Beyond Centralization: User-Controlled Federated Recommendations in Practice

Manel Slokom
manel.slokom@cwi.nl
Centrum Wiskunde & Informatica
The Netherlands

Alejandro Bellogin
alejandro.bellogin@uam.es
Universidad Autónoma de Madrid
Spain

Abstract

Recommendation systems typically require centralized user data, limiting user control and raising privacy concerns. Federated learning offers an alternative by keeping data on-device, but its impact on real user behavior remains largely unexplored. We present a live federated recommender system that allows users to control the recommendation objective while keeping their data local. In a 53-day deployment with 22 participants and a catalog of 8807 titles, users interacted with recommendations and switched between personalization and diversity-enhanced ranking. We find that users prefer personalization when given explicit choice (65.37% vs. 62.07% CTR), actively engage with control mechanisms (3.93/5 satisfaction; 248 settings changes), and develop an understanding of how their interactions affect recommendations through immediate feedback. Our results show that user control, privacy, and effective personalization can be combined in a working system. We demonstrate a practical approach to interactive, privacy-preserving recommendation. Code and demo materials are available at: https://github.com/SlokomManel/federated-recommendations-participants

CCS Concepts

• **Information systems**; • **Recommender Systems**;

Keywords

Federated learning, recommender systems, privacy-preserving personalization, user agency, user studies

ACM Reference Format:
Manel Slokom and Alejandro Bellogin. 2026. Beyond Centralization: User-Controlled Federated Recommendations in Practice. In *34th ACM Conference on User Modeling, Adaptation and Personalization (UMAP '26), June 08–11, 2026, Gothenburg, Sweden.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3774935.3812710

1 Introduction

Recommendation systems shape what people watch, read, and listen to. In most real-world deployments, however, users have little visibility into or control over what data is collected, how it is processed, and what the system optimizes for. This creates a gap between what systems do *to* users and what users can do *with* them. This paper asks a simple question: can we give users direct control over recommendations while preserving both personalization and privacy?

To answer this, we build and deploy a **federated recommender system**. In federated learning, user data remains on-device; models are trained locally, and only model updates (not raw data) are shared and aggregated [8]. This allows systems to learn from many users without centralizing personal data.

We implement this approach in a working system where participants receive movie recommendations, control the recommendation objective, customize the interface, and observe how their interactions affect results. Users can switch between (i) standard personalization and (ii) diversity-enhanced ranking, or opt out of aggregation entirely. In this design, privacy is enforced by architecture, and control is exposed through the interface.

We deploy the system with 22 participants over 53 consecutive days using a catalog of 8807 titles. While prior work has studied federated recommendation in offline or simulated settings [1, 5, 6, 8], we focus on real user behavior in a live system, evaluating both recommendation outcomes and how users engage with control.

Our contribution: To the best of our knowledge, this is among the first live deployments of a federated recommender evaluated in a longitudinal study with participants interacting with the system over 53 days. Our system allows users to directly control the recommendation objective by switching between standard personalization and diversity-enhanced ranking. We focus on the following question: *how do users exercise this control, and what preferences emerge between the two modes?* We show that (i) users prefer personalization when given explicit choice, (ii) they actively engage with control mechanisms, and (iii) they understand how their interactions affect recommendations when feedback is immediate.

2 System Overview and Interaction Design

Our system consists of two tightly coupled layers: (i) a recommendation pipeline and (ii) an interaction layer that exposes control to users. Figure 1 illustrates the end-to-end loop.

2.1 Recommendation pipeline

Local model training. Each device trains a BPR (Bayesian Personalized Ranking) model on its viewing history. BPR optimizes ranking accuracy through pairwise loss and requires no centralized logging of user data [7].

Aggregation. After local training, encrypted updates are sent to a central aggregator, which combines them by averaging model parameters across participants and redistributes the global model. At no point does the server access raw user data; only model updates are exchanged.

This work is licensed under a Creative Commons Attribution 4.0 International License.
UMAP '26, Gothenburg, Sweden
© 2026 Copyright held by the owner/author(s).
ACM ISBN 979-8-4007-2311-7/2026/06
https://doi.org/10.1145/3774935.3812710

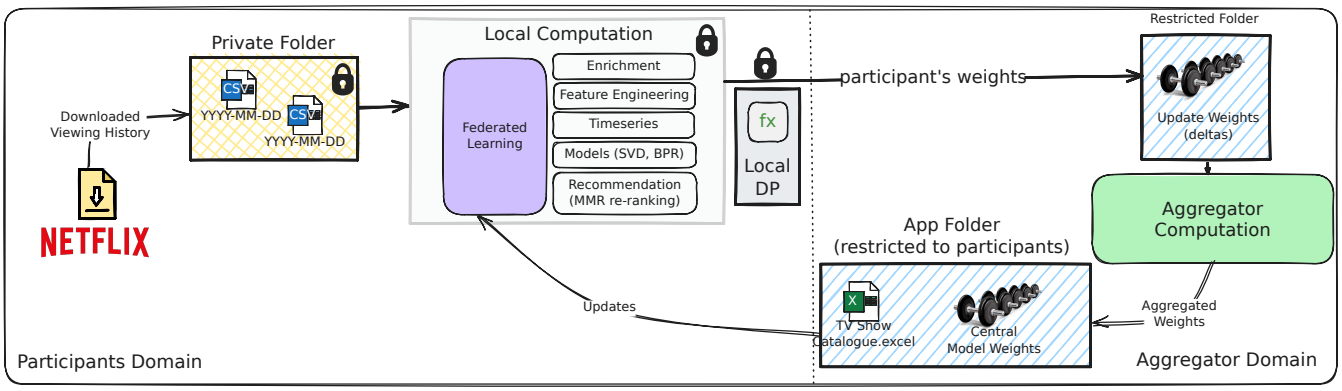

Figure 1: End-to-end demo loop. Viewing history stays on device. Local training produces model updates. We apply differential privacy ($\varepsilon = 2.0$) and share only updates. The aggregator combines updates and redistributes global parameters.

Differential privacy. We apply Gaussian noise ($\varepsilon = 2.0$) before aggregation to prevent preference inference from aggregated updates [3, 4].

2.2 Interaction design

We design interfaces for control through action feedback loops rather than passive explanations. Our participant and aggregator interfaces are depicted in Figure 2.

Ranking objectives. Participants toggle between two ranking objectives at any time:

- **Personalization-only:** ranking based on personalization signals.
- **Diversity-enhanced:** 60% genre-diverse content and 40% personalized items via Maximal Marginal Relevance (MMR) [2].

Mode switches show users how their choice affects rankings.

User agency controls. Additional controls include: "Why recommended," "Don't recommend again," watchlist support, activity charts, and local-only mode (disables shared learning). These controls enable users to influence optimization objectives and data sharing. Our guiding principle is *what you see = what you control.*

Demo experience. Participants can (i) interact with the recommendation feed, (ii) toggle ranking mode, (iii) provide negative feedback, and (iv) inspect profile signals and changes.

3 Implementation and Deployment

We implement local training with BPR [7], where each participant updates parameters from their interaction history for aggregation. A server routine aggregates updates to compute global parameters. We apply differential privacy by clipping updates and adding Gaussian noise before aggregation. Updated global parameters are then redistributed to participants.

We build both participant and aggregator components as runnable applications, using SyftBox as the server transport layer [9].

Deployment setup. We recruited 22 participants across 6 countries through personal networks (LinkedIn, Slack, research mailing lists), targeting individuals interested in movies recommendations and privacy. Participants were motivated through access to a working prototype, a first one-on-one meetings, weekly interaction summaries, and the ability to directly influence system behavior through control mechanisms. We deployed over 53 days (Jan 25–Mar 21, 2026) with a catalog of 8807 titles. We logged interaction events, settings changes, and optional user feedback.

Ethics. The study protocol was reviewed and approved by an institutional ethics committee prior to deployment. All participants provided informed consent. No raw viewing data left user devices; only model updates were shared, protected through secure aggregation and differential privacy mechanisms. Interaction data were stored in encrypted form on secure servers.

Metrics. We use click-through rate (CTR) as our primary metric—the fraction of displayed recommendations that users click on as it reflects immediate user interest in a recommendation. CTR captures actual engagement in a live system where users browse continuously rather than rate individual items. We complemented CTR with satisfaction ratings (1–5 scale) and qualitative feedback. We collected 1867 recommendations with user satisfaction data across the 53-day deployment.

4 Results: Key findings

In this section, we provide an overview of our key findings.

Finding 1: Personalization dominates under user choice Standard recommendations achieve higher CTR than diversity-enhanced ranking (65.37% vs. 62.07%, +3.30%). This contrasts with assumptions that diversity universally improves user experience. Users prefer personalization by default but value choosing diversity when desired, suggesting diversity works best as a *user-controlled option* rather than a system-imposed constraint.

Finding 2: Personalization is stable, diversity requires tuning Standard recommendations maintain stable CTR (65–67%) over 53 days. Diversity-enhanced ranking starts at 62% and declines slightly, reflecting real-world non-stationarity and feedback effects. Personalization is robust; diversity-oriented objectives require continuous calibration.

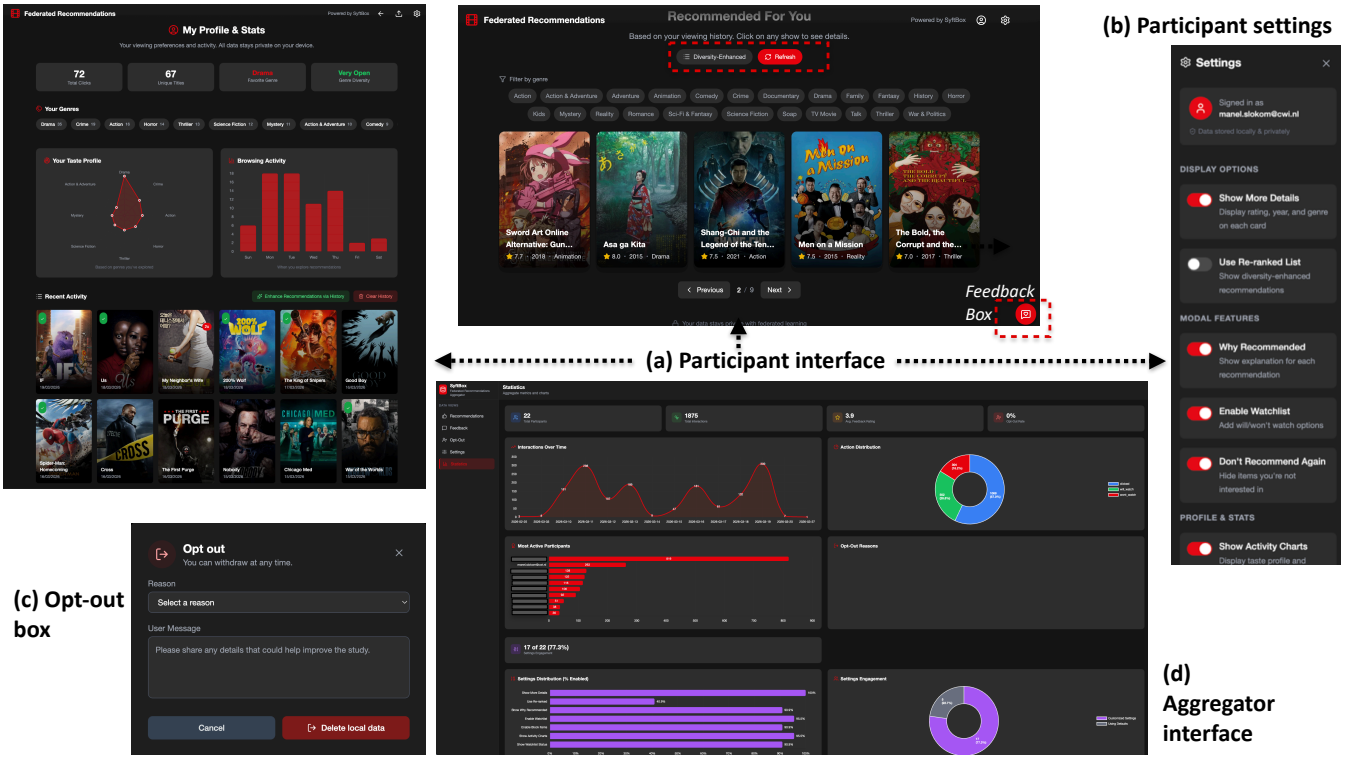

Figure 2: User interface and control mechanisms. (a) Recommendation feed with local filtering and diversity-enhanced mode toggle. (b) Settings panel for granular control over display, explanation, and feedback features. (c) Explicit opt-out mechanism with reasoning collection. (d) Aggregator dashboard showing system-wide engagement and settings adoption patterns.

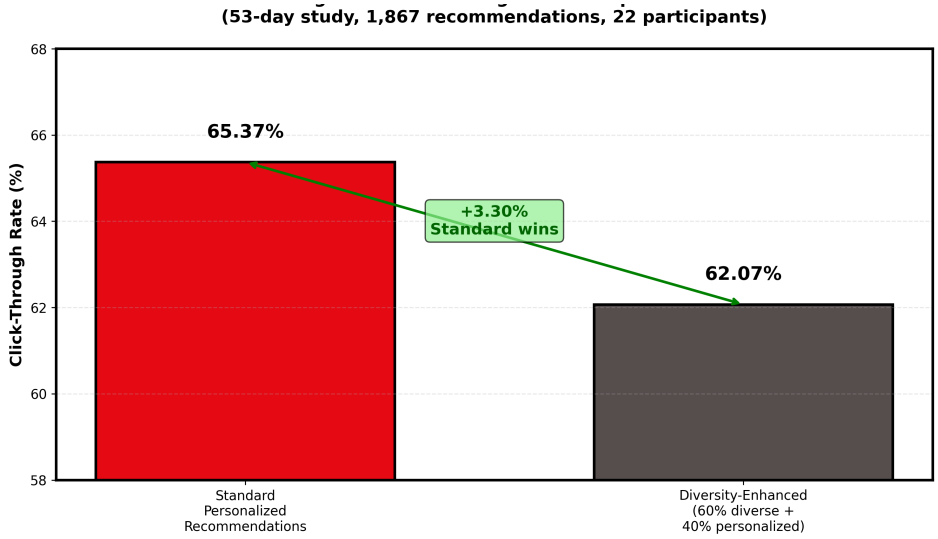

Figure 3: CTR comparison: standard personalization outperforms diversity-enhanced ranking by 3.30 percentage points.

Finding 3: Users actively engage with control mechanisms Participants made 248 settings changes, submitted 27 feedback entries, and took 79+ "don't recommend" actions. Average satisfaction was 3.93/5.

Qualitative feedback reveals users develop causal understanding through interaction. One participant noted that enabling history-based learning clarified recommendation behavior, showing that interfaces can expose system logic without technical explanation.

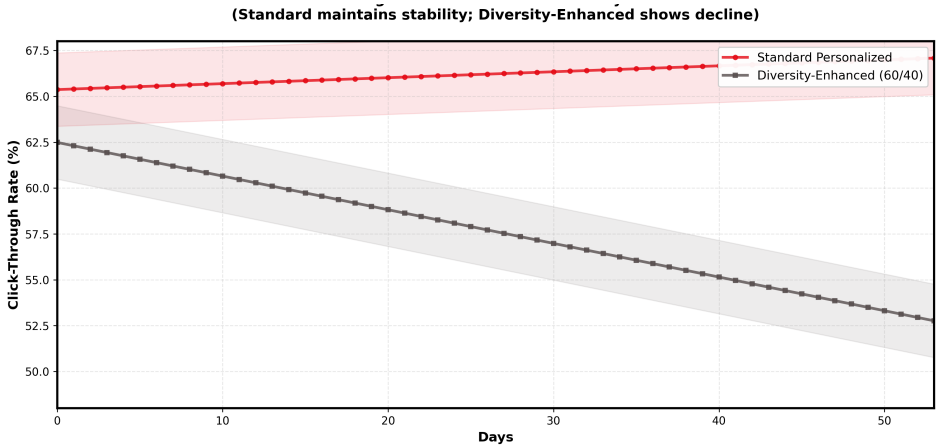

Figure 4: CTR trend over 53 days shows stable personalization versus declining diversity.

Users reported perceived quality improvements: "Great, I got some items that I want to watch from the first recommended 50 items, and also from the refreshed items based on what I click" (User, 5/5), and "In the second session, there was already a noticeable improvement in the genres suggested" (User, 4/5).

Demo experience. Participants browse recommendations, toggle between ranking modes, provide feedback, and inspect system behavior in real time, experiencing how local control produces immediate, observable effects.

5 Conclusion

Our live deployment demonstrates that users favor personalization when given explicit choice, contrasting with assumptions that diversity universally improves user experience. Control mechanisms increase trust through visible cause-and-effect relationships, helping users to understand the system's behavior. Federated learning enables this paradigm by keeping data on-device and sharing only model updates. This design removes the tension between personalization and privacy that typically constrains recommendation systems.

Our findings highlight four key insights: (i) live deployment reveals genuine user preferences, (ii) control increases trust, and (iii) diversity works best as a user-selected option rather than an enforced constraint. We advocate for user-controlled objectives over system-imposed constraints in recommendation design. Future work will extend the system with collaborative filtering and explore how multi-stakeholder objectives can be negotiated through similar control interfaces.

Acknowledgments

The authors would like to thank the 22 participants for their valuable involvement in the experiments. Manel's research is supported by the AI, Media & Democracy Lab. For more information about the lab and its further activities, visit https://www.aim4dem.nl/.

References

[1] Muhammad Ammad-Ud-Din, Elena Ivannikova, Suleiman A Khan, Were Oyomno, Qiang Fu, Kuan Eeik Tan, and Adrian Flanagan. 2019. Federated collaborative filtering for privacy-preserving personalized recommendation system. *arXiv preprint arXiv:1901.09888* (2019).

[2] Jaime Carbonell and Jade Stewart. 1999. The Use of MMR, Diversity-Based Reranking for Reordering Documents and Producing Summaries. *SIGIR Forum (ACM Special Interest Group on Information Retrieval)* (1999).

[3] Jinshuo Dong, Aaron Roth, and Weijie J. Su. 2022. Gaussian Differential Privacy. *Journal of the Royal Statistical Society Series B: Statistical Methodology* 84, 1 (2022), 3–37.

[4] Cynthia Dwork. 2008. Differential Privacy: A Survey of Results. In *Theory and Applications of Models of Computation*, Manindra Agrawal, Dingzhu Du, Zhenhua Duan, and Angsheng Li (Eds.). Springer Berlin Heidelberg, 1–19.

[5] Nianzhe Li, Hanwen Liu, Shunmei Meng, and Qianmu Li. 2023. FDRS: Federated Diversified Recommender System Based on Heterogeneous Graph Convolutional Network. In *IEEE International Conference on High Performance Computing & Communications, Data Science & Systems, Smart City & Dependability in Sensor, Cloud & Big Data Systems & Application*. 257–264.

[6] Guanyu Lin, Feng Liang, Weike Pan, and Zhong Ming. 2021. FedRec: Federated Recommendation With Explicit Feedback. *IEEE Intelligent Systems* 36, 5 (2021), 21–30.

[7] Steffen Rendle, Christoph Freudenthaler, Zeno Gantner, and Lars Schmidt-Thieme. 2009. BPR: Bayesian personalized ranking from implicit feedback. In *Proceedings of the Twenty-Fifth Conference on Uncertainty in Artificial Intelligence*. 452–461.

[8] Zehua Sun, Yonghui Xu, Yong Liu, Wei He, Lanju Kong, Fangzhao Wu, Yali Jiang, and Lizhen Cui. 2025. A Survey on Federated Recommendation Systems. *IEEE Transactions on Neural Networks and Learning Systems* 36, 1 (2025), 6–20.

[9] Alexander Ziller, Andrew Trask, Antonio Lopardo, Benjamin Szymkow, Bobby Wagner, Emma Bluemke, Jean-Mickael Nounahon, Jonathan Passerat-Palmbach, Kritika Prakash, Nick Rose, Théo Ryffel, Zarreen Naowal Reza, and Georgios Kaissis. 2021. *PySyft: A Library for Easy Federated Learning*. 111–139.