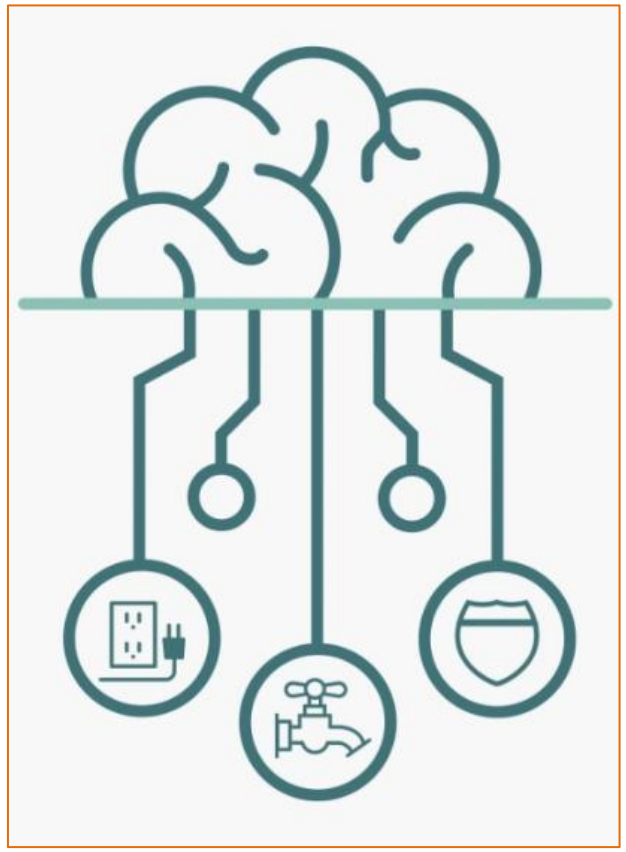


http://icicle.ai

NSF-Funded AI Institute

ICICLE

Intelligent CyberInfrastructure with Computational Learning in the Environment

Dhabaleswar K. (DK) Panda (OSU)

Vipin Chaudhary (CWRU)

Amit Majumdar (SDSC)

Joe Stubbs (TACC)

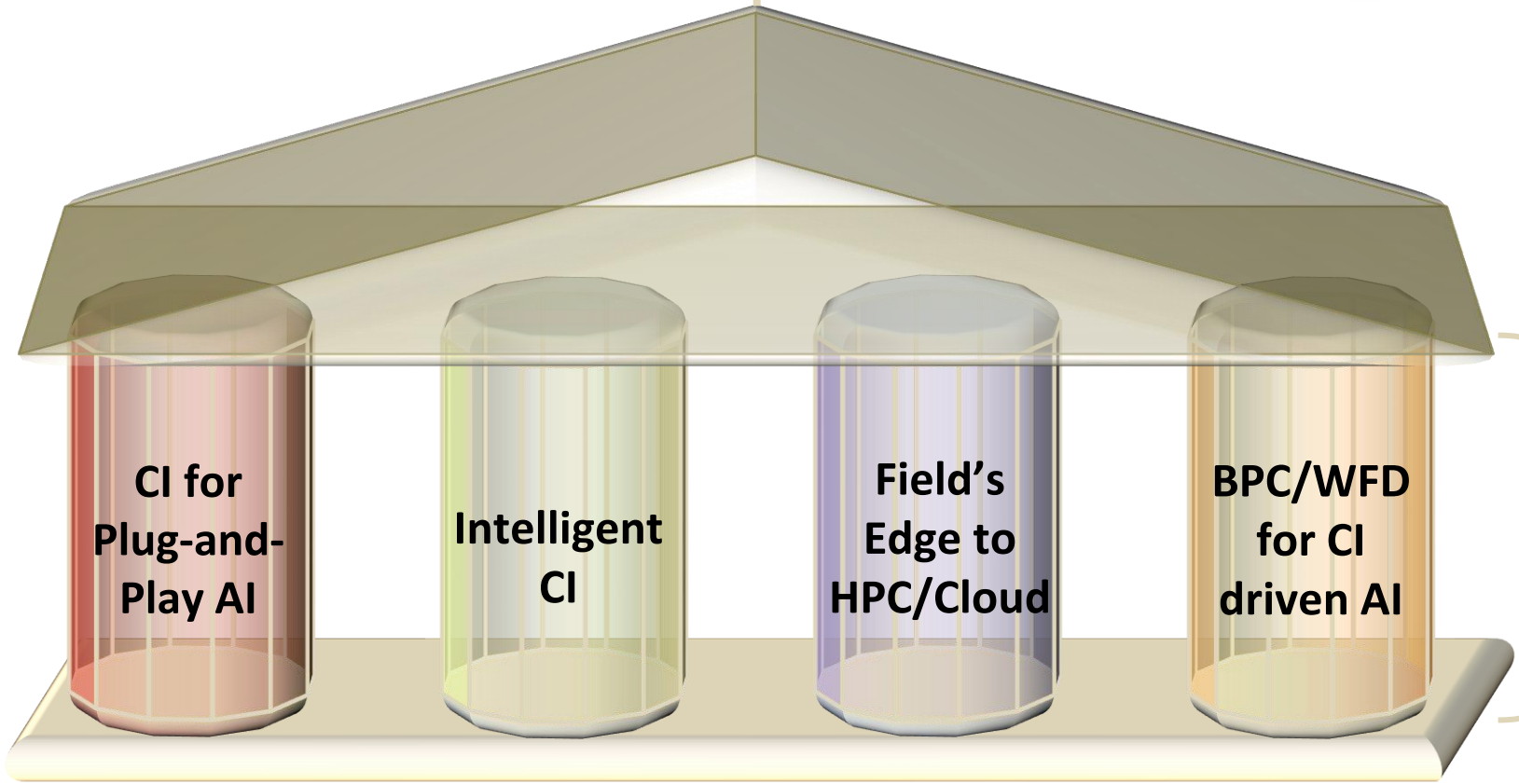
Summary and Vision

A **national infrastructure** that enables artificial intelligence at the flick of a switch: ICICLE will transform today's AI landscape from a narrow set of privileged disciplines to one where **democratized AI** empowers domains broadly through **integrated plug-and-play AI**. Converging under one virtual roof, ICICLE will foster and advance communities: catalyzing **foundational AI and CI** and transforming application domains. ICICLE will enable a **transparent and trustworthy** national infrastructure for an AI-enabled future, addressing pressing **societal problems** and enabling decisions for **national priorities**. ICICLE will grow **new generations of workforce** and incubate innovative companies with **sustained diversity and inclusion** at all levels.

ICICLE will build a brighter future for all.

Main Thrusts

Use-Inspired Science
(Smart Foodsheds,
Animal Ecology, Digital Agriculture)



Education and Outreach

Collaboration and Knowledge Transfer

Participating Institutes

Govt. Agencies & National Labs

Lawrence Livermore National Laboratory

U.S. Fish & Wildlife Service

Brookhaven National Laboratory

Fermilab

International

Max-Planck-Gesellschaft

University of Stuttgart
Germany

Industry

CNH Industrial

ADC

WildMe

Microsoft

Farm 2 Facts

The Findings Group, LLC
Research & Evaluation

Research Institutes

national center for women & information technology

MolSSI

SGCI
Science Gateways Community Institute

Trusted CI
The NSF Cybersecurity Center of Excellence

ECEP

CAHSI
Computing Alliance of Hispanic-Serving Institutions
A National INCLUDES Alliance

SCEC
An NSF+USGS Center

NSF

neon
Operated by Battelle

Jetstream

NSF AI Institutes

AIFS
AI Institute Food Systems

Hospitals & Universities

Nationwide Children's

University of North Carolina Chapel Hill

ICICLE Leadership Team

OVERALL LEADS


Fosler-Lussier (Co-PI) (OSU)


Machiraju (Co-PI) (OSU)


Panda (PI) (OSU)


Chaudhary (Co-PI) (CWRU)

Managing Director (TBD)

EDUCATION & WORKFORCE DEVELOPMENT


Lange (IC-FOODS)

BROADENING PARTICIPATION IN COMPUTING


Biggers (IU)


Thomas (SDSC)

FOUNDATIONAL SYSTEMS AI


Fosler-Lussier (OSU)


Su (OSU)

INTELLIGENT CYBERINFRASTRUCTURE

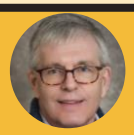
CI FOR AI


Blanas (OSU)


Zhang (TACC)

AI FOR CI-FOR-AI


Chaudhary (CWRU)


Eigenmann (UDel)

SOFTWARE ARCHITECTURE AND DESIGN


Stubbs (TACC)


Zhang (TACC)

VISUAL ANALYTICS FOR CI AND AI EXPLAINABILITY


Machiraju (OSU)


Shen (OSU)

PRIVACY, ACCOUNTABILITY AND DATA INTEGRITY


Ayday (CWRU)

CO-DESIGN FOR USE INSPIRED SCIENCE


Berger-Wolf (OSU)


Hoy (OSU)


Shearer (OSU)

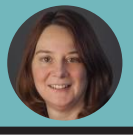
COLLABORATION & KNOWLEDGE TRANSFER


Lange (IC-FOODS)


Ramnath (OSU)

CI DEVELOPMENT FOR INTEGRATION WITH SHARED RESEARCH FACILITIES


Majumdar (SDSC)

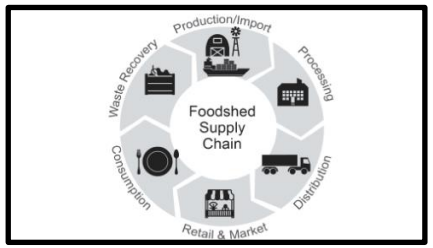

Tomko (OSU)

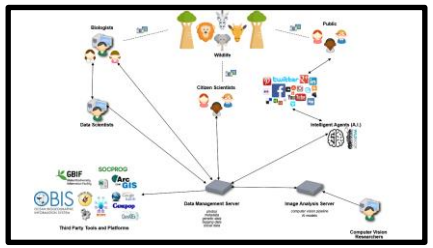
An Overview of ICICLE

- *Plug-and-play characteristics exists in current generation electricity and power grid*
- *No such plug-and-play AI exists for modern scenarios (e.g., food distributor, farmer)*
- ***Can we democratize AI in these modern scenarios?***
- *How do we assemble computational and human resources needed to build next generation CIs (ca. **2030**) capable of supporting AI across diverse use cases*?
- Democratizing AI will require **new Cyberinfrastructure** that enables:
  - **Plug-and-play AI** capabilities that are accessible to diverse stakeholders
  - Intelligent Cyberinfrastructure – **CI4AI & AI for CI4AI**
  - Advances in **foundational AI** that support
    - **Model commons**
    - **Conversational AI**
    - **Privacy-preserving AI**
    - **Adaptive AI**
  - Use-inspired research for **co-design in select target domains**
- Inclusive growth of **next generation of AI-capable workforce** trained in **transdisciplinary** settings

ICICLE: Computational Learning in the Environment

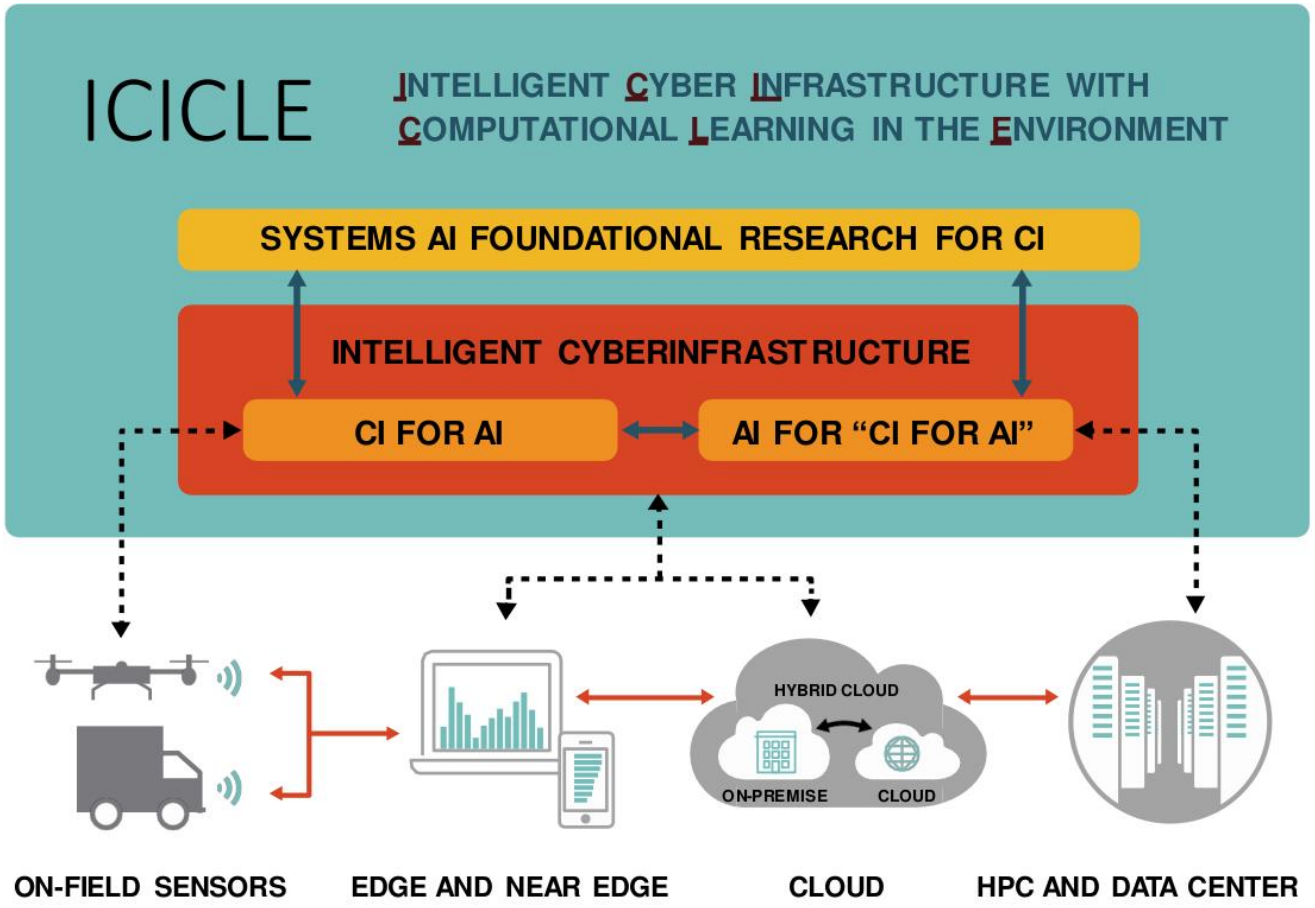
Driving Use Cases from Smart Foodsheds, Animal Ecology, and Digital Agriculture

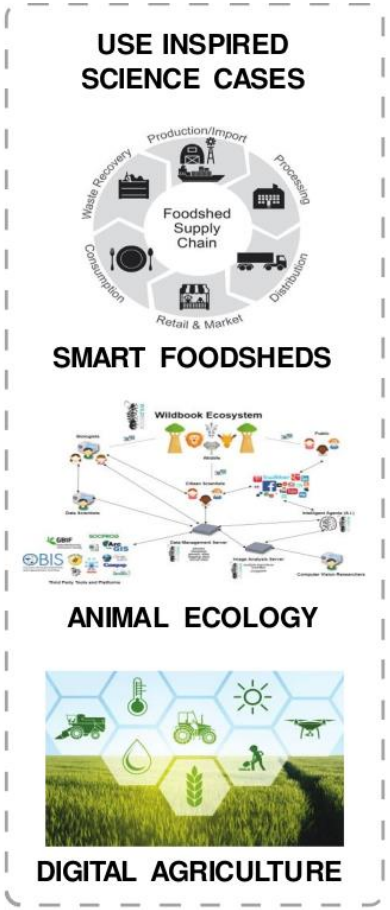




- Systems provide a continuum of **field-to-edge-to-cloud/HPC centers**
- Provide very **large, complex, heterogeneous data** for a plethora of scientific and operational questions
- Targeted domains are not well supported by AI-CI
- ICICLE will augment current cloud-based AI models by:
  - Facilitating **decision-making in the field** plagued by low network bandwidth
  - Allowing data to be **private** but facilitating **collaborative intelligence**
  - Provisioning models to use and **plan computation and data movement**

The Rationale of ICICLE





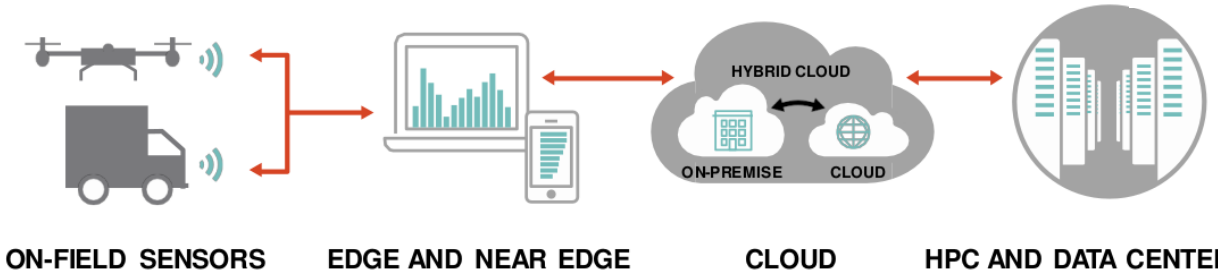
Integrating a broad range of

- Scientists-in-the-field
- Engineers
- Educators
- Collaborative partners
- Institutions

under one roof enables **democratized, adaptable, plug-and-play AI** and **long-tail science.**

Key Value Propositions of ICICLE

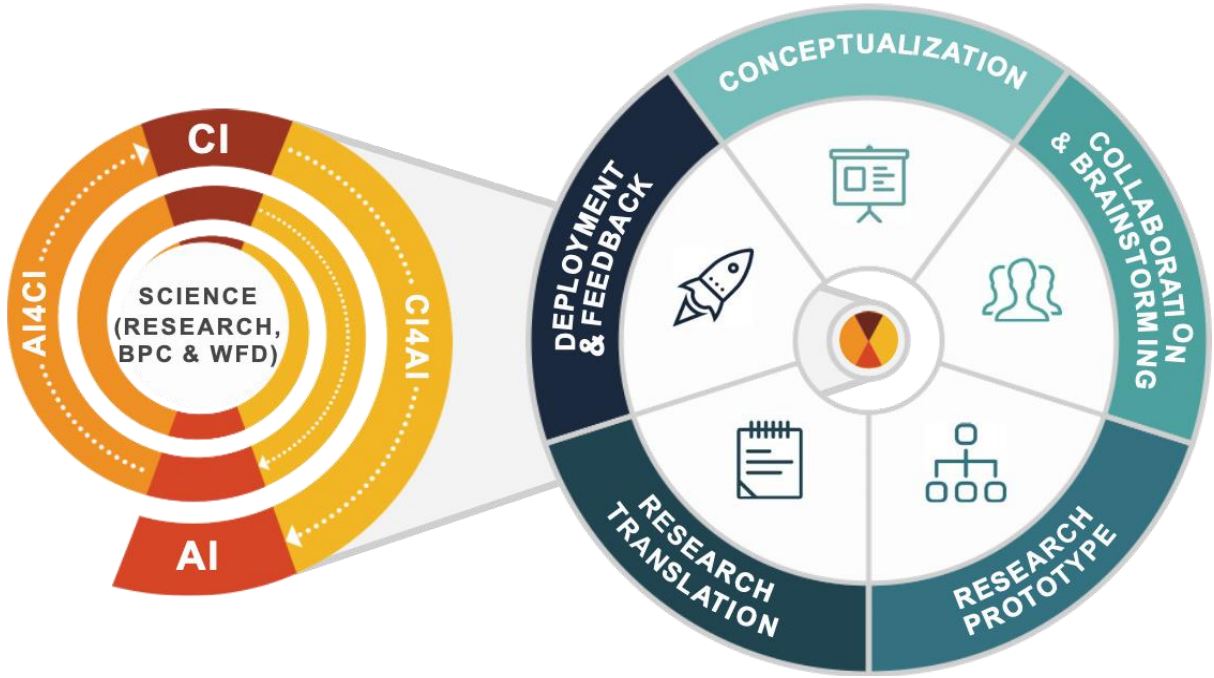
Democratizing AI



- Establish a **national cyberinfrastructure** for AI
- Develop **rubrics for training** next generation of researchers who can translate from use cases to AI-powered CI
- Integrate **emerging AI technologies** with advanced CI capabilities
- Design a **roadmap** for future AI-driven science and cyberinfrastructure
- Build a **nexus of collaboration** among AI, CI, and domain sciences

Realizing Key Value Propositions of ICICLE

The virtuous cycle: today's AI is tomorrow's CI



Innovation cycle **driven by use-inspired co-design**

Transdisciplinary teams will:
- **conceptualize problems** across use cases
- **collaboratively generalize solutions**
- **co-develop prototypes** for CI

Generalized AI infrastructure will **provide increasing research support** to other AI institutes

Research Plan: Overall Vision

How could dozens of supercell tornadoes spreading over the Great Plains and Midwest impact food supply chains over the coming months?

Which model is appropriate?
How do I even start navigating the choices?
Can I customize the model?
How do I get models to work together?
Can I keep my data secure in using a model?
Can I use these models in the field?

Current AI Infrastructure

8 Weather Forecasting Models
210 Crop Prediction Models

Cloud AI/HPC

aws
Microsoft Azure
IBM Watson

Research Plan: Overall Vision

How could dozens of supercell tornadoes spreading over the Great Plains and Midwest impact food supply chains over the coming months?

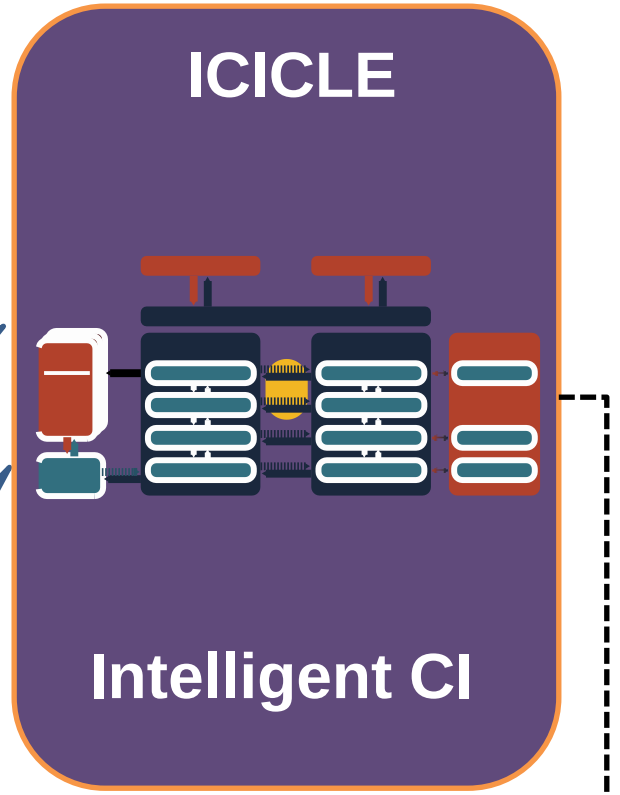
How did you arrive at this conclusion?

Can we update this model with new information we'll get onsite with TruckingCo tomorrow?

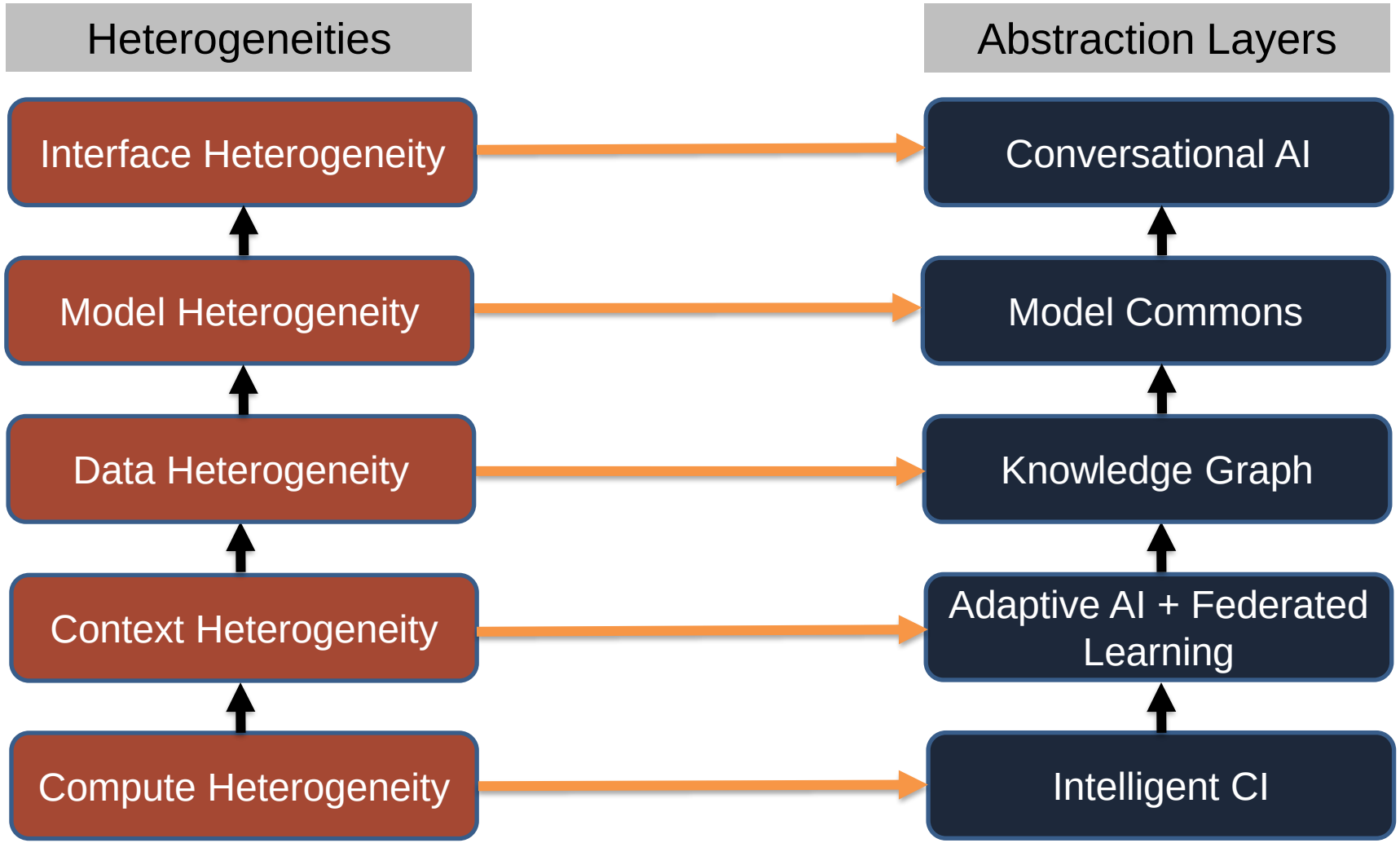
Happy to help! I see that you are in *the Midwest* and *strawberry* is in your supply chain. According to my weather forecasting and crop yield prediction models, I think strawberry yield will likely see a drop up to 50%.

The intermediate reasoning results and prediction rationale are visualized below.

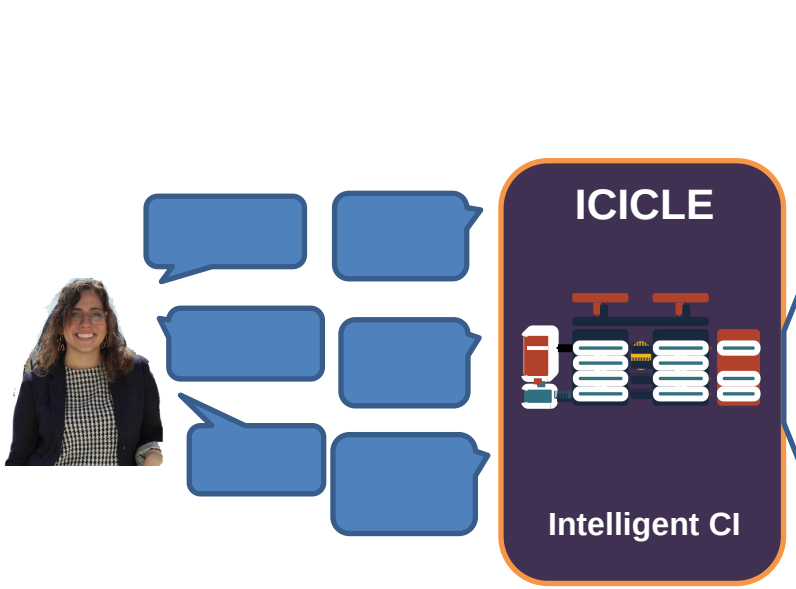
Sure. I'll have a smaller model ready for your laptop that can be updated with federated data from TruckingCo.



Research Plan: Overall Vision



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

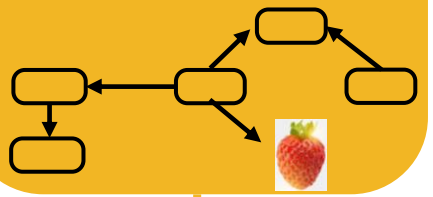
PRIVACY, ACCOUNTABILITY AND DATA INTEGRITY

CO-DESIGN FOR USE INSPIRED SCIENCE

Research Plan: Foundational Systems AI

Knowledge Graphs

- Multimodal
- Spatio-temporal
- Auto construction
- Knowledge-based reasoning and pre-training



Azad (IU), Gao (Wisc), Lange (IC-FOODS), Riloff (Utah), Salimi (UCSD), Su (OSU), Sun (OSU)

Model Commons

- KG-supported
- Precise profiling
- Flex Composition
- Versioning and provenance

Crop Yield Model
Midwest
Corn, Strawberry…

Hyder (OSU), Gao (Wisc), Machiraju (OSU), Su (OSU)

Adaptive AI

- Context-aware
- Interactive
- Continual learning
- Distillation-based compression

Chao (OSU), Gao (Wisc), Su (OSU), Salimi (UCSD), Sun (OSU)

Federated Learning

- Heterogeneity
- Applicability to a variety of models
- Context-aware
- Privacy-preserving and robustness

Ayday (CWRU), Chao (OSU), Su (OSU)

Conversational AI

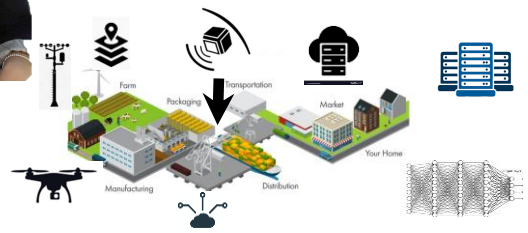
- KG- and model-commons-aware
- Bootstrapping and adaptivity
- Multimodal contextual response

Fosler-Lussier (OSU), Riloff (Utah), Su (OSU), Sun (OSU)

https://pixabay.com/photos/strawberry-field-fruit-plant-3630814/
https://commons.wikimedia.org/wiki/File:Strawberry_field_at_Bedugul.jpg

Research Plan: CI4AI: Interface

How could dozens of supercell tornadoes spreading over the Great Plains and Midwest impact food supply chains over the coming months?



ICICLE

- Conversational AI
- Model Commons
- Knowledge Graph
- Adaptive AI
- Federated Learning

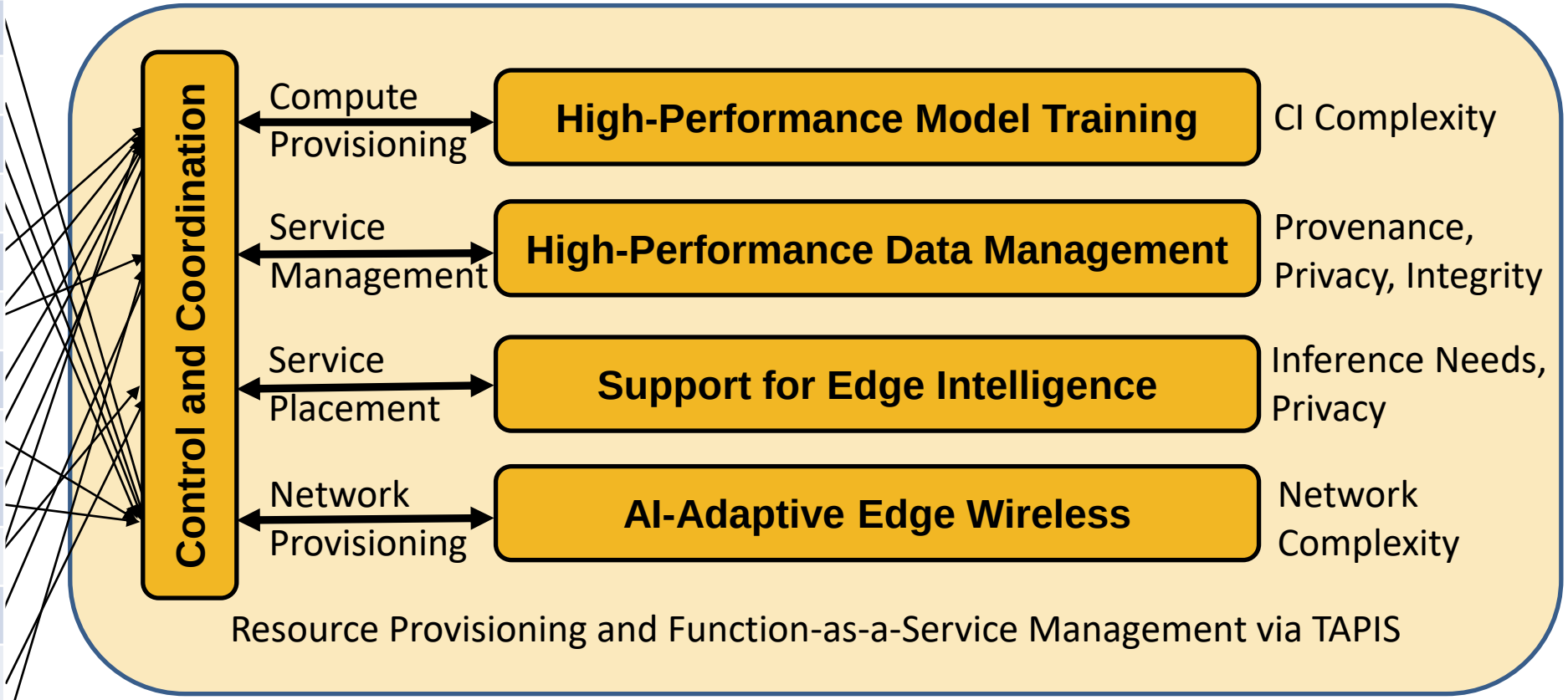
Predicting future states of your regional supply chain under forecasted weather scenarios.

REQUIREMENT	VALUE
Data	Data://user/satellite-images
Data Size	16 GB
Data Update Int.	Every night
Inventory Data	Data://producers/; Data://processors;
Pre-trained model	Weather Forecast Model
Model definition	Model-commons://path/to/model
Model to train	Disruption Prediction and Mitigation
Training Freq.	Every two days
Training Deadline	2 hours
Val Accuracy	80%
Infer. Latency	10 ms
Privacy	Private Data Access
Provenance	Enabled

CI4AI

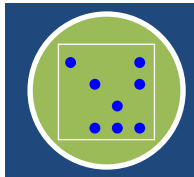
Research Plan: CI4AI: Resource Management

REQUIREMENT	VALUE
Data	Data://user/…
Data Size	16 GB
Data Update Int.	Every night
Inventory Data	Data://produ…
Pre-trained model	Weather …
Model definition	Model-c …
Model to train	Rein. Lng. …
Training Freq.	Every two …
Training Deadline	2 hours
Val Accuracy	80%
Infer. Latency	10 ms
Privacy	Private …
Provenance	Enabled



Research Plan: AI4CI

- ***Efficient* plug-and-play:** Constantly adapt and optimize heterogenous (cloud, HPC, and edge) CI to meet requirements of ICICLE applications, including digital agriculture and wildlife detection



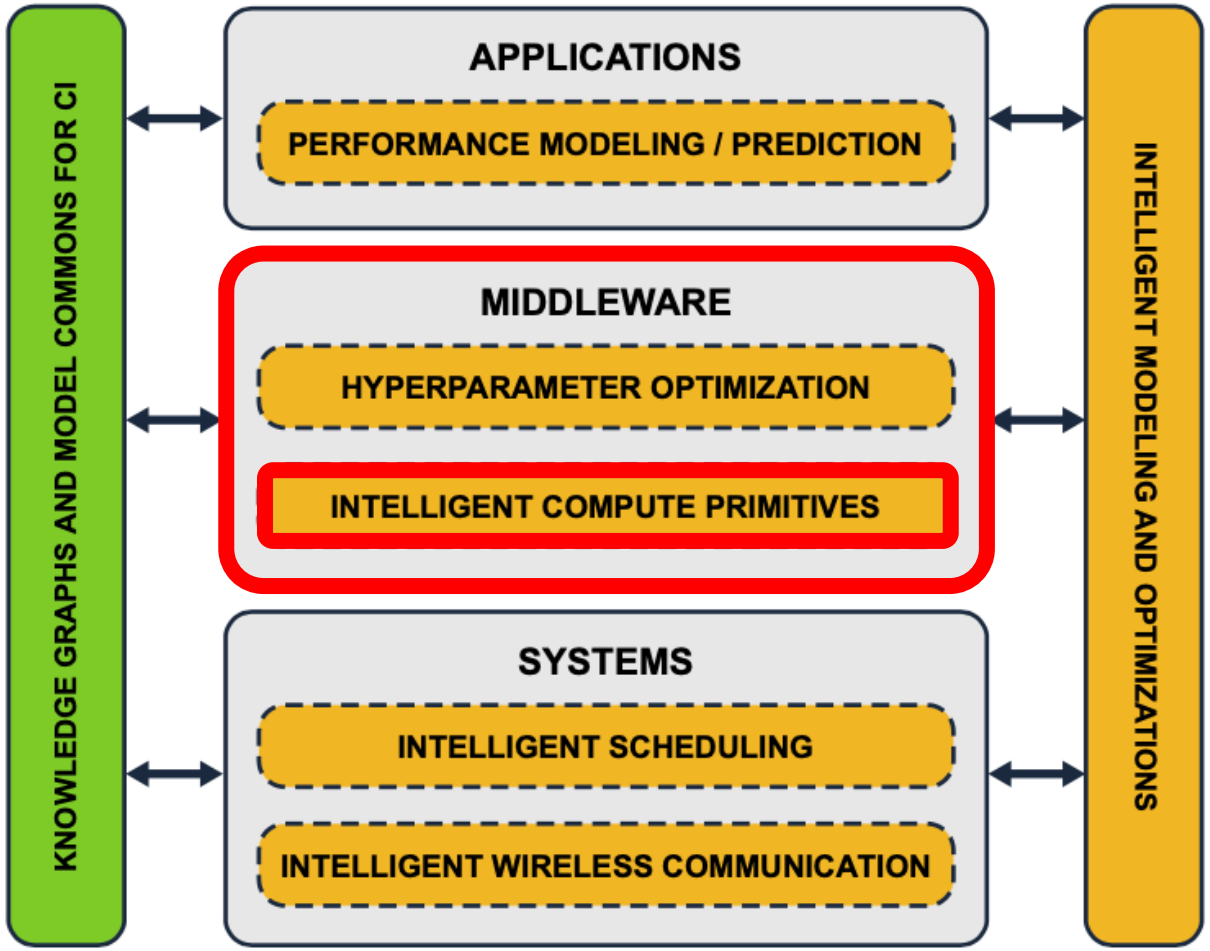
Intelligent Primitives (sparse/dense)
Product: High Performance Library
(Powered by AI to maximize the utilization of CI)

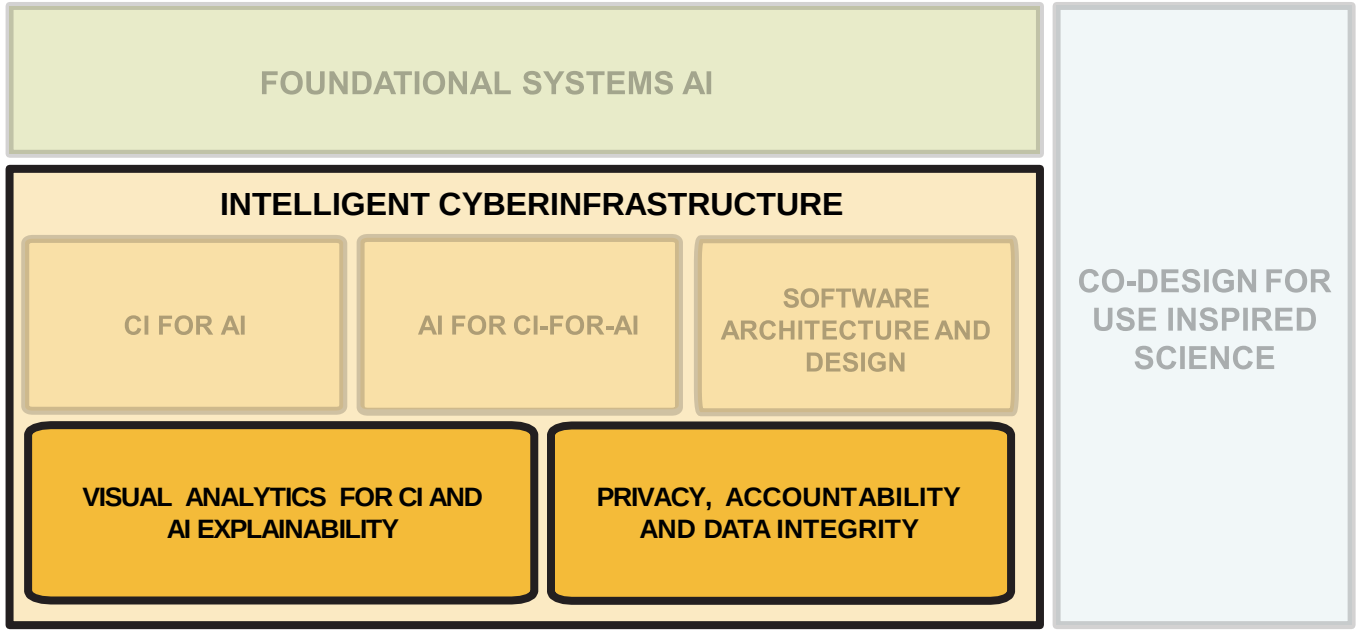
Innovation: **Learn** from hardware (e.g., SIMD width), network (e.g., bandwidth/latency), and data sparsity to **extract best attainable performance**

Utility: (1) **Portable** high performance on diverse HPC systems (2) **Usable** at the backend of any other AI system

Heterogeneous and rapidly evolving platform (CPUs, GPUs, Supercomputers, edge devices)

Chaudhary (CWRU), Zhang Z. (TACC), Azad, Sharma (IU), Duarte (UCSD), Majumdar, Tatineni (SDSC), Zhang H. (ISU), Chao, Ramnath, Shafi, Su (OSU), Tomko (OSC) Eigenmann (U. Delaware), Sadayappan (U. Utah)





RESEARCH PLAN → PRIVACY, ACCOUNTABILITY, AND DATA INTEGRITY (PADI) & EXPLAINABILITY

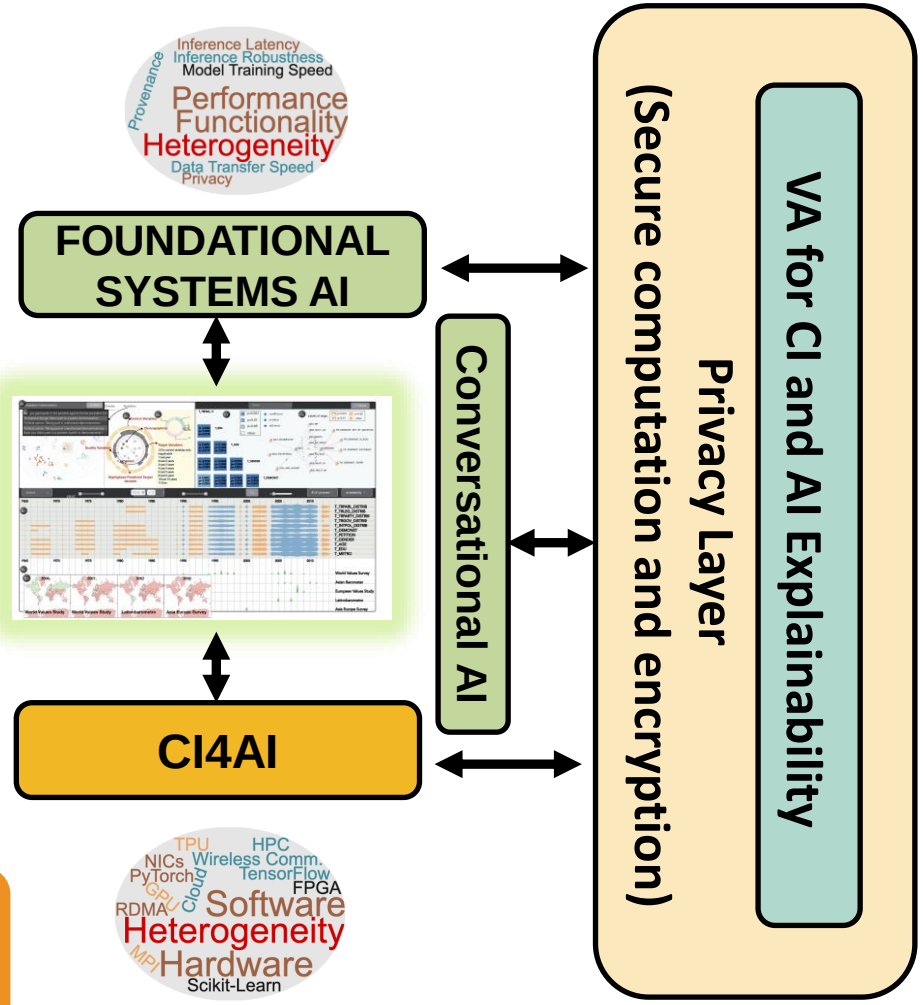
Research Plan: Privacy, Accountability, and Data Integrity (PADI)

- **Privacy**
  - **Personal clouds** to provide data owners full control of their datasets
  - **Fine-grained access control** during data exchange
  - **Privacy-enhancing techniques** for data movement and sharing
  - **Privacy risk quantification tools** to guide data sharing decisions
- **Accountability**
  - **Deep AI audit trails** to expose information about potential causal flaws
  - Integration into the ICICLE KGs for rich contextualization
- **Data Integrity**
  - Techniques to ensure the **credibility/integrity** of the data
  - Techniques to transform diverse/noisy data streams into reliable data

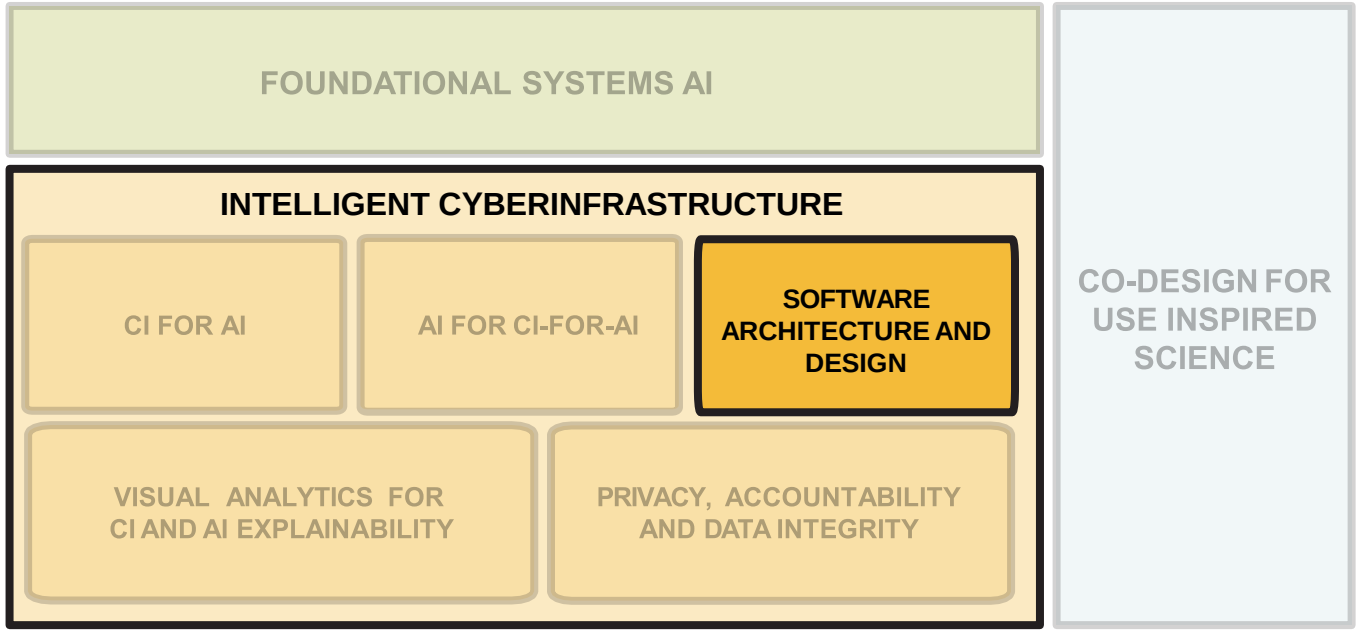
Ayday (CWRU), Su (OSU), Stubbs (TACC), Tomko (OSC)

Research Plan: Visual Analytics (VA) for CI and AI Explainability

- VA focuses on
  - **Performance** of AI and CI components
  - **Interaction** of AI and CI
  - **Decision path** of AI models
  - **Adapting to environment** of AI and CI components
- VA approach
  - **Multi-level:** performance from entire ICICLE system
  - **Multi-faceted:** heterogeneous and dynamic environments
  - **Multi-modal:** interface for a wide range of displays
  - Interface with **conversational AI**
- Tightly coupled with the **privacy layer**



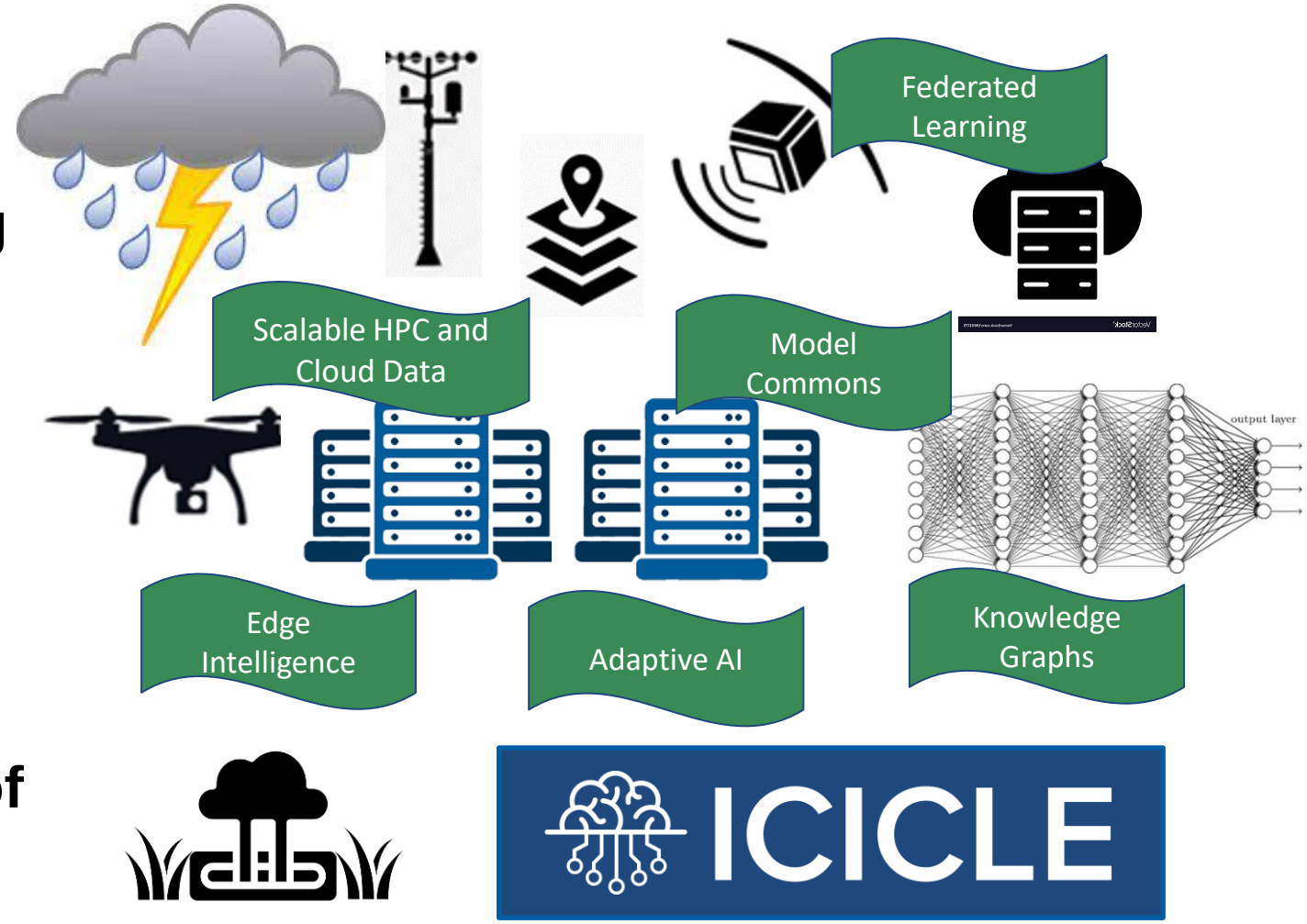
Berger-Wolf (OSU), Fosler-Lussier (OSU), Hyder (OSU), Machiraju (OSU), Salimi (UCSD), Shen (OSU), Su (OSU), Subramoni (OSU), Zhang Z. (TACC)



RESEARCH PLAN → SOFTWARE

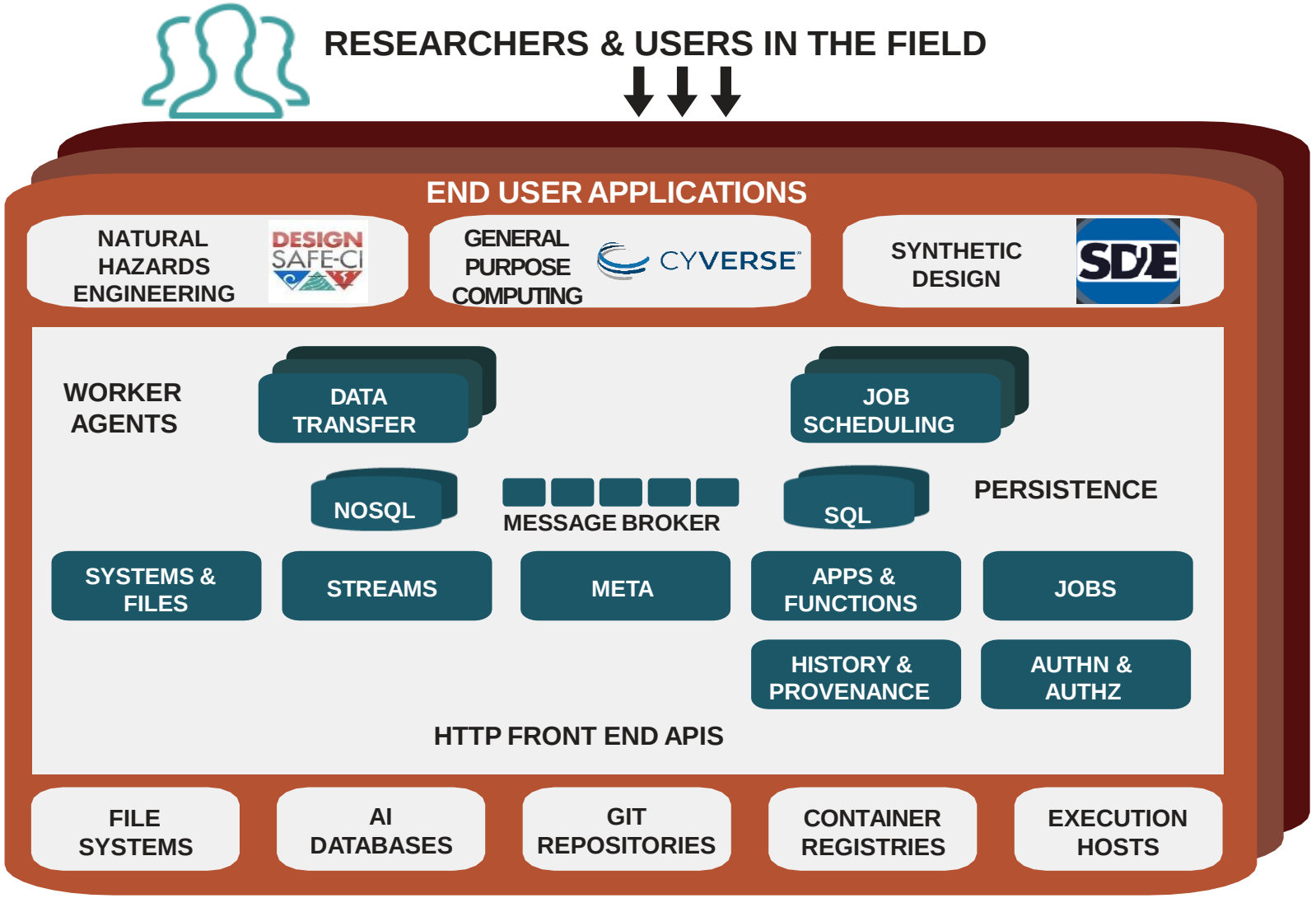
Research Plan: ICICLE Platform and Requirements

- **Scalable** cloud and HPC data storage and movement
- High performance model training for ML and DL with **dynamic resource provisioning**
- Interfaces for Model Commons, Knowledge Graphs, and Conversational AI
- Federated security architecture enabling users, projects and institutions to maintain **control of digital assets**



Oottikkal (OSC), Ramnath (OSU), Subramoni (OSU), Stubbs (TACC), Tatineni (SDSC), Zhang Z. (TACC)

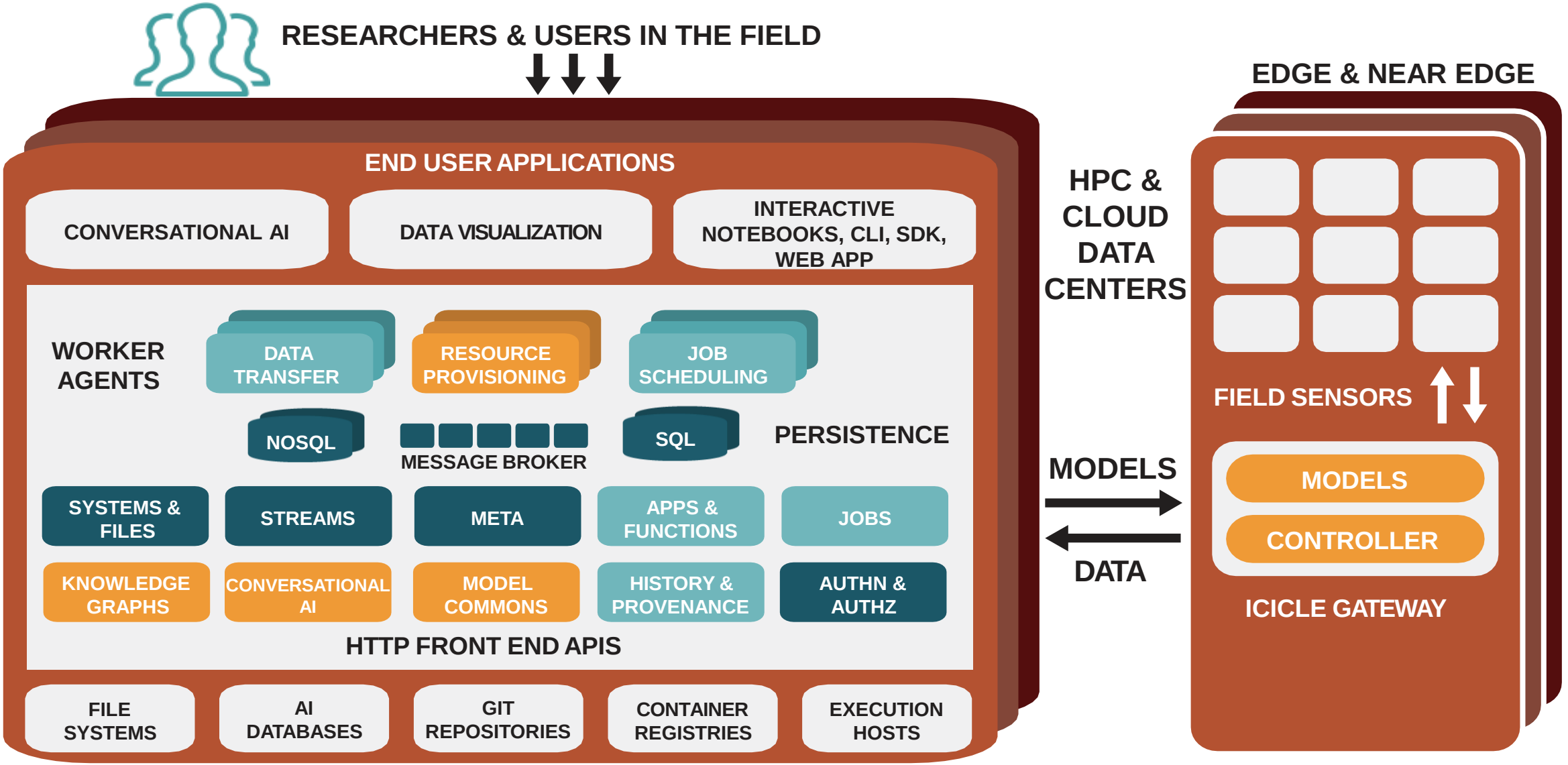
Research Plan: TAPIS API Framework



tapis

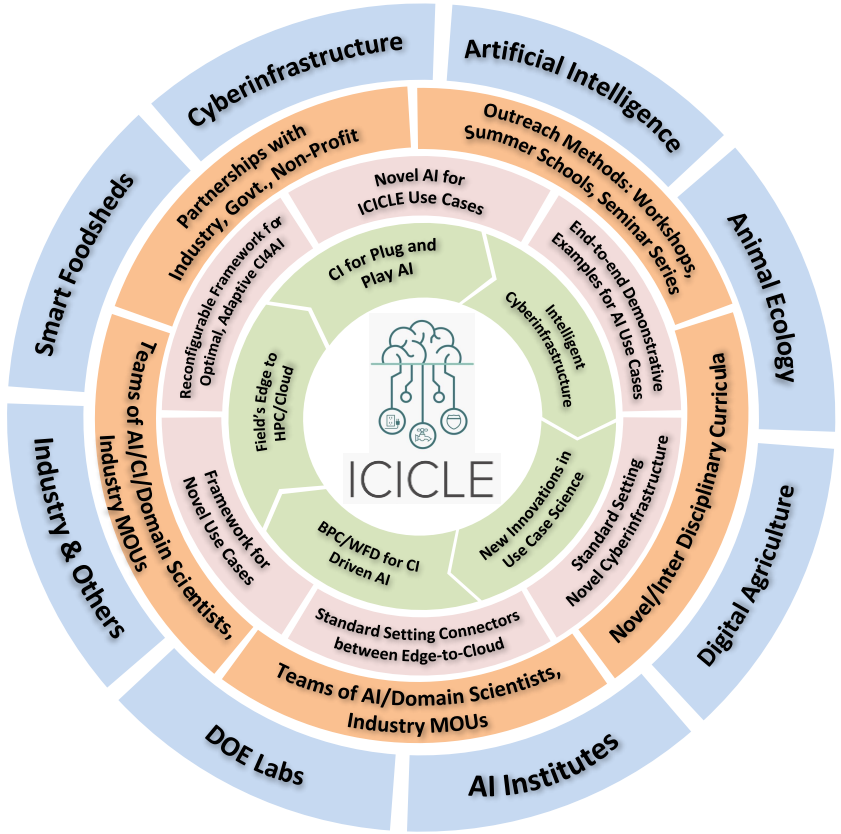
- API Framework for reproducible research computing
- Used by dozens of projects, including large CIs
- ~1M jobs run, 3B files transferred, 50K OAuth clients registered since 2015
- NSF CSSI Framework grant, 2019-2024
- Meets many requirements of ICICLE platform

Research Plan: Complete ICICLE Software Architecture



CONCLUSION

ICICLE will enable global leadership



- **Integrates into the National CI Ecosystem**
  - Short head and long tail science
  - Emerging applications
  - Maintains global leadership
- **Integrative and Interoperable**
  - Support across all CI components and emerging technologies
  - CI elements seamlessly composed
- **Leverages existing recognized capabilities**
  - Centers of Excellence, AI Institutes, Large Facilities
- **Collaborative**
  - Actively engage CI experts and domain scientists/CI users.
- **Sustainable**
  - Workforce Development, Broadening Participation, Collaboration and Knowledge Transfer
  - Benefits other institutes, large facilities, and all sciences beyond lifetime of award

Thank You