The Rise of Federated Intelligence: From Federated Foundation Models Toward Collective Intelligence

Guodong Long
Australian Artificial Intelligence Institute,
Faculty of Engineering and IT, University of Technology Sydney, Australia
guodong.long@uts.edu.au

Abstract

The success of foundation models advances the development of various intelligent and personalized agents to handle intricate tasks in their daily lives, however finite resources and privacy concerns from end users limit the potential of customizing the large intelligent agents for personal use. This paper explores the preliminary design of federated intelligence that paves the way toward personalized intelligent agents in large-scale collaboration scenarios. In Federated Intelligence, agents can collaboratively augment their intelligence quotient (IQ) by learning complementary knowledge and fine-grained adaptations. These personalized intelligent agents can also co-work together to jointly address complex tasks in the form of collective intelligence. The paper will highlight federated intelligence as a new pathway for tackling complex intelligent tasks by refining and extending centralized foundation models to an open and collaborative paradigm.

1 Introduction

Recent progress of Foundation Models, e.g., ChatGPT, enhanced public beliefs about implementing Artificial General Intelligence (AGI) [Heaven, 2023]. More and more investments have been allocated towards this direction across many application scenarios, from natural language processing [Achiam *et al.*, 2023] to video generation [Brooks *et al.*, 2024]. More initiatives have been established to transfer this capability to many other domains, including science, medicine, robotics, recommendations, and many other data types, including graphs, time series, and tabular data.

So far, the cost of computation and the volume of model parameters are critical bottlenecks that hinder the deployment and further development of those foundation models. Although, with the help of recent developments in parameter-efficient fine-tuning, users can continuously integrate new and task-focused knowledge into a pre-trained open-source model like Llama [Touvron *et al.*, 2023], they cannot make much refinement on the essential components of those pre-trained models or address underlying risks of catastrophic forgetting and hallucinations [Maynez *et al.*, 2020]. Consequently, the centralized foundation model still suffers the open challenge of improving the agent's intelligence level and understanding the insight logic of the physical world, which is an essential part towards AGI.

A collective intelligence architecture leveraging many intelligent agents can improve the performance of tackling intelligent recognition tasks. It is a promising direction that many individual intelligent agents to collaboratively evolve, thus their collective intelligent IQ might be increased towards AGI, and they still can preserve their personality by updating the intelligent model using their private dataset. Therefore, these intelligent agents can collectively enhance intelligent capability and also jointly tackle complex tasks.

The paper aims to propose a new concept named federated intelligence which is an integration of existing federated foundation models framework and collective intelligence [Ha and Tang, 2022]. Compared to traditional collective intelligence methods, federated intelligence focuses on a collaborative learning process of foundation model(s) on many intelligent participants, and how to develop personalized and(or) generalist agents from these foundation model(s). The scale of federated intelligence will be much larger than previous methods. The coordination is more challenging due to heterogeneity among participants.

This paper aims to discuss that the federated foundation model [Ren *et al.*, 2024; He *et al.*, 2024; Zhuang *et al.*, 2023] is a technical solution to implement this objective by developing a new type of collective intelligence, namely federated artificial intelligence. Specifically, the small agents will have compatible Transformer structures but are trained independently with heterogeneous data and tasks. A central server will coordinate the learning process and consolidate common knowledge shared across agents. Compared to conventional distributed machine learning [Verbraeken *et al.*, 2020], the proposed federated intelligence enables each client to be equipped with heterogeneous data to ensure the client-specific agent has unique expertise and capability. The recent boosting of Hugging Face and the release of personalized ChatGPT have paved some way towards utilizing more agents to capture the diverse world and thus towards AGI ultimately. However, these customized LLMs are separated from each other without a careful design of collaboration. Federated intelligence is a promising direction to enable these customized LLMs to collaborate evolving and co-working.

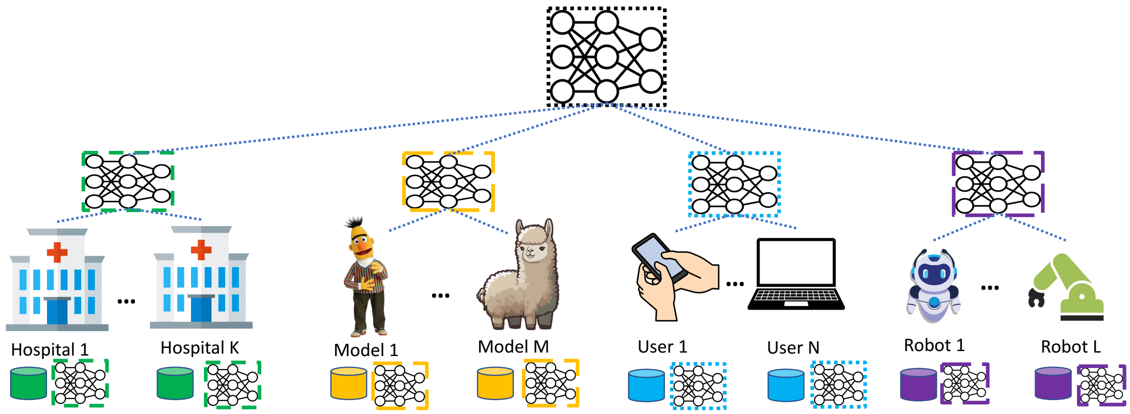

Figure 1: Architecture of Federated Intelligence with heterogeneous participants and hierarchical collaboration

2 Architecture of Federated Intelligence

As illustrated in Figure 1, in a collaboration environment, there are many participants with heterogeneous settings on hardware, software, and application scenarios. A federated intelligence system enables these participants to collaboratively share knowledge to enhance the capability of intelligent agents. Moreover, a hierarchical knowledge-sharing mechanism is designed to ensure each participant collaborates with similar peers rather than random peers.

Each participant is equipped with a pre-trained foundation model with the basic capability to tackle various tasks. These foundation models could be fine-tuned or upgraded by the participant via collaboration with other participants. In alignment with the neural architecture of foundation models, the communication of knowledge sharing could be implemented via partial model updating, client-specific embedding vectors, and other parameter-efficient fine-tuning methods.

3 Problem Formulation

In a collective environment, there are M intelligent agents equipped with intelligent models powered by foundation models parameterized by W. Each agent can has a dataset D to record its personal experience and memory. Each individual agent needs to tackle two types of intelligent tasks: recognition tasks and planning tasks.

For **recognition tasks**, e.g., image recognition, the intelligent agent can update the intelligent model by learning from its personal experience using below optimization objective.

$$W^{(i)} = \arg\min_{W_r} L(D; W_r) \tag{1}$$

in where the L is an empirical loss function for intelligent recognition tasks on dataset D.

For **planning tasks**, the intelligent agent is to maximize the reward for a sequence of steps.

$$W^{(i)} = \arg\max_{\{W_r, W_p\}} R(E; \{W_r, W_p\}) \tag{2}$$

in where the E is the environment that the agent interacts with, and R is the function to calculate the reward of a sequence of actions generated by the agent's planning model parameterized by the recognition module W_r and the policy module W_p.

In a **collective environment**, there are many ways to define the loss of reward of the entire system, e.g., the overall objective could be minimizing the average loss of all agents.

$$\min_{\{W^{(i)}\}} \mathcal{L}_g(D; W),\ \mathcal{L}_g(D; W) = \sum_{i=1}^{M} \frac{1}{m} L(D_i; W^{(i)}) \tag{3}$$

where $\mathcal{L}_g$ is the global loss of the entire system that may take various forms with different weighting metrics, such as dataset scale as an importance weight per agent, and fairness highlighted weighting. It is also a usual way to add regularization terms to enhance the model's generalization capability of in-distribution and out-of-distribution scenarios.

Nowadays, a **pre-trained foundation model** parameterized by W_f can be equipped to each agent to enable them with the basic capacity towards AGI in specific application scenarios. Then, many client-specific customized decision-making logic or a secondary intelligent model can be learned via W_l to cooperate with the foundation model W_f. The overall objective function could be formulated accordingly.

$$\min_{\{W_l^{(i)}\}} \mathcal{L}_g(D; W_f; \{W_l^{(i)}\}) \tag{4}$$

in where $\{W_l^{(i)}\}$ is a set of learnable parameters from customized models of many agents indexed by i.

It is worth noting that the current foundation model works very well on most intelligent recognition tasks. Although there is good progress in this direction, it is still an open challenge to use the foundation models to tackle complex planning tasks in the general world.

4 From Federated Learning toward Federated Intelligence

4.1 Heterogeneity is the Key Challenge

Federated Learning was proposed as a collaborative machine learning framework to tackle privacy-preservation and heterogeneous settings. The first heterogeneous setting is statistical heterogeneous, also named non-IID data or horizontal FL. To tackle this, the following development of federated learning has been transited to robust FL [Yan and Long, 2024], multi-center FL [Long *et al.*, 2023], and personalized FL [Mansour *et al.*, 2020; Tan *et al.*, 2022a]. This transition enables the centralized intelligence to be decomposed toward multiple groups or many personalized individual intelligence.

The second is system heterogeneous including different settings on hardware capacity and software configuration. Many heterogeneous FL methods have been designed to tackle this issue. The third is data format heterogeneous, also named vertical FL, in which the contained information is inconsistent in format or feature space. It usually occurs in cross-silo FL applications that focus on collaboration across multiple organizations.

4.2 Foundation Models Uplift Agents' IQ

The success of foundation models, especially the large language models (LLM), can uplift an individual agent's intelligence capability to a new level. For example, in the NLP domain, a LLM can be used to tackle various downstream tasks with general intelligent ability at a certain level. Moreover, these LLMs can be customized with different domain-specific data or functionalities, for example, numerous open-sourced LLMs in Hugging Face can be applied to tackle various tasks in various domains. Thus, various ready-to-use customized foundation models can be equipped with agents to improve their intelligent IQ significantly. In Federated Intelligence, these agents can jointly tackle complex tasks as collective intelligence and also enhance their own intelligent IQ through a collaborative learning process.

4.3 Federated Intelligence

Existing Federated Foundation Models (FFM) focus on fine-tuning or training the foundation model for domain-specific machine learning tasks while paying less attention to developing effective coordination and collaboration mechanisms to tackle complex intelligent tasks. The recent development of the LLM agent [Zhao *et al.*, 2024] indicates a new trend to further enhance and expand the foundation models' intelligent capability to a more comprehensive environment and more challenging tasks. To this end, we would like to highlight federated intelligence as a new domain to tackle complex intelligent tasks using the technology derived from FFM

Federated intelligence has a well-built infrastructure for implementing collective intelligence. The federated foundation model technique has been widely applied in many scenarios including Smartphone APPs, Artificial Intelligence of Things (AIoT) [Liu *et al.*, 2020; Tiwari *et al.*, 2023], hospitals [Chenthara *et al.*, 2019; Molaei *et al.*, 2024; Chen *et al.*, 2023], and finance organisations [Long *et al.*, 2020]. Based on existing application scenarios, federated intelligence can be easily evolved into a collective intelligence for AGI.

4.4 The relation between Federated Intelligence and Collective Intelligence

Collective intelligence enables many individual intelligent agents to collaboratively accomplish a complex task. The capability of collective intelligence is stronger than each individual. In addition to this, federated intelligence also focuses on enhancing each individual through the learning process by leveraging the knowledge from the collaborative environment with many agents. Therefore, the ultimate goal of federated intelligence is to implement AGI in a decentralized manner. Therefore, we don't need to worry about a super AGI to take over everything, instead many individual intelligent agents controlled by different users will tackle many tasks in a relatively independent manner. The architecture of decentralized AGI enables every user to have a tiny part of power in a collaborative environment that is similar to a vote in a democratic society.

The most traditional collective intelligence and multi-agent systems have to tackle the competition among agents, thus the intelligence capability has been decreased. Federated intelligence focuses on collaboration rather than debating and competing. Federated intelligence is to embody a new type of collective intelligence towards AGI. Federated intelligence can use debating and competing as a mechanism to enhance the agent's intelligent IQ and robustness rather than to ensure the agent gains the most reward in the setting of game theory.

5 Future Directions of Federated Intelligence

To implement a federated intelligence framework, the below components need to be carefully designed.

5.1 Heterogeneous Agents

Each agent in federated intelligence should be equipped with a foundation model to tackle various tasks. There are two ways to preserve the heterogeneous across agents or tasks. The first way is to preserve the heterogeneous into a vector, e.g. Class-specific Prototypes [Tan *et al.*, 2022b] and User embeddings [Zhang *et al.*, 2024b] to tackle the heterogeneous. Specifically, the foundation model can use agent-specific information in the form of prompts on contexts.

$$\min_{\{W_l;W_v\}} L(D;\{W_l;W_v\}) \tag{5}$$

where W_v is a model to project each agent's unique characteristics or requirements to an embedding vector.

The second way to tackle heterogeneous tasks is to apply parameter-efficient fine-tuning to pre-trained foundation models by updating partial layers or adapters, thus we can split the model into frozen parameters W_f and learnable parameters W_l.

$$min_{W_l} L(D;\{W_f,W_l\}) \tag{6}$$

where we only update the learnable parameters W_l of the model while freezing the w_f.

5.2 Customizing Foundation Models

Although the pre-trained foundation models can tackle various downstream tasks via prompt engineering and in-context learning, the user can further enhance the foundation model using parameter-efficient fine-tuning (PEFT) to customize the foundation model. For example, Low-Rank Adaptation (LoRA) [Hu *et al.*, 2021] significantly reduces the cost of customizing a foundation model. They show us the feasibility of continuously aggregating new knowledge into foundation models by agents from heterogeneous environments .

The LoRA is a type of parameter-efficient fine-tuning technology to be applied to foundation models. As stated in Eq. 6, a foundation model is decomposed into a freezing part W_f and a learnable part W_l. Specifically, fine-tuning the model W_f is to add a task-specific gradient ΔW_l. The scale of ΔW_l is the same as W_f which is huge. LoRA method use low-rank decomposition to replace W_l with two vectors A and B which are about 0.01% of W_l.

$$h = W_f x + \Delta W_l x = W_0 X + BAx \tag{7}$$

Several Federated LoRA [Babakniya *et al.*, 2023; Sun *et al.*, 2024b; Yi *et al.*, 2023] have been developed. [Yang *et al.*, 2024] used a Dual-Personalizing Adaptation mechanism to improve the out-of-distribution generalization ability of the Federated LoRA system.

5.3 Structured Collaboration for Heterogeneous

Collaboration among heterogeneous participants is a critical challenge in implementing federated intelligence. Upon the gradually increasing heterogeneity, the corresponding solution varies. If the participants are from two different domains, e.g. hospitals and banks, the collective intelligence system should have two different global foundation models to serve different domains respectively. Thus, a clustering-based knowledge-sharing mechanism [Ma *et al.*, 2023] is required. If the participants are from many different domains with a hidden relation graph, a graph-guided collaboration mechanism [Chen *et al.*, 2022] is required to implement the collective intelligence.

Take the ultimate scenario described in Figure 1, the participants are from many different domains in an open-set world. One possible solution is to use a hierarchical integration framework [Chen *et al.*, 2024] to approximate the complex relationship among all participants.

5.4 Tackling Sequential Decision

Recent research suggests that an Agent co-worked with foundation models is a practical framework to tackle various tasks in an open world. In the domain of NLP, LLM-based agents [Sun *et al.*, 2024a] can tackle complex tasks that need sequential reasoning. Federated intelligence enables many LLM agents to collaboratively share common knowledge and also preserve the ability to generate personalized reasoning.

In the domain of robotics [Liu *et al.*, 2019], embodied AI [Duan *et al.*, 2022] is proposed to enable the AI agent to learn and evolve through the interaction between their embodied body and the environment. Given the large-scale use of robots, federated intelligence could be a collaboration mechanism to enable the robots to jointly explore the environment by sharing their experience and knowledge, and also to collaboratively complete complex tasks.

5.5 Enhanced Intelligent by Shared Memory

One agent's memory and experience are limited, thus they can co-work together to jointly evolve the shared knowledge and shared dataset, and also upgrade each agent's intelligent capability by leveraging both shared knowledge and highlighting the personal experience and interests stored in its dataset D. Assuming many agents are willing to contribute part of their non-sensitive raw data to a shared proxy dataset, the central coordinator will maintain the shared dataset by selecting the most valuable knowledge and updating each agent's proxy dataset. For example, A new Retrieval Augmented Generation (RAG) [Lewis *et al.*, 2020] technique can be applied to further store information to augment the agent's intelligence IQ by leveraging more raw data that could be from external sources or accumulated shared historical data among agents.

$$\min_{W_l} L(\{D, D_r\}; \{W_f, W_l, W_r\}) \tag{8}$$

where W_r is the retrieval model to be applied to augment the query or demands ($X \in D$) by leveraging the retrieved information from the proxy dataset D_r.

5.6 Interpretability for Trustworthy AI

Interpretability is critical to building a trustworthy AI system. There are many existing works to enhance the Transformer and deep learning models with model-level interpretability. A new challenge of collective intelligence is to interpret the differences among participants. [Yan and Long, 2023] proposed to decompose the shared knowledge and personal knowledge via VAE. Further, [Yan and Long, 2024] uses a set of concept vectors to explain the personalized knowledge of participants in federated settings. [Cui *et al.*, 2024; Wang, 2019] studied to evaluate the contribution of each participant to maintain trust among participating entities, ensuring equitable resource sharing, and fostering a sustainable collaboration framework.

5.7 Most Promising Application Scenarios

Compared to image and NLP tasks, the recommendation system is a more promising scenario for implementing federated intelligence. The recommendation has a more frequent interaction between the model and end users and also occurs in a dynamic and complex environment with various scenarios including short videos, news, shopping items, social posts, and business advertisements. Privacy concern is also an important factor in promoting on-device recommendations. The recent research demonstrated that the Transformer-based recommendation system [Zhai *et al.*, 2024] can achieve superior performance in various recommendation tasks. In a Federated setting, using a traditional recommendation system with dual personalization [Zhang *et al.*, 2023; Li *et al.*, 2024] can significantly improve the performance. [Zhang *et al.*, 2024a] proposed the first federated foundation model for recommendation by leveraging the LoRA-based parameter-efficient fine-tuning. Other application scenarios could be IoT-based applications including smart homes, self-driving cars, and smartphone APPs.

6 Conclusions

This paper proposes federated intelligence which is a new domain to leverage many foundation models towards solving comprehensive problems and tackling complex environments. Based on the problem formulation, we introduced future directions to pave the way for implementing federated intelligence and its applications.

References

[Achiam *et al.*, 2023] Josh Achiam, Steven Adler, Sandhini Agarwal, Lama Ahmad, Ilge Akkaya, Florencia Leoni Aleman, Diogo Almeida, Janko Altenschmidt, Sam Altman, Shyamal Anadkat, et al. Gpt-4 technical report. *arXiv preprint arXiv:2303.08774*, 2023.

[Babakniya *et al.*, 2023] Sara Babakniya, Ahmed Roushdy Elkordy, Yahya H Ezzeldin, Qingfeng Liu, Kee-Bong Song, Mostafa El-Khamy, and Salman Avestimehr. Slora: Federated parameter efficient fine-tuning of language models. *arXiv preprint arXiv:2308.06522*, 2023.

[Brooks *et al.*, 2024] Tim Brooks, Bill Peebles, Connor Holmes, Will DePue, Yufei Guo, Li Jing, David Schnurr, Joe Taylor, Troy Luhman, Eric Luhman, Clarence Ng, Ricky Wang, and Aditya Ramesh. Video generation models as world simulators. 2024.

[Chen *et al.*, 2022] Fengwen Chen, Guodong Long, Zonghan Wu, Tianyi Zhou, and Jing Jiang. Personalized federated learning with a graph. In Lud De Raedt, editor, *Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence, IJCAI-22*, pages 2575–2582. International Joint Conferences on Artificial Intelligence Organization, 7 2022. Main Track.

[Chen *et al.*, 2023] Yiqiang Chen, Wang Lu, Xin Qin, Jindong Wang, and Xing Xie. Metafed: Federated learning among federations with cyclic knowledge distillation for personalized healthcare. *IEEE Transactions on Neural Networks and Learning Systems*, 2023.

[Chen *et al.*, 2024] Shutong Chen, Tianyi Zhou, Guodong Long, Jie Ma, Jing Jiang, and Chengqi Zhang. Multi-level additive modeling for structured non-iid federated learning. *arXiv preprint arXiv:2405.16472*, 2024.

[Chenthara *et al.*, 2019] Shekha Chenthara, Khandakar Ahmed, Hua Wang, and Frank Whittaker. Security and privacy-preserving challenges of e-health solutions in cloud computing. *IEEE access*, 7:74361–74382, 2019.

[Cui *et al.*, 2024] Yue Cui, Chung-ju Huang, Yuzhu Zhang, Leye Wang, Lixin Fan, Xiaofang Zhou, and Qiang Yang. A survey on contribution evaluation in vertical federated learning. *arXiv preprint arXiv:2405.02364*, 2024.

[Duan *et al.*, 2022] Jiafei Duan, Samson Yu, Hui Li Tan, Hongyuan Zhu, and Cheston Tan. A survey of embodied ai: From simulators to research tasks. *IEEE Transactions on Emerging Topics in Computational Intelligence*, 6(2):230–244, 2022.

[Ha and Tang, 2022] David Ha and Yujin Tang. Collective intelligence for deep learning: A survey of recent developments. *Collective Intelligence*, 1(1):26339137221114874, 2022.

[He *et al.*, 2024] Yuanqin He, Yan Kang, Lixin Fan, and Qiang Yang. Fedeval-llm: Federated evaluation of large language models on downstream tasks with collective wisdom. *arXiv preprint arXiv:2404.12273*, 2024.

[Heaven, 2023] Will Douglas Heaven. Google deepmind wants to define what counts as artificial general intelligence, 2023.

[Hu *et al.*, 2021] Edward J Hu, Yelong Shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, Lu Wang, and Weizhu Chen. Lora: Low-rank adaptation of large language models. *arXiv preprint arXiv:2106.09685*, 2021.

[Lewis *et al.*, 2020] Patrick Lewis, Ethan Perez, Aleksandra Piktus, Fabio Petroni, Vladimir Karpukhin, Naman Goyal, Heinrich Küttler, Mike Lewis, Wen-tau Yih, Tim Rocktäschel, et al. Retrieval-augmented generation for knowledge-intensive nlp tasks. *Advances in Neural Information Processing Systems*, 33:9459–9474, 2020.

[Li *et al.*, 2024] Zhiwei Li, Guodong Long, and Tianyi Zhou. Federated recommendation with additive personalization. In *The Twelfth International Conference on Learning Representations*, 2024.

[Liu *et al.*, 2019] Boyi Liu, Lujia Wang, and Ming Liu. Lifelong federated reinforcement learning: a learning architecture for navigation in cloud robotic systems. *IEEE Robotics and Automation Letters*, 4(4):4555–4562, 2019.

[Liu *et al.*, 2020] Yi Liu, Sahil Garg, Jiangtian Nie, Yang Zhang, Zehui Xiong, Jiawen Kang, and M Shamim Hossain. Deep anomaly detection for time-series data in industrial iot: A communication-efficient on-device federated learning approach. *IEEE Internet of Things Journal*, 8(8):6348–6358, 2020.

[Long *et al.*, 2020] Guodong Long, Yue Tan, Jing Jiang, and Chengqi Zhang. Federated learning for open banking. In *Federated Learning: Privacy and Incentive*, pages 240–254. Springer, 2020.

[Long *et al.*, 2023] Guodong Long, Ming Xie, Tao Shen, Tianyi Zhou, Xianzhi Wang, and Jing Jiang. Multi-center federated learning: clients clustering for better personalization. *World Wide Web*, 26(1):481–500, Jan 2023.

[Ma *et al.*, 2023] Jie Ma, Tianyi Zhou, Guodong Long, Jing Jiang, and Chengqi Zhang. Structured federated learning through clustered additive modeling. *Advances in Neural Information Processing Systems*, 36, 2023.

[Mansour *et al.*, 2020] Yishay Mansour, Mehryar Mohri, Jae Ro, and Ananda Theertha Suresh. Three approaches for personalization with applications to federated learning. *arXiv preprint arXiv:2002.10619*, 2020.

[Maynez *et al.*, 2020] Joshua Maynez, Shashi Narayan, Bernd Bohnet, and Ryan McDonald. On faithfulness and factuality in abstractive summarization. *arXiv preprint arXiv:2005.00661*, 2020.

[Molaei *et al.*, 2024] Soheila Molaei, Anshul Thakur, Ghazaleh Niknam, and et al. Federated learning for heterogeneous electronic health records utilising augmented temporal graph attention networks. In *International Conference on Artificial Intelligence and Statistics*, pages 1342–1350. PMLR, 2024.

[Ren *et al.*, 2024] Chao Ren, Han Yu, Hongyi Peng, Xiaoli Tang, Anran Li, Yulan Gao, Alysa Ziying Tan, Bo Zhao, Xiaoxiao Li, Zengxiang Li, et al. Advances and open challenges in federated learning with foundation models. *arXiv preprint arXiv:2404.15381*, 2024.

[Sun *et al.*, 2024a] Chuanneng Sun, Songjun Huang, and Dario Pompili. Llm-based multi-agent reinforcement learning: Current and future directions. *arXiv preprint arXiv:2405.11106*, 2024.

[Sun *et al.*, 2024b] Youbang Sun, Zitao Li, Yaliang Li, and Bolin Ding. Improving loRA in privacy-preserving federated learning. In *The Twelfth International Conference on Learning Representations*, 2024.

[Tan *et al.*, 2022a] Alysa Ziying Tan, Han Yu, Lizhen Cui, and Qiang Yang. Towards personalized federated learning. *IEEE Transactions on Neural Networks and Learning Systems*, 2022.

[Tan *et al.*, 2022b] Yue Tan, Guodong Long, Lu Liu, Tianyi Zhou, Qinghua Lu, Jing Jiang, and Chengqi Zhang. Fedproto: Federated prototype learning across heterogeneous clients. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 36, pages 8432–8440, 2022.

[Tiwari *et al.*, 2023] Prayag Tiwari, Abdullah Lakhan, and et al. Consumer-centric internet of medical things for cyborg applications based on federated reinforcement learning. *IEEE Transactions on Consumer Electronics*, 69(4):756–764, 2023.

[Touvron *et al.*, 2023] Hugo Touvron, Thibaut Lavril, Gautier Izacard, Xavier Martinet, Marie-Anne Lachaux, Timothée Lacroix, Baptiste Rozière, Naman Goyal, Eric Hambro, Faisal Azhar, Aurelien Rodriguez, Armand Joulin, Edouard Grave, and Guillaume Lample. Llama: Open and efficient foundation language models. *ArXiv*, abs/2302.13971, 2023.

[Verbraeken *et al.*, 2020] Joost Verbraeken, Matthijs Wolting, Jonathan Katzy, Jeroen Kloppenburg, Tim Verbelen, and Jan S Rellermeyer. A survey on distributed machine learning. *ACM Computing Surveys (CSUR)*, 53(2):1–33, 2020.

[Wang, 2019] Guan Wang. Interpret federated learning with shapley values. *arXiv preprint arXiv:1905.04519*, 2019.

[Yan and Long, 2023] Peng Yan and Guodong Long. Personalization disentanglement for federated learning. In *2023 IEEE International Conference on Multimedia and Expo (ICME)*, pages 318–323. IEEE, 2023.

[Yan and Long, 2024] Peng Yan and Guodong Long. Client-supervised federated learning: Towards one-model-for-all personalization. *arXiv preprint arXiv:2403.19499*, 2024.

[Yang *et al.*, 2024] Yiyuan Yang, Guodong Long, Tao Shen, Jing Jiang, and Michael Blumenstein. Dual-personalizing adapter for federated foundation models. *arXiv preprint arXiv:2403.19211*, 2024.

[Yi *et al.*, 2023] Liping Yi, Han Yu, Gang Wang, and Xiaoguang Liu. Fedlora: Model-heterogeneous personalized federated learning with lora tuning. *arXiv preprint arXiv:2310.13283*, 2023.

[Zhai *et al.*, 2024] Jiaqi Zhai, Lucy Liao, Xing Liu, Yueming Wang, Rui Li, Xuan Cao, Leon Gao, Zhaojie Gong, Fangda Gu, Michael He, et al. Actions speak louder than words: Trillion-parameter sequential transducers for generative recommendations. *arXiv preprint arXiv:2402.17152*, 2024.

[Zhang *et al.*, 2023] Chunxu Zhang, Guodong Long, Tianyi Zhou, Peng Yan, Zijian Zhang, Chengqi Zhang, and Bo Yang. Dual personalization on federated recommendation. In Edith Elkind, editor, *Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence, IJCAI-23*, pages 4558–4566. International Joint Conferences on Artificial Intelligence Organization, 8 2023. Main Track.

[Zhang *et al.*, 2024a] Chunxu Zhang, Guodong Long, Hongkuan Guo, Xiao Fang, Yang Song, Zhaojie Liu, Guorui Zhou, Zijian Zhang, Yang Liu, and Bo Yang. Federated adaptation for foundation model-based recommendations. In *IJCAI 2024*, 2024.

[Zhang *et al.*, 2024b] Chunxu Zhang, Guodong Long, Tianyi Zhou, Peng Yan, Zijjian Zhang, and Bo Yang. Graph-guided personalization for federated recommendation. In *ACM SIG KDD*, volume 2024, pages 1–9, 2024.

[Zhao *et al.*, 2024] Andrew Zhao, Daniel Huang, Quentin Xu, Matthieu Lin, Yong-Jin Liu, and Gao Huang. Expel: Llm agents are experiential learners. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38, pages 19632–19642, 2024.

[Zhuang *et al.*, 2023] Weiming Zhuang, Chen Chen, and Lingjuan Lyu. When foundation model meets federated learning: Motivations, challenges, and future directions. *arXiv preprint arXiv:2306.15546*, 2023.