NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

Towards Trustworthy Federated Learning

TrustFUL

Trustworthy Federated Ubiquitous Learning

Han Yu

Nanyang Assistant Professor
School of Computer Science and Engineering
Nanyang Technological University, Singapore

Self-Introduction

Han Yu

Nanyang Assistant Professor
School of Computer Science and Engineering
Nanyang Technological University, Singapore

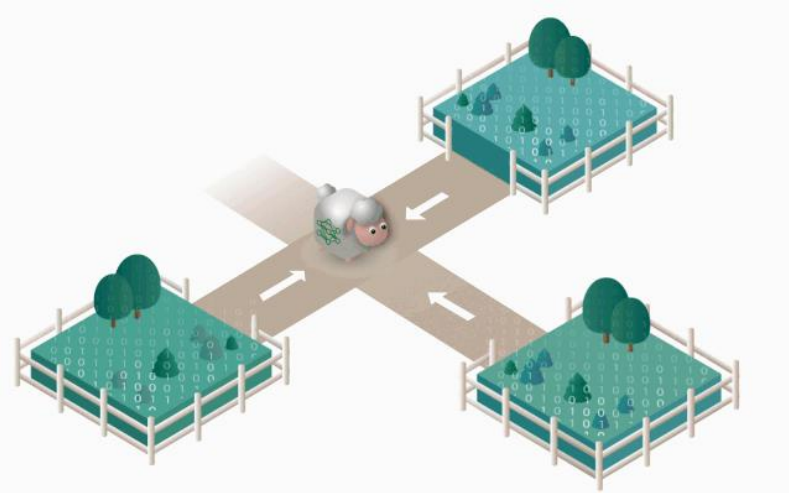
http://hanyu.sg/

han.yu@ntu.edu.sg

Research Areas:

- Federated Learning
- Multi-Agent Systems

Federated Learning – Privacy-Preserving Machine Learning

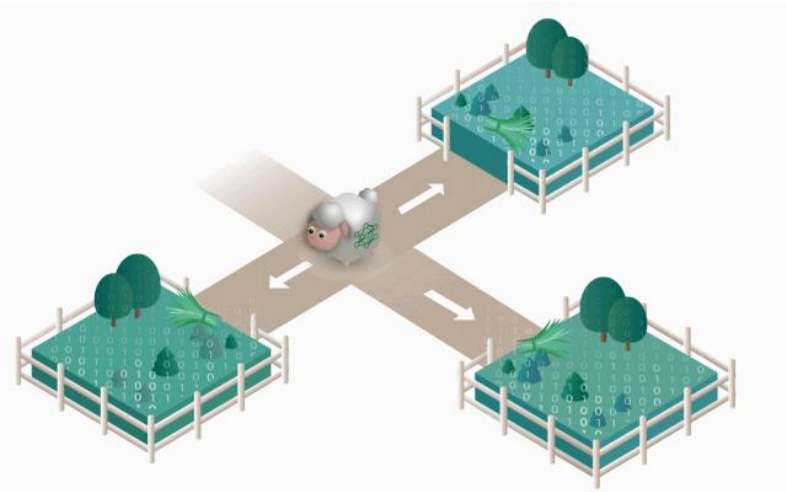


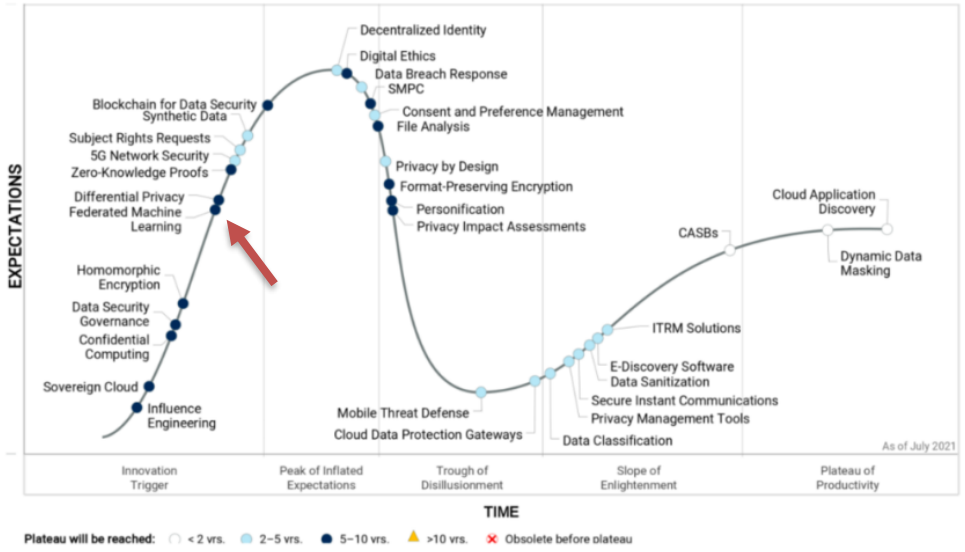
Traditional Machine Learning:

- Moving data to a centralized entity for model training
- Privacy often exposed

Federated Learning:

- Moving model training to where data originate
- Privacy is preserved





An Overview of TrustFUL

https://trustful.federated-learning.org/

SCAN ME

Trustworthy Federated Ubiquitous Learning (TrustFUL) – building trust to enable data providers to participate in AI model co-creation, while protecting their sensitive data.

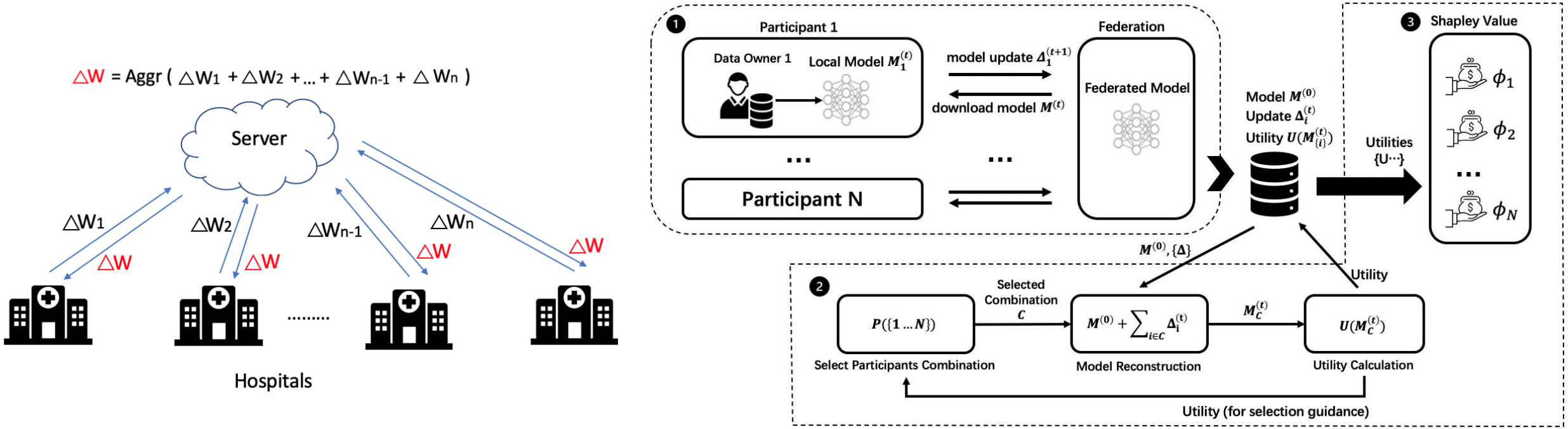
Achieving Trust through:

- Interpretability
  - Data, features, models
- Fairness
  - Opportunities, payoffs
- Robustness
  - Security, scalability

Achieving Ubiquity through:

- Personalizability of models
  - Resource & data heterogeneity
- Transferrability of knowledge
  - Cross country, cross sector, cross tasks

Explainable & Fair Federated Learning

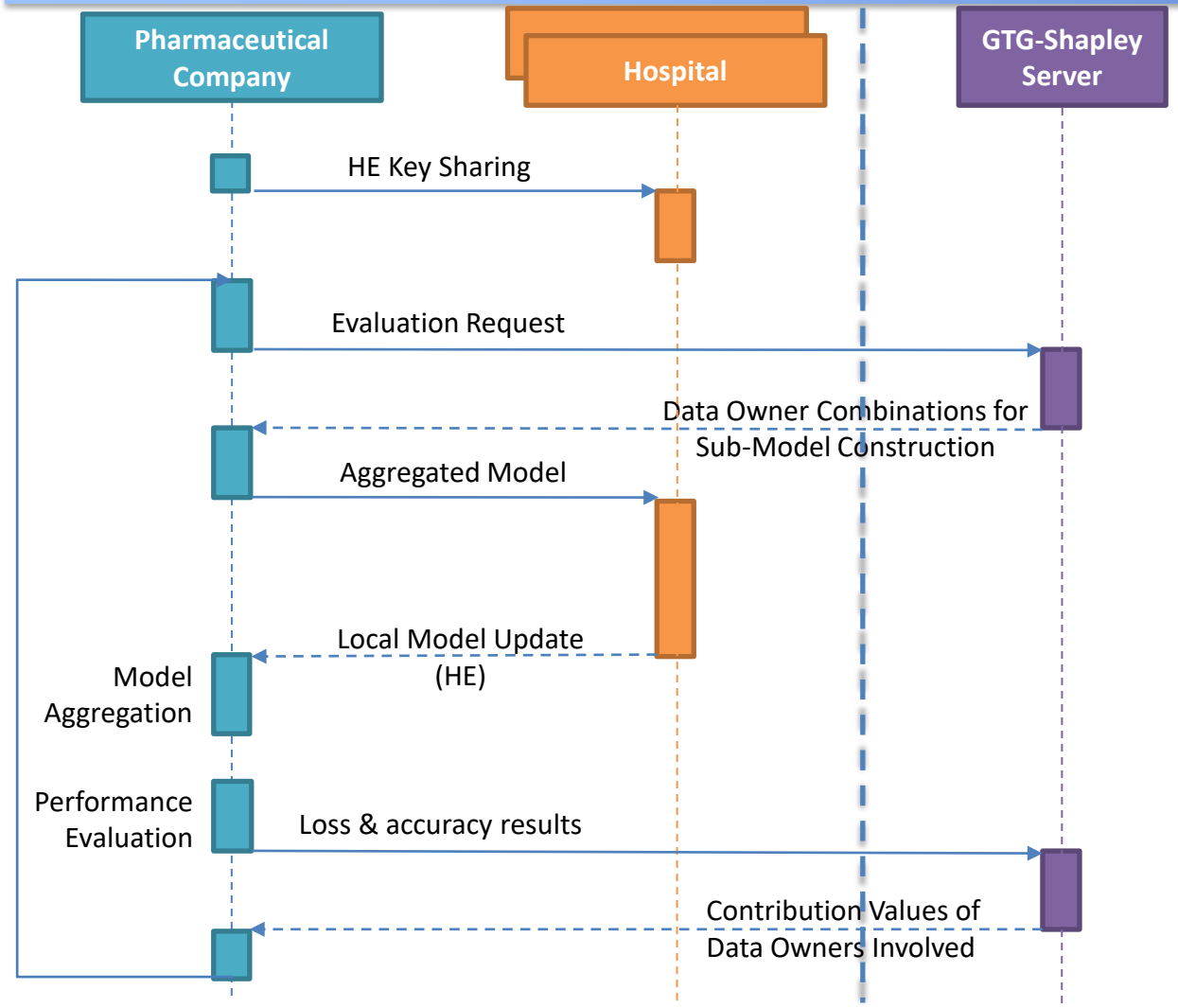


Fair and Efficient FL Participant Contribution Evaluation

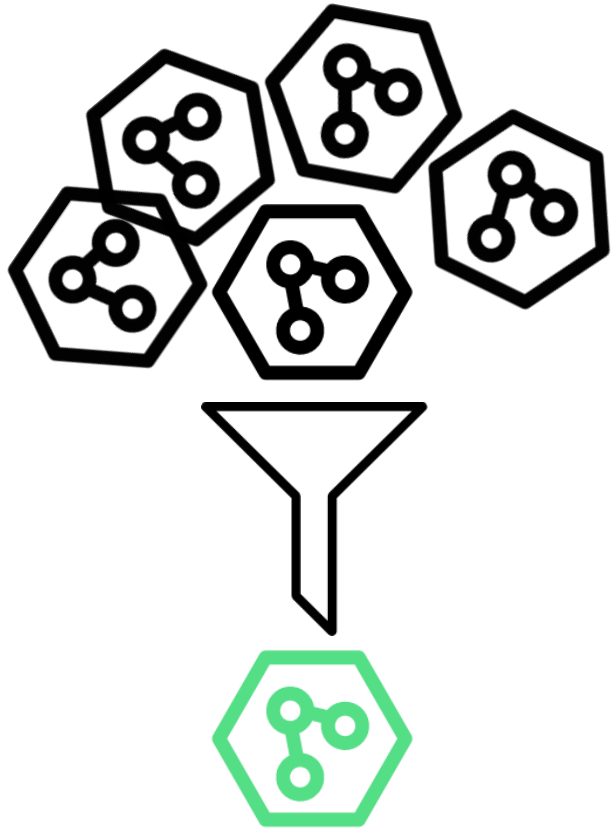
- Developed a fair and efficient algorithm to evaluation FL data owner contributions.
- Significantly enhanced the scalability of Shapley value-based data valuation.

Zelei Liu, Yuanyuan Chen, Han Yu, Yang Liu & Lizhen Cui. GTG-Shapley: Efficient and accurate participant contribution evaluation in federated learning. *ACM Transactions on Intelligent Systems and Technology*, vol. 13, no. 4, pp. 60:1-60:21, ACM (2022).

CAreFL – Contribution-Aware Federated Learning



FL sub-models

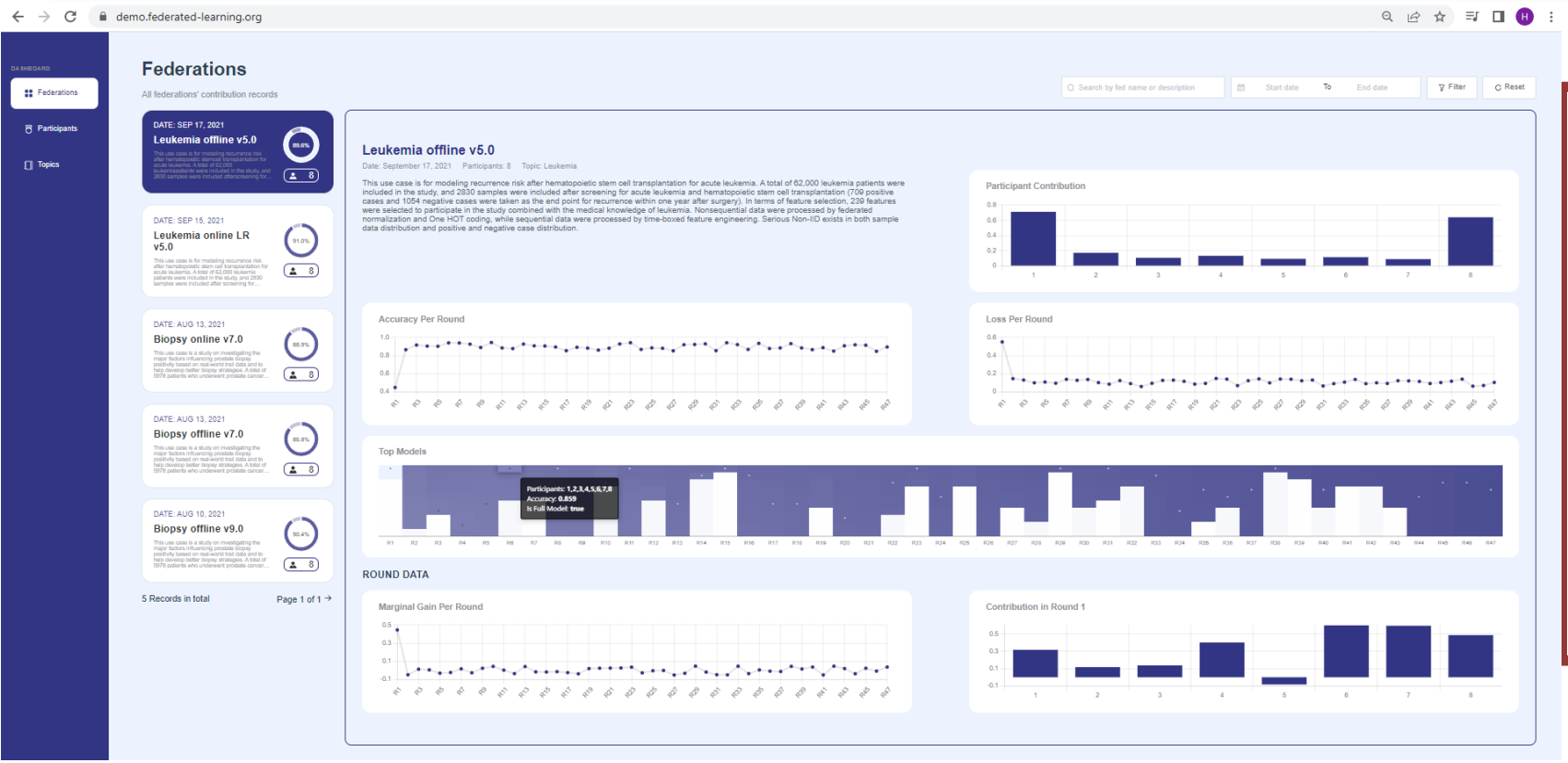


"Best Subset" FL Model

Deployment in the Healthcare Industry

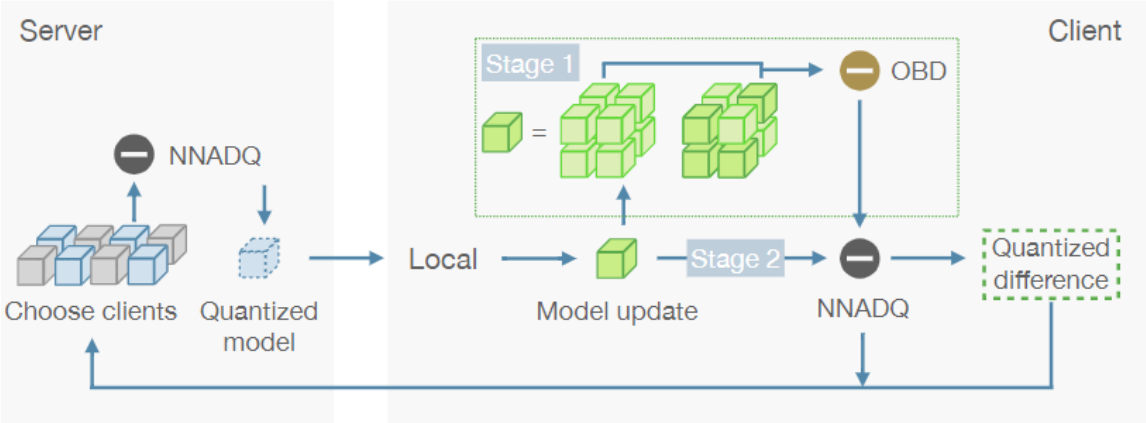
https://demo.federated-learning.org/

SCAN ME



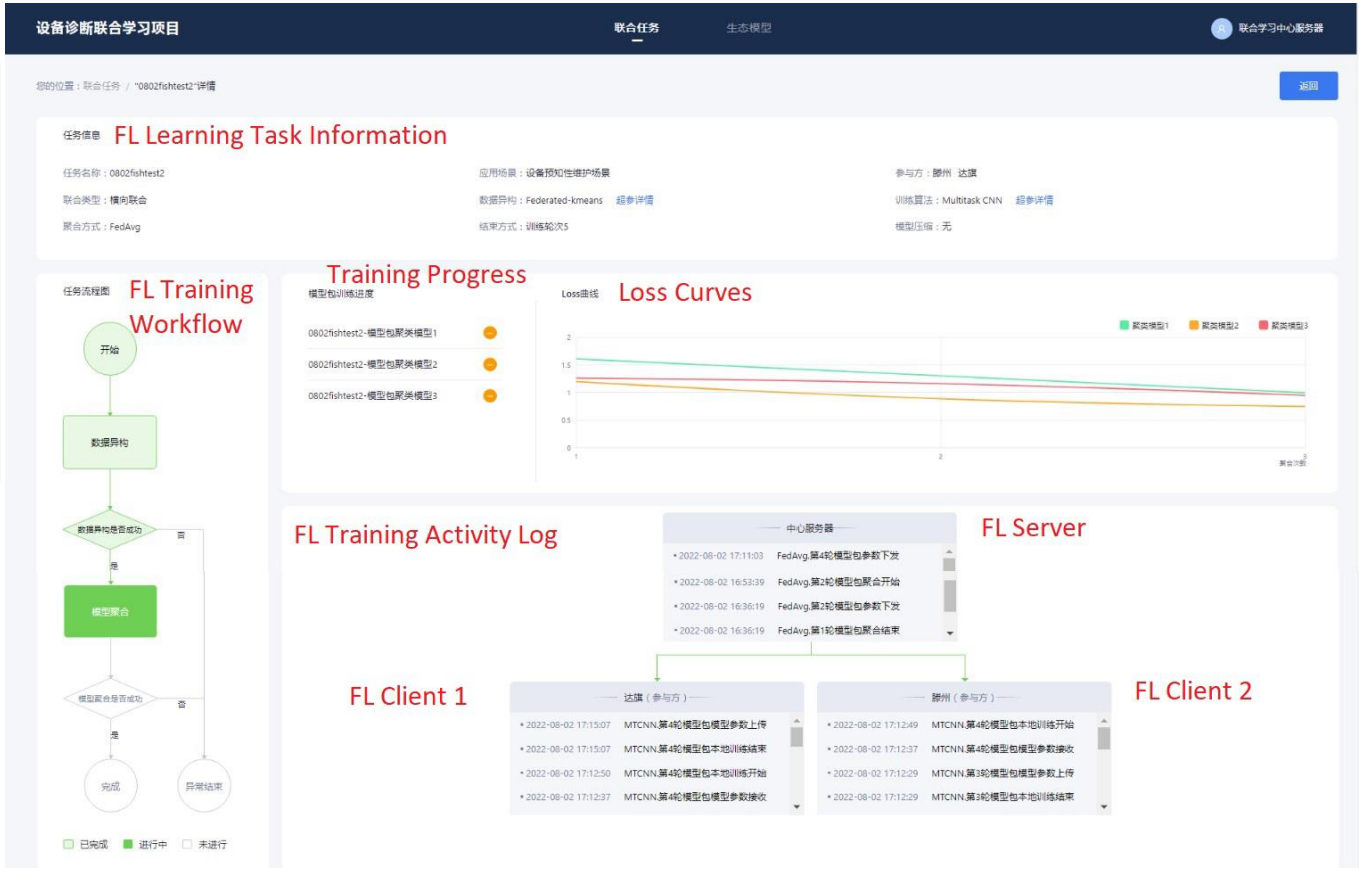
Z. Liu, Y. Chen, Y. Zhao, H. Yu, Y. Liu, R. Bao, J. Jiang, Z. Nie, Q. Xu & Q. Yang, "Contribution-Aware Federated Learning for Smart Healthcare," in *Proceedings of the 34th Annual Conference on Innovative Applications of Artificial Intelligence (IAAI-22)*, pp. 12396-12404, 2022. (**Innovative Application of AI Award**)

Efficient Large-Scale Federated Learning

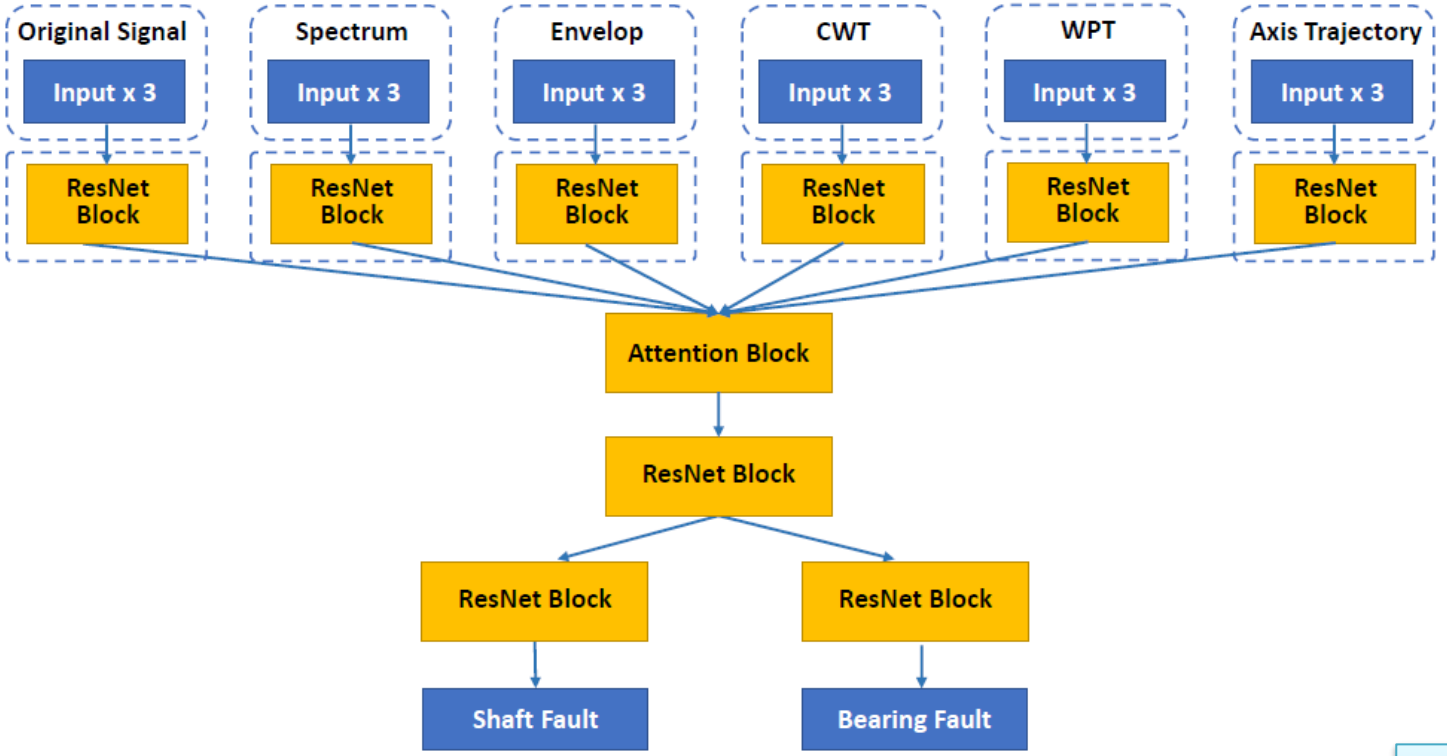


Opportunistic Block Dropout for Scalable FL

- A unique opportunistic semantic block dropout approach to enable only important model blocks to be transmitted.
- Enables efficient training of high-performance large-scale deep FL models.
- Y. Chen, Z. Chen, P. Wu & H. Yu, "FedOBD: Opportunistic Block Dropout for Efficiently Training Large-scale Neural Networks through Federated Learning," *IJCAI*, pp. 3541–3549, 2023.



Deployment with ENN Group



- Training a model with **30 million** parameters.
- Reduced total communication cost from **368 GB** to **104 GB**, while maintaining model performance at 85% F1 Score.
- Reduced model retraining time from **52 hours** to **14.5 hours** (at a limit of 2MB/sec bandwidth allowable for FL training).

Y. Chen, Z. Chen, S. Guo, Y. Zhao, Z. Liu, P. Wu, C. Yang, Z. Li & H. Yu, "Efficient Training of Large-scale Industrial Fault Diagnostic Models through Federated Opportunistic Block Dropout," in *Proceedings of the 35th Annual Conference on Innovative Applications of Artificial Intelligence (IAAI-23)*, pp. 15485–15493, 2023. (**Innovative Application of AI Award**)

Promising FL Research Directions

Generative AI for Visual Persuasion

- Chang Liu & Han Yu, "AI-empowered persuasive video generation: A survey," *ACM Computing Surveys*, 2023.

Fairness-Aware Federated Learning

- Yuxin Shi, Han Yu & Cyril Leung, "Towards Fairness-Aware Federated Learning," *IEEE Transactions on Neural Networks and Learning Systems*, 2023.

Quantum Federated Learning

- Chao Ren, Han Yu, Rudai Yan, Minrui Xu, Yuan Shen, Huihui Zhu, Dusit Niyato, Zhao Yang Dong & Leong Chuan Kwek, "Towards Quantum Federated Learning," *arXiv preprint arXiv:2306.09912*, 2023.

Verifiable Federated Learning

- Yanci Zhang & Han Yu, "Towards Verifiable Federated Learning," in *Proceedings of the 31st International Joint Conference on Artificial Intelligence (IJCAI'22)*, pp. 5686-5693, 2022.

Privacy and Robustness in Federated Learning

- Lingjuan Lyu, Han Yu, Xingjun Ma, Lichao Sun, Jun Zhao, Qiang Yang & Philip S. Yu, "Privacy and Robustness in Federated Learning: Attacks and Defenses," *IEEE Transactions on Neural Networks and Learning Systems*, 2022.

Interpretable Federated Learning

- Anran Li, Rui Liu, Ming Hu, Luu Anh Tuan & Han Yu, "Towards Interpretable Federated Learning," *arXiv preprint arXiv:2302.13473*, 2023.

Domain Adaption in LLM

- Xu Guo & Han Yu, "On the Domain Adaptation and Generalization of Pretrained Language Models: A Survey," *arXiv preprint arXiv:2211.03154*, 2022.

Personalized Federated Learning

- Alysa Ziying Tan, Han Yu, Lizhen Cui & Qiang Yang. Towards personalized federated learning. *IEEE Transactions on Neural Networks and Learning Systems*, 2022.

Federated **Graph Neural Networks**

- Rui Liu, Pengwei Xing, Zichao Deng, Anran Li, Cuntai Guan & Han Yu, "Federated Graph Neural Networks: Overview, Techniques and Challenges," *arXiv preprint arXiv:2202.07256*, 2022.

Auction-based Federated Learning

- Xiaoli Tang & Han Yu, "Towards Trustworthy AI-Empowered Real-Time Bidding for Online Advertisement Auctioning," *arXiv preprint arXiv:2210.07770*, 2022.

Open Collaborative Federated Learning Ecosystem

https://trustful.federated-learning.org/

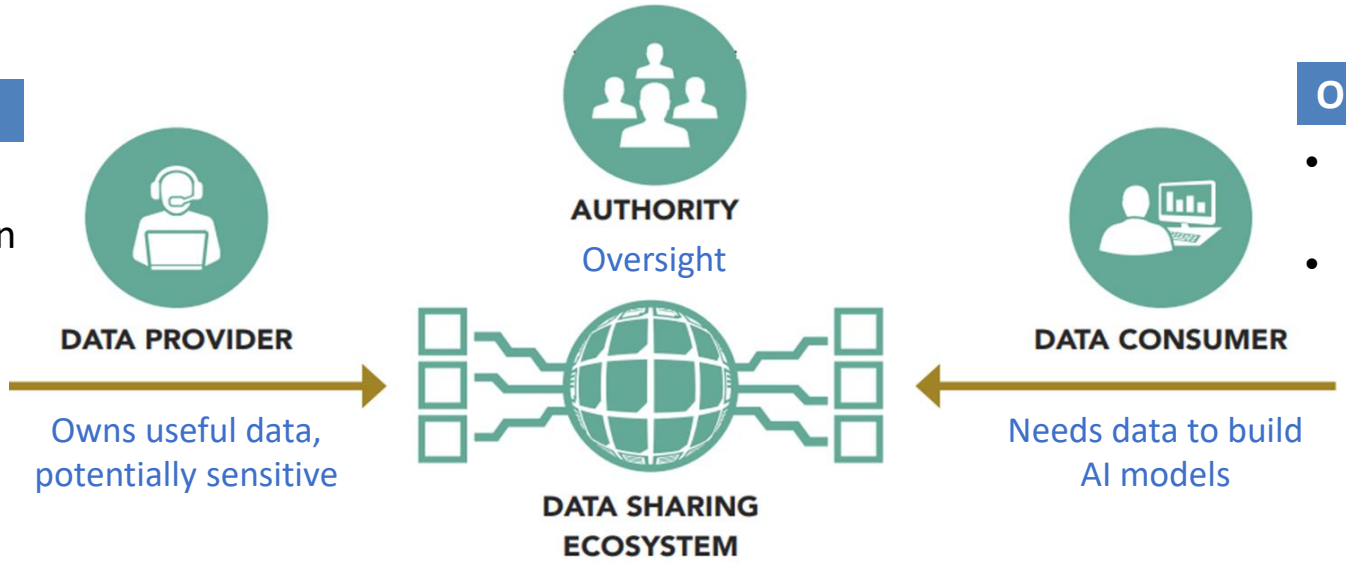
SCAN ME

Governance

- Privacy-Preserving Data Valuation/Auditing
- Interpretable FL Training Visualization Monitoring
- Fair Participant Contribution Assessment
- Participant Behaviour Modelling
- Misbehaviour Deterrence Optimization

Operation

- Dynamic FL Collaboration Formation



Operation

- Monetary Incentivization for FL Participants
- Non-Monetary Incentivization for FL Participants

TrustFUL

Thank you!

SCAN ME

NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

WeBank 微众银行

SMU SINGAPORE MANAGEMENT UNIVERSITY

SiT SINGAPORE INSTITUTE OF TECHNOLOGY

AI SINGAPORE

EPFL

中国科学院 CHINESE ACADEMY OF SCIENCES

UBC THE UNIVERSITY OF BRITISH COLUMBIA