Bielefeld University

Seminar Title: AI Know You so Well - Personalization and User Modeling in Intelligent Systems

# Privacy & Security in Federated Learning

**Author:**

Syed Zain Ali

Data Science's Student

4373556

**Date:** April 9, 2025

# Contents

# Declaration of Consent (Affidavit)

“Hereby I declare that I wrote this essay independently and that I prepared all graphical representations and tables included independently. I did not use any sources other than those indicated and always marked those parts of the essay, which are taken from other works, as borrowed from a specific clearly-stated source, including tables and illustrations. This essay was personally written by me and not generated by any largescale language model, such as Chat-GPT or similar.”

Syed Zain Ali
Bielefeld, April 9, 2025

Syed Zain Ali

# 1 Abstract

In this essay, the role of privacy and security in Federated Learning (FL) for accurate and reliable user modeling and personalization is described. Privacy and security threats in FL, such as data poisoning, data leakage, and malicious attacks, along with their solutions, are explained in the context of how they affect user profiling, personalization, and data safety. Mainly methodological, framework-based, and survey papers were studied to explore how privacy-preserving techniques impact system and the personalized models. This essay highlights the importance of privacy and security for achieving reliable user profiling and personalization. Developers can adopt strategies based on this research and implement built-in security and privacy methods in devices, which will enhance user safety and performance. Future research should emphasize user transparency, scrutable modeling, and adaptive frameworks with the involvement of users to support secure and scalable personalization [1, 2].

# 2 Introduction

Internet of thing (IoT) is one of the biggest examples where physical devices like smartphones, smartwatches and large-scale industrial machines interact with the environment to get data like global positioning systems (GPS) location information, user's electrocardiogram (ECG) etc., companies require this data to get the assessment of devices and to improve their machine learning (ML) models so usage also [3]. To train an ML model, data is required and typically training data needs to be stored in a device or by an organization to train and run the learning algorithm [4]. In today's modern era one of the biggest threats to all the people is their private data, which is used by many companies and devices, mainly on the internet without any consent. Companies need this data for user modeling and personalization for better experience. This increasing demand for data always raises concerns about data hacking and misuse of data, along with data storage and transmission is challenging with growing users [3]. Data regulation authorities are working on this but still, only the regulating authorities are not enough to handle this scenario, companies who are making these devices and using the data will also need to behave ethically to avoid misuse of data. This increasing usage of technology, growing needs for user profiling, and security aspects of user's personal data raise the need to analyze and review current methods and threats in privacy-preserving systems. User modeling is the process to build and keep a current user model, an adaptive system gathers data from multiple sources, which can involve both implicitly monitoring user interactions and explicitly asking users for direct input[2]. The term personalization can have different meanings in general it is used to deliver and receive information, main objective of personalization is to deliver information that is relevant for specific users in specific time [5]. Personalization is the act of tailoring a service or product to suit an individual user, a concept that underlies many user modeling and recommender systems[6]. To overcome privacy and security issues, we use federated learning (FL), where data holders collaborate during the learning process instead of relying on storing data [4, 7, 8]. However, it is a balancing trade-off as described in [9] that achieving high fairness, accuracy is compromised [10]. The core concept of FL is training a model without storing or transferring the data centrally [10]. FL has several formal definitions in the literature. One such definition describes it as: *"FL is a distributed machine learning process, which allows multiple nodes to work together to train a shared model without exchanging raw data. It offers several key advantages, such as data privacy, secu-*

*rity, efficiency, and scalability, by keeping data local and only exchanging model updates through the communication network."* [11]. In [12] it is said that in recent years FL has been one of the spreading fields of ML because of its security and privacy features that meet user data protection laws [13, 14]. This essay explains privacy and security concerns of FL and advantages of FL over ML, how privacy and security is achieved in FL considering users profiling and personalization, what are the threats in FL, and how to identify and handle these issues. FL systems need data to model users, which can be hacked. This essay provides an overview of privacy and security techniques. Common attacks including poisoning, inference, backdoor attacks, and system downtime. Additionally, several methods to enhance security are presented to protect user profiles by reducing the risk of data breaches and enabling model training directly on users' devices. Without strong precautions against these attacks, systems could not safely learn from individual users. In this way, privacy and security are the foundation that makes user-specific modeling and customization both effective and ethically responsible. The paper is structured as follows: After the introduction, a brief background on Federated Learning (FL) is provided. This is followed by a section that explains how user modeling and personalization relates to FL. The next two sections examine the privacy and security aspects of FL in detail. Subsequently, the challenges in FL are discussed. The paper concludes with a comparison and discussion section, followed by the conclusion that highlights the role of FL in the context of personalization and user modeling, including my own findings and point of view.

# 3 Methodology

This essay is based on a literature review of experimental and research papers. Relevant works were identified using academic databases such as Google Scholar, ResearchGate, and IEEE Xplore, with keywords including "federated learning privacy," "user modeling FL," and "FL personalization." Sources were selected for their relevance to privacy, security, user modeling, and personalization in the context of federated learning. The review is descriptive in nature, comparing approaches, challenges, and solutions, with a focus on how FL supports secure and ethical personalization.

# 4 Overview & Background

It is found in [10] that Federated Learning was originally introduced in 2016 by researchers at Google [15, 16]. As introduces in the previous section, the exponential growth of technology has increased the demand for data and its applications. However, due to modern software, handling of large amounts of data is not a problem anymore but as the data increases, concerns with data also become severe, especially privacy and security. In in [17] it is stated that data leakage isn't a small problem, and now the attention of the public to their data security is growing [18, 19, 20]. As ML algorithms need to be trained on data and keeping data security and privacy while the training model is crucial because of the ML algorithm's working demand and data availability at centralized location. This problem makes artificial intelligence (AI) performance is limited due to user concerns. At this point, FL comes in use and avoids data leakage and hacking thus maintains the privacy and security of data without losing much accuracy. FL offers different frameworks to achieve privacy without compromising on the accuracy of models. Furthermore, FL enables systems to take advantage of ML in smaller domains, when we don't have enough data to train an ML model [12]. User modeling is important for personalized systems because it helps understand things like user preferences, behavior, and habits etc and have been used in different applications [21, 22]. This makes user modeling a central component in FL-driven personalized systems, where privacy must be preserved without losing individual adaptation. But most traditional methods collect all the data in one place, which can lead to privacy issues. Federated Learning solves this by keeping the data on users' devices while still allowing systems to learn and improve [21]. FL components can be divided into three main parts: clients, server, and communication protocol. However, this is not a strict rule, for example in [23] they described that FL-systems are consist of two parts and in [24], owners are referred as clients or agents and main server is referred as aggregator which receive the individual models and computes the overall single model for high performance [25]. Because of client's own data we may have problems with data such as whether the data of each agent is biased or not balanced [24]. In [10] few categories of biased-ness are described that cause the algorithm to unfair, biases caused by missing data, biases from algorithmic goals, biases resulting from the use of proxy variables and biases included in the datasets. Fairness is broad perspective however in FL, it involves either FL framework should give priority to users with more data, or it should also consider the users with less samples [3],[26]. In [10] it is explained in detail that FL set off itself through four main principles:

- Data sets are not shared
- Training is not centralized; however, it is distributed and collaborative
- Local model training is decentralized
- Privacy preservation is inbuilt as sensitive user data is not shared

FL architecture is described in different ways depending on the use cases. For example, in [10], [27] two main types of FL architectures are defined as Horizontal and vertical, both are different in structure and deal differently with user data and model learning is also different in both. On the other hand, in [28], [3] FL is classified in two different structures: centralized FL (CFL) and decentralized FL (DFL). Furthermore, a different iterative approach classifies FL accomplishment into three steps: Model selection, Local model training, and aggregation of local models, it is a continuous repetitive learning process and repeats model training and aggregation to keep the model updated [12].

# 5 Challenges in FL

As explained in the literature, FL can be implemented using two primary architectures: CFL and DFL. In CFL, the server receives model parameters from all agents, assembles them, and then reassign the collected model [3]. Model parameters are collected and redistributed by one or more primary elements, Data holders in FL exchange these parameters natively, which may not be sufficient enough to ensure data privacy [4]. Although FL has the solution of sensitive data privacy in ML. However, sharing the parameters and multiple repetitions and communication introduces new security risks and challenges of hacking in FL [12]. According to [21] FL still faces several practical challenges and FL still have privacy leakage risks, and it is noted that privacy and accuracy cannot be achieved simultaneously. FL is still new, emerging and underexplored especially in the aspect of privacy and security, It is under advancement and and continues to evolve through different methods and practical implementations [12]. FL is mainly decentralized; however, despite this decentralization it still needs a main server to control clients associated with FL space [12]. In addition to this, decentralized approach still does not cover all the concerns, there are some more aspects that makes FL challenging.

In DFL, clients need to make a communication protocol between each other and need to share their local knowledge also to perform the system accurately [3]. One more concern addressed in [12] is that due to the distributed nature of FL handling a large number of clients may in millions across different regions are challenging as compared to ML models which may rely on a few nodes [10]. Moreover, heterogeneity is also one of the major challenges in FL. In CFL, clients and servers have identical but fundamentally distinct objectives, selfish behavior of network can cause the game dynamics among clients due to heterogeneity, which may further distributed in 3 categories: intergroup heterogeneity, intragroup heterogeneity and system heterogeneity [29]. Models in the FL system can be biased because of malignant attacks from vicious clients, poisoning attacks are one the most appropriate example of these attacks [10]. These challenges directly affect the effectiveness of user modeling and personalization in FL systems. For instance, heterogeneity and communication delays can reduce the accuracy and fairness of personalized models, while security threats can compromise sensitive user behavior data that models rely on.

# 6 User Modeling and Federated Learning

User modeling requires data to analyze user preferences or behavior. Currently, the majority of user modeling techniques rely on centralized data collection, which is problematic for user privacy, data can be hacked or leaked easily. This limits practical implementation and poses challenges for real-world deployment [21]. To address these limitations, the authors in [21] proposed federated user modeling, which aims to perform user modeling across decentralized and non-uniform clients using federated learning techniques. This approach enhances user profiling while considering ethical aspects, as well as the privacy and security of user data. Furthermore, the same authors proposed a novel framework, Hierarchical Personalized Federated Learning (HPFL), specifically designed for decentralized user modeling.

Similarly, the researchers in [30] described user profiling techniques within federated learning, where each user relies exclusively on their local dataset. In this approach, user profiles are created using only local data or global statistics collected through a secure aggregation strategy. Additionally, the authors of [31] proposed a method called Federated User Representation Learning (FURL), which they describe as a resource-efficient, simple, scalable, and privacy-preserving method that offers advantages over conventional centralized training models.

Federated user modeling is expected to enable secure collaboration across multiple clients while maintaining data confidentiality [21]. Yang et al. (2019) explain that Federated Learning supports the development of detailed user profiles across diverse domains such as smart home environments and autonomous vehicles. By enabling on-device data processing, these systems continuously refine user models in real time while preserving user privacy [32]. In this paper, the focus is on examining how the privacy and security mechanisms of FL contribute to more ethical and secure user modeling and personalization.

# 7 Personalization and Federated Learning

Authors in [33] explained that in traditional centralized machine learning (CML), user personalization is achieved by collecting and aggregating data from all users onto a central server. However, this approach raises serious privacy risks, especially when the data is sensitive or governed by strict data protection laws. In contrast, Federated Learning (FL) is a privacy-preserving alternative, allowing model training to occur across multiple user devices without moving raw data from the local environment [33]. A major challenge in FL is that user data is typically non-independent and identically distributed (non-IID), meaning it varies significantly between users. This non-IID nature makes training a single global model ineffective, as it cannot accurately capture individual preferences or behavior [33]. Personalized Federated Learning (PFL) enables tailoring models to individual users, which supports user modeling while ensuring that user data remains private on their local devices [33]. These concerns raise privacy and security aspects of FL methods. Authors in [34] One efficient personalization method is partial model personalization. They proposed that splitting the model into shared parameters and user-specific parameters, allowing each client to train only part of the model locally. Each client keeps their personal parameters locally and only shares the updated global parameters with the server [34]. Additionally, personalization may create fairness concerns, especially when users have unequal data availability or computing resources [34]. However, user-specific updates may still leak private patterns if the shared components are not protected with proper security techniques [33]. Proposed methods discussed in [24] like Fair and Private Federated Learning (FPFL) attempt to balance personalization with fairness and privacy using local differential privacy (LDP) and fairness objectives. Personalization in FL is essential for adapting models to user needs, but it must be tightly integrated with privacy and fairness measures to avoid compromising user security or ethical standards [33, 34]. These issues can be fixed by analyzing the privacy and security aspects in FL.

It is stated in [35] that personalization in FL is important due to poor convergence on heterogeneous data and the limitations of using a single global model for all users.

# 8 Privacy in Federated Learning

FL refers to constructing and combining user models while ensuring that personal data remains separate, thereby maintaining data security[36, 21]. As I have explained in FL, we may have data leakage due to malicious attacks. To avoid these attacks different methods are used like cryptographic solutions or differential privacy [24]. Privacy itself is a broad term and takes multiple factors into account. In [35] it is concluded that its not possible to have a uniform solution for privacy that meets both user preferences and legal conditions, so we have to customize privacy considering user profile. Researchers [24] proposed novel based framework called fair and private federated learning (FPFL) to join both fairness and privacy, their goal is to maintain the privacy of data and sensitive characteristics along with fairness. They have described this model training in two main steps, in the first step every client on its own dataset trains the model privately to ensure unbiased and accurate predictions, and in the second step, to imitate the fair predictions from the first model clients train a differentially private model, and in FPFL only the model trained second step is sent to the central server [24]. The effectiveness of the model is achieved by considering two important aspects, fairness, and privacy, fairness is achieved by considering demographic party (DemP) and equalized odds (EO), and privacy is quantified with the idea of local differential privacy (LDP) [24]. Experimentally they worked on three datasets, Adult, Dutch, and Bank, where the sensitive attribute for the first two data sets is gender and for Bank is age, they used two fully connected neural networks for each client with two hidden layers, and for DemP they considered 5 agents and also estimated EO, to examine the result they have explained two baselines, in first baseline clients train the model without considering fairness and to achieve maximum accuracy, in second baseline every client trains the model with taking into account the accuracy and fairness with DemP or EO loss [24]. Differential privacy (DP) is a data analysis approach that is designed to ensure that modifying a single entry in the database does not significantly impact the overall results of the analysis [10], [37]. In [38] author proposes a FL framework by using LDP, blockchain technology, and zero-knowledge proof to overcome the challenges of privacy [10]. DP is a strict mathematical framework that sets a mathematical limit on how much a single individual's data can affect the results, making it harder for an attacker to determine whether someone is in the dataset [4]. Researchers in [4] proposed a novel federated learning system and claims guaranteed privacy for different trust cases with better accuracy as compared to other available approaches, the idea is that data stays on the participants, and using secure multiparty computations (SMC) privacy is assured. In the explanation [4], they described that they have considered two types of inference attacks, insider attacks that are caused by all the participants and outsider attacks that are caused by intruders. Usually, SMC protocols allow n participants to get their output of a function over n inputs and avoid sharing any other information [4]. In the proposed method [4], they combine SMC and DP to develop a mechanism that assures privacy without compromising on accuracy, in their method they have defended an extra input which consists of three parameters: fm that states the training algorithm, $\epsilon$ is a parameter to guaranty the privacy and t indicates a minimum number of honest participants. A fundamental element in this approach is the capability of reducing noise by using SMC while incorporating a flexible trust parameter [4]. Furthermore, for a customized model, they introduce a fully private federated learning system that combines SMC and DP to generate highly accurate models, in addition to this, they also claimed that this is the first paper to show the usage of these combined techniques to get high accuracy with given privacy as compared to other available methods [4]. These privacy measures play a vital role in personalized systems by securing user profiles against misuse, ensuring both data protection and the ability to refine accurate user models. FL may suffer different kinds of attack and in [12] these threats are categorized into following:

- Membership inference attacks: Intruders get information of other users' training data by a global model, In these cases, the attacker guesses details about the training data by making assumptions and training the predictive model to reproduce the original data[12].

- Unintentional data leakage and reconstruction through inference: This is a situation where at the main server unintentional information is leaked by agents' updates [12].

- GANs-based inference attacks: It investigates privacy risks in federated learning, not just from malicious servers but also from dishonest clients, Some clients might share old data just to get the global model and once they have it, they could try to figure out other users' private information [12]. It's hard to catch this behavior because there's so little known about the clients and how trustworthy they are [12].

Furthermore, they've also explained some techniques to identify and prevent attacks which are:

- Secure multi-party computation: It is guarded with cryptographic methods and has been used to secure updates from agents in the FL environment, due to handling of parameters only it increases computing efficiency as compared to traditional SMC [12].

- Differential privacy: In DP noise is added to existing parameters and data loss is comparatively low as compared to user privacy protection, it is brought in to add noises to the parameters uploaded by users[12].

- VerifyNet[39]: It's a secure and trustworthy federated learning framework [39, 12]. By using a double-masking technique, it makes it much harder for attackers to access or guess the training data [12].

- Adversarial training

It is found in [12] that some attacks are more threatening for FL than ML and one of these attacks are GANs based attacks that are more suitable for FL and not so common in distributed machine learning methods which makes it specific for FL. Privacy preserving is not about only considering the vicious attacks prevention but we also need to consider the trade-off and cost to avoid these attacks. The cost here represents an additional burden or result arising from the adopted improvement strategy. SMC and DP enhance the privacy safeguarding ability of FL, but this comes at a greater cost in conjunction with accuracy and efficiency [12]. If the FL model ensures stronger privacy preservation, it sacrifices accuracy and needs more time to reach convergence, Conversely, if the model aims to maintain a specific level of accuracy it must evaluate whether the degree of privacy protection is sufficient or not [12]. In another practical study [40] linked with the cost analysis of FL, researchers simulated experiments using Reddit datasets to evaluate the accuracy of a global model with Differential Privacy. Their results showed that in datasets with similar vocabulary sizes, the accuracy of DP-FL and non-DPFL models remained nearly the same, as the DP technique is based on adding noise to preserve privacy but on the other hand this noise can affect the accuracy of the model, and have the effect of overall convergence [12]. In this essay, only a few privacy aspects are explained as FL is still under development and it needs more research.

# 9 Security in FL

Security and privacy are two terms that seems to be identical, but they have different meanings. Security is more related to securing data and connections; it mainly focuses on protecting data and connections to ensure safe communication. Security is defined as, "Security is an older concept that deals with several issues such as access control for data, encryption, and secure connectivity and securing data from hacking and alteration by cyber criminals [41, 42], [43]". It is explained [43] that security and privacy are different but still, there are some integrations, for security: confidentiality, integrity, availability, authentication, trust, policy, preventing threats, and hacking attacks on systems or data are important aspects. Security in FL isn't just about data—it also affects personalization. If an intruder tampers with the data, the system may build incorrect user models and deliver poor or biased results, for example wrong product recommendation. In the FL environment there are some vulnerabilities, a vulnerability can be described as a flaw security gap within a system that provides an opportunity for a hacker/attacker to obtain unauthorized entry [12], [44]. Researchers [12] classified the vulnerabilities in five different origins:

- Communication Protocol: As FL is a repetitive learning process and needs a greater number of epochs to train a model, hence if the communication channel is weak, it can cause vulnerability.

- Client Data Manipulations: Large-scale FL systems are more vulnerable to attacks due to the numerous clients, which can create opportunities for attackers to exploit data and model parameters.

- Compromised Central Server: The main server must be secure because it handles the aggregation, distribution, and updating of models to all agents. Any vulnerabilities in the server could lead to security risks.

- Weak Aggregation Algorithm: Collection algorithms have a main role in FL, it should be intelligent enough to get irregularity with agents and should be capable of dropping doubtful agents. A few proposed algorithms that work better for FL are: FedAvg, SMC-Avg, FedProx, FedMA, Scaffold, tensor Factorization and personalized-based algorithms.

- Implementors of FL environment: The implementation team itself can pose a security risk, either intentionally or unintentionally, due to neglecting security measures, concealing facts, or lacking awareness or understanding of the sensitivity involved.

Like vulnerabilities, an FL system also needs to handle certain security attacks. The security challenges discussed here need to be addressed to ensure that FL-driven personalization remains both accurate and privacy-preserving. Researchers in [12] explained that these security attacks are classified into the following categories:

- Poisoning: Poisoning attacks has maximum chances to occur in FL model [45, 46], because each agent has access to training data so there are maximum chances that this data will be altered. As it occurs during training so it can influence the overall model performance [12]. Furthermore, poisoning can be sub-classified into following categories [12]:
  - Data Poisoning
  - Model Poisoning
  - Data Modification

- Backdoor Attacks: Backdoor attacks are difficult to identify, they adds a harmful function to an existing model without affecting its accuracy on the main task. Detecting such attacks is challenging and time-consuming since they do not immediately affect the model's performance on its primary task [12].

- System disruption IT downtime: Highly secure applications often face downtime due to expected or unexpected server activities. In FL, this risk is lower since each client maintains a local-global model, enabling training to continue after an outage. However, downtime can still be a strategic attack to extract sensitive data [12].

- Inference: Inference attacks are just as serious as poisoning attacks since they can easily come from either participants or a compromised central server in the FL process.

- GANs

- Communication Bottlenecks

- Malicious Server

These listed attacks are the most common, but other types of threats also exist. However, there are some protective measures available to counter these attacks. It is explained [12] that some defense strategies are used to protect against known threats and minimize risks, these are divided into two types:

I. **Proactive defense:** It focuses on identifying potential risks in advance and taking preventive measures.

II. **Reactive defense:** It activates after an attack is detected, implementing solutions to address the issues within the system.

Several techniques are outlined in [12] to reduce risks, including:

- **Sniper:** Authors in [47] proposed the Sniper approach, which is capable of distinguishing legitimate users and greatly lowering the effectiveness of poisoning attacks, even in the presence of multiple attackers.

- **Anomaly Detection:** An efficient anomaly detection system requires a baseline of normal behavior or events to recognize attacks as deviations from this profile [12]. In [48, 49], the authors suggest using Autoencoders for anomaly detection, which helps identify harmful updates to local models.

- **Knowledge Distillation:** Knowledge distillation is a model compression method where a trained neural network transfers its knowledge to a smaller model, reducing training costs. In [50], the authors proposed a federated distillation framework that enables personalized models and uses translators to share knowledge with clients [48].

- **Trusted Execution Environment**

- **Federated Multitask Learning**

- **Moving Target Defense**

- **Data Sanitization**

- **Pruning**

# 10 Comparison & Discussion

The methods introduced in current findings show that privacy can be improved by combining different methods, as authors in [4] combined SMC and DP to gain more privacy and authors in [24] combined fairness and proposed a new method. Combining fairness with privacy improves user accurate profiling and helps to build more real personalization. Overall, they both showed that privacy can be improved without sacrificing accuracy but on the other hand complexity of the model increases which may affect the model on a large scale. This makes personalization complicated and we need to make each personalized model specifically for that user only. According to authors of [35] we need to adjust privacy according to user profile .Researchers in [12] provide a comprehensive analysis of privacy and security as compared to other research papers which is enough to understand FL's main concept and challenges. The methods proposed in [12] are effective to the extent, but still most of the things lead to some tradeoffs that must be consider and decide according to model requirements and needs.

As FL is still almost new, so we need more surveys and proposals to explore FL aspects in detail. Only a few of the available overviews and background of FL are discussed, current challenges, threats in privacy & security, and techniques to minimize the threats and enhance privacy. Moreover, detailed analysis on the basis of different architectures of FL and their impact on user profiling and personalization is not explained here, like horizontal and vertical architecture, rather than exploring all the methods, some methods and issues are described here, which make this research limited. Furthermore, the current discussed methods and suggested techniques are not implemented on different scales, we can know about these methods by implementing these methods on different data sets and conclude some results based on comparisons.

## 11 Conclusion

In conclusion, the essay is on the basis of a few papers only which makes it limited to spotlight on all the security and privacy threats and their solutions which can enhance user profiling and personalization. According to authors [12] there are some areas in FL that can open research in the future and some of these are: Trusted traceability, client selection and training plan in FL, building FL privacy protection enhanced frameworks in practice, and optimization techniques for different ML algorithms. FL also opens research regarding ethics and user consents. Only a few aspects of privacy and security are explained here, as most of the content in this essay is based on only five research papers, which makes it limited. As the awareness of users towards privacy and data protection is growing, this makes developers to develop new techniques and solutions for their protection. This opens new research gates for both users and companies. Researchers can adapt these methodologies for better efficiency based on these analyses, which will enhance model accuracy, improve data privacy, and increase the overall effectiveness and scalability of Federated Learning systems and better user profiling. Moreover in future user modeling and personalization need more privacy and security aspects. According to [2], future user models should let people see and control how and what the system understands them. Moreover, in [1] authors explained few challenges of user modeling and personalization, one of them is about lack of repeatability in personalization results. For example due to heterogeneity in FL systems, models may perform well in one context but fail to generalize across different user groups or domains. On the basis of current research and environment it can be concluded that, cooperation between researchers, manufacturers, data protection authorities, and also considering users will be helpful in implementing privacy-preserving techniques. This will further improve user trust, market growth, better personalization and user profiling, and data protection.

# 12 Abbreviations

| | |
|---|---|
| **I** | IoT – Internet of Things |
| **II** | GPS – Global Positioning System |
| **III** | ECG – Electrocardiogram |
| **IV** | ML – Machine Learning |
| **V** | FL – Federated Learning |
| **VI** | AI – Artificial Intelligence |
| **VII** | CFL – Centralized Federated Learning |
| **VIII** | DFL – Decentralized Federated Learning |
| **IX** | FPFL – Fair and Private Federated Learning |
| **X** | DemP – Demographic Party |
| **XI** | EO – Equalized Odds |
| **XII** | LDP – Local Differential Privacy |
| **XIII** | DP – Differential Privacy |
| **XIV** | SMC – Secure Multiparty Computations |

# References

[1] Paul De Bra. "Challenges in User Modeling and Personalization". In: *IEEE Intelligent Systems* 32.5 (2017), pp. 76–80. DOI: 10.1109/MIS.2017.3711638.

[2] Peter Brusilovsky and Eva Millán. "User Models for Adaptive Hypermedia and Adaptive Educational Systems". In: *The Adaptive Web*. Ed. by Peter Brusilovsky, Alfred Kobsa, and Wolfgang Nejdl. Vol. 4321. Lecture Notes in Computer Science. Springer Berlin Heidelberg, 2007, pp. 3–53. DOI: 10.1007/978-3-540-72079-9_1. URL: https://doi.org/10.1007/978-3-540-72079-9_1.

[3] Liangqi Yuan, Ziran Wang, and Christopher G. Brinton. "Digital Ethics in Federated Learning". In: *IEEE Internet Computing* 28.5 (2024), pp. 66–74. DOI: 10.1109/MIC.2024.3370408. URL: https://doi.org/10.1109/MIC.2024.3370408.

[4] Stacey Truex et al. "A Hybrid Approach to Privacy-Preserving Federated Learning". In: *Proceedings of the 12th ACM Workshop on Artificial Intelligence and Security (AISec)*. 2019, pp. 1–11. DOI: 10.1145/3338501.3357370. URL: https://doi.org/10.1145/3338501.3357370.

[5] Won Kim. "Personalization: Definition, Status, and Challenges Ahead". In: *Journal of Object Technology* 1.1 (2002). Accessed: 2025-04-02, pp. 29–40. URL: http://www.jot.fm/issues/issue_2002_05/column3/.

[6] Gediminas Adomavicius and Alexander Tuzhilin. "Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions". In: *IEEE Transactions on Knowledge and Data Engineering* 17.6 (2005), pp. 734–749. DOI: 10.1109/TKDE.2005.99.

[7] Keith Bonawitz et al. "Practical Secure Aggregation for Privacy-Preserving Machine Learning". In: *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security*. ACM, 2017, pp. 1175–1191.

[8] Reza Shokri and Vitaly Shmatikov. "Privacy-Preserving Deep Learning". In: *Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security*. ACM, 2015, pp. 1310–1321.

[9] Jon Kleinberg, Sendhil Mullainathan, and Manish Raghavan. "Inherent Trade-Offs in the Fair Determination of Risk Scores". In: *8th Innovations in Theoretical Computer Science Conference (ITCS 2017)*. Ed. by Christos H. Papadimitriou. Vol. 67. Leibniz International Proceedings in Informatics (LIPIcs). Dagstuhl, Germany: Schloss Dagstuhl–Leibniz-Zentrum für Informatik, 2017, 43:1–43:23. ISBN: 978-3-95977-029-3. DOI: 10.4230/LIPIcs.ITCS.2017.43. URL: https://drops.dagstuhl.de/entities/document/10.4230/LIPIcs.ITCS.2017.43.

[10] Sean Vucinich and Qiang Zhu. "The Current State and Challenges of Fairness in Federated Learning". In: *IEEE Access* 11 (2023), pp. 80903–80914. DOI: 10.1109/ACCESS.2023.3295412.

[11] B. Yurdem et al. "Federated learning: Overview, strategies, applications, tools and future directions". In: *Heliyon* 10.19 (Sept. 2024), e38137. DOI: 10.1016/j.heliyon.2024.e38137.

[12] Viraaji Mothukuri et al. “A survey on security and privacy of federated learning”. In: *Future Generation Computer Systems* 115 (2021), pp. 619–640. ISSN: 0167-739X. DOI: https://doi.org/10.1016/j.future.2020.10.007. URL: https://www.sciencedirect.com/science/article/pii/S0167739X20329848.

[13] White House. “Consumer Data Privacy in a Networked World: A Framework for Protecting Privacy and Promoting Innovation in the Global Digital Economy”. In: *J. Priv. Confident.* (2013).

[14] European Union. *General Data Protection Regulation.* Web. URL: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679. 2018.

[15] Jakub Konečný et al. *Federated Learning: Strategies for Improving Communication Efficiency.* arXiv preprint. Not peer-reviewed. 2016. arXiv: 1610.05492 [cs.LG]. URL: https://arxiv.org/abs/1610.05492.

[16] Jakub Konečný et al. *Federated Optimization: Distributed Machine Learning for On-Device Intelligence.* arXiv preprint. Not peer-reviewed. 2016. arXiv: 1610.02527 [cs.LG]. URL: https://arxiv.org/abs/1610.02527.

[17] Chen Zhang et al. “A survey on federated learning”. In: *Knowledge-Based Systems* 216 (2021), p. 106775. ISSN: 0950-7051. DOI: https://doi.org/10.1016/j.knosys.2021.106775. URL: https://www.sciencedirect.com/science/article/pii/S0950705121000381.

[18] Cheng Zhang et al. “Understanding on toughening mechanism of bioinspired bulk metallic glassy composites by thermal spray additive manufacturing”. In: *Scripta Materialia* 177 (2020), pp. 112–117. ISSN: 1359-6462. DOI: https://doi.org/10.1016/j.scriptamat.2019.10.017. URL: https://www.sciencedirect.com/science/article/pii/S1359646219306001.

[19] Maoguo Gong, Jialun Feng, and Yu Xie. “Privacy-enhanced multi-party deep learning”. In: *Neural Networks* 121 (2020), pp. 484–496. ISSN: 0893-6080. DOI: https://doi.org/10.1016/j.neunet.2019.10.001. URL: https://www.sciencedirect.com/science/article/pii/S0893608019303235.

[20] Yu Xie et al. “Secure collaborative few-shot learning”. In: *Knowledge-Based Systems* 203 (2020), p. 106157. ISSN: 0950-7051. DOI: https://doi.org/10.1016/j.knosys.2020.106157. URL: https://www.sciencedirect.com/science/article/pii/S0950705120304019.

[21] Jiahao Wu et al. “Hierarchical Personalized Federated Learning for User Modeling”. In: *Proceedings of the Web Conference 2021.* Association for Computing Machinery, 2021, pp. 957–968. DOI: 10.1145/3442381.3449926. URL: https://doi.org/10.1145/3442381.3449926.

[22] Ingrid Zukerman and David W. Albrecht. “Predictive Statistical Models for User Modeling”. In: *User Modeling and User-Adapted Interaction* 11.1-2 (2001), pp. 5–18. DOI: 10.1023/A:1011187500863. URL: https://doi.org/10.1023/A:1011187500863.

[23] Qiang Yang et al. “Federated Machine Learning: Concept and Applications”. In: *ACM Trans. Intell. Syst. Technol.* 10.2 (Jan. 2019). ISSN: 2157-6904. DOI: 10.1145/3298981. URL: https://doi.org/10.1145/3298981.

[24] Manisha Padala, Sankarshan Damle, and Sujit Gujar. *Federated Learning Meets Fairness and Differential Privacy.* 2021. arXiv: 2108.09932 [cs.LG]. URL: https://arxiv.org/abs/2108.09932.

[25] O.A. Wahab et al. "Federated Machine Learning: Survey, Multi-Level Classification, Desirable Criteria and Future Directions in Communication and Networking Systems". In: *IEEE Communications Surveys Tutorials* 23.2 (2021), pp. 1342–1397.

[26] Tian Li et al. "Fair Resource Allocation in Federated Learning". In: *Proceedings of the International Conference on Learning Representations (ICLR).* 2020. URL: https://openreview.net/forum?id=ByexElSYDr.

[27] Q. Yang et al. "Federated Machine Learning: Concept and Applications". In: *ACM Transactions on Intelligent Systems and Technology* 10.2 (2019), pp. 1–19.

[28] Vishnu Pandi Chellapandi et al. "A Survey of Federated Learning for Connected and Automated Vehicles". In: *2023 IEEE 26th International Conference on Intelligent Transportation Systems (ITSC).* 2023, pp. 1–8. DOI: 10.1109/ITSC57777.2023.10421974. URL: https://doi.org/10.1109/ITSC57777.2023.10421974.

[29] Liangqi Yuan et al. "Decentralized Federated Learning: A Survey and Perspective". In: *IEEE Internet of Things Journal* 11.21 (2024), pp. 34617–34638. DOI: 10.1109/JIOT.2024.3407584. URL: https://doi.org/10.1109/JIOT.2024.3407584.

[30] Tiago Brandão. "Prediction of Privacy Preferences with User Profiles: A Federated Learning Approach". MA thesis. University of Porto, 2021. URL: https://repositorio-aberto.up.pt/bitstream/10216/139326/2/527563.pdf.

[31] Duc Bui et al. *Federated User Representation Learning.* arXiv preprint. Not peer-reviewed. 2019. arXiv: 1909.12535 [cs.LG]. URL: https://arxiv.org/abs/1909.12535.

[32] Sandeep A. Awachar. "Federated Learning: Enhancing Privacy and Efficiency in Decentralized Machine Learning Systems". In: *Educational Administration: Theory and Practice* 30.1 (2024), pp. 3842–3852. DOI: 10.53555/kuey.v30i1.7593. URL: https://kuey.net/index.php/kuey/article/download/7593/5677/14786.

[33] Alysa Ziying Tan et al. "Towards Personalized Federated Learning". In: *IEEE Transactions on Neural Networks and Learning Systems* 34.12 (2023), pp. 9587–9606.

[34] Krishna Pillutla et al. "Federated Learning with Partial Model Personalization". In: *International Conference on Machine Learning.* PMLR. 2022, pp. 17715–17752.

[35] Alfred Kobsa. "Tailoring Privacy to Users' Needs 1". In: *User Modeling 2001.* Ed. by Mathias Bauer, Piotr J. Gmytrasiewicz, and Julita Vassileva. Berlin, Heidelberg: Springer Berlin Heidelberg, 2001, pp. 301–313. ISBN: 978-3-540-44566-1.

[36] H. Brendan McMahan et al. "Communication-Efficient Learning of Deep Networks from Decentralized Data". In: *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics (AISTATS).* Vol. 54. Proceedings of Machine Learning Research. PMLR, 2017, pp. 1273–1282. URL: https://proceedings.mlr.press/v54/mcmahan17a.html.

[37] Cynthia Dwork. "Differential Privacy: A Survey of Results". In: *Theory and Applications of Models of Computation.* Xi'an, China: Springer, 2008, pp. 1–19.

[38] T. Rückel, J. Sedlmeir, and P. Hofmann. "Fairness, Integrity, and Privacy in a Scalable Blockchain-Based Federated Learning System". In: *Computers Networks* 202 (2022), Art. no. 108621.

[39] G. Xu et al. "VerifyNet: Secure and Verifiable Federated Learning". In: *IEEE Transactions on Information Forensics and Security* 15 (2020), pp. 911–926. DOI: 10.1109/TIFS.2019.2929409. URL: http://dx.doi.org/10.1109/TIFS.2019.2929409.

[40] E. Bagdasaryan, O. Poursaeed, and V. Shmatikov. "Differential Privacy Has Disparate Impact on Model Accuracy". In: *Advances in Neural Information Processing Systems*. Ed. by H. Wallach et al. Vol. 32. Curran Associates, Inc., 2019, pp. 15479–15488.

[41] S. Sicari et al. "Security, privacy and trust in Internet of Things: The road ahead". In: *Computer Networks* 76 (2015), pp. 146–164. ISSN: 1389-1286. DOI: https://doi.org/10.1016/j.comnet.2014.11.008. URL: https://www.sciencedirect.com/science/article/pii/S1389128614003971.

[42] Hassan Takabi, James B.D. Joshi, and Gail-Joon Ahn. "Security and Privacy Challenges in Cloud Computing Environments". In: *IEEE Security Privacy* 8.6 (2010), pp. 24–31. DOI: 10.1109/MSP.2010.186.

[43] Adnan Ahmed Abi Sen and Abdullah M. Basahel. "A Comparative Study between Security and Privacy". In: *2019 6th International Conference on Computing for Sustainable Global Development (INDIACom)*. 2019, pp. 1282–1286.

[44] OWASP. *OWASP Definition for Vulnerability*. Web. 2018. URL: https://www.owasp.org/index.php/Category:Vulnerability.

[45] Ji Feng, Qi-Zhi Cai, and Zhi-Hua Zhou. "Learning to Confuse: Generating Training Time Adversarial Data with Auto-Encoder". In: *Advances in Neural Information Processing Systems 32 (NeurIPS 2019)*. 2019, pp. 11971–11981. URL: https://proceedings.neurips.cc/paper/2019/hash/0a1bf96c6f1a4b4fa5fdf10e6e8c9c5d-Abstract.html.

[46] L. Muñoz-González et al. "Towards Poisoning of Deep Learning Algorithms with Back-Gradient Optimization". In: *Proceedings of the 10th ACM Workshop on Artificial Intelligence and Security (AISec '17)*. ACM Press, 2017. DOI: 10.1145/3128572.3140451. URL: http://dx.doi.org/10.1145/3128572.3140451.

[47] D. Cao et al. "Understanding Distributed Poisoning Attack in Federated Learning". In: *2019 IEEE 25th International Conference on Parallel and Distributed Systems (ICPADS)*. IEEE, 2019, pp. 233–239. DOI: 10.1109/ICPADS47876.2019.00045.

[48] Suyi Li et al. *Learning to Detect Malicious Clients for Robust Federated Learning*. arXiv preprint. Not peer-reviewed. 2020. arXiv: 2002.00211 [cs.LG]. URL: https://arxiv.org/abs/2002.00211.

[49] Minghong Fang et al. "Local Model Poisoning Attacks to Byzantine-Robust Federated Learning". In: *29th USENIX Security Symposium (USENIX Security 20)*. Peer-reviewed conference paper. 2020, pp. 1605–1622. URL: https://www.usenix.org/conference/usenixsecurity20/presentation/fang.

[50] Daliang Li and Junpu Wang. “FedMD: Heterogeneous Federated Learning via Model Distillation”. In: *NeurIPS 2019 Workshop on Federated Learning for Data Privacy and Confidentiality*. Workshop paper. Peer-review status may vary. 2019. URL: `https://arxiv.org/abs/1910.03581`.