Federated Learning for AI and LLMs for Science: Overview of the Activities at the NSF-AI Institute ICICLE

Talk at TPC Track C Workshop (@SC-Asia ‘24)

by

http://icicle.ai

Dhabaleswar K. (DK) Panda
The Ohio State University
E-mail: panda@cse.ohio-state.edu
http://www.cse.ohio-state.edu/~panda
Credits to all ICICLE Team Members!!

ICICLE Members Attending All-Hands-Meeting In-Person (Nov ‘23)
Special Credits (working in Federated Learning)!!

Wei-Lin Chou, Distinguished Assistant Professor of Engineering Inclusive Excellence, Computer Science and Engineering, The Ohio State University

Song Gao, The Ohio State University, Associate Professor, Director of Geospatial Data Science Lab, University Wisconsin-Madison

Yu Su, Distinguished Assistant Professor of Engineering Inclusive Excellence, Computer Science and Engineering, The Ohio State University

Huan Sun, Associate Professor and Endowed CoE Innovation Scholar, Computer Science and Engineering, The Ohio State University
Outline

• ICICLE Vision and Goals
• Research Challenges Addressed
• Highlights of Federated Learning Activities
• How to get engaged?
Computing has been evolving over the last three decades with multiple phases:

- Phase 1 (1975-): Scientific Computing/HPC
- Phase 2 (2000-): HPC + Big Data Analytics
- Phase 3: (2010-): HPC + AI (Machine Learning/Deep Learning)
Emergence of the Computing Continuum

1. Scientific Computing
2. Big Data & Data Science
3. Artificial Intelligence

HPC Systems & Data Centers

Clouds

On Field Sensors

Edge & Near Edge

HPC Systems & Data Centers

Computing Continuum
AI-Driven Digital Agriculture

https://ccag.tamu.edu/research-project/digital-agriculture/

AI-Driven Foodshed Supply Chain Management?

Which food supply chains will likely be affected by an approaching storm?
AI-Driven Animal Ecology

- **Basic science**: The focus of Animal Ecology is understanding the functioning and behavior of animals individually and in groups *in the context of environment* and evolution.

- **Science + translational**: 
  - Monitoring, understanding, and protecting biodiversity of the planet
  - Monitoring and understanding the impact of changing habitats on animals that live in them

- **Translational**: biodiversity conservation and mitigating the impact of climate change
Broad Challenge

Designing the next-generation intelligent cyberinfrastructure for a computing continuum with heterogenous resources that is usable in a plug-and-play manner by stakeholders to solve societal challenges?
Objectives: Intelligent CyberInfrastructure for Computing Continuum

Use Inspired Science Domains

Digital Agriculture
Smart Foodsheds
Animal Ecology

ICICLE: Intelligent CyberInfrastructure with Computational Learning in the Environment
Systems AI Foundational Research for CI

Intelligent Cyber Infrastructure
CI for AI
AI for “CI for AI”

On Field Sensors
Edge & Near Edge
Clouds
HPC Systems & Data Centers

Emerging Computing Continuum
Participation:
14 Organizations, 33 faculty, 41 staff, (58 PhD, 16 MS, 16 undergrad, 6 K-12) students & many Collaborators
Outline

- ICICLE Vision and Goals
- Research Challenges Addressed
- Highlights of Federated Learning Activities
- How to get Engaged?
Research Plan: Overall Vision

ICICLE
Intelligent CI

FOUNDATIONAL SYSTEMS AI

INTELLIGENT CYBERINFRASTRUCTURE

CI FOR AI
AI FOR CI-FOR-AI
SOFTWARE ARCHITECTURE AND DESIGN
VISUAL ANALYTICS FOR CI AND AI EXPLAINABILITY
PRIVACY, ACCOUNTABILITY AND DATA INTEGRITY

CO-DESIGN FOR USE INSPIRED SCIENCE
Thrust: Foundational Systems AI

Components address CI complexity and heterogeneity for plug-and-play

Knowledge Graphs
- Multimodal KG to encode & reason rich data modalities (e.g., camera trap)
- Auto construction
- Interplay with LLM and knowledge-based QA

Model Commons
- MINT to support ICICLE use cases, KG, and models
- Precise profiling
- Flex composition
- Versioning and provenance

Adaptive AI
- Context-aware
- Efficient update
- User-friendly adaptation process
- Adaptation of foundation models, conversational AI

Federated Learning
- Heterogeneity
- Context-aware
- Privacy-preserving and robustness
- Going beyond classification (GNN, foundation models)

Conversational AI
- KG- and model-commons-aware
- LLM-powered
- Grounding LLMs to the context
- Hallucination reduction
- Complex reasoning
Outline

• ICICLE Vision and Goals
• Research Challenges Addressed
• Highlights of Federated Learning Activities
• How to get Engaged?
Federated learning (FL) in ICICLE use cases

• Data in animal ecology, digital agriculture, and smart foodsheds are usually collected by individual stakeholders using distributed devices (e.g., drones, camera traps).

• There is a need to train powerful AI models, potentially using all the data and available computational resources.

• There is a need for data privacy, protection, and ownership.
Focus of ICICLE’s federated learning (FL) research

• **A versatile framework to address:**
  • Data heterogeneity (e.g., non-IID distribution) across users
  • Communication constraints
  • Training powerful but deeper models
  • Personalization needs

• **Integration with ICICLE use cases:**
  • Visualization and diagnosis
  • Analyzing geospatial resilience of multicommodity food flows
  • Conversational AI parser for effective question-answering
  • Training models for camera traps and digital agriculture
ICICLE’s FL framework

1. Bayesian ensemble [ICLR 21]
   - Model ensemble
   - Knowledge distillation
   - "Unlabeled" data

2. Decoupled training for personalization [ICLR 22]
   - Feature Extractor
   - G-Head
   - P-Head
   - Robust risk
   - Empirical risk
   - User’s context
   - Hypernetwork

3. Fixed BN for training DNNs with BNs in FL [submitted, ICML 24]
   - Batch Norm
   - Personalized Test
   - Balanced Test

4. Making pre-training applicable in FL [ICLR 23, AAAI 23]
   - Model
   - Predictor
   - Shared Weights
   - Similarity

5. Efficient update [CVPR 23]
   - Efficient update

   - Dataset
   - Fractal Pair Generation

7. Efficient adaptation [submitted, IJCAI 24]
   - Global aggregation
   - Local training
Highlight: FixBN for training deeper models in FL

- Our FixBN practice makes it possible to train deep neural nets with Batch Normalization (BN) in FL!

Results on ImageNet

<table>
<thead>
<tr>
<th>Method</th>
<th>Network</th>
<th>Acc. $\Delta_{BN}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>GN</td>
<td>ResNet18 (He et al., 2016)</td>
<td>33.33± 0.57</td>
</tr>
<tr>
<td>BN</td>
<td></td>
<td>48.30± 1.21</td>
</tr>
<tr>
<td>FixBN</td>
<td></td>
<td>52.43± 0.68 (+4.1)</td>
</tr>
</tbody>
</table>

Results on CIFAR

<table>
<thead>
<tr>
<th>Normalization Layer</th>
<th>Acc (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BN (Ioffe &amp; Szegedy, 2015)</td>
<td>53.97 ± 4.18</td>
</tr>
<tr>
<td>GN (Wu &amp; He, 2018)</td>
<td>59.69 ± 0.76</td>
</tr>
<tr>
<td>GN +WN (Qiao et al., 2019)</td>
<td>66.90 ± 0.81</td>
</tr>
<tr>
<td>LN (Ba et al., 2016)</td>
<td>54.54 ± 1.21</td>
</tr>
<tr>
<td>IN (Ulyanov et al., 2016)</td>
<td>59.76 ± 0.43</td>
</tr>
<tr>
<td>FIXUP (Zhang et al., 2019)</td>
<td>70.66 ± 0.24</td>
</tr>
<tr>
<td>FixBN (Ours)</td>
<td>76.56 ± 0.66</td>
</tr>
</tbody>
</table>

Results on Cityscapes (segmentation)

<table>
<thead>
<tr>
<th>Method</th>
<th>Backbone</th>
<th>Mean IoU $\Delta_{BN}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>GN</td>
<td>MobileNet-v2</td>
<td>43.2±0.33</td>
</tr>
<tr>
<td>BN</td>
<td>(Sandler et al., 2018)</td>
<td>48.9±0.36</td>
</tr>
<tr>
<td>FixBN</td>
<td></td>
<td>54.0±0.29 (+5.1)</td>
</tr>
<tr>
<td>GN</td>
<td>ResNet50 (He et al., 2016)</td>
<td>47.8±0.30</td>
</tr>
<tr>
<td>BN</td>
<td></td>
<td>52.6±0.38</td>
</tr>
<tr>
<td>FixBN</td>
<td></td>
<td>57.2±0.32 (+4.6)</td>
</tr>
</tbody>
</table>

ICICLE’s FL visualization tool

- A novel visualization tool to uncover the FL training trajectory, to help users monitor training progress

- The tool leads to new understandings:
  - Pre-training makes FL more robust to data heterogeneity, enabling convergence to the same minima

[Understanding Federated Learning through Loss Landscape Visualizations: A Pilot Study, NeurIPS 2022 Workshop]
ICICLE’s FL work for the resilience of food flows

• Understanding and measuring the **resilience of food supply networks** is a global imperative to tackle increasing food insecurity.

• However, the **complexity** of these networks, with their **multidimensional interactions and decisions**, presents significant challenges.
ICICLE’s FL work for the resilience of food flows

• We propose **FLEE-GNN** to enhance food supply network resilience analysis across geographical regions.

• **FLEE-GNN** combines the robustness and adaptability of GNNs with the privacy-conscious and decentralized aspects of FL.

[FLEE-GNN: A Federated Learning System for Edge-Enhanced Graph Neural Network in Analyzing Geospatial Resilience of Multicommodity Food Flows, ACM SIGSPATIAL International Workshop on AI for Geographic Knowledge Discovery]
ICICLE’s FL work for the resilience of food flows

Federated learning architecture for the edge-enhanced graph neural networks

Global EE-GNN
Secure Multi-Party Computation, Federated Averaging

#1 EE-GNN
#2 EE-GNN
#3 EE-GNN
#4 EE-GNN

West Region
South Region
Midwest Region
Northeast

[FLEE-GNN: A Federated Learning System for Edge-Enhanced Graph Neural Network in Analyzing Geospatial Resilience of Multicommodity Food Flows, ACM SIGSPATIAL International Workshop on AI for Geographic Knowledge Discovery]
ICICLE’s FL work for the resilience of food flows

Traditional methods: underestimate the global resilience (bluer)

Proposed methods: better performance (close to white colors)

[FLEE-GNN: A Federated Learning System for Edge-Enhanced Graph Neural Network in Analyzing Geospatial Resilience of Multicommodity Food Flows, ACM SIGSPATIAL International Workshop on AI for Geographic Knowledge Discovery]
ICICLE’s federated semantic parsing for QA

Proposed solutions (Lorar)

For ours (Lorar):

\[ p_i^t = \frac{\left| D_i \right| \Delta L_i^t}{\sum_{i \in C_t} \left| D_i \right| \Delta L_i^t} \]

Lorar adjusts each client’s contribution to the global model update based on its training loss reduction

Proposed benchmarks

<table>
<thead>
<tr>
<th>Domain</th>
<th>SQL Pattern count</th>
<th>Questions / unique query count</th>
<th>Unique tables / query</th>
<th>SELECTS / query</th>
</tr>
</thead>
<tbody>
<tr>
<td>Advising</td>
<td>2629</td>
<td>174</td>
<td>3.0</td>
<td>1.23</td>
</tr>
<tr>
<td>ATIS</td>
<td>4347</td>
<td>751</td>
<td>5.6</td>
<td>1.79</td>
</tr>
<tr>
<td>GeoQuery</td>
<td>549</td>
<td>17</td>
<td>1.1</td>
<td>1.77</td>
</tr>
<tr>
<td>Restaurants</td>
<td>228</td>
<td>17</td>
<td>2.3</td>
<td>1.17</td>
</tr>
<tr>
<td>Scholar</td>
<td>499</td>
<td>146</td>
<td>3.2</td>
<td>1.02</td>
</tr>
<tr>
<td>Academic</td>
<td>120</td>
<td>92</td>
<td>3.0</td>
<td>1.04</td>
</tr>
<tr>
<td>IMDB</td>
<td>78</td>
<td>52</td>
<td>1.5</td>
<td>1.01</td>
</tr>
<tr>
<td>Yelp</td>
<td>78</td>
<td>89</td>
<td>1.2</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Improved results

Reference

- On the Importance and Applicability of Pre-Training for Federated Learning, ICLR 2023
- Visual Query Tuning: Towards Effective Usage of Intermediate Representations for Parameter and Memory Efficient Transfer Learning, CVPR 2023
- FLEE-GNN: A Federated Learning System for Edge-Enhanced Graph Neural Network in Analyzing Geospatial Resilience of Multicommodity Food Flows, ACM SIGSPATIAL Workshop 2023
- Federated Learning of Shareable Bases for Personalization-Friendly Image Classification, arXiv:2304.07882
- On Bridging Generic and Personalized Federated Learning for Image Classification, ICLR 2022
- Understanding Federated Learning through Loss Landscape Visualizations: A Pilot Study, NeurIPS 2022 Workshop
- FedBE: Making Bayesian model ensemble applicable to federated learning, ICLR 2021
Outline

• ICICLE Vision and Goals
• Research Challenges Addressed
• Highlights of Federated Learning Activities
• How to get Engaged?
Multiple Levels of Collaboration and Engagement

• Using the Released Software/CI components
  – Available at https://icicle.osu.edu/cyberinfrastructure/software
  – Get engaged as a member in the Stakeholder Roundtable (more details below)

• Become a part of ICICLE (multiple options)
  – Student Associate
  – Visiting Research Fellow
  – Academic Collaborator
  – Industry Partner
  – Stakeholder Roundtable Member
  – More details at: https://icicle.osu.edu/engagement/join-us

• Join the ICICLE mailing lists (https://icicle.osu.edu/engagement/mailing-lists)
  – icicle-announce
  – icicle-discuss
An Expanded Presentation on ICICLE

Keynote Talk in the

AI-driven Infrastructure Track
Room C3.2
1:30 pm (Today, 02/21/24)
### Acknowledgments to all ICICLE Participants (Faculty, Students and Staffs)

#### Current Faculty
- E. Ayday, CWRU
- S. Blanais, OSU
- R. Machiraju, OSU
- Y. Su, OSU
- A. Ahmad, Uni Stuttgart
- S. Sanders, IU
- T. Jiang, UU
- J. Duarte, UC San Diego
- M. Norman, UC San Diego
- M. A. Rahman, IU
- T. Zhang, ISU
- C. Garcia, TACC
- P. Huber, UC Davis
- R. Estanislao, SDSC
- J. Yao, OSU
- S. Zac, TIH ITTB

#### Current Staff
- M. Lange, IC-Foods
- M. Abduljabbar, OSU – A. Shafi, OSU
- T. Ruemping, IC-Foods – K. Armstrong, OSU
- D. Siedband, IC-Foods – J. Chan, OSU
- M. Biggers, IUS – K. Tomko, OSC
- R. Ping, IU – C. Guzman, OSU
- B. Plale, IU – D. Choi, SDSC
- J. Wernert, IU – M. Kandes, SDSC
- N. Savardekar, OSU – A. Majumdar, SDSC – Z. Zhang, TACC

#### Current Ph.D. Students
- P. Kousha, OSU – C. Tu, OSU
- Z. Li, OSU – Y. Tu, OSU
- V. Pahuja, OSU – S. Vallabhabhajosula, OSU – X. Yue, OSU
- R. Qiu, OSU – L. Waltz, OSU
- E. Romero, OSU – B. Wang, OSU
- Z. Wang, OSU – Z. Zhang, OSU – J. Yan, IU
- M. Rosas, UD – J. Rao, UW-Madison
- T. Zhang, OSU – T. Jiang, IU
- K. Zhang, OSU – Y. Xu, OSU
- G. Ubbiali, IC-Foods

#### Current Masters Students
- R. Danhi, IC-Foods – C. Wang, OSU – S. Suresh, UW Madison
- J. Cheng, OSU – J. Yang, OSU – G. Wilkins, UW Madison
- S. Deshmukh, OSU – Q. Ding, TACC
- M. Han, OSU – V. deBella, UW Madison
- A. O’Quinn, OSU – M. Krempeley, UW Madison

#### Current Undergraduate Students
- T. Chen, OSU
- KA. Irizarry, OSU
- M. Lieber, OSU
- E. Luo, OSU
- D. Venkatakrishnan, OSU
- S. Shah, UT Austin
- A. Karunakaran, UW Madison
- M. Kuhn, UW Madison
- Y. Qu, UW Madison
- K. Sung, UW Madison

#### Current K-12 Students
- R. Estanislaos, SDSC
- D. Lee, SDSC
- S. Samar, SDSC
- J. Karpinski, SDSC

#### Past Staff
- C. Campbell, IU
- S. Sanders, IU
- A. Ivanovic, OSU
- P. Rose, UCSD
- K. Pierce, TACC

#### Past Faculty
- C. Hoy, OSU
- T. Tomich, UC Davis
- J. Duarte, UC San Diego
- M. Norman, UC San Diego
- A. Jain, OSU
- D. Suresh, OSU
- S. Raji, IU
- H. Park, UW Madison

#### Current International Faculty TIH – IITB
- M. Baghani, IITB
- Chalapathigiri, IITB
- A. Sinha, IITB
- R. Velmurugan, IITB
- S. Paramane, TIH ITTB

#### Current Institute Evaluators (WFD)
- T. McKlin, TFG
- C. Wise, TFG

#### Past Ph.D. Students
- FB Saravi, CWRU
- MK. Rahman, IU
- T. Zhang, ISU
- H. Ahn, OSU
- P. Chawla, OSU
- E. Goetz, OSU
- Y. Gu, OSU
- A. Jain, OSU
- D. Suresh, OSU
- S. Raji, IU
- H. Park, UW Madison

#### Past Masters Students
- SR. Kallal, OSU
- H. Panday, OSU
- RR. Loka, UW Madison
- D. Sykes, UW Madison

#### Past UG Students
- S. Okerman, OSU
- KP. Saha, OSU
- C. Washington, OSU
- J. Kim, TACC
- C. Skewfili, TACC
- S. Wegner, UW Madison

---

**SC-Asia ’24**
Thank You!