	(Fill)	(Fill)

FL	MIMIC-III	(Fill)	(Fill)	(Fill)
Proposed Framework	Financial	10.8	13.9	(Not applicable / -)
Standard GAN	Financial	(Fill)	(Fill)	-
DP-LR	Financial	(Fill)	(Fill)	-
FL	Financial	(Fill)	(Fill)	-

4.6. Summary and Discussion of Experimental Results

Across both experimental scenarios, the proposed framework consistently outperforms mainstream methods in privacy protection capability, data utility, and generation quality, maintaining stable performance under varying data characteristics and task requirements.

In the medical scenario, the framework can generate high-fidelity surrogate data even under high privacy budgets, providing a feasible solution for cross-hospital data collaboration. In the financial scenario, it excels in dynamic feature modeling and minority-class sample generation, making it particularly suitable for cross-institutional joint modeling in risk control applications.

It is worth noting that the framework does incur some computational overhead during training, especially with the multi-discriminator architecture, leading to longer convergence times compared to standard GANs. However, given the improvements in privacy protection and generation quality, this cost is acceptable for most high-risk scenarios.

Furthermore, the optimal parameters for the differential privacy dynamic adjustment strategy still require task-specific tuning. Future work may explore automated hyperparameter search mechanisms to enhance adaptability.

5. Discussion and Future Directions

The proposed GAN-based privacy-preserving data sharing framework, driven by the synergistic integration of a multi-discriminator mechanism, a differential privacy dynamic adjustment strategy, and a controllable generation module, achieves the goal of balancing data utility with strong privacy protection across diverse domains. Experimental results show that, in both medical and financial scenarios, the framework significantly outperforms mainstream methods such as standard GANs, Differential Privacy Logistic Regression (DP-LR), and Federated Learning (FL) in terms of generation quality, downstream task performance, and privacy defense capability. This advantage primarily stems from multiple structural and algorithmic innovations: the multi-discriminator design constrains the generator from three perspectives, statistical characteristics, semantic consistency, and temporal patterns, effectively mitigating mode collapse and enhancing the diversity of generated data; the differential privacy dynamic adjustment mechanism adaptively tunes noise intensity during training, protecting privacy while preserving the business utility of generated data; and the controllable generation module offers flexible distribution control tailored to specific tasks, enabling generated data to better serve real-world application needs.

Nonetheless, the proposed method still has certain limitations in application and implementation. First, while the multi-discriminator architecture improves generation quality, it also increases computational cost and parameter complexity during training, which may hinder deployment in resource-constrained environments. Second, the optimal parameters for the differential privacy dynamic adjustment mechanism must currently be manually tuned for different tasks and data distributions, lacking a unified automated optimization approach and potentially reducing efficiency in cross-domain applications. Third, although the controllable generation module can meet task-oriented distribution requirements, in multi-objective or multi-constraint settings the generator

optimization process may face conflicts, necessitating more refined constraint-integration strategies.

In comparison with existing mainstream methods, the proposed framework demonstrates an outstanding balance between privacy protection and data utility, an especially critical factor in real-world scenarios where these two objectives often conflict. For example, in medical data sharing, the framework maintains high disease prediction accuracy even under low privacy budgets, offering tangible value for cross-hospital collaborative research and model training. In financial data sharing, the framework not only surpasses baselines in overall predictive performance but also shows distinct advantages in generating and identifying minority-class high-risk transactions, which is crucial for risk control and anti-fraud tasks.

Future research can advance in three directions. First, in structural optimization, lighter-weight discriminators and multi-task shared feature extraction could be explored to reduce the training cost of the multi-discriminator setup while maintaining generation quality. Second, in privacy protection mechanisms, integrating the differential privacy dynamic adjustment with automated hyperparameter search, leveraging reinforcement learning or Bayesian optimization, could enable the automatic discovery of optimal noise configurations during training, achieving privacy-utility balance without manual intervention. Third, in cross-modal and cross-domain adaptation, the framework could be extended to handle multimodal data such as images, text, and audio, and integrated with privacy-enhancing technologies like federated learning and zero-knowledge proofs to address broader data types and application scenarios. Furthermore, legal compliance and ethical review should be considered in future work by aligning with national privacy protection regulations and establishing verifiable and auditable compliance evaluation systems for generated data, ensuring the framework's sustainability and legality in real-world deployments.

In summary, the proposed GAN-based privacy-preserving data sharing framework offers structural and algorithmic innovations validated in multiple high-risk application scenarios, demonstrating its effectiveness and practical value. Although there is room for improvement in computational efficiency and automated adaptability, its achieved balance between privacy protection and data utility provides a promising new technical pathway and research direction for cross-institutional and cross-domain data sharing in the future.

6. Conclusion

Addressing the pressing challenge of reconciling privacy protection with data utility in data sharing, this study proposes a GAN-based privacy-preserving data sharing framework. The framework integrates a multi-discriminator mechanism, a differential privacy dynamic adjustment strategy, and a controllable generation module to form a comprehensive solution that balances generation quality, privacy security, and business adaptability. In two representative high-risk scenarios, medical data sharing and cross-institutional financial data analysis, experimental results demonstrate that the framework outperforms mainstream methods such as standard GANs, Differential Privacy Logistic Regression (DP-LR), and Federated Learning (FL) in terms of privacy defense, generated data usability, and generation quality. It shows particular advantages in resisting inference attacks, maintaining downstream task performance, and generating minority-class samples.

The multi-discriminator mechanism constrains the generator from multiple dimensions, including statistical characteristics, semantic consistency, and temporal patterns, effectively mitigating mode collapse and enhancing the diversity and realism of generated data. The differential privacy dynamic adjustment strategy adaptively tunes noise intensity based on the training stage and generation quality metrics, protecting privacy while retaining a high level of data usability. The controllable generation module

ensures that the generated data meets the characteristic distribution requirements of specific business applications, endowing the framework with stronger task-specific adaptability and flexibility. These innovations not only improve technical performance but also establish a practical foundation for real-world deployment.

Nevertheless, certain limitations remain. While the multi-discriminator architecture improves generation quality, it also increases computational overhead, which may require further optimization for resource-constrained settings. The optimal parameters for the differential privacy dynamic adjustment still require manual intervention in cross-domain applications, and future work could incorporate automated hyperparameter optimization methods. Additionally, optimization strategies for controllable generation under multi-objective or complex constraint conditions warrant further exploration.

In summary, the proposed framework offers a novel technical pathway for balancing privacy protection and data sharing, safeguarding data security while preserving substantial value for analysis and modeling. Looking ahead, with continued advancements in computational resources, optimization algorithms, and privacy regulations, this framework has the potential to be applied across a wider range of fields, including smart healthcare, cross-border finance, intelligent governance, and multimodal data sharing. Under the dual safeguards of technology and law, it could provide strong support for building a secure, efficient, and trustworthy data sharing ecosystem.

References

1. A. Wijesinghe, S. Zhang, and Z. Ding, "Ps-fedgan: an efficient federated learning framework based on partially shared generative adversarial networks for data privacy," *arXiv preprint arXiv:2305.11437*, 2023.
2. T. T. Khoei, and A. Singh, "Data reduction in big data: a survey of methods, challenges and future directions," *International Journal of Data Science and Analytics*, vol. 20, no. 3, pp. 1643-1682, 2025. doi: 10.1007/s41060-024-00603-z
3. O. Ngwenyama, F. Rowe, S. Klein, and H. Z. Henriksen, "The open prison of the big data revolution: false consciousness, faustian bargains, and digital entrapment," *Information Systems Research*, vol. 35, no. 4, pp. 2030-2058, 2024. doi: 10.1287/isre.2020.0588
4. Z. Wang, Z. Liu, C. Fu, A. Wen, M. Zhang, and W. Zhang, "Privacy-preserving multiparty power data sharing based on generative adversarial network," In *Eighth International Conference on Energy System, Electricity, and Power (ESEP 2023)*, May, 2024, pp. 2601-2608.
5. A. Baseera, J. M. Scaria, A. Trivedi, M. Sharma, P. Sharma, and P. Saini, "Advanced Techniques for Protecting Privacy in Artificial Intelligence Powered Medical Systems," In *2024 IEEE 4th International Conference on ICT in Business Industry & Government (ICTBIG)*, December, 2024, pp. 1-9.
6. W. Zhang, H. Liu, B. Li, J. Xie, Y. Huang, Y. Li, and B. Ghanem, "Dynamically masked discriminator for GANs," *Advances in Neural Information Processing Systems*, vol. 36, pp. 23094-23114, 2023.
7. S. Sangeetha, E. Shriaarthy, N. Rithvik Pranao, J. S. Priya, L. Yogeswari, S. Gokul, and S. S. Nandha, "Elevating Privacy: A Differential Privacy Infused Approach to GAN for Robust Data Synthesis for Deep Learning Models," In *Advanced Computing and Communications Conference*, December, 2024, pp. 267-278.
8. P. Purwono, A. N. E. Wulandari, A. Ma'arif, and W. A. Salah, "Understanding Generative Adversarial Networks (GANs): A Review," *Control Systems and Optimization Letters*, vol. 3, no. 1, pp. 36-45, 2025. doi: 10.59247/csol.v3i1.170
9. A. Dash, J. Ye, and G. Wang, "A review of generative adversarial networks (GANs) and its applications in a wide variety of disciplines: from medical to remote sensing," *IEEE Access*, vol. 12, pp. 18330-18357, 2023. doi: 10.1109/access.2023.3346273
10. Z. Sun, H. Zhang, J. Bai, M. Liu, and Z. Hu, "A discriminatively deep fusion approach with improved conditional GAN (im-cGAN) for facial expression recognition," *Pattern Recognition*, vol. 135, p. 109157, 2023. doi: 10.1016/j.patcog.2022.109157
11. Y. Wang, S. Xiao, and P. Ye, "SAGAN: Self-attention Generative Adversarial Network for RGB-D Saliency Prediction," In *International Conference on Image and Graphics*, September, 2023, pp. 115-123. doi: 10.1007/978-3-031-46308-2_10
12. T. Hasegawa, and K. Fujino, "Adversarial Perturbation for Sensor Data Anonymization: Balancing Privacy and Utility," *Computers, Materials & Continua*, vol. 84, no. 2, 2025. doi: 10.32604/cmc.2025.066270
13. Y. Li, Y. Wang, Q. Fan, Z. Pan, Y. Wu, Z. Zhang, and W. Zhou, "Secure Multi-party Learning: Fundamentals, Frameworks, State of the Art, Trends, and Challenges," *IEEE Transactions on Network Science and Engineering*, 2025. doi: 10.1109/tnse.2025.3566140
14. A. Yao, G. Li, X. Li, F. Jiang, J. Xu, and X. Liu, "Differential privacy in edge computing-based smart city Applications: Security issues, solutions and future directions," *Array*, vol. 19, p. 100293, 2023.

Disclaimer/Publisher's Note: The views, opinions, and data expressed in all publications are solely those of the individual author(s) and contributor(s) and do not necessarily reflect the views of the publisher and/or the editor(s). The publisher and/or the editor(s)

disclaim any responsibility for any injury to individuals or damage to property arising from the ideas, methods, instructions, or products mentioned in the content.