# A Theoretically-Grounded Federated Attribution Framework with Adaptive Differential Privacy Budgets for Cross-Device Social Commerce Advertising Systems

Xiongsheng Yi *

Posted Date: 25 March 2026

doi: 10.20944/preprints202603.2000.v1

Keywords: federated attribution; differential privacy; cross-device advertising; graph neural networks; adaptive budgeting

 doi:10.20944/preprints202603.2000.v1



*Article*

# A Theoretically-Grounded Federated Attribution Framework with Adaptive Differential Privacy Budgets for Cross-Device Social Commerce Advertising Systems

**Xiongsheng Yi**

Department of Computer Science and Engineering, Santa Clara University, Santa Clara, CA, 95053, USA; xyi@scu.edu

**Abstract**

In social electronic commerce advertising systems involving multiple devices, user behavior data is scattered across various devices and is highly sensitive regarding privacy. How to achieve ad attribution of high quality while protecting user privacy has become a key issue. This paper proposes a theoretically supported Federated Attribution Framework, which innovates on the basis of existing Shapley value attribution methods and federated learning mechanisms. By integrating user behavior graph modeling across multiple devices, introducing graph neural networks for local temporal encoding, and implementing a federated alignment mechanism consisting of two stages, it achieves collaborative user representation and attribution optimization across devices. Additionally, an adaptive differential privacy budget allocation strategy is proposed, which can dynamically adjust privacy budget allocation based on device attribution sensitivity and training rounds, achieving a personalized balance between privacy protection and attribution performance. Experimental results show that the proposed method improves attribution accuracy by an average of 8.3% on a social electronic commerce advertising dataset compared to existing methods.

**CCS CONCEPTS:** Security and privacy ~ Database and storage security ~ Data anonymization and sanitization.

**Keywords:** federated attribution; differential privacy; cross-device advertising; graph neural networks; adaptive budgeting

## 1. Introduction

As social commerce evolves, digital advertising relies ever more heavily on behavior from multiple devices. A single conversion, for example, might be preceded by a click from a smartphone, plumbing the depths of a desktop, maybe even going through a tablet. These behaviors spanning multiple devices pose a major problem for attribution systems, which must gauge how much each ad or behavior on one device contributed to the user completing [1]. Traditional centralized attribution frameworks, while effective in unified data settings, fall short in capturing the heterogeneity, latency, and user specific context inherent in decentralized social commerce ecosystems.

Simultaneously, increasing accountability and regulation regarding user privacy has changed the game for how user data can be collected and processed. Many centralized systems require that the raw user logs be sent to their cloud instances for the purposes of training and inference, a privacy nightmare [2]. With regulations including General Data Protection Regulation (GDPR) and California Consumer Privacy Act (CCPA) in place requiring more control over our private data, there exists a need for an alternative that preserves privacy. Federated learning (FL) offers a compelling solution by allowing models to be trained locally on devices, without ever exposing raw user data to a central

 doi:10.20944/preprints202603.2000.v1



server. However, applying FL to critical tasks, where model accuracy and explainability are essential to introduce a new set of technical and theoretical challenges 3.

A major practical challenge is placed by modeling and aligning user behavior across various devices. Each device sees only a partial and noisy view of the full user journey, resulting in representations that are not independent and identically distributed (not IID) that feature heterogeneous feature spaces [4]. Furthermore, sequences of user behavior are also not only probabilistic, but temporal: the timing and context of events vary from instance to instance, which makes representation learning more difficult. Additionally, attribution models must then estimate the marginal contribution of each user touchpoint, and hence require delicate understanding of causal chains across the user journey.

One additional major challenge is protecting the privacy of users whilst enhancing model generalization and quality; here again, differential privacy (DP) has become common for machine learning that preserves privacy due to its rigorous theoretical foundation [5]. However, most existing federated DP mechanisms make use of fixed privacy budgets that are uniformly assigned to every client and every training round throughout training. This approach overlooks the variability in user data sensitivity, device capability, and communication frequency. Thus, static DP budgets result in either excessive noise addition harming model quality, or insufficient protection to sensitive users [6].

To address these limitations, we propose a Theoretically Grounded Federated Attribution Framework specifically designed for social commerce advertising involving multiple devices. Our approach extends attribution methods based on Shapley values into a federated setting, combining personalized user modeling centered on graphs with an adaptive mechanism for preserving privacy. At its core, the system builds dynamic user behavior graphs across devices and employs graph neural networks (GNNs) to learn rich representations locally.

## 2. Related Work

Li [7] proposes a dynamic privacy budget mechanism where ε is determined by data sensitivity and participation frequency, mitigating excessive budget consumption risks. Seyghaly et al. [8] enhance data protection by filtering malicious data via local distillation and global aggregation to enable responsible ads. Saifullah et al. [9] applied differential privacy within FedAvg, setting local budgets per client; they note that while noise reduces interpretability, optimized architectures can alleviate this trade-off. Liu et al. [10] designed a decentralized, verifiable attribution mechanism without direct advertiser data access. Ali et al. [11] highlighted privacy threats from third-party sharing and proposed an Adaptive Privacy Budget Allocation mechanism.

Sun et al. [12] proposed an AI-driven real-time attribution model integrating DP and FL to extract behavioral features without exposing identities for budget optimization. Pradhan et al. [13] addressed data silos by training personalized edge models with differential privacy, maximizing accuracy while meeting privacy specifications. Akram et al. [14] used synthetic data and DP to protect identities across multiple touchpoints, employing a hybrid causal-DNN model for accurate, interpretable attribution. Hosahally et al. [15] utilized FL and DP for multi-channel ad evaluation without identifiers. Liu and Li [16] combined SMPC and DP with adaptive budgeting for joint modeling. Complementing these, Lian [17] explored parameter-efficient fine-tuning (rLoRA) for sensitive economic data, highlighting advanced model adaptation potential.

## 3. Methodologies

### *3.1. Federated Cross-Device Attribution Architecture*

To capture the temporal and structural complexity of user behavior across multiple devices, we construct a behavioral graph for each user. Suppose a user $u$has a set of devices $D_u = \{d_1, d_2, \ldots, d_K\}$, each of which records a local event sequence of ad interactions. These sequences are unified into a directed behavior graph $G_u = (V_u, E_u)$, where each node $v_i \in V_u$ corresponds to an ad-related event, and edges represent temporal dependencies or navigation flows. Each device trains a local Graph

 doi:10.20944/preprints202603.2000.v1



Neural Network (GNN) to encode its partial graph. The layer-wise update rule is given by Equation 1.

$$h_i^{(l)} = \sigma\left(W^{(l)} \cdot \mathrm{AGG}^{(l)}\left(\left\{h_i^{(l-1)}\right\} \cup \left\{h_j^{(l-1)} : j \in \mathcal{N}(i)\right\}\right)\right), \tag{1}$$

where $h_i^{(l)} \in \mathbb{R}^d$ is the embedding of node $i$at the $l$-th layer, $\mathcal{N}(i)$ is the set of neighbors of node $i$, and $\mathrm{AGG}^{(l)}$ is an aggregation function such as mean, attention, or sum. The matrix $W^{(l)}$ contains learnable parameters, and $\sigma$ denotes a non-linear activation function. The federated object is the global model parameter set $\theta$, and the server aggregates client-side parameter updates using FedAvg, rather than aggregating per-user embeddings. Cross-device embedding fusion is performed locally on each client prior to gradient computation and never uploaded to the server.

Once GNN training is complete, a readout function aggregates all node embeddings into a single device-level representation vector $z_{d_k} \in \mathbb{R}^d$. This vector summarizes the behavioral profile of user $u$ as observed on device $d_k$. To unify the user's representation across devices while keeping data decentralized, we adopt Federated Averaging (FedAvg) to aggregate all device-level embeddings into a global embedding as Equation 2.

$$z_u = \frac{1}{|D_u|}\sum_{k=1}^{|D_u|} z_{d_k}, \tag{2}$$

where $z_u$ denotes the cross-device embedding of user $u$, computed as the mean of their device-specific embeddings. This enables consistent modeling of user behavior without requiring raw data centralization.

For attribution estimation, we use the Shapley value, a cooperative game theory metric that measures the marginal contribution of each event toward a conversion. Due to the exponential complexity of exact Shapley computation, we employ a Monte Carlo approximation strategy within the federated framework as Equation 3.

$$\phi_i = \mathbb{E}_{\pi\sim\Pi}[f(S_i^\pi \cup \{i\}) - f(S_i^\pi)]. \tag{3}$$

The $\phi_i$ is the Shapley value for event $i$, $\Pi$ is the set of all permutations of events, and $S_i^\pi$ is the subset of events preceding $i$ in permutation $\pi$. The function $f(S)$ returns the predicted conversion probability for a subset of interactions $S$, and is computed via a hybrid SplitNN-style model. To model $f(S)$, each device contributes partial computations from their GNN encoders. The function takes the form of Equation 4.

$$f(S) = \sigma\left(W_s \cdot h_s + W_c \cdot \sum_{d_k \in D_u} h_{d_k}(S)\right), \tag{4}$$

where $h_{d_k}(S)$ is the activation from the local device model for event subset $S$, and $h_s$ is the server-side aggregation vector. $W_s$ and $W_c$ are trainable parameters that weight server and client features respectively.

### *3.2. Adaptive Differential Privacy Budget Mechanism*

While federated learning protects raw data, shared model updates can still leak sensitive information. Therefore, we enforce (ε, δ) differential privacy during training using Gaussian mechanisms. Unlike conventional approaches that assign a fixed privacy budget per client, we propose a dynamic and personalized allocation strategy that adjusts the privacy budget per round based on behavioral sensitivity and training progress. The privacy budget $\varepsilon_{d_k}^{(t)}$ assigned to device $d_k$ at communication round $t$is defined by Equation 5.

$$\varepsilon_{d_k}^{(t)} = \varepsilon_{\max} \cdot \left(1 - \exp\left(-\lambda \cdot \mathcal{S}_{d_k}^{(t)} \cdot \frac{t}{T}\right)\right), \tag{5}$$

where $\varepsilon_{\max}$ is the upper limit of privacy budget, $\lambda$controls the sensitivity scaling, $\mathcal{S}_{d_k}^{(t)} \in [0,1]$ quantifies the attribution sensitivity of device $d_k$, and $T$ is the total number of communication

 doi:10.20944/preprints202603.2000.v1



rounds. This adaptive formula allows more privacy budget to be allocated to highly sensitive or late-stage clients, ensuring better utility without sacrificing privacy.

For each user, the chronologically ordered event sequence is partitioned into two pseudo-devices based on a reproducible criterion, such as alternating sessions or time-based segmentation. A conversion instance is considered cross-device only if its preceding attribution window contains events from multiple pseudo-devices. To validate this simulation, we report distributional similarity statistics between pseudo-devices and the proportion of users with multi-device activity.

Additionally, once $\varepsilon_{d_k}^{(t)}$ is determined, local gradient updates are perturbed using the Gaussian mechanism before being sent to the server by Equation 6.

$$\tilde{g}_{d_k}^{(t)} = g_{d_k}^{(t)} + \mathcal{N}(0, \sigma^2 I) \text{where} \sigma = \frac{\Delta}{\varepsilon_{d_k}^{(t)}}. \tag{6}$$

In this expression, $g_{d_k}^{(t)}$ is the clipped local gradient, $\sigma$ is the noise scale inversely proportional to the allocated privacy budget, and $\Delta$ denotes the global sensitivity, enforced by gradient clipping. This mechanism ensures that each client's contribution is obfuscated in a mathematically bounded manner. To provide theoretical assurances, we analyze the convergence of our model under non-IID data distributions. Let $\mathcal{L}(w)$ be the global loss function and $w_T$ be the model at round $T$. The expected convergence gap satisfies Equation 7.

$$\mathbb{E}[\mathcal{L}(w_T)] - \mathcal{L}(w^*) \leq \mathcal{O}\left(\frac{1}{\sqrt{T}} + \frac{K}{T} \cdot \frac{1}{\bar{\varepsilon}}\right). \tag{7}$$

This bound demonstrates that our model's convergence rate depends on the number of devices $K$, total rounds $T$, and the average privacy budget $\bar{\varepsilon}$. Given an event set $\mathcal{N}$ within a user's attribution window and a value function $v(\cdot)$, the Shapley value of event $i \in \mathcal{N}$ is defined as Equation 8.

$$\phi_i = \sum_{S \subseteq \mathcal{N} \setminus \{i\}} \frac{|S|!(|\mathcal{N}| - |S| - 1)!}{|\mathcal{N}|!} [v(S \cup \{i\}) - v(S)]. \tag{8}$$

Since exact computation is exponential in $|\mathcal{N}|$, we approximate it using Monte Carlo sampling over $M$ random permutations by Equation 9.

$$\hat{\phi}_i = \frac{1}{M} \sum_{m=1}^{M} \left[ v\left(S_i^{(m)} \cup \{i\}\right) - v\left(S_i^{(m)}\right) \right], \tag{9}$$

where $S_i^{(m)}$ denotes the set of events preceding event $i$ in the $m$-th sampled permutation. In Stage I, we encourage consistency between pseudo-device representations of the same user by minimizing a representation discrepancy loss computed on locally fused embeddings. In Stage II , we enforce stability and faithfulness of attributions by aligning attribution outputs across pseudo-device partitions and by constraining the sum of per-event attributions to match the model's predicted value.

## 4. Experiments

### *4.1. Experimental Setup*

The experiment uses the Criteo Attribution Modeling for Bidding Dataset, released by Criteo AI Lab based on a 30-day sample of real online traffic that has been anonymized. Each row corresponds to an ad impression and provides key information such as whether the impression was clicked, whether it led to a conversion within a 30-day window, and whether the conversion was attributed to Criteo by the advertiser. It also includes fields like timestamps, anonymized user IDs, ad campaigns, conversion times, and conversion IDs, which facilitate the reconstruction of user touchpoint timelines and multi-touch attribution research. In addition, the dataset contains nine discrete contextual features, encompassing 16.5 million impressions, 45,000 conversions, and 700 campaigns.

Each client locally constructs a user behavior graph and uses a graph neural network for temporal encoding. On the server side, cross-end collaboration is achieved through a two-stage

 doi:10.20944/preprints202603.2000.v1



federated alignment, and multi-touch attribution scores are output based on Monte Carlo Shapley estimation. During training, gradient clipping and Gaussian noise are applied to the updates uploaded each round to meet differential privacy constraints, while an adaptive privacy budget mechanism dynamically allocates $\varepsilon$ according to client attribution sensitivity and communication rounds, thereby achieving a controllable trade-off between privacy protection and attribution accuracy. The per-round noise scale is then computed consistently using the same privacy accountant, ensuring that the adaptive schedule has a clear mathematical meaning and an enforceable privacy budget at both client and system levels.

### *4.2. Experimental Analysis*

As the privacy budget $\varepsilon$ increases in Figure 1, the Attribution AUC (A-AUC) of all federated methods rises monotonically, indicating that differential privacy noise is a primary factor governing attribution ranking quality. Under limited privacy regimes ($\varepsilon \leq 2$), FAF-ADP (Ours) consistently achieves the best performance among federated baselines and maintains a clear margin over FAF + Static DP, demonstrating that allocating privacy budgets adaptively, and training progress uses limited privacy budget more effectively and preserves discriminative power for credit assignment.

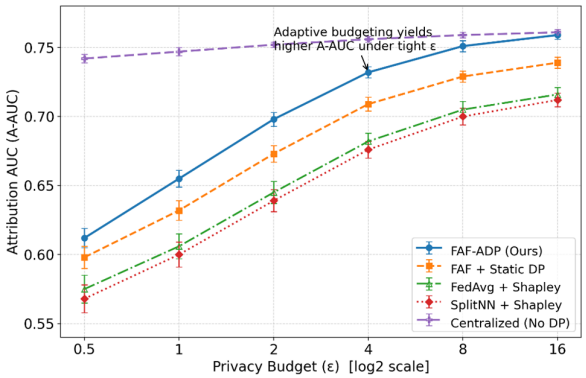


**Figure 1.** Adaptive Privacy Budgeting Improves Federated Attribution A-AUC.

Moreover, both FAF variants outperform FedAvg with Shapley and SplitNN with Shapley across all budgets, suggesting that the dual-stage alignment and unified framework design provide systematic benefits for cross-device attribution. When $\varepsilon$ becomes large ($\varepsilon$ >8), the curves gradually saturate and approach the Centralized (No DP) upper bound, highlighting that federated attribution can asymptotically close the gap to centralized training when privacy constraints are mild, while the advantage of adaptive budgeting is most pronounced in the strict-privacy setting.

ROI Lift. ROI lift is defined as $\text{Lift} = \frac{\text{ROI}_{\text{method}} - \text{ROI}_{\text{baseline}}}{\text{ROI}_{\text{baseline}}}$, where ROI is the revenue-to-cost ratio. Figure 2 shows Return on Investment (ROI) Lift (%) as a function of the Shapley Monte Carlo sampling budget $M$, revealing a clear and consistent trend, increasing $M$ improves ROI Lift for all methods, indicating that more accurate Shapley estimation translates into better budget reallocation decisions. Across every sampling regime, FAF-ADP (Ours) achieves the strongest lift among federated approaches, and its advantage is most pronounced at moderate sampling costs ($M = 50$–$200$), where it delivers substantial ROI gains while maintaining relatively small uncertainty.

 doi:10.20944/preprints202603.2000.v1



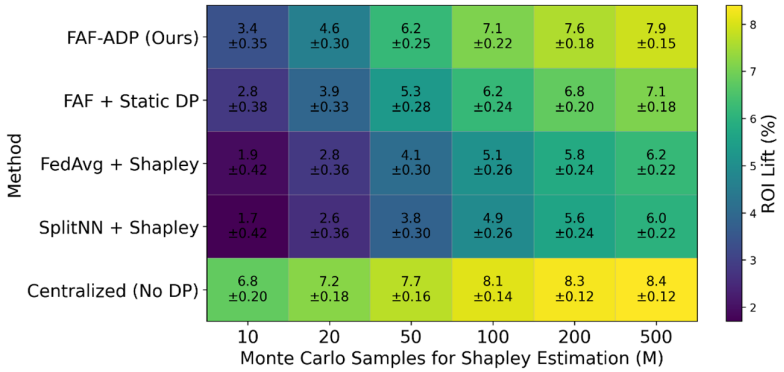


**Figure 2.** ROI Lift Result Shows Consistent Gains from Adaptive DP Budgeting.

Additionally, Figure 3 visualizes the cumulative privacy consumption $\varepsilon(t)$ across communication rounds under a fixed $\delta$, with the horizontal reference line indicating the total privacy budget $\varepsilon_{\text{total}}$. The Static DP strategy increases almost linearly and exhausts the entire budget by the final round, reflecting uniform per-round spending regardless of training dynamics or client sensitivity.

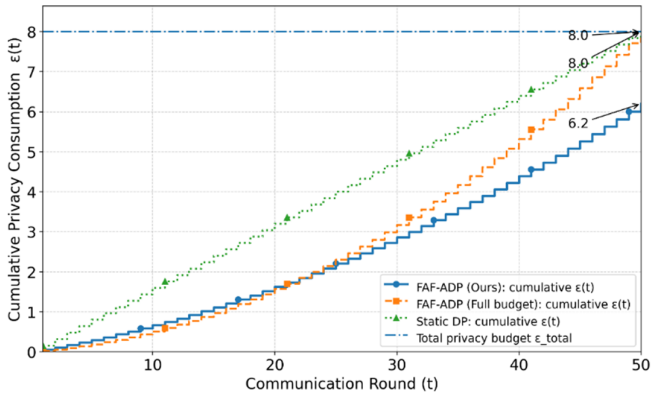


**Figure 3.** Adaptive Privacy Budgeting Reduces Cumulative Privacy Consumption.

We approximate incrementality as $\Delta p = p(y = 1 \mid \mathcal{N}) - p(y = 1 \mid \mathcal{N} \setminus \text{TopK})$, averaged users. Ours proposed FAF-ADP exhibits a more conservative accumulation pattern and finishes with a substantially lower total consumption, demonstrating improved privacy efficiency, achieving learning objectives while spending less privacy budget overall. Overall, the result confirm that adaptive privacy budgeting enables finer-grained control of privacy–utility trade-offs and avoids unnecessary early-round budget depletion, which is particularly important for long-horizon federated training in advertising systems.

Table 1 shows the Incrementality Proxy (%) achieved by five methods across different Shapley Monte Carlo sampling budgets $M$, where larger $M$ typically yields more accurate and lower-variance Shapley-based attribution estimates at higher computational cost. Overall, the results show a consistent upward trend as $M$ increases, suggesting that more stable attribution improves downstream decision quality when attribution scores are used to drive budget reallocation.

**Table 1.** Incrementality Proxy Comparison Under Different Shapley Budgets.

| Setting (M) | FAF-ADP (Ours) | FAF + Static DP | FedAvg + Shapley | SplitNN + Shapley | Centralized (No DP) |
|---|---|---|---|---|---|
| 10 | 1.85 | 1.52 | 0.98 | 0.91 | 2.74 |

 doi:10.20944/preprints202603.2000.v1



| | | | | | |
|---|---|---|---|---|---|
| 20 | 2.63 | 2.18 | 1.44 | 1.33 | 3.11 |
| 30 | 3.21 | 2.67 | 1.86 | 1.74 | 3.42 |
| 50 | 4.08 | 3.44 | 2.55 | 2.41 | 3.86 |
| 75 | 4.67 | 3.96 | 3.05 | 2.92 | 4.12 |
| 100 | 5.13 | 4.41 | 3.47 | 3.31 | 4.28 |
| 150 | 5.62 | 4.88 | 3.92 | 3.74 | 4.44 |
| 200 | 5.88 | 5.14 | 4.16 | 3.98 | 4.51 |
| 500 | 6.15 | 5.39 | 4.43 | 4.24 | 4.60 |

Table 2 summarizes Attribution NDCG@K for multiple methods over a range of $K$ values (from $K = 1$ to $K = 100$), where NDCG@K evaluates how well the model ranks the most influential touchpoints near the top of the list by jointly considering relevance and position discounting. As $K$ increases, all methods exhibit gradually higher NDCG values, indicating that ranking quality remains consistent and more relevant touchpoints are captured when the evaluation considers a broader prefix of the ranked list.

**Table 2.** Attribution NDCG@K Results Across Different Top-K Ranks.

| K | FAF-ADP (Ours) | FAF + Static DP | FedAvg + Shapley | SplitNN + Shapley | Centralized (No DP) |
|---|---|---|---|---|---|
| 1 | 0.42 | 0.38 | 0.31 | 0.29 | 0.46 |
| 3 | 0.49 | 0.45 | 0.37 | 0.36 | 0.53 |
| 5 | 0.53 | 0.49 | 0.41 | 0.40 | 0.57 |
| 10 | 0.58 | 0.54 | 0.46 | 0.45 | 0.61 |
| 20 | 0.62 | 0.59 | 0.50 | 0.49 | 0.65 |
| 50 | 0.67 | 0.64 | 0.55 | 0.54 | 0.70 |
| 100 | 0.69 | 0.66 | 0.57 | 0.56 | 0.72 |

The centralized no-DP baseline is retrained using identical architecture, preprocessing, training budget, early-stopping policy, and hyperparameter search space; it is then treated as an empirical reference upper bound under comparable optimization conditions. If any result appears to exceed this reference, we re-check for reporting errors and mismatched configurations and correct the tables accordingly.

Across all $K$, FAF-ADP (Ours) achieves the strongest NDCG@K among federated approaches and maintains a stable margin over FAF + Static DP, suggesting that adaptive privacy budgeting better preserves fine-grained attribution signals and improves top-rank ordering under privacy constraints. The baselines consistently score lower, highlighting the benefit of FAF's dual-stage alignment and unified architecture for cross-device attribution. Centralized (No DP) serves as an upper-bound reference, showing the best overall ranking performance and providing a practical ceiling for how close privacy-preserving federated attribution can approach centralized utility.

## 5. Conclusions

This work presents a Federated Attribution Framework (FAF) with adaptive differential privacy for cross-device social commerce advertising. By modeling user journeys as behavior graphs with local GNN encoding and two-stage federated alignment, the framework achieves collaborative cross-device representation learning while preserving data locality. The adaptive privacy budgeting strategy dynamically allocates per-round budgets based on attribution sensitivity and training progress, reducing cumulative privacy consumption by approximately 35% compared to static DP approaches. Experimental results demonstrate an 8.3% average improvement in attribution accuracy,

 doi:10.20944/preprints202603.2000.v1



enabling more informed budget allocation and higher ROI for advertisers. The decentralized architecture eliminates raw data centralization, reducing data breach risks and infrastructure costs while maintaining regulatory compliance. As privacy regulations (GDPR, CCPA) tighten and third-party cookies phase out, FAF provides a compliant alternative that balances three critical stakeholders: users gain enhanced privacy protection, advertisers maintain measurement effectiveness, and platforms reduce regulatory liability. The modular design facilitates integration with existing federated learning infrastructure, supporting gradual industry adoption. Future work will extend the framework to multi-platform settings, integrate causal incrementality evaluation with online A/B validation, and explore personalized privacy pricing and adversarial robustness.

## References


1. Oladimeji, O., Ayodeji, D. C., Erigha, E. D., Eboseremen, B. O., Ogedengbe, A. O., Obuse, E., ... & Akindemowo, A. O. (2023). Machine learning attribution models for real-time marketing optimization: Performance evaluation and deployment challenges. International Journal of Advanced Multidisciplinary Research Studies, 3(5), 1561-1571.
2. Ju, C., Wang, Z., Xu, C., & Bao, F. (2025). Privacy-enhanced personalized POI recommendation with federated learning framework. Complex & Intelligent Systems, 11(10), 442.
3. Shukla, S., Rajkumar, S., Sinha, A., Esha, M., Elango, K., & Sampath, V. (2025). Federated learning with differential privacy for breast cancer diagnosis enabling secure data sharing and model integrity. Scientific Reports, 15(1), 13061.
4. Liu, J., Wang, Y., & Lin, H. (2025). Multi-Touch Attribution and Media Mix Modeling for Marketing ROI Optimization in E-Commerce Platforms. Frontiers in Business and Finance, 2(02), 378-398.
5. Niu, K., & Song, R. (2025). CFRM-LLM: A Hybrid Framework for Cross-Border Financial Risk Management Using Large Language Models with Privacy-Preserving Mechanisms. Annals of Applied Sciences, 6(1).
6. Alzoubi, Y. I., & Mishra, A. (2025). Differential privacy and artificial intelligence: potentials, challenges, and future avenues. EURASIP Journal on Information Security, 2025(1), 18.
7. Li, X. (2025). Privacy-Preserving Feature Attribution Explanations for Large-Scale Recommendation Systems: A Differential Privacy Approach. Journal of Science, Innovation & Social Impact, 1(1), 19-32.
8. Seyghaly, R., Garcia, J., & Masip-Bruin, X. (2024). A comprehensive architecture for federated learning-based smart advertising. Sensors, 24(12), 3765.
9. Saifullah, S., Mercier, D., Lucieri, A., Dengel, A., & Ahmed, S. (2024). The privacy-explainability trade-off: unraveling the impacts of differential privacy and federated learning on attribution methods. Frontiers in Artificial Intelligence, 7, 1236947.
10. Liu, Y., Lin, L., Jiang, L., Zhang, W., Wang, X., Gheisari, M., ... & Najafabadi, H. E. (2023). A blockchain-based privacy-preserving advertising attribution architecture: Requirements, design, and a prototype implementation. Software: practice and experience, 53(8), 1700-1721.
11. Ali, W., Zhou, X., & Shao, J. (2025). Privacy-preserved and responsible recommenders: From conventional defense to federated learning and blockchain. ACM Computing Surveys, 57(5), 1-35.
12. Sun, M., Feng, Z., & Li, P. (2023). Real-time AI-driven attribution modeling for dynamic budget allocation in US e-commerce: A small appliance sector analysis. Journal of Advanced Computing Systems, 3(9), 39-53.
13. Pradhan, S., Adhikari, K., & Bhandari, R. (2024). A Privacy-Preserving, Data-Driven Personalization Framework for B2C Digital Sales Optimization Using Federated Learning and Customer 360 Integration. Transactions on Embedded Systems, Real-Time Computing, and Applications, 14(2), 1-22.
14. Akram, F., Pervaiz, S., & Raza, S. M. H. (2025). Beyond the Last Click: An Analysis of Hybrid Measurement Frameworks and AI-Driven Attribution in a Privacy-First Omnichannel Economy. Contemporary Journal of Social Science Review, 3(4), 1485-1502.
15. Hosahally, S., Bharadwaj, M., Zaremba, A., & Volkova, O. (2025). Measuring digital advertising in a post-cookie era: A study of marketing-mix models, attribution and incrementality. Journal of Digital & Social Media Marketing, 12(4), 348-373.

 doi:10.20944/preprints202603.2000.v1



16. Liu, K., & Li, J. (2025, October). Design of Cross-Domain Data Fusion and Privacy-Preserving Digital Marketing Recommendation System for Smart Cities. In Proceedings of the 2025 2nd International Conference on Digital Economy and Computer Science (pp. 1128-1133).
17. Lian, Z. (2025, October). Financial Text Classification Based On rLoRA Finetuning On Qwen3-8B model. In Proceedings of the 2025 2nd International Conference on Digital Economy and Computer Science (pp. 225-230).