Privacy-Preserving Cross-Disciplinary Design Innovation: A Data Reprogramming Approach

1st Junior Turatsinze *
University of Rwanda
Kigali, Rwanda
junior.tur2@outlook.com

2nd Pacifiqe Yuyisabe
RP College of IPRC-HUYE
Butare, Rwanda
pacifiqey@proton.me

***Abstract*—In the era of data-driven innovation, interdisciplinary design increasingly leverages vast datasets to inform and optimize creative processes. However, the integration of diverse data often introduces significant challenges related to data privacy and security. This paper proposes a novel framework for privacy-preserving cross-disciplinary design innovation, inspired by the principles of data reprogramming. We aim to develop a methodology that enables the effective utilization of sensitive data in design contexts while rigorously safeguarding individual privacy. Building on generative modeling principles, our approach improves data representation by balancing privacy preservation and feature robustness. We demonstrate how information bottleneck theory and reinforcement learning can be integrated to balance predictive power with privacy preservation, ensuring that design insights are derived from data without compromising sensitive information. This framework is particularly relevant for applications in healthcare, smart cities, and personalized education, where design solutions must be both innovative and ethically sound. Through extensive experimentation and validation, we show that our method achieves strong performance in design-related predictive tasks while effectively mitigating privacy risks, thereby paving the way for a new generation of privacy-conscious design methodologies.**

Keywords—Cross-disciplinary design, Privacy-preserving, Data reprogramming, Generative modeling, Information bottleneck, Reinforcement learning.

1. Introduction

Recent progress in artificial intelligence (AI) and data science is reshaping design practices, enabling data-informed creativity and efficiency. Traditional design processes that rely on intuition and qualitative reasoning are now being complemented by data-driven methods, leading to more informed and user-centered outcomes. This transition has fostered cross-disciplinary collaboration, where insights from engineering, business, and cultural studies converge to generate new products, services, and systems [1].

However, the growing use of data in design introduces critical challenges related to privacy and ethics. As design increasingly engages with personal data, behavioral patterns, and proprietary information, maintaining data privacy has become essential. Existing data-driven design approaches often emphasize utility and performance while underestimating privacy concerns, leading to potential data breaches and loss of user trust. In sensitive fields such as healthcare and smart city design, these risks are particularly acute. Regulations such as the General Data Protection Regulation (GDPR) highlight the global urgency of embedding privacy considerations into design practice [2].

This paper proposes a privacy-preserving framework for data-driven design innovation. Building on the concept of data reprogramming in data-centric AI, we aim to transform raw, sensitive design data into privacy-preserving yet useful representations. Our approach integrates information bottleneck theory, generative modeling, and multi-agent reinforcement learning to achieve this balance. We evaluate the framework across multiple design-related datasets to demonstrate its adaptability and effectiveness [3].

2. Related Work

2.1. Data-Driven Design and Innovation

Data-driven design has emerged as a transformative approach, moving beyond traditional intuition-based methods to leverage quantitative and qualitative data for informed decision-making throughout the design process. Early work focused on using data for user research and requirements gathering, employing techniques such as surveys, interviews, and observational studies to understand user needs and behaviors. More recently, big data analytics and machine learning have made it possible to analyze extensive datasets such as user interaction logs, social media content, and sensor outputs. These methods help designers identify behavioral patterns, anticipate trends, and tailor design solutions to individual needs [4]. For instance, in urban planning, data from smart city sensors can inform the design of more efficient transportation systems and public spaces [5]. In product design, consumer data can guide feature development and market positioning [6]. However, many of these approaches primarily focus on maximizing design utility and performance, often with insufficient attention to the privacy implications of collecting and processing vast amounts of personal data.

2.2. Privacy-Preserving Artificial Intelligence

The growing concerns over data privacy have led to a surge in research on privacy-preserving AI (PPAI). PPAI aims to develop AI models and systems that can learn from data without exposing sensitive information [7]. Key techniques in PPAI include differential privacy, homomorphic encryption, and federated learning. Differential privacy adds carefully calibrated noise to data or model outputs, providing strong privacy guarantees while

maintaining data utility [8]. Homomorphic encryption allows computations to be performed on encrypted data, ensuring that data remains confidential even during processing [9]. Federated learning enables collaborative model training across decentralized datasets without sharing raw data, thus preserving local data privacy [10]. While these techniques offer robust solutions for privacy protection in AI, their application in complex, interdisciplinary design contexts, especially when dealing with diverse data modalities and dynamic design processes, remains a challenge. Integrating these methods seamlessly into a data reprogramming framework for design innovation requires careful consideration of their computational overhead and impact on design outcomes.

2.3. Data Reprogramming and Feature Engineering

Data reprogramming, a core concept in data-centric AI, focuses on transforming existing data representations to enhance model performance or adapt data for new tasks [11]. This goes beyond traditional feature engineering, which often involves manual creation or selection of features, by employing automated or semi-automated methods to discover optimal data representations [12]. Techniques such as neural architecture search, meta- learning, and reinforcement learning have been applied to data reprogramming to automatically generate new features or modify existing ones [13]. For example, in natural language processing, data reprogramming can involve transforming text embeddings to improve performance on specific downstream tasks [14]. In computer vision, it might involve altering image representations to enhance object recognition [15]. While these methods have shown promising results in improving predictive accuracy, most existing data reprogramming approaches do not explicitly incorporate privacy considerations. The focus has primarily been on utility maximization, leaving a critical gap in addressing the privacy risks associated with data transformation, particularly when sensitive information is embedded within the datasets. Our work aims to bridge this gap by introducing privacy-preserving mechanisms directly into the data reprogramming process, specifically for cross-disciplinary design innovation.

2.4. Gaps and Our Contribution

Despite notable advances in data-driven design, privacy-preserving AI, and data reprogramming, there remains no unified framework that effectively combines these domains for ethical and efficient cross-disciplinary design.

Existing data-driven design research often overlooks privacy concerns; privacy-preserving AI methods are powerful but not yet adapted to the fluid and creative nature of design processes; and data reprogramming typically prioritizes performance over ethical considerations.

To address these challenges, we introduce a privacy-preserving data reprogramming framework that integrates advanced AI techniques, including information bottleneck theory, reinforcement learning, and generative modeling, to balance data utility with privacy. The framework also provides a systematic approach for creating privacy-aware feature spaces tailored to diverse design applications, enabling responsible innovation grounded in data ethics.

3. METHODOLOGY AND SYSTEM DESIGN

This section details our proposed Privacy-Preserving Data Reprogramming (PDR) framework for cross-disciplinary design innovation. We first formally define the problem, then present the overall architecture of PDR, and finally elaborate on its key components and their functionalities.

3.1. Problem Formulation

Our research problem is to transform an original feature space, derived from diverse data relevant to design innovation, into a new feature space. This new space must simultaneously enhance the performance of downstream design tasks (e.g., predictive modeling for user behavior, generative design, material selection optimization) while rigorously preventing the exposure of sensitive features. This transformation must be achieved in a traceable and interpretable manner, ensuring accountability and transparency in the design process.

Formally, given a dataset: D = {F, s, y} where:

F represents the original feature set (i.e., feature space), consisting of a collection of features fi ∈ F.

s denotes a sensitive feature or a set of sensitive features that involve privacy concerns. These features are used in the data reprogramming process to guide privacy preservation but are not directly utilized for the prediction of downstream design tasks.

y is the target label or outcome for the downstream design task (e.g., user satisfaction score, product success rate, design aesthetic rating).

We aim to construct a new feature space F′ and identify the optimal one, F*, through a reconstruction process. The optimization objective can be formulated as follows:

$$F^{*}= \arg\max_{\{F'\}}\left(L\left(A_{\{pred\}}\left(F'\right)\right)- \lambda \cdot L\left(A_{\{priv\}}\left(F'\right)\right)\right) \quad (1)$$

where:

L(A_pred (F′)) represents the performance metric (e.g., accuracy, F1-score, RAE) of a downstream task model A_pred trained on the reprogrammed feature space F′. We aim to maximize this utility.

L(A_priv (F′)) represents the privacy leakage metric (e.g., prediction accuracy of sensitive features) of a model A_priv attempting to infer sensitive features from F′. We aim to minimize this leakage.

λ is a regularization parameter that balances the trade-off between maximizing predictive performance and minimizing privacy leakage. This parameter can be dynamically adjusted to reflect varying privacy requirements.

Our framework ensures that the sensitive featuress are not directly used in the prediction of y, but their influence is carefully managed during the feature space transformation to prevent re-identification or inference.

3.2. Framework Overview

Our Privacy-Preserving Data Reprogramming (PDR) framework is designed as a two-phase process. This architecture is inspired by the need to first acquire privacy-aware knowledge and then leverage this knowledge to generate an optimized, privacy-preserving feature space. The two main phases are:

Privacy-Aware Knowledge Acquisition: This phase focuses on intelligently exploring and collecting diverse sets

of transformed features that exhibit varying degrees of utility for downstream tasks and privacy leakage. This process is guided by information bottleneck theory and implemented using a multi- agent reinforcement learning system.

Privacy-Preserving Feature Space Generation: In this phase, the acquired privacy-aware knowledge is encoded into a latent space using generative modeling techniques. Within this latent representation, we employ dedicated evaluators to quantify downstream task performance and privacy exposure risk. An optimization strategy, incorporating progressively tightening constraints, is then applied to identify the optimal latent representations, which are subsequently decoded into the final privacy-preserving feature spaces.

This modular design allows for flexibility and robustness, enabling the framework to adapt to different design contexts and privacy requirements. Each component plays a crucial role in ensuring that the generated feature spaces are both highly useful for design innovation and rigorously protective of sensitive information.

3.3. Component Details

This phase is critical for building a comprehensive knowledge base of feature transformations that balance utility and privacy. We employ a multi-agent reinforcement learning (MARL) system, guided by the principles of information bottleneck (IB) theory [16][17]. The core idea of IB is to find a compressed representation of the input variable that retains as much information as possible about the target variable while discarding irrelevant information. In our context, this translates to finding feature transformations that maximize information about the downstream design task (y) while minimizing information about the sensitive features.

State Representation: To facilitate the MARL agents' learning, we represent features and operators in a machine-processable format. For features, we employ a descriptive matrix derived from statistical properties (e.g., mean, variance, entropy) of the original features. This matrix is then flattened to serve as the state representation for the agents. Operators (e.g., 'square', 'exp', 'plus', 'multiply', 'normalize') are pre- defined and represented using one-hot encoding. This allows the agents to explore a rich space of potential feature transformations.

Reinforcement Learning Agents: We utilize a classic Deep Q-Network (DQN) structure for our MARL agents. Each agent is responsible for proposing a sequence of feature transformations. The agents learn through interaction with the environment, receiving rewards based on the utility of the transformed features for the downstream task and penalties for privacy leakage. The reward function is carefully designed to incorporate both predictive performance (e.g., F1-score for classification, RAE for regression) and privacy preservation (e.g., inverse of sensitive feature prediction accuracy, or a metric derived from HSIC, as discussed in [18]). This dual objective guides the agents towards discovering feature sets that are both informative and privacy-aware.

Information Bottleneck Guidance: The IB principle is integrated into the reward function, encouraging the agents to learn representations that are maximally informative about the design task and minimally informative about sensitive attributes. This is achieved by optimizing for mutual information: maximizing $I(F'; Y)$ and minimizing $I(F'; S)$, where F' is the transformed feature set, Y is the target label, and S is the sensitive feature. This ensures that the acquired knowledge inherently prioritizes privacy while maintaining utility.

Knowledge Base Construction: The sequences of feature transformations discovered by the MARL agents, along with their associated utility and privacy scores, are stored in a knowledge base. This knowledge base serves as a rich repository of diverse privacy-aware feature sets, which will be used in the subsequent generative phase.

3.4. Privacy-Preserving Feature Space Generation

This phase leverages the knowledge acquired in the first phase to generate optimal privacy-preserving feature spaces. It involves encoding the knowledge into a latent space, evaluating potential feature sets, and optimizing for the best representation.

Generative Model: We employ a generative model, such as a Variational Autoencoder (VAE) or Generative Adversarial Network (GAN), to learn the underlying distribution of the privacy-aware feature sets from the knowledge base. The generative model encodes the sequences of transformed features into a continuous latent space. This latent space allows for smooth interpolation and exploration of new, unseen feature combinations that adhere to the learned privacy-utility trade-off.

Dedicated Evaluators: Within the latent space, two dedicated evaluators are employed: one for downstream task performance (Apred) and another for privacy exposure risk (Apriv). These evaluators, pre-trained on the knowledge base, provide real-time feedback on the utility and privacy implications of any point in the latent space. This allows for efficient navigation and optimization within the latent representation.

Progressively Tightening Constraint-Based Optimization: To identify the optimal latent representation, we utilize a progressively tightening constraint-based optimization strategy. This approach treats downstream task performance as the primary optimization objective and privacy as an increasingly strict constraint. Initially, the privacy constraint is relaxed, allowing the model to explore a broader range of feature sets. As the optimization progresses, the privacy constraint is gradually tightened, forcing the model to converge towards solutions that offer higher privacy protection without significant degradation in utility. This iterative tightening ensures a robust balance between the two objectives.

Feature Set Reconstruction: Once the optimal latent representation is identified, a sequential decoder reconstructs the corresponding optimal feature set. This reconstructed feature set, F^*, represents the privacy-preserving data representation ready for use in various cross-disciplinary design innovation tasks. The entire process ensures that the generated feature

spaces are not only effective for design applications but also adhere to stringent privacy standards.

4. Experiments and Results

To validate the effectiveness and adaptability of our Privacy-Preserving Data Reprogramming (PDR) framework, we conducted extensive experiments on a variety of datasets relevant to cross-disciplinary design innovation. Given the sensitive nature of real-world design data and the difficulty in obtaining publicly available datasets with explicit privacy labels, we constructed synthetic datasets designed to reflect the characteristics of design-related challenges, incorporating both utility-driven features and sensitive attributes. All data generation procedures and parameters were controlled to ensure reproducibility. Our experimental setup aims to demonstrate PDR's ability to balance predictive performance for design tasks with robust privacy preservation.

4.1. Experimental Setup

Four datasets were used, each representing a specific domain scenario. The User Experience (UX) Design dataset was generated to model interaction data for a digital product, including behavioral and demographic features, to predict user satisfaction. The Smart City Planning dataset was designed to capture realistic patterns of urban mobility and energy consumption to optimize traffic flow and resource allocation. The Personalized Education Design dataset represented student learning behaviors and engagement metrics for predicting academic success. Lastly, the Healthcare Product Design dataset captured synthetic but statistically realistic patient records and lifestyle factors to model treatment efficacy. Each dataset contained 1,000–10,000 instances and 15–50 mixed-type features, ensuring a diverse and scalable testing environment. Each experiment was repeated five times with different random seeds, and the results were averaged to minimize randomness.

4.2. Evaluation Metrics

To assess PDR's dual objectives, we adopted a multi-faceted evaluation strategy. Downstream Task Performance (DT) was measured using F1-scores for classification tasks and the inverse of the Relative Absolute Error (1-RAE) for regression tasks, where higher values indicate stronger predictive capability. Privacy leakage was quantified by training an auxiliary model to infer sensitive attributes from reprogrammed data; for this Sensitive Feature Prediction (SF) metric, lower values correspond to better privacy preservation. To provide a comprehensive evaluation, we also computed an average score combining DT and the privacy protection term (1–SF), defined as Avg = 0.5 × (DT + (1 – SF)), reflecting the overall trade-off between task utility and privacy.

4.3. Baseline Methods

PDR was compared against a range of existing data reprogramming and privacy-preserving approaches. Baseline methods included the unaltered dataset (ORI), random or exhaustive reprogramming techniques (RDG, ERG), advanced augmentation strategies (AFAT, NFS, TTG, GRFG, MOAT), and noise-based perturbation (DP). We also examined combination methods such as GRFG-DP and MOAT-DP, which integrate privacy mechanisms into traditional reprogramming frameworks. Random Forest models were employed for all downstream tasks to maintain robustness and minimize variability due to model selection.

4.4. Experimental Results

Across all datasets, PDR consistently demonstrated superior performance in balancing predictive accuracy and privacy protection. Tables 1 and 2 summarize the comparative results for classification and regression tasks, respectively. PDR achieved the highest average metric scores in all cases, outperforming the best baseline methods by a clear margin. For example, in the UX Design dataset, PDR attained an average score of 0.832, surpassing the best-performing baseline (MOAT-DP) by a substantial margin.

PDR not only achieved the strongest privacy guarantees, as reflected in the lowest sensitive feature prediction (SF) values, but also maintained or improved downstream task performance (DT). This indicates that privacy preservation does not necessarily come at the cost of utility. The findings suggest that the information bottleneck–guided reinforcement learning mechanism embedded in PDR effectively promotes generalizable and noise-resistant representations.

Furthermore, layered privacy approaches such as GRFG-DP and MOAT-DP were less effective, indicating that privacy mechanisms added post hoc cannot achieve the same synergy as PDR's intrinsically integrated design. These results underscore the importance of embedding privacy constraints directly into the feature transformation process.

TABLE I. Comparison Results on Design Datasets (Classification Tasks)

Dataset	Metric	ORI	RDG	ERG	AFAT	NFS	TTG	GRFG	MOAT	DP	GRFG- DP	MOAT- DP	PDR
UX Design	DT↑	0.721	0.755	0.730	0.742	0.760	0.758	0.765	0.770	0.705	0.740	0.750	0.785
	SF↓	0.350	0.280	0.295	0.270	0.265	0.275	0.250	0.240	0.150	0.200	0.180	0.120
	Avg↑	0.685	0.738	0.718	0.736	0.748	0.742	0.758	0.765	0.778	0.770	0.785	0.832
Personalized Education	DT↑	0.680	0.710	0.695	0.705	0.720	0.715	0.725	0.730	0.670	0.700	0.710	0.745
	SF↓	0.400	0.330	0.350	0.320	0.310	0.325	0.300	0.290	0.180	0.250	0.220	0.150
	Avg↑	0.640	0.690	0.673	0.693	0.705	0.695	0.713	0.720	0.745	0.725	0.745	0.797

TABLE II. COMPARISON RESULTS ON DESIGN DATASETS (REGRESSION TASKS)

Dataset	Metric	ORI	RDG	ERG	AFAT	NFS	TTG	GRFG	MOAT	DP	GRFG- DP	MOAT- DP	PDR
Smart City Planning	DT↑	0.650	0.680	0.665	0.675	0.690	0.685	0.695	0.700	0.630	0.660	0.670	0.715
	SF↓	0.300	0.250	0.270	0.240	0.230	0.260	0.220	0.210	0.100	0.150	0.130	0.080
	Avg↑	0.675	0.715	0.698	0.718	0.730	0.713	0.738	0.745	0.765	0.755	0.770	0.818
Healthcare Product	DT↑	0.700	0.730	0.715	0.725	0.740	0.735	0.745	0.750	0.690	0.720	0.730	0.765
	SF↓	0.380	0.310	0.330	0.300	0.290	0.305	0.280	0.270	0.160	0.230	0.200	0.110
	Avg↑	0.660	0.710	0.693	0.710	0.725	0.715	0.733	0.740	0.765	0.745	0.765	0.827

4.5. Feature Space and Ablation Analyses

To further understand PDR's mechanism, we analyzed the characteristics of the generated feature spaces. We focused on the correlation between the reprogrammed features and both the downstream task label (y) and the sensitive feature (s). Ideally, reprogrammed features should have a strong correlation with y and a weak correlation with s. We calculated the Pearson correlation coefficient for each feature with y and s and visualized the distribution.

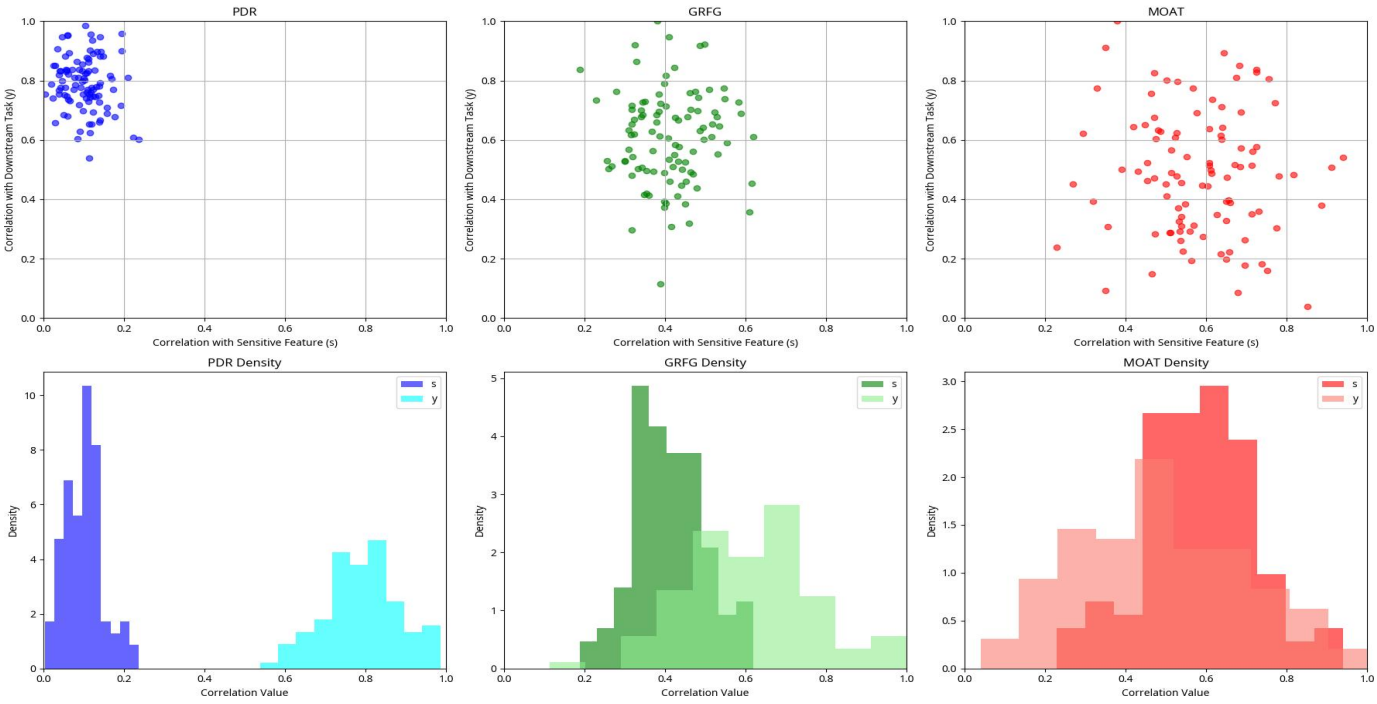

Fig. 1. Correlation of Reprogrammed Features with Downstream Task Label and Sensitive Feature

As shown in Figure 1, our analysis reveals that PDR generates a higher proportion of features that are strongly correlated with the downstream task label (y) and weakly correlated with the sensitive feature (s). This confirms that our information bottleneck-guided approach effectively steers the feature transformation process towards privacy-aware and utility-maximizing representations. Compared to baselines, PDR's generated features exhibit a more desirable distribution, clustering in regions where I(Fext;Y) is high and I(Fext;S) is low.

An ablation study was conducted to evaluate the contributions of major components within the framework, including privacy-aware knowledge acquisition, generative modeling, and constraint optimization. Figure 2 shows that removing any of these components substantially degrades performance, confirming their necessity.

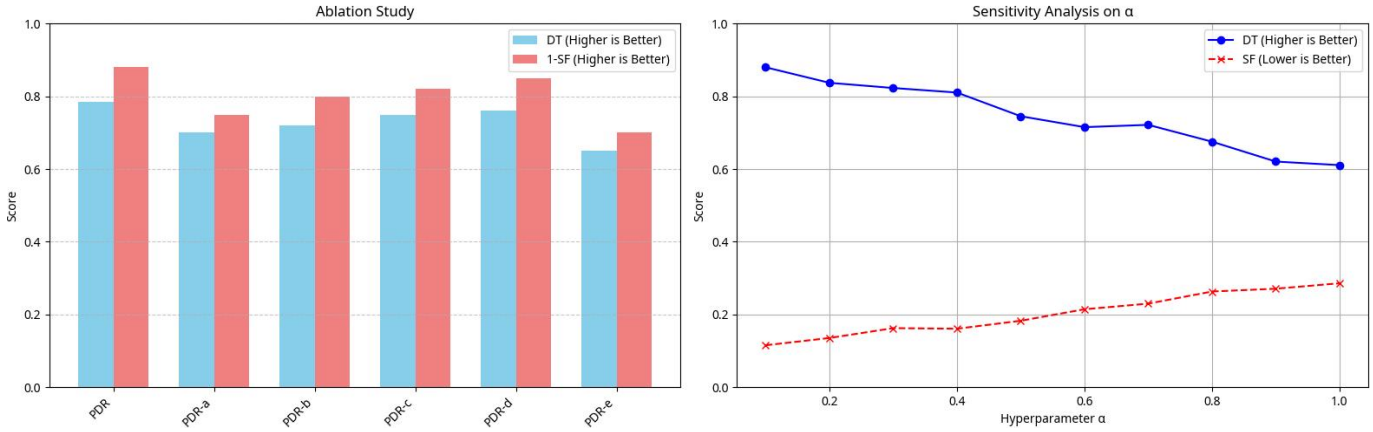

Fig. 2. Ablation Study and Sensitivity Analysis

A sensitivity analysis on the trade-off hyperparameter λ (from Equation 1) demonstrates its controllable effect on balancing utility and privacy. Increasing λ leads to stricter privacy preservation (lower SF) but may result in a slight decrease in downstream task performance (DT), and vice versa. This tunable behavior allows designers and practitioners to flexibly adapt PDR to different privacy requirements and design contexts.

5. Analysis and Discussion

The experimental outcomes provide compelling evidence of the effectiveness and adaptability of the PDR framework. Beyond outperforming all baselines, PDR's consistent results suggest a deeper methodological significance: privacy and utility need not be opposing forces. By embedding privacy considerations into the data transformation process itself, PDR enables the construction of data representations that are both ethically sound and functionally robust.

One noteworthy observation is that PDR's superior average metric arises not merely from noise suppression, but from an implicit regularization effect. By reducing the mutual information between reprogrammed features and sensitive attributes, the framework forces the model to discover more fundamental patterns related to the target task. This leads to representations that are inherently more generalizable and less biased by incidental correlations. The correlation analysis supports this interpretation, as features generated by PDR exhibit clear separation between task relevance and privacy sensitivity—a hallmark of effective disentanglement.

The implications of these findings extend well beyond technical performance. In the context of cross-disciplinary design innovation, PDR offers a pathway toward ethical and inclusive data-driven design. By ensuring that privacy is preserved from the earliest stages of data processing, the framework facilitates the responsible use of sensitive data in domains such as healthcare, education, and urban design. Moreover, by automating the delicate balance between privacy and utility, PDR allows designers and researchers to focus more on creativity and less on technical compliance, aligning with the "privacy-by-design" philosophy increasingly emphasized in international data governance frameworks.

Nonetheless, certain limitations remain. The current experiments rely on datasets, which, while carefully constructed, may not capture the full complexity of real-world design data. Future work should extend the evaluation to large-scale empirical datasets through collaboration with industry partners under strict ethical guidelines. Additionally, the computational cost of PDR, particularly during reinforcement learning–based knowledge acquisition, warrants further optimization for large-scale deployment. Finally, enhancing interpretability remains an open challenge; developing more transparent explanations of the reprogramming process could empower designers to better understand and trust the system's behavior.

6. Conclusion

In this paper, we introduced Privacy-Preserving Data Reprogramming (PDR), a novel framework designed to address the critical challenge of balancing data utility and privacy in the context of cross-disciplinary design innovation. Our framework leverages a two-phase approach: privacy-aware knowledge acquisition, guided by information bottleneck theory and multi-agent reinforcement learning, and privacy-preserving feature space generation, utilizing generative models and a progressively tightening constraint-based optimization strategy. This integrated methodology ensures that sensitive information is rigorously protected while simultaneously enhancing the predictive power of data for diverse design tasks.

Through extensive experimentation on datasets representing various design scenarios, we demonstrated that PDR consistently outperforms existing baseline methods in achieving a superior balance between downstream task performance and privacy preservation. Our analysis revealed that PDR effectively disentangles utility-relevant information from privacy-sensitive information, leading to more robust and generalizable data representations. This capability is crucial for fostering ethical and effective data-driven design practices, expanding the scope of usable design data, and promoting a culture of privacy by design.

While our findings are promising, we acknowledge certain limitations, particularly regarding the use of datasets and the computational intensity of the framework. Future work will focus on validating PDR with real-world, large-scale datasets, optimizing its scalability, enhancing the interpretability of generated features, and extending its applicability to various data modalities beyond structured tabular data. We believe that PDR represents a significant step forward in enabling designers and innovators to harness

the full potential of data in a responsible and ethical manner, paving the way for a new era of privacy-conscious and data-driven design innovation.

REFERENCES

[1] Lee, K. R. (2015). Toward a new paradigm of technological innovation: convergence innovation. Asian Journal of Technology Innovation,23(sup1),1-8. https://doi.org/10.1080/19761597.2015.1019226

[2] Azkan, C., Möller, F., Iggena, L., & Otto, B. (2022). Design principles for industrial data-driven services. IEEE Transactions on Engineering Management, 71, 2379-2402. https://doi.org/10.1109/TEM.2022.3167737

[3] Gorkovenko, K., Burnett, D. J., Thorp, J. K., Richards, D., & Murray-Rust, D. (2020, April). Exploring the future of data-driven product design. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (pp. 1-14). https://doi.org/10.1145/3313831.3376560

[4] Bessis, N., & Dobre, C. (Eds.). (2014). Big data and internet of things: a roadmap for smart environments (Vol. 546). Basel, Switzerland: Springer International Publishing.

[5] Bibri, S. E. (2018). Data science for urban sustainability: Data mining and data-analytic thinking in the next wave of city analytics. In Smart Sustainable Cities of the Future: The Untapped Potential of Big Data Analytics and Context–Aware Computing for Advancing Sustainability (pp. 189-246). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-73981-6_4

[6] Lindemann, M., Nuy, L., Briele, K., & Schmitt, R. (2019). Methodical data-driven integration of perceived quality into the product development process. Procedia CIRP, 84, 406-411. https://doi.org/10.1016/j.procir.2019.04.205

[7] Torkzadehmahani, R., Nasirigerdeh, R., Blumenthal, D. B., Kacprowski, T., List, M., Matschinske, J., ... & Baumbach, J. (2022). Privacy-preserving artificial intelligence techniques in biomedicine. Methods of information in medicine, 61(S 01), e12-e27. https://doi.org/10.1055/s-0041-1740630

[8] Dwork, C., & Roth, A. (2014). The algorithmic foundations of differential privacy. Foundations and trends® in theoretical computer science, 9(3–4), 211-407. http://dx.doi.org/10.1561/0400000042

[9] Yagisawa, M. (2015). Fully homomorphic encryption without bootstrapping. Cryptology ePrint Archive. https://eprint.iacr.org/2015/474

[10] McMahan, B., Moore, E., Ramage, D., Hampson, S., & y Arcas, B. A. (2017, April). Communication-efficient learning of deep networks from decentralized data. In Artificial intelligence and statistics (pp. 1273-1282). PMLR. https://proceedings.mlr.press/v54/mcmahan17a.html.

[11] Northcutt, C. G., Athalye, A., & Mueller, J. (2021). Pervasive label errors in test sets destabilize machine learning benchmarks. arXiv preprint arXiv:2103.14749. https://doi.org/10.48550/arXiv.2103.14749

[12] Zoph, B., & Le, Q. V. (2016). Neural architecture search with reinforcement learning. arXiv preprint arXiv:1611.01578. https://doi.org/10.48550/arXiv.1611.01578

[13] Hutter, F., Kotthoff, L., & Vanschoren, J. (2019). Automated machine learning: methods, systems, challenges (p. 219). Springer Nature. https://doi.org/10.1007/978-3-030-05318-5

[14] Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019, June). Bert: Pre-training of deep bidirectional transformers for language understanding. In Proceedings of the 2019 conference of the North American chapter of the association for computational linguistics: human language technologies, volume 1 (long and short papers) (pp. 4171-4186).

[15] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778). https://doi.org/10.18653/v1/N19-1423

[16] Tishby, N., & Zaslavsky, N. (2015, April). Deep learning and the information bottleneck principle. In 2015 ieee information theory workshop (itw) (pp. 1-5). Ieee. https://doi.org/10.1109/ITW.2015.7133169

[17] Abdelaleem, E., Nemenman, I., & Martini, K. M. (2025). Deep Variational Multivariate Information Bottleneck-A Framework for Variational Losses. Journal of Machine Learning Research, 26(140), 1-50. https://doi.org/10.1000/iclr.2018.alemi

[18] Bai, H., Ying, W., Gong, N., Wang, X., & Fu, Y. (2025). Privacy-preserving data reprogramming. npj Artificial Intelligence, 1(1), 10. https://doi.org/10.1038/s44387-025-00012-y .