INTERNATIONAL JOURNAL OF DATA SCIENCE AND MACHINE LEARNING (ISSN: 2692-5141)
VOL.06 ISSUE 01 (2026)

AMERICAN ACADEMIC PUBLISHER

Research Article

Beyond Global Models: Personalized and Clustered Federated Learning for Distributed Intelligence

Lukas Reinhardt [1]

[1]Department of Biomedical Engineering, University of Zurich, Switzerland

check for updates

Received: 12 December 2025
Revised: 2 January 2025
Accepted: 01 February 2026
Published: 23 March 2026

DOI:
https://doi.org/10.55640/ijdsml-06-01-03

Abstract

The rapid expansion of distributed intelligent systems across edge, cyber–physical, and metaverse-enabled environments has elevated federated learning from a privacy-preserving optimization protocol into a foundational paradigm for large-scale, heterogeneous, and highly personalized artificial intelligence. Classical federated learning, which relies on the aggregation of locally trained models into a single global representation, increasingly fails to capture the reality of non-identically distributed data, divergent device capabilities, and highly contextual user behaviors. In response, the research community has turned toward personalized and cluster-aware federated learning as a means to reconcile privacy, efficiency, and accuracy under extreme heterogeneity. This article develops a comprehensive theoretical and methodological synthesis of this emerging paradigm, situating it within broader traditions of inductive bias learning, multi-task learning, meta-learning, and representation alignment. Anchored in the recent theoretical consolidation of personalized and cluster-aware federated learning provided by Moreno (2026), this study extends existing perspectives by embedding personalization into a unified framework that connects statistical heterogeneity, communication-efficient optimization, contrastive representation learning, and edge-enabled intelligence.

Through an extensive critical analysis of prior work, this article demonstrates that personalization in federated systems is not a single algorithmic choice but a multilayered epistemological commitment to modeling diversity. From early formulations of mixture models and local-global decomposition to contemporary cluster-based, meta-learned, and generative approaches, personalization reflects the need to encode client-specific inductive biases without sacrificing the benefits of collective learning. Building on this foundation, the methodology presented here articulates a text-based, theoretically grounded framework for cluster-aware personalization in federated networks, integrating self-knowledge distillation, optimal transport, contrastive learning, and hierarchical edge coordination. Rather than proposing a new numerical algorithm, the study constructs an interpretive and methodological architecture that allows existing approaches to be understood as special cases of a broader paradigm.

The results section synthesizes evidence from the literature to show that cluster-aware personalization improves robustness, fairness, and generalization in heterogeneous federated environments, particularly in Internet of Things, healthcare, cybersecurity, and metaverse-linked digital twin systems. These improvements are not merely technical but epistemic, enabling models to respect local data semantics while participating in global knowledge production. The discussion further explores the implications of this shift, addressing tensions between privacy and personalization, the risk of overfitting local biases, and the emerging role of representation learning as the connective tissue of personalized federation. By situating personalized federated learning within a larger socio-technical and theoretical landscape, this article argues that the future of distributed intelligence lies not in

universal models but in adaptive, cluster-aware ecosystems that learn with and from diversity.

Keywords: Personalized federated learning, data heterogeneity, cluster-aware learning, edge intelligence, contrastive representation learning, meta-learning, distributed AI

INTRODUCTION

The Federated learning emerged as a response to a fundamental contradiction in modern artificial intelligence: the need to train powerful models on massive, sensitive, and distributed data without centralizing that data in ways that compromise privacy, security, or ownership. Early formulations of federated learning were shaped by the assumption that a single global model could adequately represent the statistical structure of data distributed across many clients, an assumption that reflected both technical convenience and the historical legacy of centralized machine learning (Kairouz et al., 2019). However, as federated systems have expanded into domains such as healthcare, mobile computing, cyber–physical systems, and the metaverse, the heterogeneity of client data has become too pronounced to ignore, giving rise to the field now known as personalized federated learning (Tan et al., 2022; Moreno, 2026).

Data heterogeneity in federated learning is not merely a statistical inconvenience but a reflection of deep socio-technical diversity. Devices operate in different physical contexts, users exhibit distinct behavioral patterns, and sensors capture environment-specific signals that cannot be meaningfully averaged without losing crucial information (Hsu et al., 2019; Huang et al., 2020). In healthcare, for example, patient populations differ in demographics, lifestyle, and disease prevalence, rendering a single diagnostic model suboptimal for many individuals (Wu et al., 2020). In cybersecurity and Internet of Things ecosystems, device types, network conditions, and threat profiles vary widely, necessitating adaptive and personalized defense mechanisms (Ghimire and Rawat, 2022). These realities expose the limitations of classical federated averaging and motivate the search for models that can learn collectively while remaining locally specialized.

Theoretical work on inductive bias learning has long recognized that learning systems must encode assumptions about the structure of the tasks they face (Baxter, 2000). In a federated setting, each client can be understood as facing a related but distinct task, transforming federated learning into a form of distributed multi-task learning (Mills et al., 2021; Dinh et al., 2021). Personalized federated learning thus inherits the epistemological challenge of multi-task learning: how to share information across tasks without erasing their individuality. Moreno (2026) advances this perspective by proposing that personalization and clustering are not ad hoc extensions of federated learning but foundational responses to data heterogeneity, providing theoretical justification for grouping clients into latent clusters that share statistical and semantic properties.

Historical approaches to personalization in federated learning initially focused on simple model decompositions. Mixture models of global and local parameters allowed each client to adapt a shared representation to its own data (Hanzely and Richtarik, 2020; Arivazhagan et al., 2019). While effective in mild heterogeneity, such approaches struggled when client distributions diverged significantly, leading to slow convergence and degraded performance (Haddadpour and Mahdavi, 2019). Meta-learning-based methods then reframed personalization as rapid adaptation from a shared initialization, drawing on the theoretical foundations of model-agnostic meta-learning (Finn et al., 2017; Fallah et al., 2020a). These methods demonstrated improved flexibility but introduced new challenges related to stability, communication cost, and sensitivity to task distribution shifts (Fallah et al., 2020b).

Parallel to these developments, representation learning underwent a revolution driven by contrastive and self-supervised methods, which showed that robust, transferable features could be learned without explicit labels (Chen et al., 2020; Wang and Qi, 2022). The integration of contrastive learning into federated settings has opened new avenues for personalization, as clients can align their representations through shared latent spaces while preserving local semantic nuances (Wu et al., 2022; Liu et al., 2021). Moreno (2026) situates these advances within a broader theoretical framework, arguing that cluster-aware federated learning enables systems to discover and exploit latent similarities among clients, effectively performing unsupervised or semi-supervised task clustering in a privacy-preserving manner.

Despite this progress, the literature remains fragmented across domains, methodologies, and theoretical commitments. Some works emphasize communication efficiency and optimization (Basu et al., 2020; Dai et al., 2019), others focus on statistical modeling and regularization (Dinh et al., 2020; Deng et al., 2020), and still others explore generative, graph-based, or reinforcement learning approaches to personalization (Cao et al., 2022; Zhu et al., 2022). What is missing is a unifying conceptual and methodological account that explains how these diverse strands cohere within a single paradigm of personalized and cluster-aware federated learning. This gap is particularly evident in emerging application domains such as the edge-enabled metaverse, where digital twins, autonomous agents, and immersive environments demand both global coordination and extreme local adaptation (Xu et al., 2022; Han et al., 2022).

This article addresses that gap by developing a comprehensive, theoretically grounded framework for personalized and cluster-aware federated learning under data heterogeneity. Building on Moreno (2026), it integrates insights from meta-learning, optimal transport, contrastive representation learning, and edge intelligence to articulate how federated systems can simultaneously respect diversity and harness collective knowledge. Rather than proposing a single algorithm, the study offers an interpretive synthesis that allows existing approaches to be understood as complementary components of a broader architecture. In doing so, it aims to provide both conceptual clarity and practical guidance for researchers and practitioners navigating the increasingly complex landscape of distributed intelligent systems.

METHODOLOGY

The methodological orientation of this study is fundamentally integrative and theoretical, reflecting the complexity and interdisciplinary nature of personalized and cluster-aware federated learning. Rather than constructing a numerical experiment or proposing a single algorithmic pipeline, the methodology is designed to synthesize, interpret, and systematize a wide range of existing approaches into a coherent conceptual framework. This approach aligns with the epistemological stance articulated by Moreno (2026), who emphasizes that understanding personalization under data heterogeneity requires a unification of theoretical, methodological, and application-driven perspectives.

The first pillar of the methodology is a structured analytical reading of the federated learning literature, focusing specifically on works that address personalization, clustering, or heterogeneity. These include meta-learning approaches that frame personalization as rapid task adaptation (Fallah et al., 2020a; Finn et al., 2017), regularization-based methods that balance global and local objectives (Dinh et al., 2020; Deng et al., 2020), and cluster-based techniques that group clients according to latent similarities (Li et al., 2022; Farnia et al., 2022). Each of these strands is treated not as a competing solution but as a partial articulation of a deeper principle: that federated systems must learn a structured family of models rather than a single universal one.

The second pillar involves conceptual abstraction. By examining how different methods decompose models into shared and private components, or how they align representations across clients, the methodology identifies recurring patterns that transcend specific implementations. For example, hypernetworks and personalization layers (Ha et al., 2017; Arivazhagan et al., 2019) can be interpreted as mechanisms for encoding client-specific inductive biases within a shared parameterization, while optimal transport approaches (Farnia et al., 2022) provide a mathematical lens for aligning heterogeneous distributions in a common latent space. These abstractions allow the formulation of a cluster-aware personalization framework in which clients are dynamically associated with latent groups that define how knowledge is shared and adapted.

The third pillar is domain contextualization. Personalized federated learning is not an abstract mathematical exercise but a response to concrete challenges in edge computing, Internet of Things, cybersecurity, and metaverse-enabled digital twins (Ghimire and Rawat, 2022; Xu et al., 2022; Han et al., 2022). By situating theoretical constructs within these domains, the methodology ensures that the resulting framework remains grounded in practical realities. For instance, energy-aware clustering in vehicular networks (Li et al., 2022) illustrates how personalization interacts with resource constraints, while personalized intrusion detection systems (Huang et al., 2022) highlight the security implications of heterogeneity.

Finally, the methodology incorporates critical evaluation. Each class of approaches is examined in terms of its assumptions, strengths, and limitations, drawing on convergence theory, communication efficiency analyses, and empirical findings reported in the literature (Haddadpour and Mahdavi, 2019; Basu et al., 2020; Kairouz et al., 2019). This critical lens is essential for identifying the trade-offs inherent in personalization, such as the tension between local overfitting and global generalization, or between privacy preservation and representation alignment (Duchi et al., 2014; Moreno, 2026).

Through these four methodological pillars, the study constructs a comprehensive, text-based analytical framework that can accommodate the diversity of personalized federated learning research while revealing its underlying coherence. This methodology does not aim to replace empirical evaluation but to provide the theoretical and conceptual infrastructure necessary for interpreting and guiding such evaluations in future work.

RESULTS

The synthesis of the literature through the methodological framework described above yields several robust and interrelated findings about the nature and impact of personalized and cluster-aware federated learning. Across diverse application domains and methodological traditions, a consistent pattern emerges: models that incorporate personalization and clustering mechanisms outperform purely global federated models in terms of accuracy, stability, and robustness under heterogeneity (Tan et al., 2022; Moreno, 2026). This finding is not merely a matter of improved predictive metrics but reflects a deeper alignment between the learning architecture and the structure of distributed data.

One central result is that cluster-aware personalization effectively reduces negative transfer between clients whose data distributions are fundamentally dissimilar. Classical federated averaging implicitly assumes that all clients contribute to a single task, leading to the contamination of local models by irrelevant or misleading gradients (Hsu et al., 2019). By contrast, clustering-based approaches identify subsets of clients that share statistical or semantic properties, allowing them to exchange information more productively (Li et al., 2022; Farnia et al., 2022). Moreno (2026) provides theoretical grounding for this observation by showing that clustering approximates an optimal partition of the task space, minimizing within-cluster variance while preserving cross-cluster diversity.

Another significant result concerns the role of representation learning in personalization. Methods that integrate contrastive or self-supervised learning into federated settings demonstrate a remarkable ability to align heterogeneous data into shared latent spaces without requiring explicit labels or centralized data (Chen et al., 2020; Wu et al., 2022). This alignment enables clients to benefit from global knowledge while retaining local semantic distinctions, a balance that is particularly valuable in domains such as human activity recognition and medical imaging (Yu et al., 2021; Zhu et al., 2022). The literature consistently shows that such representation-level personalization improves both convergence speed and final model performance, supporting Moreno's (2026) claim that personalization operates most effectively at the level of latent features rather than raw parameters.

A further result is the demonstrated synergy between personalization and edge intelligence. In edge and IoT environments, where devices have limited energy, computation, and communication capacity, cluster-aware personalization allows systems to allocate resources more efficiently by tailoring model complexity and communication patterns to the needs of specific client groups (Wu et al., 2020; Li et al., 2022). This not only improves performance but also extends the operational lifetime of distributed systems, an outcome that is crucial for large-scale deployments in smart homes, autonomous vehicles, and digital twins (Han et al., 2022; Meng et al., 2022).

Finally, the literature reveals that personalized federated learning enhances resilience and security. In cybersecurity applications, personalized models are better able to detect locally specific attack patterns that would be diluted in a global model (Ghimire and Rawat, 2022; Huang et al., 2022). Cluster-aware approaches further improve this resilience by isolating anomalous or adversarial clients into separate groups, reducing the risk of widespread model corruption (Moreno, 2026). These results underscore the broader implication that personalization is not only a performance optimization but a structural safeguard for distributed intelligence.

DISCUSSION

The results synthesized above invite a deep theoretical and critical reflection on the nature of learning in distributed, heterogeneous environments. Personalized and cluster-aware federated learning represents a paradigm shift from the universalist aspirations of classical machine learning toward a pluralistic epistemology that recognizes diversity as a fundamental property of data and tasks. Moreno (2026) articulates this shift by framing personalization as a response to the intrinsic heterogeneity of real-world systems, a perspective that resonates with earlier theories of inductive bias and multi-task learning (Baxter, 2000; Mills et al., 2021).

At a theoretical level, cluster-aware personalization can be understood as a form of latent structure discovery. Rather than imposing a predefined taxonomy of tasks, federated systems infer clusters based on observed data and model behavior, effectively performing an unsupervised decomposition of the global learning problem (Farnia et al., 2022; Li et al., 2022). This process parallels developments in contrastive and representation learning, where models learn to organize data into meaningful latent spaces without explicit supervision (Chen et al., 2020; Wang and Qi, 2022). The convergence of these traditions suggests that personalization is most powerful when it operates at the level of representations, enabling flexible adaptation without sacrificing shared understanding.

However, this paradigm is not without tensions. One persistent concern is the risk of over-personalization, where models become so finely tuned to local data that they lose the ability to generalize or benefit from collective learning (Deng et al., 2020; Dinh et al., 2020). Cluster-aware frameworks mitigate this risk by embedding local models within a broader structure of shared clusters, but the balance between individuality and collectivity remains delicate (Moreno, 2026). This balance is further complicated by privacy constraints, as richer representation sharing can inadvertently leak sensitive information if not carefully designed (Duchi et al., 2014; Kairouz et al., 2019).

Another important dimension is the socio-technical implication of personalization. In domains such as healthcare and smart cities, personalized models can improve fairness and inclusivity by accounting for the needs of diverse populations (Wu et al., 2020; Han et al., 2022). At the same time, clustering mechanisms risk reinforcing existing biases if they align too closely with demographic or socioeconomic divisions, an issue that requires careful ethical and methodological consideration (Moreno, 2026). The literature thus calls for a reflexive approach to personalization, one that integrates technical innovation with normative awareness.

Looking forward, the integration of personalized federated learning with emerging metaverse and digital twin technologies opens new horizons for adaptive, immersive, and context-aware intelligence (Xu et al., 2022; Meng et al., 2022). In such environments, users, devices, and virtual agents interact in real time, generating streams of heterogeneous data that demand both global coordination and hyper-local adaptation. Cluster-aware personalization provides a conceptual and methodological foundation for navigating this complexity, suggesting that future distributed systems will be organized not around monolithic models but around dynamic, evolving constellations of specialized learners.

CONCLUSION

Personalized and cluster-aware federated learning represents a fundamental evolution in the theory and practice of distributed artificial intelligence. By acknowledging and embracing data heterogeneity, it moves beyond the limitations of universal models toward a more nuanced and adaptive understanding of collective learning. Grounded in the theoretical insights of Moreno (2026) and enriched by advances in meta-learning, representation learning, and edge intelligence, this paradigm offers a pathway toward more robust, fair, and context-sensitive AI systems. As distributed technologies continue to permeate every aspect of society, the ability to learn with and from diversity will define the next generation of intelligent systems.

REFERENCES

1. Dinh, C. T., Vu, T. T., Tran, N. H., Dao, M. N., and Zhang, H. Fedu: A unified framework for federated multi-task learning with laplacian regularization. arXiv preprint arXiv:2102.07148, 2021.

2. Wu, Q., Chen, X., Zhou, Z., and Zhang, J. Fedhome: Cloud-edge based personalized federated learning for in-home health monitoring. IEEE Transactions on Mobile Computing, 21(8):2818–2832, 2020.

3. Basu, D., Data, D., Karakus, C., and Diggavi, S. N. Qsparselocal-sgd: Distributed sgd with quantization, sparsification, and local computations. IEEE Journal on Selected Areas in Information Theory, 1(1):217–226, 2020.

4. Moreno, A. Toward personalized and cluster-aware federated learning under data heterogeneity: Theoretical foundations, methodological advances, and emerging

paradigms. International Journal of Data Science and Machine Learning, 6(1), 2026.

5. Chen, T., Kornblith, S., Norouzi, M., and Hinton, G. A simple framework for contrastive learning of visual representations. In International conference on machine learning, pages 1597–1627. PMLR, 2020.

6. Ghimire, B. and Rawat, D. B. Recent advances on federated learning for cybersecurity and cybersecurity for federated learning for internet of things. IEEE Internet of Things Journal, 9(11):8229–8249, 2022.

7. Farnia, F., Reisizadeh, A., Pedarsani, R., and Jadbabaie, A. An optimal transport approach to personalized federated learning. IEEE Journal on Selected Areas in Information Theory, 3(2):162–171, 2022.

8. Li, Y., Qin, X., Chen, H., Han, K., and Zhang, P. Energy-aware edge association for cluster-based personalized federated learning. IEEE Transactions on Vehicular Technology, 71(6):6756–6761, 2022.

9. Tan, A. Z., Yu, H., Cui, L., and Yang, Q. Towards personalized federated learning. IEEE Transactions on Neural Networks and Learning Systems, 2022.

10. Hsu, T. H., Qi, H., and Brown, M. Measuring the effects of non-identical data distribution for federated visual classification. arXiv preprint arXiv:1909.06335, 2019.

11. Xu, M., Ng, W. C., Lim, W. Y. B., Kang, J., Xiong, Z., Niyato, D., Yang, Q., Shen, X. S., and Miao, C. A full dive into realizing the edge-enabled metaverse: Visions, enabling technologies, and challenges. IEEE Communications Surveys and Tutorials, 2022.

12. Han, Y., Niyato, D., Leung, C., Kim, D. I., Zhu, K., Feng, S., Shen, X., and Miao, C. A dynamic hierarchical framework for iot-assisted digital twin synchronization in the metaverse. IEEE Internet of Things Journal, 10(1):268–284, 2022.

13. Arivazhagan, M. G., Aggarwal, V., Singh, A. K., and Choudhary, S. Federated learning with personalization layers. arXiv preprint arXiv:1912.00818, 2019.

14. Baxter, J. A model of inductive bias learning. Journal of Artificial Intelligence Research, 12:149–198, 2000.

15. Fallah, A., Mokhtari, A., and Ozdaglar, A. Personalized federated learning: A meta-learning approach. arXiv preprint arXiv:2002.07948, 2020.

16. Finn, C., Abbeel, P., and Levine, S. Model-agnostic meta-learning for fast adaptation of deep networks. arXiv preprint arXiv:1703.03400, 2017.

17. Haddadpour, F. and Mahdavi, M. On the convergence of local descent methods in federated learning. arXiv preprint arXiv:1910.14425, 2019.

18. Huang, X., Liu, J., Lai, Y., Mao, B., and Lyu, H. Eefed: Personalized federated learning of execution and evaluation dual network for cps intrusion detection. IEEE Transactions on Information Forensics and Security, 18:41–56, 2022.

19. Kairouz, P., McMahan, H. B., Avent, B., Bellet, A., Bennis, M., Bhagoji, A. N., Bonawitz, K., Charles, Z., Cormode, G., and Cummings, R. Advances and open problems in federated learning. arXiv preprint arXiv:1912.04977, 2019.

20. Wu, M., Pan, S., and Zhu, X. Attraction and repulsion: Unsupervised domain adaptive graph contrastive learning network. IEEE Transactions on Emerging Topics in

Computational Intelligence, 6(5):1079–1091, 2022.

21. Zhu, Z., Yu, L., Wu, W., Yu, R., Zhang, D., and Wang, L. Murcl: Multi-instance reinforcement contrastive learning for whole slide image classification. IEEE Transactions on Medical Imaging, 2022.