Trustworthy Agentic AI via Privacy-Preserving Synthetic Data: Lessons from Financial Tabular Data for PKG Systems

Oshani Seneviratne*, Michael Zuo*, Inwon Kang* and Stacy Patterson*

Rensselaer Polytechnic Institute, Troy NY 12180, USA

Abstract

Agentic AI systems increasingly rely on large volumes of personal data to support autonomous decision-making, personalization, and recommendation. However, the use of sensitive data raises fundamental challenges related to trust, privacy, and accountability. Synthetic data generation offers a promising mechanism to enable data sharing and model development without directly exposing individual records. In this paper, we revisit privacy-preserving synthetic data generation through the lens of trustworthy agentic AI systems. Although our experiments focus on financial tabular datasets, we argue that these methods naturally extend to personal knowledge graph (PKG)–driven AI systems, as tabular records can be systematically lifted into graph representations of entities, attributes, and relationships.

Keywords

privacy-preserving synthetic data, differential privacy, agentic AI, trustworthy AI, privacy-utility tradeoffs, AI accountability

1. Introduction

Agentic AI systems increasingly rely on large collections of personal, financial, and behavioral data to autonomously reason, recommend actions, and make decisions on behalf of users. These systems analyze data, recommend actions, and sometimes execute decisions autonomously on behalf of users. While this capability enables powerful new applications, it also introduces significant challenges related to *trust*, *privacy*, and *accountability*. Users and organizations must ensure that such systems can learn from sensitive data while maintaining strong guarantees that individual records are not exposed or misused.

Synthetic data generation has emerged as a promising mechanism to address this challenge. These datasets can be used for tasks such as model training, data sharing, benchmarking, research, and the simulation of rare events.

However, synthetic data does not automatically guarantee privacy. Recent research has shown that synthetic datasets may still leak information about the training data through mechanisms such as memorization by generative models [1], membership inference attacks [2, 3], and overfitting to rare records [4]. For agentic AI systems operating over sensitive personal data, these risks undermine both trust and accountability. Therefore, synthetic data generation methods must be evaluated not only for statistical fidelity but also for formal privacy guarantees.

In our prior work [5], we empirically investigated the privacy-utility tradeoffs of several representative tabular data generators on financial datasets, a domain characterized by highly sensitive attributes and severe class imbalance. Our work evaluated multiple generative approaches, including GAN-based, diffusion-based, and statistical generators, and introduced differentially private variants of generative models to mitigate potential privacy leakage. We analyzed how these methods balance three

ESWC'26: Trust, Autonomy and Accountability in PKG-Based Agentic AI (TAAPAAI) Workshop on May 10, 2026 in Dubrovnik, Croatia

*Corresponding author.

0000-0001-8518-917X (O. Seneviratne); 0009-0006-7188-9347 (M. Zuo); 0000-0001-8912-287X (I. Kang); 0000-0001-7711-6018 (S. Patterson)

critical dimensions: data fidelity, downstream model utility, and resilience to privacy attacks such as membership inference.

In this paper, we revisit the problem of privacy-preserving synthetic data generation through the lens of trustworthy agentic AI systems and summarize and contextualize the findings of [5] for the emerging ecosystem of AI systems that operate over structured personal data. Although our original study focuses on tabular financial datasets, we posit that the core insights directly translate to the emerging paradigm of *personal knowledge graph (PKG)* driven agentic AI systems. Tabular records can be systematically transformed into knowledge graph representations consisting of entities, attributes, and relationships. Consequently, privacy-preserving synthetic tabular data generation provides a practical pathway for producing synthetic PKGs or PKG-derived training data.

2. Differentially Private Synthetic Tabular Data Generation

Differential privacy (DP) provides a rigorous mathematical framework for quantifying privacy risk [6]. Informally, a differentially private mechanism ensures that the inclusion or exclusion of any single individual record does not significantly affect the output of a computation.

In the context of synthetic data generation, DP can be incorporated during generator training by adding carefully calibrated noise to gradients or other intermediate computations [7, 8]. This prevents the resulting generative model from encoding too much information about specific training examples.

Our study [5] evaluates multiple representative synthetic data generators such as Gaussian Copula [9] (statistical generator), TabDiff [10] (diffusion-based generator), CTGAN [11] (generative adversarial network based generator for tabular data), and TVAE [11] (variational autoencoder for tabular data). We further construct differentially private versions of CTGAN and TVAE by introducing privacy-preserving training mechanisms. These modifications allow us to study how privacy guarantees affect three critical dimensions of synthetic data:

1. Data quality (distributional similarity to the original data)
2. Downstream utility (performance of models trained on synthetic data)
3. Privacy risk (vulnerability to membership inference attacks and statistical similarity tests).

The results of our study demonstrate a fundamental tradeoff: stronger privacy guarantees typically reduce downstream model performance, although carefully designed generators can preserve useful signal while maintaining formal privacy bounds.

3. From Synthetic Tabular Data to Synthetic Personal Knowledge Graphs

Although our evaluation focuses on tabular financial datasets, the findings extend naturally to the domain of PKGs. Under this perspective, synthetic tabular data generation can be interpreted as generating synthetic knowledge graph instances or graph-derived training data. Because tabular-to-RDF transformations can be systematically defined through standardized mappings, synthetic tabular datasets can be systematically lifted into synthetic PKGs that preserve many structural and statistical properties of the original data while limiting exposure of individual records. In practice, the extent to which relational semantics and higher-order graph topology are preserved depends on the fidelity of the underlying synthetic tabular generation process.

Tabular → PKG Transformations: Tabular datasets can be converted into knowledge graph representations through standard transformations in the following ways. This process produces RDF triples that encode the same information in graph form while enabling richer semantic reasoning and integration with other knowledge sources.

- Rows correspond to entities (e.g., customers, accounts, or financial events)

- Columns correspond to attributes or relations
- Categorical values become linked entities
- Numerical values become datatype properties

Example Mapping: To illustrate the above transformation, consider financial datasets such as those included in the TabArena benchmark [12] for tabular machine learning. For example, datasets modeling credit risk or financial distress contain attributes describing a customer's financial history, such as income level, credit utilization, payment delays, number of open accounts, and recent delinquency events. In a PKG representation, each customer can be modeled as an entity node, with properties describing financial attributes and relationships connecting the customer to other entities such as credit accounts, lenders, or financial events. For instance, a row describing a borrower in the *Give Me Some Credit* synthetically generated dataset can be converted into a PKG as illustrated in Listing 1.

Listing 1: PKG Representation of a Sample Customer Data from the *Give Me Some Credit* dataset

```
:Customer_123
    rdf:type :Borrower ;
    :hasCreditUtilization "0.62"^^xsd:decimal ;
    :hasNumberOfOpenAccounts "7"^^xsd:int ;
    # ...
    :experiencedFinancialDistress true .
```

Additional relationships could connect customers to financial institutions, credit products, or transaction histories, yielding a richer financial knowledge graph that captures both attributes and relationships. Such graph representations are increasingly used in PKG-based recommendation systems, financial risk analysis, and personalized decision-support systems.

Standardized mappings from tabular and relational data to RDF graphs: The transformation from tabular or relational data into graph-structured knowledge representations is not ad hoc but supported by well-established standards within the Semantic Web community. The W3C has defined several specifications for systematically mapping structured datasets into RDF graphs. The *CSV on the Web* (CSVW) recommendation [13] provides a standardized mechanism for converting tabular datasets into RDF representations through CSV-to-RDF mappings (CSV2RDF) [14]. For relational databases, the W3C defines both the *Direct Mapping* specification and the more expressive *R2RML* (RDB to RDF Mapping Language) [15], which allow database schemas and tuples to be transformed into RDF triples in a principled and reproducible way.

4. Implications for Trust and Accountability in Agentic AI

Our findings in [5] highlight several important implications for PKG-based agentic AI systems that operate over sensitive personal or financial data.

Auditable data generation processes. Formal privacy guarantees such as DP provide measurable accountability mechanisms for synthetic data generation pipelines. In practice, these guarantees are typically quantified through the differential privacy parameters (ϵ, δ). By quantifying privacy loss through these privacy budget parameters, system designers can explicitly reason about the risks associated with releasing synthetic datasets. This makes the data generation process more transparent and auditable, allowing organizations to document privacy protections and evaluate whether synthetic data is appropriate for downstream tasks. In the context of PKG-driven agentic AI, such guarantees help ensure that autonomous agents trained on synthetic data cannot easily infer sensitive information about individual users represented in the original datasets.

Privacy-aware data sharing. Synthetic data generation with DP can serve as a preprocessing layer that enables safer data sharing across autonomous agents, research teams, and organizational boundaries. Instead of exposing raw datasets, organizations can release synthetic datasets that preserve important statistical patterns while reducing the risk of revealing individual records. When tabular data is lifted into knowledge graph representations, this process naturally extends to the generation of synthetic knowledge graphs derived from sensitive datasets. Such synthetic PKGs allow researchers and practitioners to experiment with new algorithms, evaluate agentic AI behaviors, and test reasoning or recommendation systems without exposing real user data. This capability is particularly valuable in domains such as finance and healthcare, where direct data sharing is often constrained by privacy concerns.

5. Conclusion

Trustworthy agentic AI systems must balance the competing demands of data utility and privacy protection. Synthetic data generation provides a promising mechanism for enabling learning over sensitive datasets while reducing direct exposure of individual records. However, without rigorous privacy guarantees and careful evaluation, synthetic data may still leak sensitive information.

In this work, we revisited our prior study on privacy-preserving synthetic tabular data generation and positioned its findings within the emerging ecosystem of PKG-based agentic AI systems. We argued that differentially private synthetic data generation can serve as a foundational privacy-preserving infrastructure layer for such systems. Because tabular data can be systematically lifted into knowledge graph representations through established standards, privacy-preserving synthetic tabular datasets can be transformed into synthetic PKGs that retain useful statistical structure while limiting disclosure risks. However, preserving higher-order graph topology and relational semantics remains an open challenge for graph-native synthetic data generation.

Looking ahead, several research directions remain important. These include developing graph-native synthetic data generators with formal privacy guarantees, designing privacy auditing techniques tailored to knowledge graph settings, and integrating privacy-preserving synthetic data pipelines into agentic AI infrastructures that operate over decentralized personal data. Advancing these capabilities will be essential for building PKG-driven AI systems that are not only autonomous and intelligent, but also trustworthy and accountable.

Supplementary Material Statement

All research artifacts, including source code, are available in our GitHub repository (https://github.com/brains-group/dpsdg), and the project code is published as a Python package at https://pypi.org/project/dpsdg. All external datasets and software dependencies used in this work are documented and linked in the repository's README.

Declaration on Generative AI

During the preparation of this work, the authors used ChatGPT and Grammarly for grammar and spelling checks, as well as for paraphrasing and rewording. After using these tools, the authors reviewed and edited the content as needed and take full responsibility for the publication's content.

Acknowledgment

Authors acknowledge support from the NSF IUCRC CRAFT center research grant (CRAFT Grant #22023) for this research. The opinions expressed in this publication do not necessarily represent the views of NSF IUCRC CRAFT.

References

[1] X. Gu, C. Du, T. Pang, C. Li, M. Lin, Y. Wang, On memorization in diffusion models, Transactions on Machine Learning Research (2023).

[2] T. Stadler, B. Oprisanu, C. Troncoso, Synthetic data – anonymisation groundhog day, in: 31st USENIX Security Symposium (USENIX Security 22), USENIX Association, Boston, MA, 2022, pp. 1451–1468.

[3] Y. Zhao, J. Zhang, Does training with synthetic data truly protect privacy?, in: The Thirteenth International Conference on Learning Representations, 2025.

[4] B. van Breugel, H. Sun, Z. Qian, M. van der Schaar, Membership inference attacks against synthetic data through overfitting detection, in: International Conference on Artificial Intelligence and Statistics, PMLR, 2023, pp. 3493–3514.

[5] M. Zuo, I. Kang, S. Patterson, O. Seneviratne, Measuring Privacy Risks and Tradeoffs in Financial Synthetic Data Generation, ACM (2026). The Second International Workshop on Transformative Insights in Multifaceted Evaluation at The Web Conference 2026 (TIME 2026).

[6] C. Dwork, K. Kenthapadi, F. McSherry, I. Mironov, M. Naor, Our data, ourselves: Privacy via distributed noise generation, in: S. Vaudenay (Ed.), Advances in Cryptology - EUROCRYPT 2006, Springer Berlin Heidelberg, Berlin, Heidelberg, 2006, pp. 486–503.

[7] J. Yoon, J. Jordon, M. van der Schaar, PATE-GAN: Generating synthetic data with differential privacy guarantees, in: International Conference on Learning Representations, 2019.

[8] M. L. Fang, D. S. Dhami, K. Kersting, DP-CTGAN: Differentially Private Medical Data Generation Using CTGANs, in: M. Michalowski, S. S. R. Abidi, S. Abidi (Eds.), Artificial Intelligence in Medicine, 2022, pp. 178–188.

[9] N. Patki, R. Wedge, K. Veeramachaneni, The synthetic data vault, in: 2016 IEEE International Conference on Data Science and Advanced Analytics (DSAA), 2016, pp. 399–410.

[10] J. Shi, M. Xu, H. Hua, H. Zhang, S. Ermon, J. Leskovec, Tabdiff: a mixed-type diffusion model for tabular data generation, in: The Thirteenth International Conference on Learning Representations, 2025.

[11] L. Xu, M. Skoularidou, A. Cuesta-Infante, K. Veeramachaneni, Modeling Tabular data using Conditional GAN, in: Advances in Neural Information Processing Systems, volume 32, 2019.

[12] N. Erickson, L. Purucker, A. Tschalzev, D. Holzmüller, P. M. Desai, D. Salinas, F. Hutter, Tabarena: A living benchmark for machine learning on tabular data, arXiv preprint arXiv:2506.16791 (2025). arXiv:2506.16791.

[13] J. Tennison, O. D. Institute, Csv on the web: A primer, W3C Working Group Note, 2016. URL: https://www.w3.org/TR/tabular-data-primer/.

[14] N. Haider, F. Hossain, et al., Csv2rdf: Generating rdf data from csv file using semantic web technologies, Journal of Theoretical and Applied Information Technology 96 (2018) 6889–6902.

[15] S. Das, S. Sundara, R. Cyganiak, R2RML: RDB to RDF mapping language, W3C Recommendation, 2012. URL: https://www.w3.org/TR/r2rml/.