



LISTIC



Symposium On Women in Services Computing

# WISC 2023 PhD Forum

## Advancing Bias Mitigation Techniques in Federated Learning

Presented by

Ferraguig Lynda

Under the supervision of

Benoit Alexandre, Loukil Faiza

In collaboration with

Lin-Kwong-Chon Christophe, Bettinelli Mickaël



# Plan

---

A- Research problem

What is Federated Learning?

What is Bias? And Why do we need to mitigate Bias in ML, FL?

B-Motivation

Why is it harder to solve this problem in Federated Learning Systems?

C-Approach

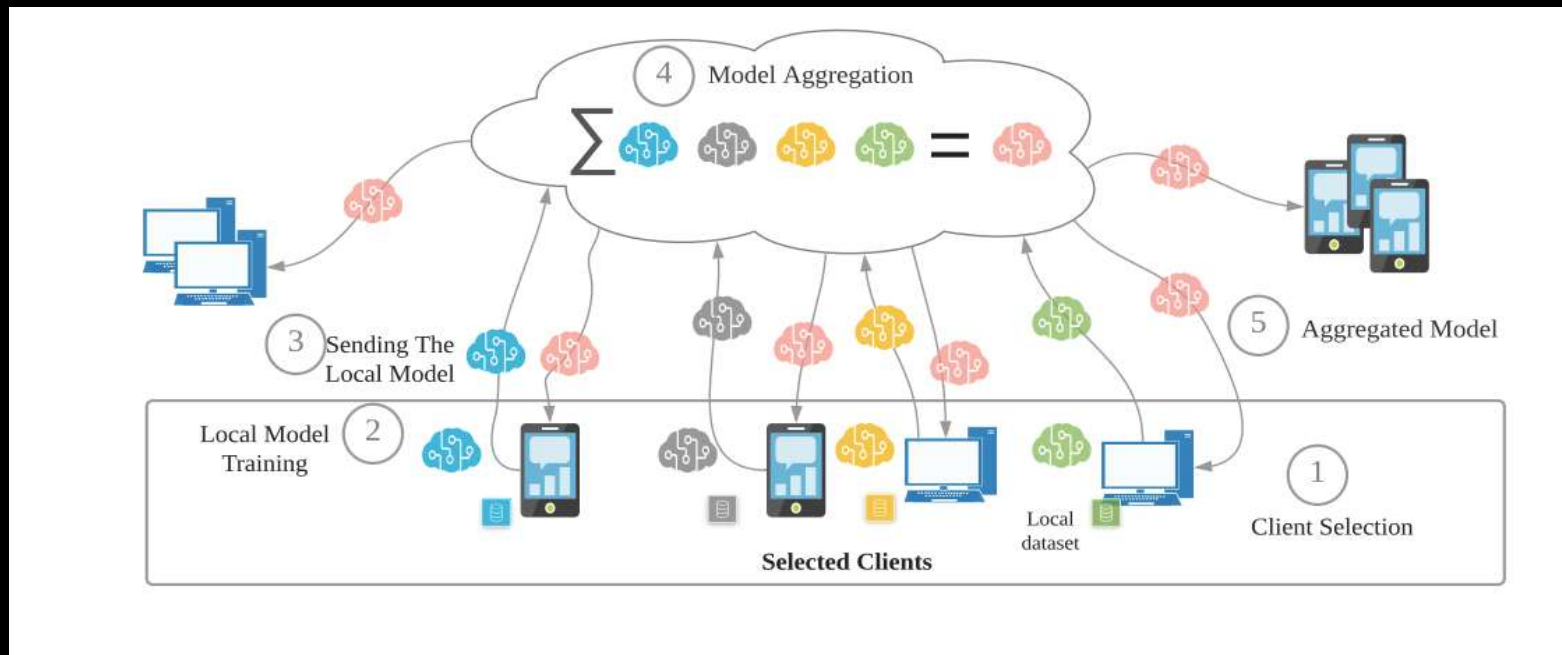
Clustered Federated Learning and Personalized Federated Learning For Bias Mitigation in FL

D-Initial Results and Next Steps

Ongoing work

What is Federated Learning?

- 1- **Machine Learning (ML)** is widely used across industries to **analyze data** and support **decision-making processes**.
- 2- With the emergence of **Big Data**, **Distributed Machine Learning (DML)** has become a prominent approach.
- 3- Researchers and businesses have swiftly adopted DML to analyze large-scale datasets.
- 4- In response to **new privacy regulations**, the direct use of **sensitive user data has become restricted**. In **2016**, Google introduced **Federated Learning (FL)** as a solution to address privacy concerns.



What is Bias? And Why do we need to mitigate Bias in ML, FL?

## Definitions (Mehrabi et al., 2019)

**Equity** is the **absence** of any **prejudice** or **favouritism** towards an individual or group of individuals due to their inherent or acquired characteristics.

-> **Sensitive attributes** are characteristics that divide the population into groups based on certain common traits, such as gender, age, etc.

**Bias** is the **inclination or prejudice** of a decision made by an ML system that is **perceived as unfair**, either for or against a person or group.

## Example



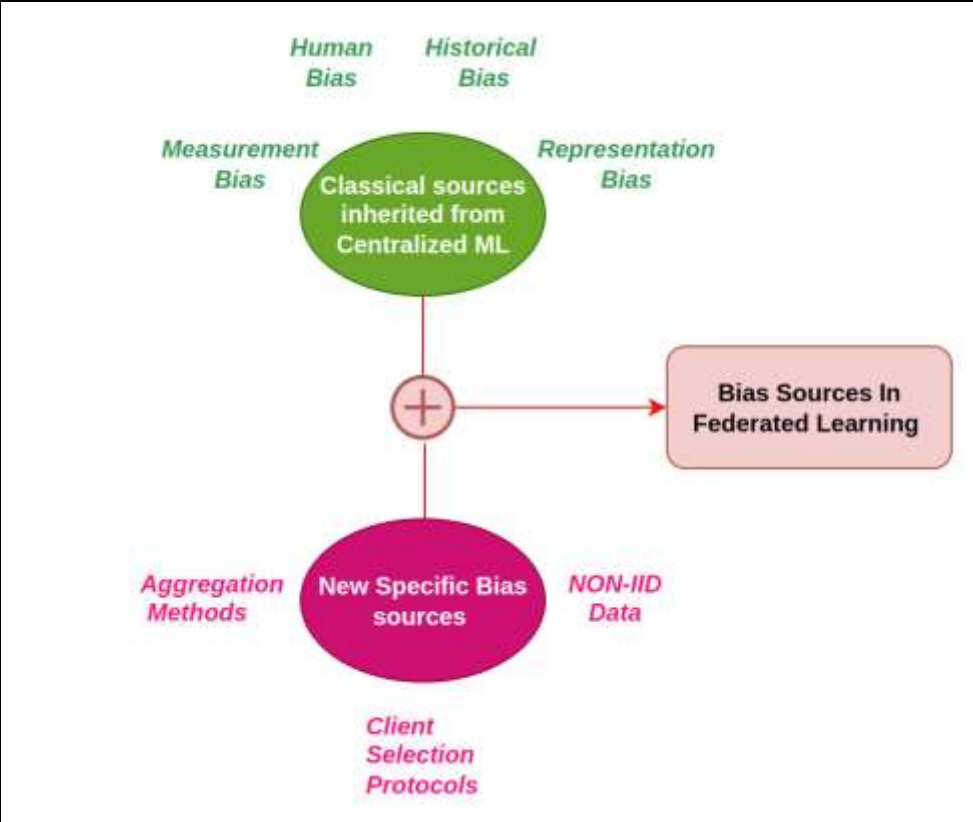
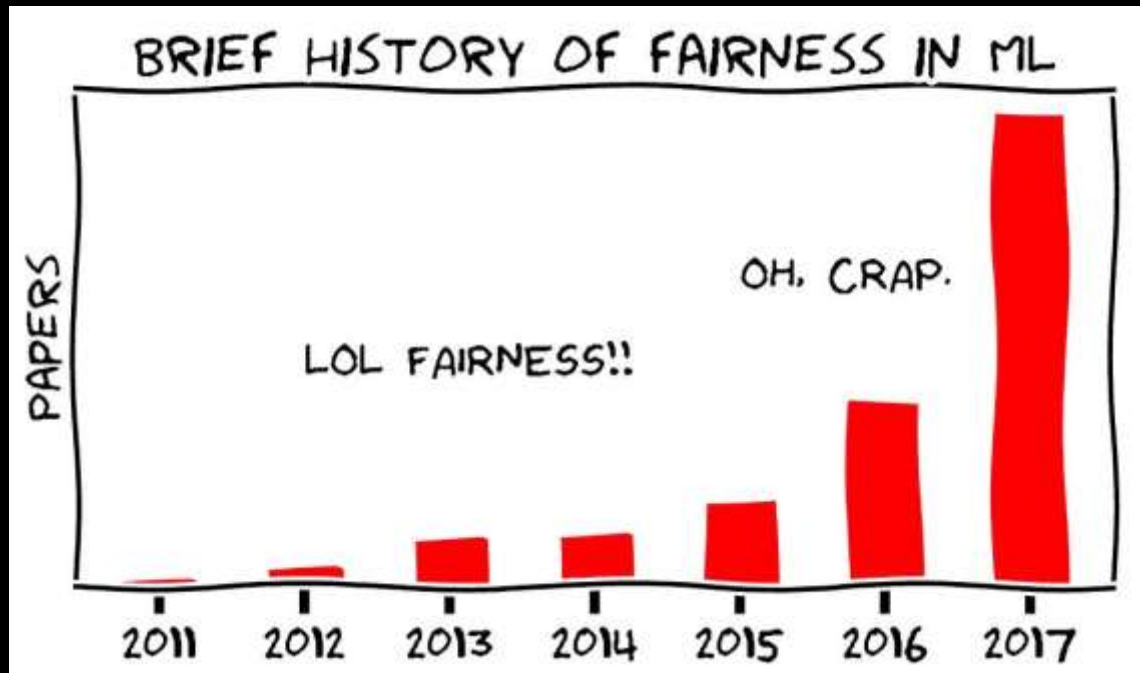
**Compas (ProPublica, 2016)** is a ML algorithm used by the Supreme Court in the USA to predict the risk of recidivism.

Sensitive attribute = Person's origin  
Group 1: African-American → High risk of recidivism  
Group 2 : White-American → Low risk

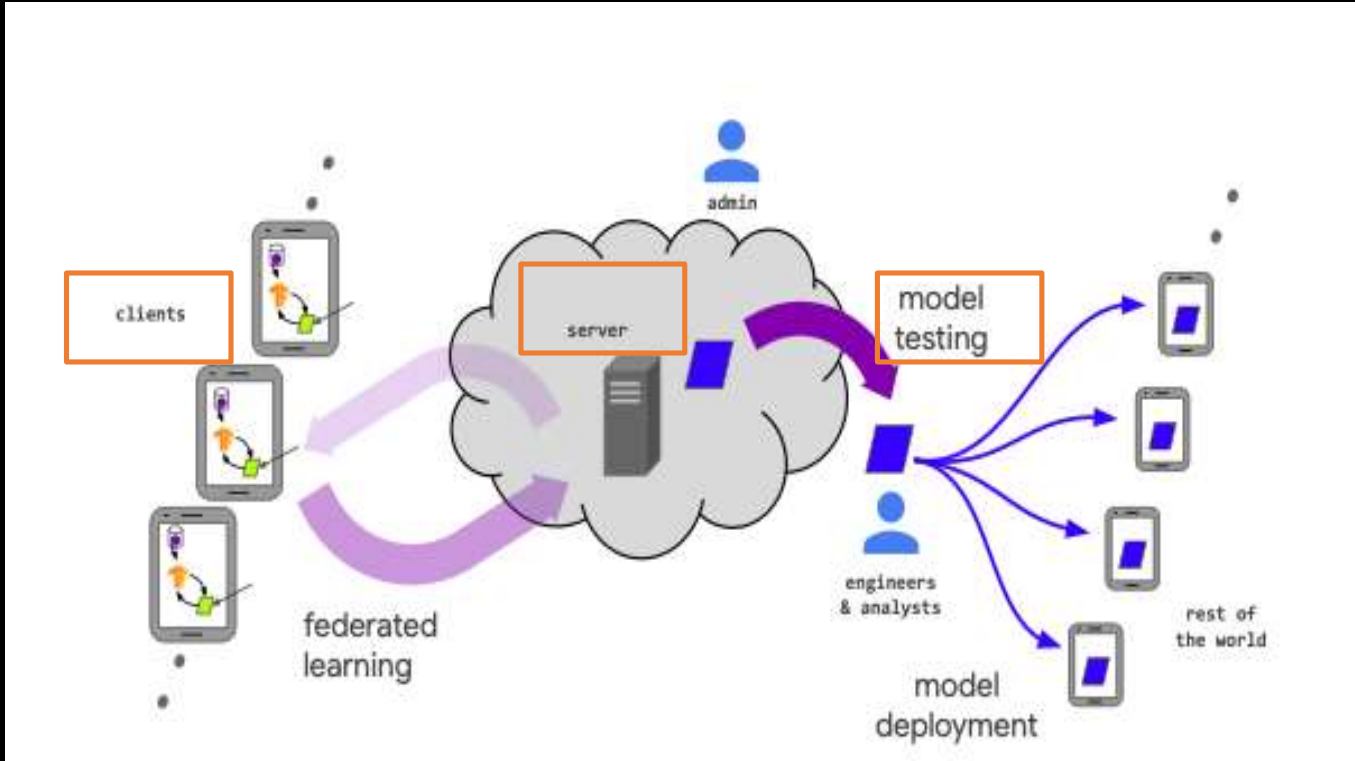
**Machine learning may learn how to be racist, sexist and discriminatory...**  
**i.e machine learning may be biased**

# Why is it harder to solve this problem in Federated Learning Systems?

- Federated Learning (FL) → **New sources of bias** [2]
- New **definitions of bias** and fairness [3]
- **Exacerbates bias** [4]
- **Inapplicability** of classical techniques to FL (due to privacy data issues) [4]



# Bias Mitigation Techniques in Federated Learning Related Works



(Peter Kairouz et al., 2019)

Level	Proposed Solution
<b>Client's</b>	-Apply classical ML techniques (pre-processing, in-processing) [5,6,7]
<b>Server's</b>	-New client selection algorithm ( FedCs [8] ) -New fusion algorithms (FMDAM [10], FeedProxy [11], FPFL[12], FairFL [13] , Fair Fed [14] )
<b>At the auditing step ( after finishing training ) - server side -</b>	-Apply post-processing techniques little bit modified [15,16]



## Clustered Federated Learning and Personalized Federated Learning For Bias Mitigation in FL

**Clustered Federated Learning (CFL).** addresses suboptimal results in FL due to diverging data distributions by grouping clients based on their similarities. [Sattler et al., 2020, Duan et al., 2021]

**Personalized Federated Learning (PFL).** the training process is customized for each client based on their preferences. This helps overcome the problem of using a single model that may not work well with different types of data. As a result, performance is improved, and the model is better aligned with each client's specific goals. [Deng et al., 2020]

*So far, these techniques are used to improve the problem of accuracy degradation in Federated Learning without considering the problem of bias ( no evaluation of any bias metric )*

**Explore Clustering, Personalization for bias detection and mitigation in Federated Learning**

### Initial Results

1- State of the art of bias mitigation techniques in Machine Learning in general and FL in particular, Clustered Federated Learning, and Personalized FL

--> Studied the limitations of state-of-the-art works.

2- Propose a novel Framework using clustering of FL clients according to bias level using bias metrics

### Next steps

1-Test the framework on a large real-world dataset following a non-iid (non-independent and identically distributed) distribution.

2-Study the impact of privacy and security protocols on our solutions.

3-Find a compromise between mitigating bias and the robustness of federated learning (FL) systems.



# Thank you!

---

Contact

Email [Lynda.ferraguig@univ-savoie.fr](mailto:Lynda.ferraguig@univ-savoie.fr)



# Some references

- [Calders et al., 2009] Calders, T., Kamiran, F., and Pechenizkiy, M. (2009). Building classifiers with independency constraints. In 2009 Ninth IEEE International Conference on Data Mining Workshops, pages 13–18.
- [Chang and Shokri, 2023] Chang, H. and Shokri, R. (2023). Bias propagation in federated learning. In The Eleventh International Conference on Learning Representations.
- [Deng et al., 2020] Deng, Y., Kamani, M. M., and Mahdavi, M. (2020). Adaptive personalized federated learning. arXiv preprint arXiv:2003.13461.
- [Duan et al., 2021] Duan, M., Liu, D., Ji, X., Wu, Y., Liang, L., Chen, X., and Tan, Y. (2021). Flexible clustered federated learning for client-level data distribution shift.
- [Ferraguig et al., 2021] Ferraguig, L., Djebrouni, Y., Bouchenak, S., and Marangozova, V. (2021). Survey of bias mitigation in federated learning. In Conférence francophone d’informatique en Parallélisme, Architecture et Système.
- [Kairouz et al., 2021] Kairouz, P., McMahan, H. B., Avent, B., Bellet, A., Bennis, M., Bhagoji, A. N., Bonawitz, K., Charles, Z., Cormode, G., Cummings, R., et al. (2021). Advances and open problems in federated learning. Foundations and Trends® in Machine Learning.
- [Kamishima et al., 2012] Kamishima, T., Akaho, S., Asoh, H., and Sakuma, J. (2012). Fairness-aware classifier with prejudice remover regularizer. In Joint European Conference on Machine Learning and Knowledge Discovery in Databases, pages 35–50. Springer.
- [Lambrech, 2018] Lambrecht, A. C. E. T. (2018). Algorithmic bias? an empirical study into apparent gender-based discrimination in the display of stem career ads. Same.
- [Louppe et al., 2017] Louppe, G., Kagan, M., and Cranmer, K. (2017). Learning to pivot with adversarial networks. In Proceedings of the 31st International Conference on Neural Information Processing Systems, pages 982–991.
- [McMahan et al., 2016] McMahan, H. B., Moore, E., Ramage, D., Hampson, S., and y Arcas, B. A. (2016). Communication-efficient learning of deep networks from decentralized data. In Proceedings of the 20th International Conference on Artificial Intelligence and Statistics (AISTATS).
- [Mehrabi et al., 2021] Mehrabi, N., Morstatter, et al. (2021). A survey on bias and fairness in machine learning. ACM Computing Surveys (CSUR), 54(6):1–35.
- [Sattler et al., 2020] Sattler, F., Müller, K.-R., and Samek, W. (2020). Clustered federated learning: Model-agnostic distributed multi-task optimization under privacy constraints. IEEE Transactions on Neural Networks and Learning Systems, 31(9):3400–3412.