# A Common Framework for Scalable Autonomous Science AI

Kyle Chard
Argonne National Laboratory

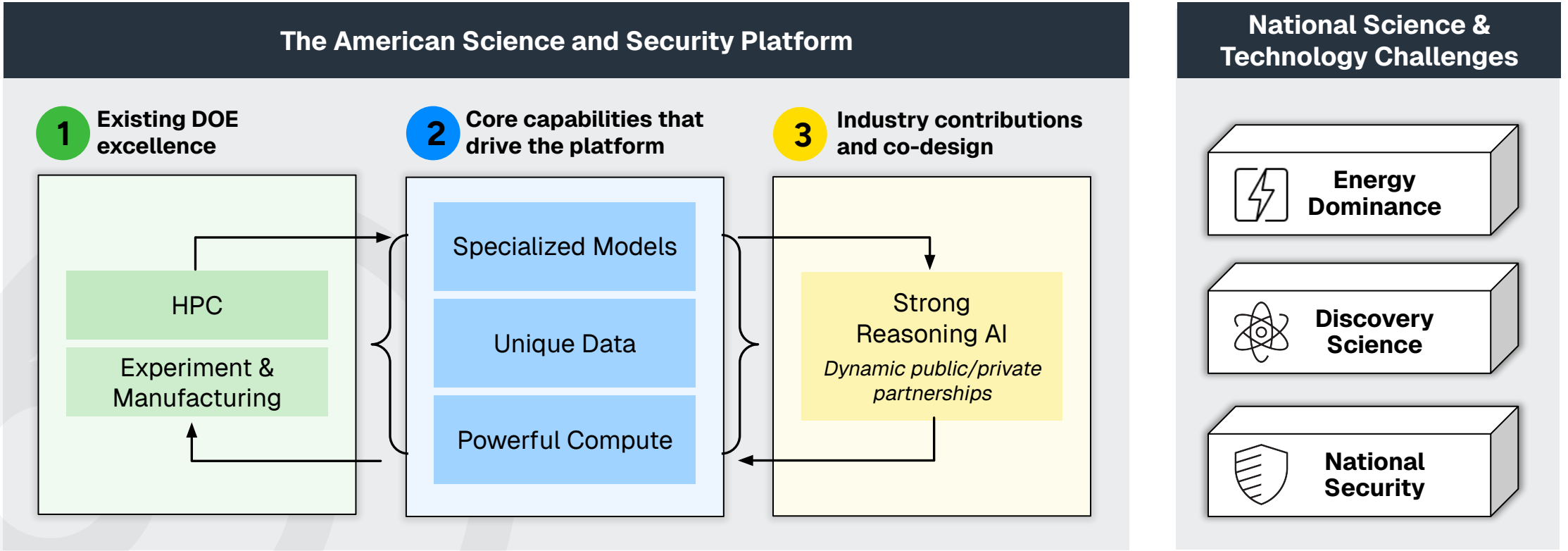
Argonne NATIONAL LABORATORY

Genesis Mission

U.S. DEPARTMENT of ENERGY

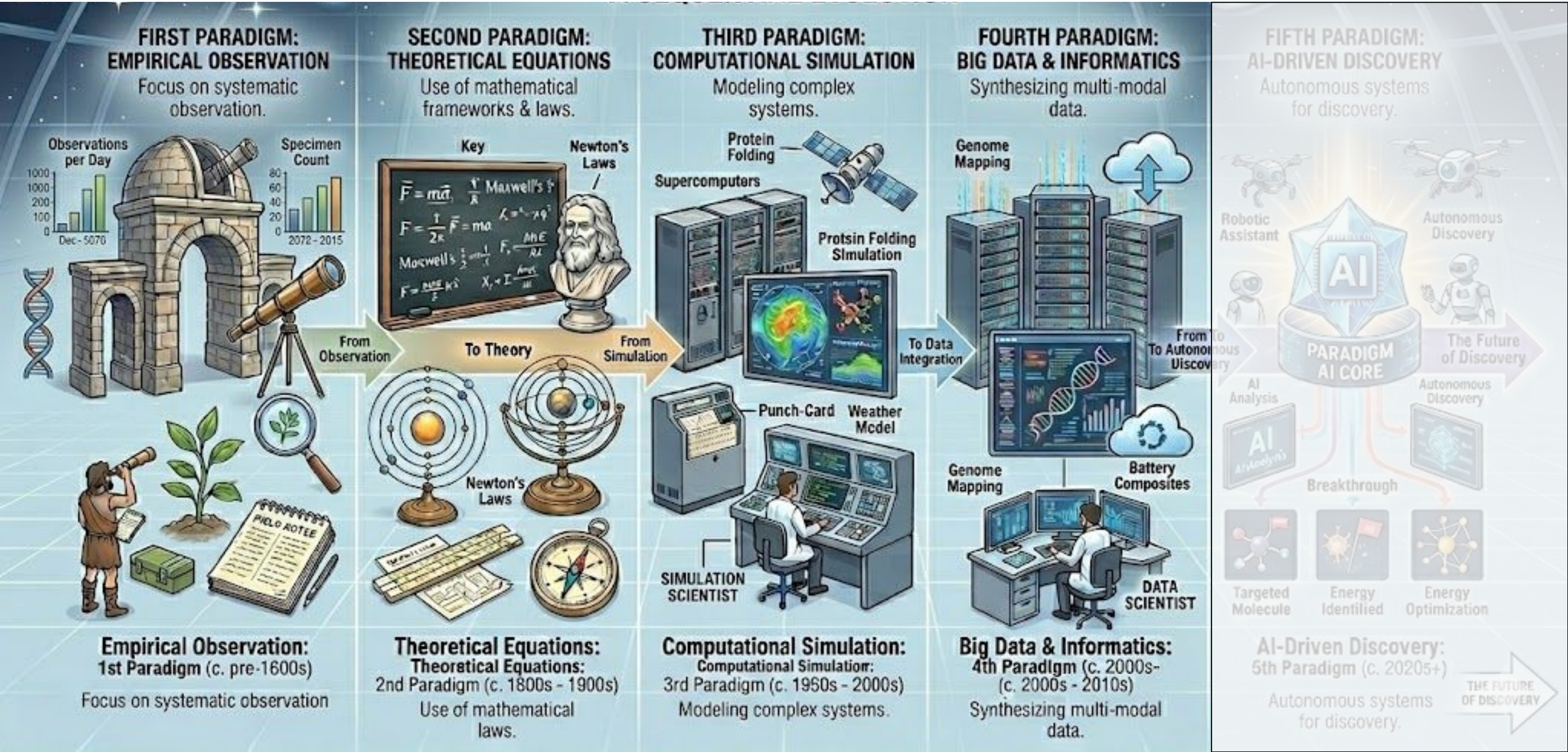
Genesis Mission

## The Genesis Mission is a DOE-led national initiative to build world-leading AI infrastructure and apply it to the nation's most pressing science and technology challenges.




## The platform enables the breakthroughs. The National Challenges define what we're solving for. Combined, we aim to double the productivity and impact of American R&D within a decade.

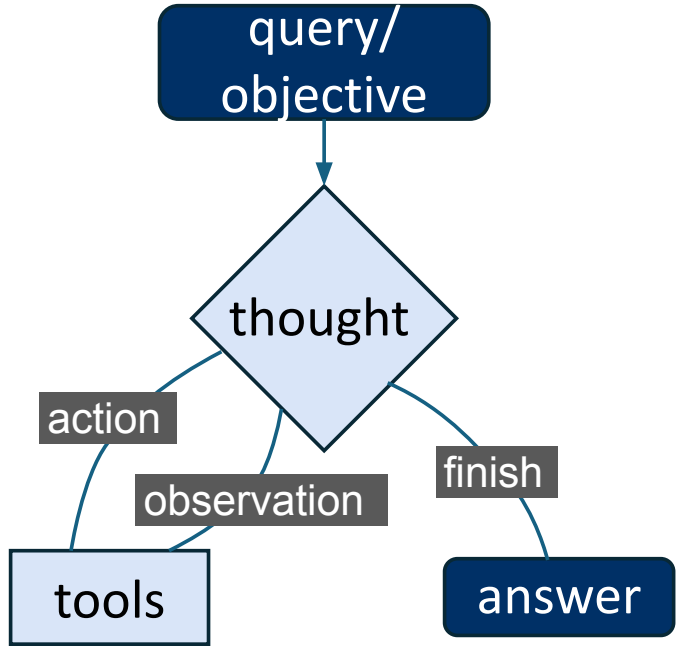
# AI is accelerating scientific discovery and innovation



# What is an “agent”?

**An agent is a persistent, stateful process that acts on behalf of a user or system**. An agent may:

- **Observe** inputs or events
- **Plan** (decide on) actions using a policy (rules or LLM)
- **Act**: Execute tools or call other agents
- **Learn**: Update state to adapt over time




*We can think of an agent as an entity that can reason, act, and coordinate on our behalf*

# Toward an agentic scientific method


*J. G. Pauloski, K. Chard and I. Foster, "Agentic Discovery: Closing the Loop With Cooperative Agents," in Computer, vol. 58, no. 10, pp. 20-27, 2025*

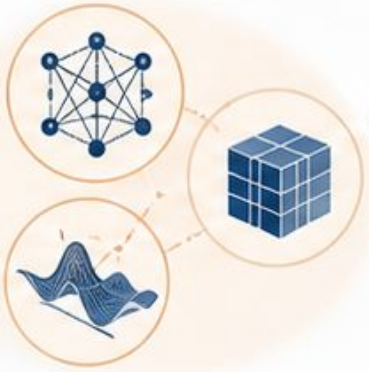
# Agents must operate at DOE science scale

**HPC SYSTEMS**
Leadership-class supercomputers, AI accelerators, quantum computers

**DIVERSE LABS & INSTRUMENTS**
A wide range of facilities, instruments, and experimental capabilities

**LARGE & DISTRIBUTED DATA**
Petabytes to exabytes of data across repositories, facilities, and domains

**SPECIALIZED MODELS & SOFTWARE**
Domain-specific models, simulations, and tools built for science



**Agentic systems today are not scalable, federated & infrastructure-aware**

**SECURITY & POLICY**
Strict policies, access controls, and compliance requirements across systems and organizations

# A core agentic framework for DOE science

### Interact with LLMs

1 Prompt, interpret responses, and manage complex dialog with diverse LLMs.

### Deploy Federated Agents

2 Persistent and stateful agents operating from edge devices to HPC centers.

### Scale

3 Run agents at exascale on cutting-edge DOE computing infrastructure.

### Call External Tools

4 Invoke remote resources, scientific APIs, and physical instruments.

### Monitor & Manage

5 Steer, pause, and engage with agents in real-time for human-in-the-loop.

### Track Provenance

6 Track lineage, decisions, and history to ensure scientific reproducibility.

*Enabling robust, autonomous, and scalable scientific discovery through a unified agentic ecosystem.*

# A core agentic framework for DOE science

### Interact with LLMs

1 Prompt, interpret responses, and manage complex dialog with diverse LLMs.

### Deploy Federated Agents

2 Persistent and stateful agents operating from edge devices to HPC centers.

### Scale

3 Run agents at exascale on cutting-edge DOE computing infrastructure.

### Call External Tools

4 Invoke remote resources, scientific APIs, and physical instruments.

### Monitor & Manage

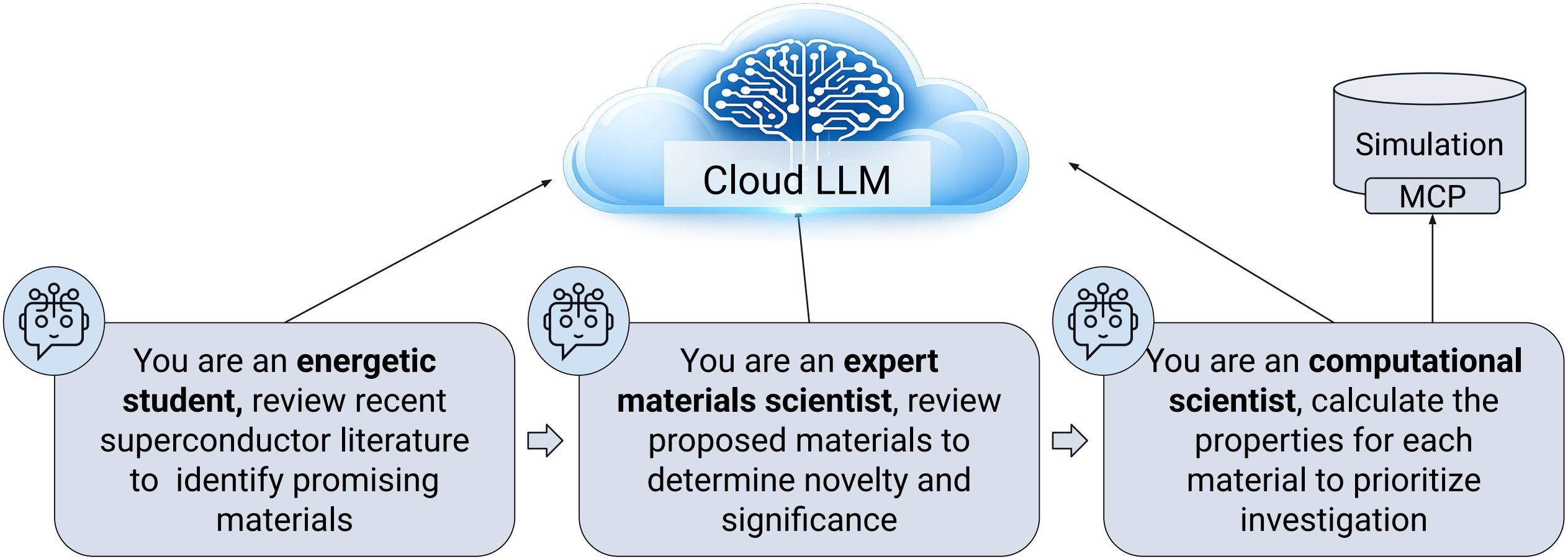
5 Steer, pause, and engage with agents in real-time for human-in-the-loop.

### Track Provenance

6 Track lineage, decisions, and history to ensure scientific reproducibility.

*Enabling robust, autonomous, and scalable scientific discovery through a unified agentic ecosystem.*

# Many agentic systems focus on centralized, single model with role-playing + remote tool calling

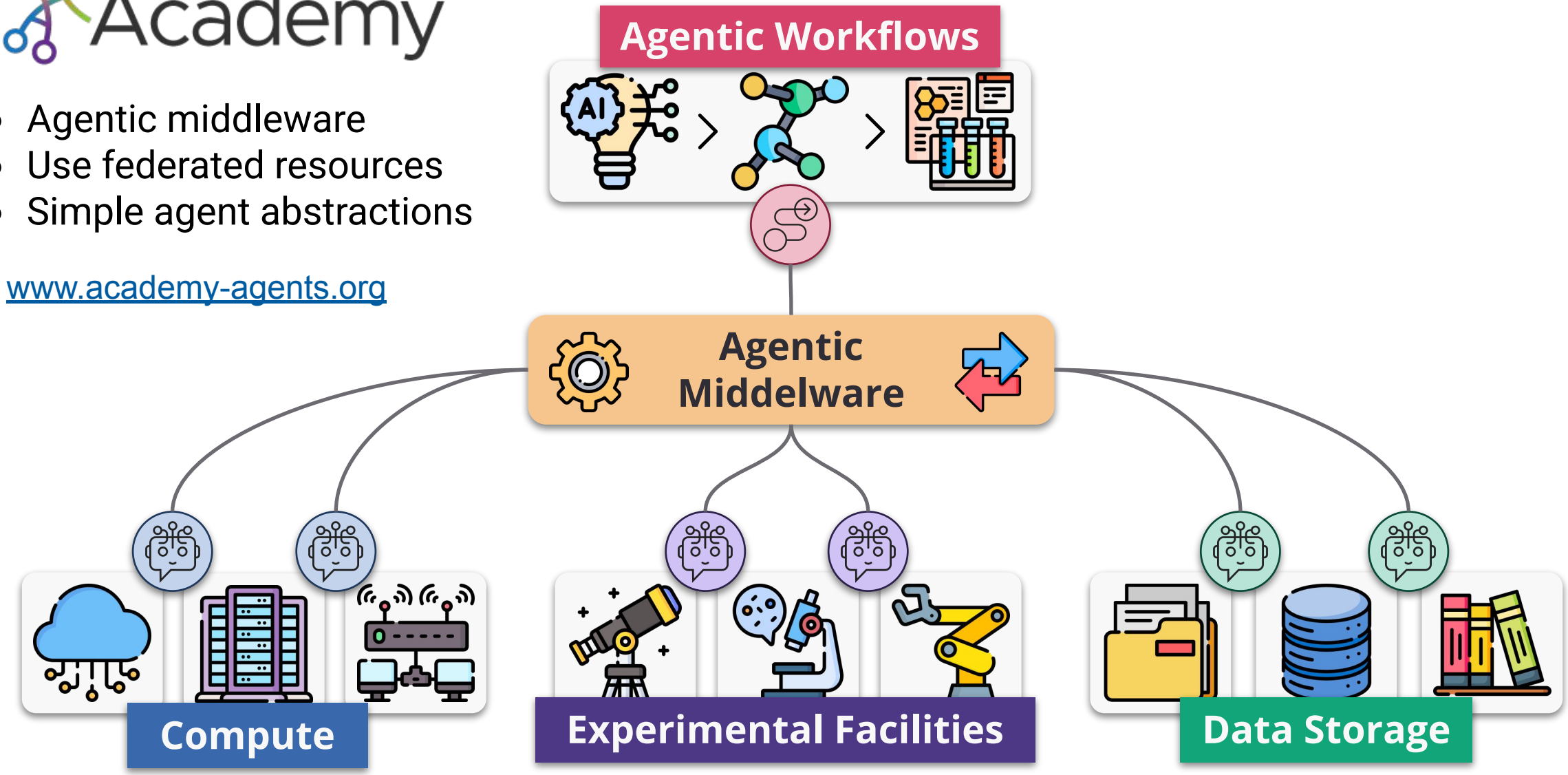


# Deploying and managing agents across the computing continuum

# Academy

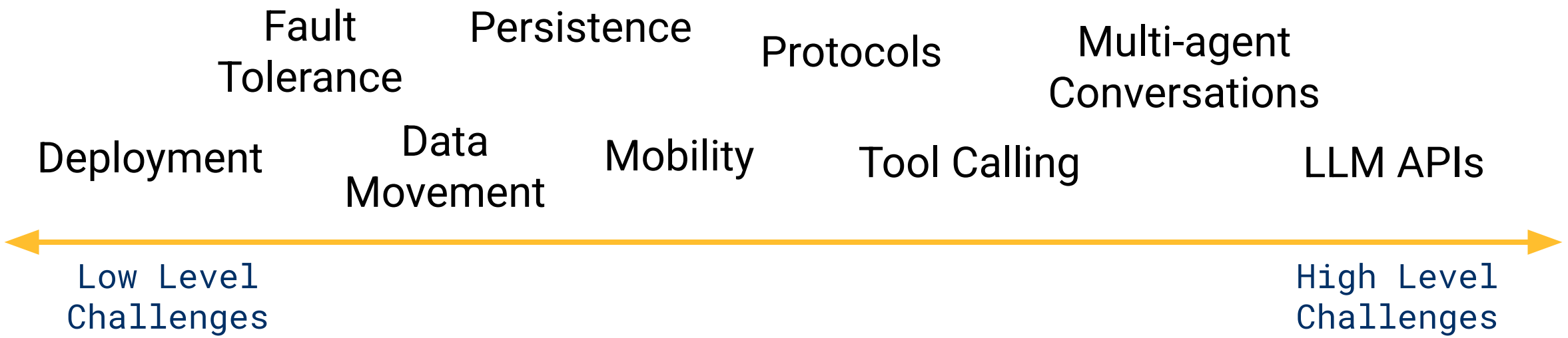
- Agentic middleware
- Use federated resources
- Simple agent abstractions

www.academy-agents.org




*Kamatar, Pauloski et al. Empowering Scientific Workflows with Federated Agents, IPDPS, 2026*

# Agentic Middleware: Scope & Challenges




**Academy** →

← **LangChain, AutoGen, Pydantic AI, etc.**

# Our initial thoughts concerning tech applicability

| Your Need | Best Fit |
|---|---|
| Deterministic pipelines | LangGraph or workflow tools |
| Multi-step reasoning with LLM | LangGraph |
| Conversational multi-agent behavior | AutoGen |
| Persistent agents that call each other | AutoGen or Academy |
| Distributed, secure, long-lived agents | Academy |
| Autonomous labs and instrument control | Academy |

In brief: **LangGraph** = controlled reasoning and workflows
**AutoGen** = conversational agent interactions
**Academy** = scientific autonomy at scale

We are developing example code to show how to combine frameworks: e.g., LangGraph for agent logic and Academy for scaling

ModCon Core Agentic Framework

# A core agentic framework for DOE science

### Interact with LLMs

1 Prompt, interpret responses, and manage complex dialog with diverse LLMs.

### Deploy Federated Agents

2 Persistent and stateful agents operating from edge devices to HPC centers.

### Scale

3 Run agents at exascale on cutting-edge DOE computing infrastructure.

### Call External Tools

4 Invoke remote resources, scientific APIs, and physical instruments.

### Monitor & Manage

5 Steer, pause, and engage with agents in real-time for human-in-the-loop.

### Track Provenance

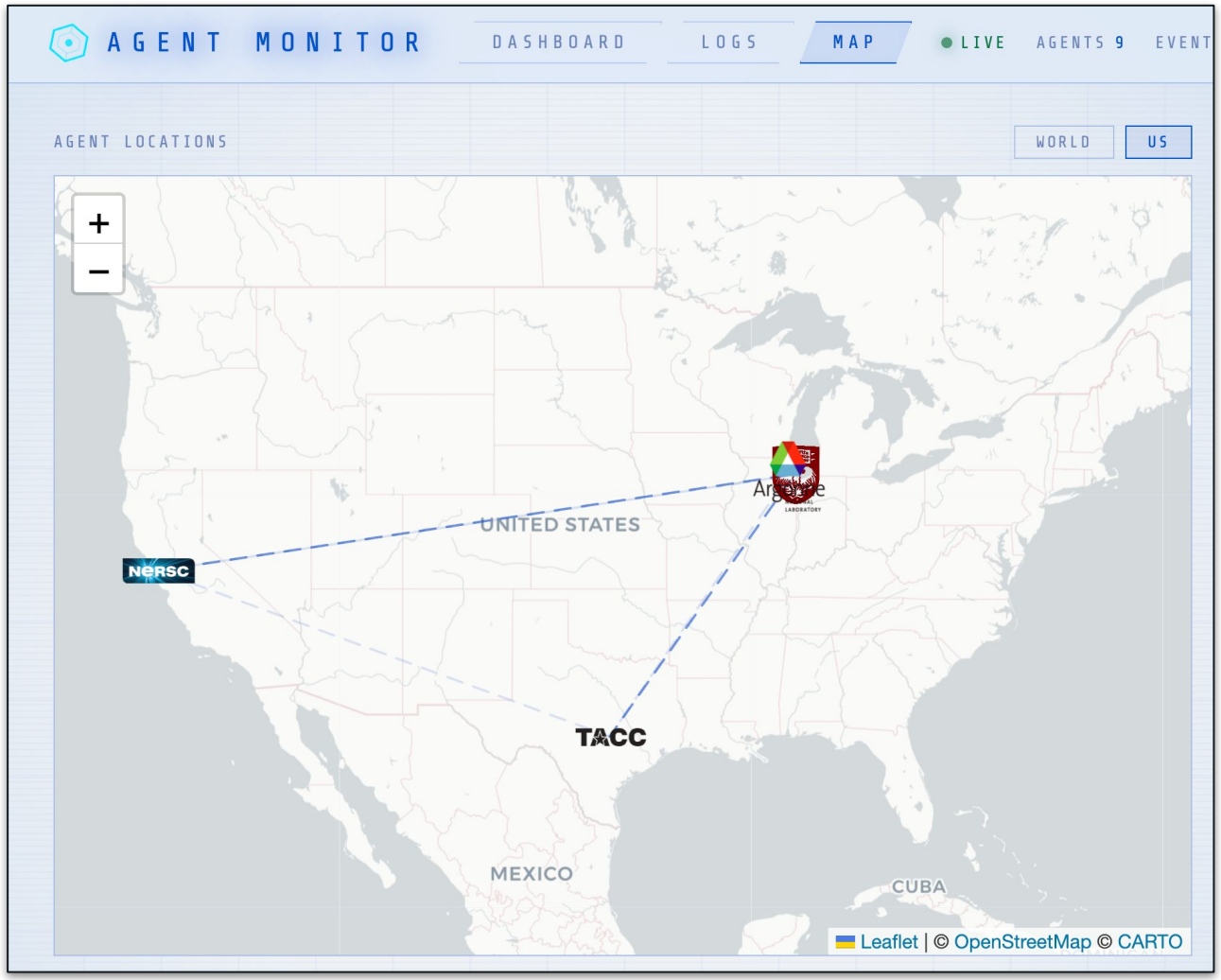
6 Track lineage, decisions, and history to ensure scientific reproducibility.

*Enabling robust, autonomous, and scalable scientific discovery through a unified agentic ecosystem.*

# Monitoring and managing (distributed) agents

- **Agent drift:** agents pursuing long-horizon tasks can deviate from intended goals → requiring efficient human intervention
- **Resource control:** unconstrained agents can consume compute and API budgets unpredictably → guardrails enforced at the system level
- **Coordination:** concurrent agents risk conflicting approaches, duplicated work, and stale state → active observation and orchestration
- **Failure & recovery:** partial failures are common → errors must be detected and resolved without halting the system

# Monitoring and managing distributed agents




# A core agentic framework for DOE science

### Interact with LLMs

1 Prompt, interpret responses, and manage complex dialog with diverse LLMs.

### Deploy Federated Agents

2 Persistent and stateful agents operating from edge devices to HPC centers.

### Scale

3 Run agents at exascale on cutting-edge DOE computing infrastructure.

### Call External Tools

4 Invoke remote resources, scientific APIs, and physical instruments.

### Monitor & Manage

5 Steer, pause, and engage with agents in real-time for human-in-the-loop.

### Track Provenance

6 Track lineage, decisions, and history to ensure scientific reproducibility.

*Enabling robust, autonomous, and scalable scientific discovery through a unified agentic ecosystem.*

# Observing and tracking agentic systems

- **PROV-AGENT** captures agentic provenance information: real-time decisions, tool use, LLM, reasoning across frameworks
- **Flowcept** provides unified provenance data capture and management framework using a simple annotation-based model and supports both workflow and agentic  frameworks
- **Streaming provenance** capture supports real-time monitoring with minimal performance impact

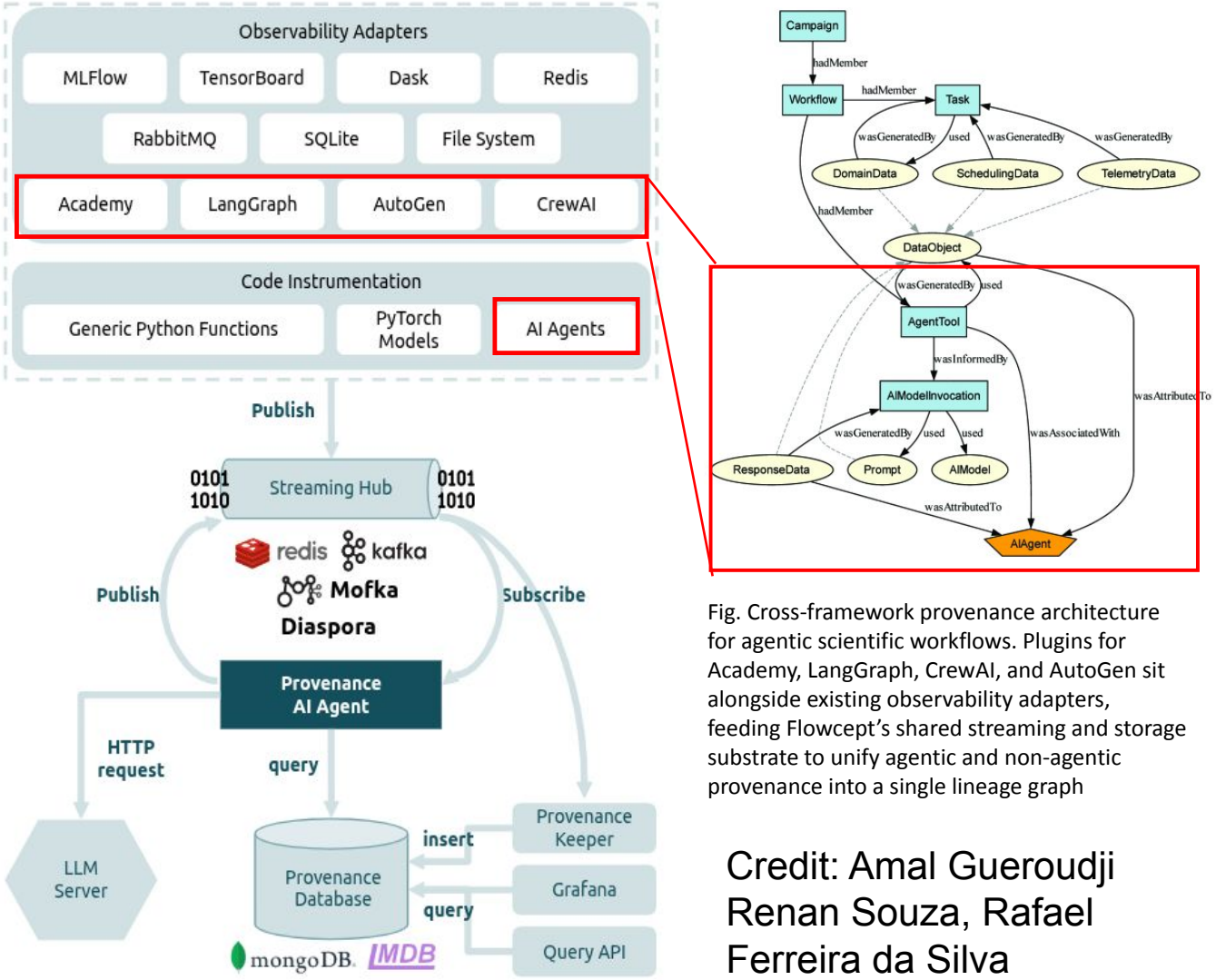
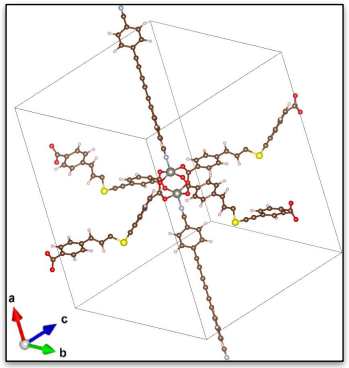


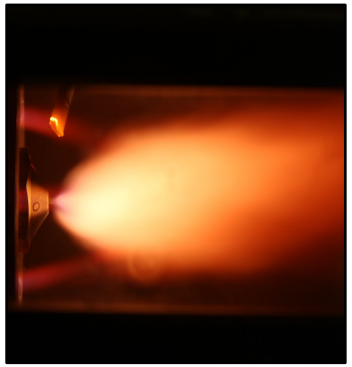

Fig. Cross-framework provenance architecture for agentic scientific workflows. Plugins for Academy, LangGraph, CrewAI, and AutoGen sit alongside existing observability adapters, feeding Flowcept's shared streaming and storage substrate to unify agentic and non-agentic provenance into a single lineage graph

Credit: Amal Gueroudji Renan Souza, Rafael Ferreira da Silva

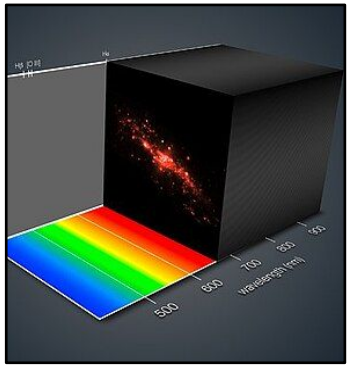
# Agentic systems in the Genesis Mission



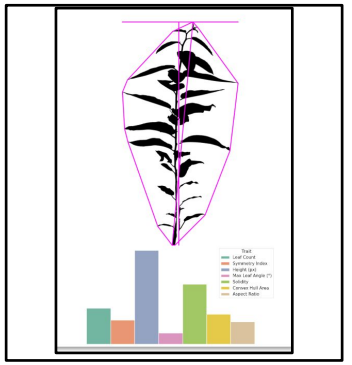
*Materials Discovery*



*Combustion*

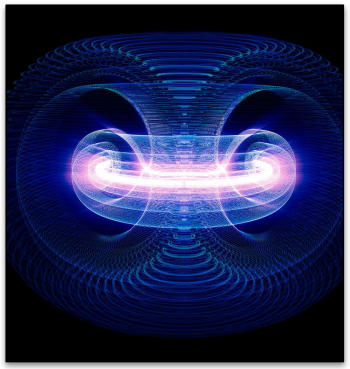


*Astrophysics*



*Plant Science*

*Engineering*

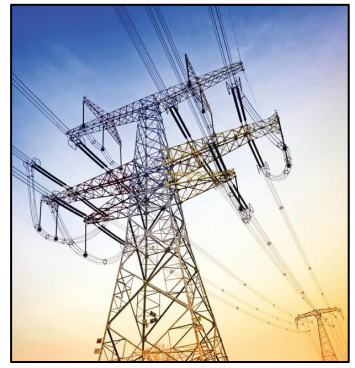


*Fusion*

*Biology*

*Lightsources*


*HPC Code*

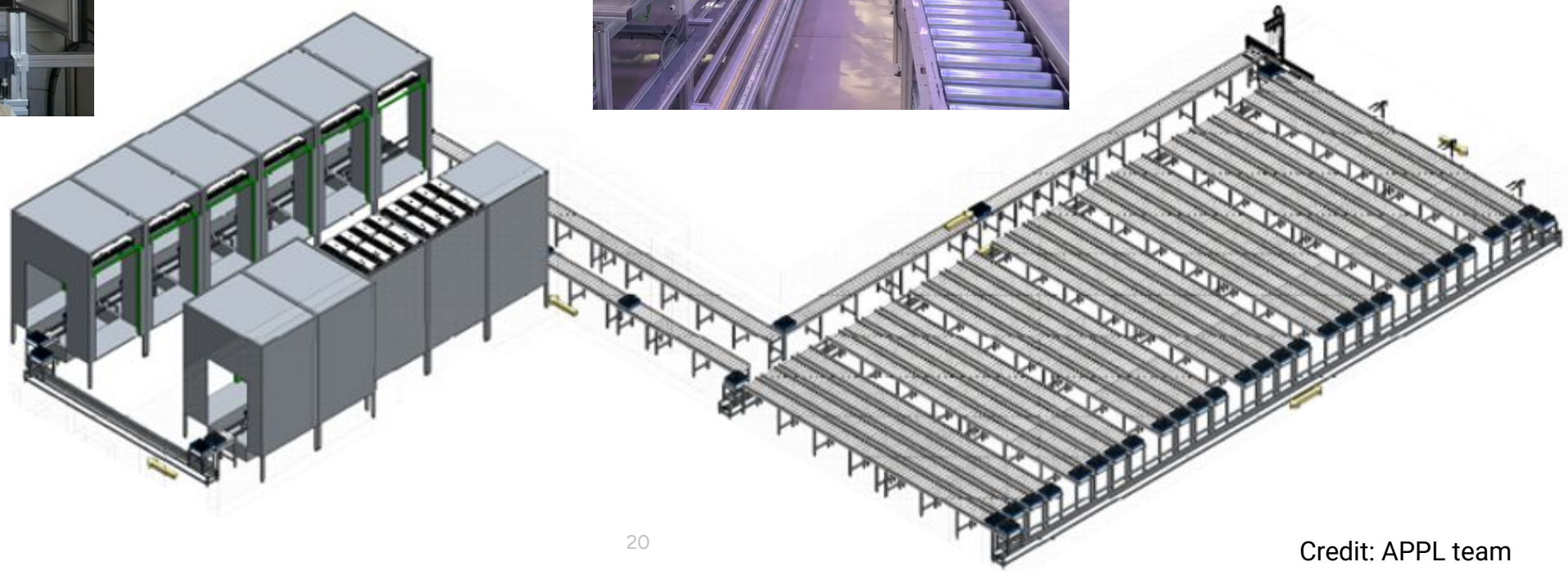


*Power Grid*

# APPL is an automated 520-tray conveyor system that continually monitors plants using multimodal imaging.

Buffer

Imaging loop

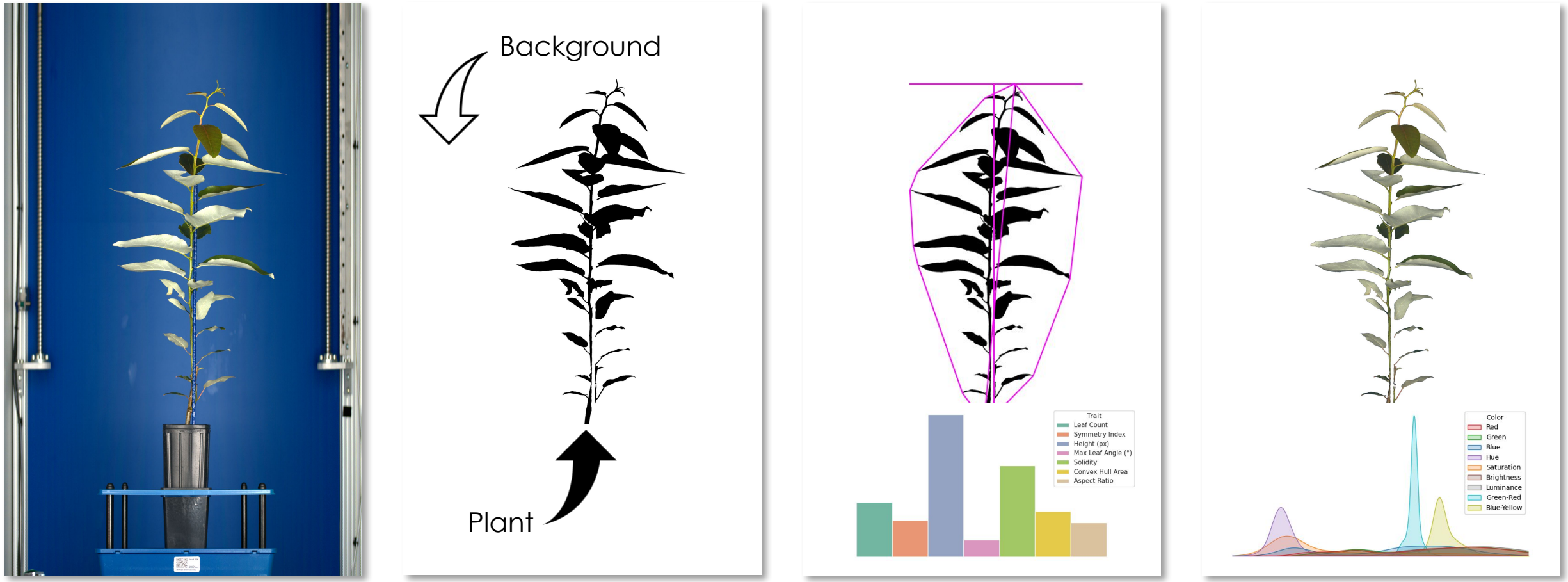




Credit: APPL team

# Accurate image segmentation enables prediction of complex plant phenotypes

Original → Segmentation → Morphology → Spectral traits






Credit: APPL team

Daniel Rosendo

Renan Souza

Scientist

# Genesis Mission
## Platform Deployment for OPAL

Cloud

Model Services
*Model Catalog*

Intelligent Int.
*LiteLLM Inference*

i2 Auth
*Globus / Google*

Metadata Services
*Data lakehouse*

S3M

Co-pilot Agent
*Academy*

Chat UI
*Chainlit*

Kubernetes cluster

aws

Provenance Service
*Flowcept*

OLCF

ViT inference
*Parsl*

Compute Agent
*Academy*

Compute node

Login node

Lustre filesystem


APPL

Plant images + experiment metadata

Platform

**Legend**

APPL components

Frontier

AmSC services

AWS services

**Demo video**
https://www.youtube.com/watch?v=7szA15ncDx0

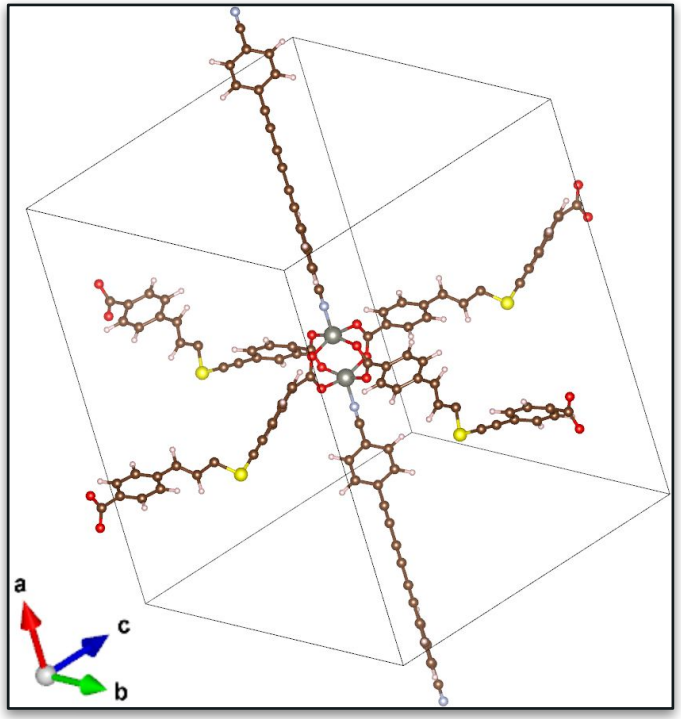
Credit: APPL team

# Searching enormous materials search spaces

## Metal Organic Frameworks (MOF)

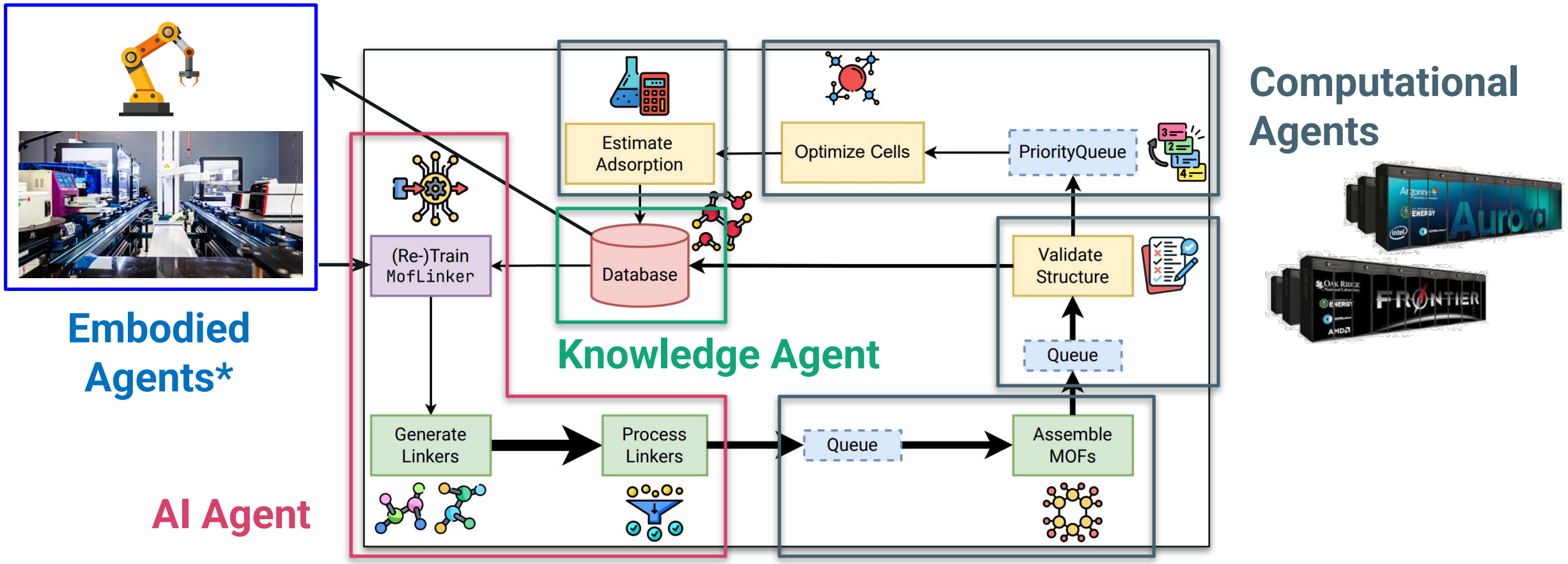
- ➔ Composed of organic molecules (ligands) and inorganic metals (nodes)
- ➔ The sponges of materials science!
- ➔ Porous structures that adsorb and store gases
- ➔ Topologies can be optimized for targeted gas storage → **Methane Capture**

**How to efficiently discover MOFs with desirable properties for target applications?**

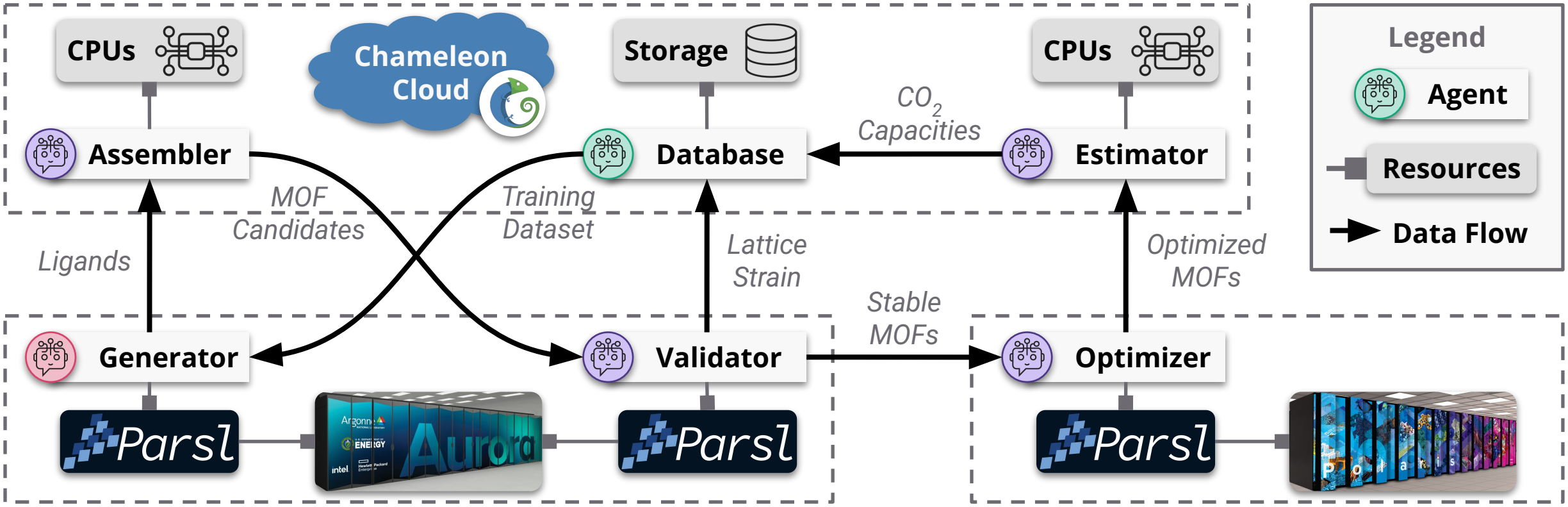



*Intractable search space of ligand, node, & geometry combinations*

# MOFA: Online learning + GenAI + Simulation



# MOFA realized with autonomous agents




*Agents executed remotely via Globus Compute*

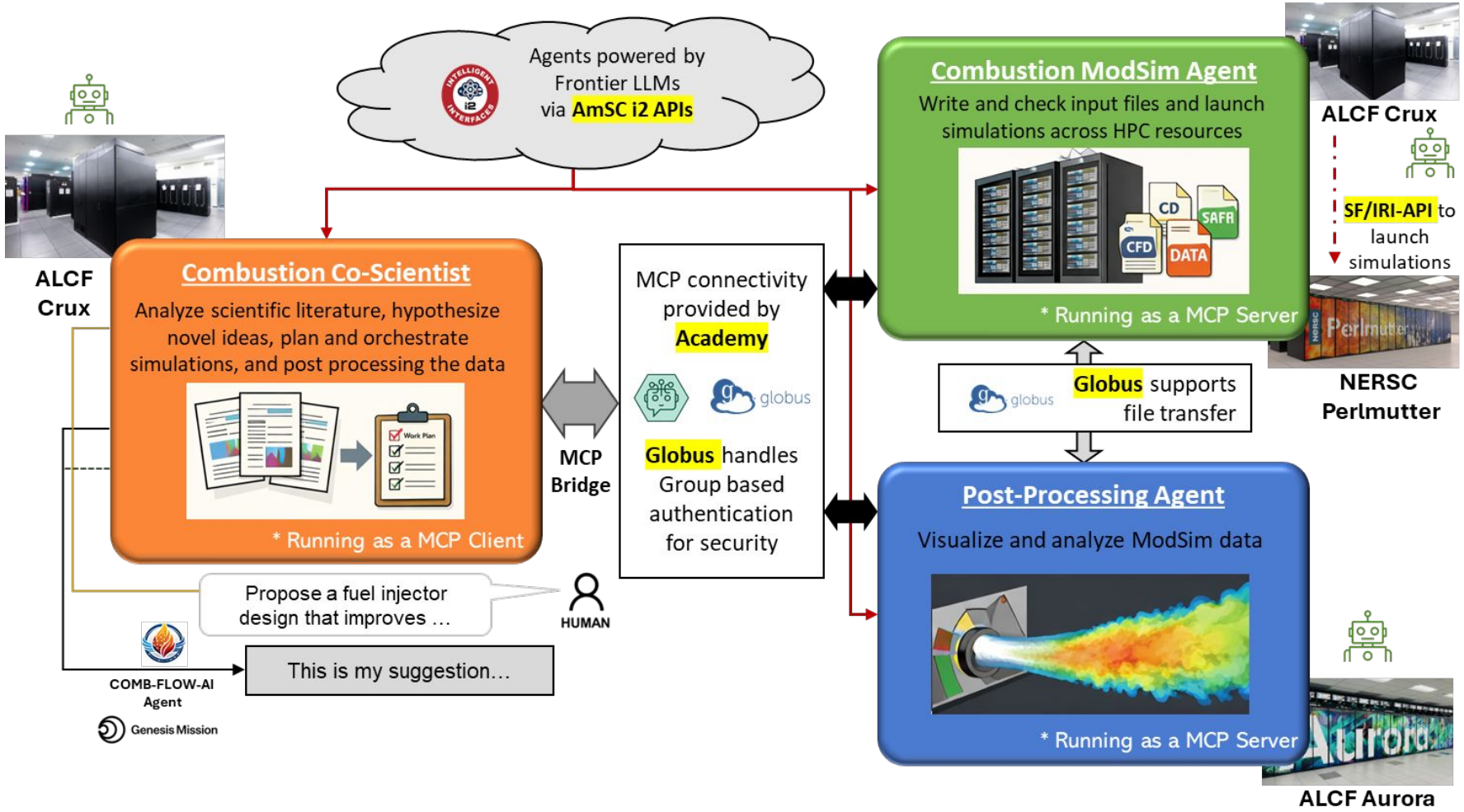
globus

# AISAC POWERS MANY DOMAIN APPLICATIONS

## Use-Case : Genesis Combustion Model Team's Combustion Co-Scientist

Chandruchur Bhattacharya

- **AISAC powers Combustion Co-scientist** that performs reasoning, knowledge synthesis, and user-interaction
- Combustion Co-scientist hands off tasks to other agents living on different HPC hardware via Academy
- Globus ensure authentication and security
- User can launch jobs on NERSC / ALCF from their laptops, running as server, or connected to a HPC-hosted instance
- Full COMB-FLOW-AI Agent system automates job launch, monitoring and post-processing
- Cross DOE-facility agent and task coordination




Chandrachur Bhattacharya, Martin Rieth, Kevin Griffin, Bruce Perry, Jean Sexton, Kiran Yalamanchi, Yixuan Sun, Sandeep Madireddy, Pei Zhang, Paul Laiu, Ann Almgren, Navin Mahto, Pinaki Pal, Jackie Chen, Shashank Yellapantula, Ramanan Sankaran, Sibendu Som

# Learn more and engage with CAF

## CAF Agent of the Week

The CAF **Agent of the Week** (AotW) is an exemplary agentic application or individual agent that demonstrates superior architecture, use of common tools, and scientific impact. This much-sought-after recognition brings, on occasion, a free drink and a CAF hat for the winning team--but the true joy in winning is the glory that comes from recognition by your peers. Needless to say, many aspire to win, but few are chosen.

### Discover the Agents of the Week

1. OPAL — Plant Phenotyping Analysis
2. IGNOS — Fusion Foundation Model
3. Foam-Agent — Automating OpenFOAM CFD
4. LABMATE-Supervisor — Accelerating catalysis research
5. A2MC-Adaptive Multi-target Calibration for Earth System Models
6. AISAC-AI Scientific Assistant Core

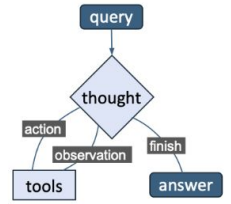
https://github.com/AI-ModCon/BaseCAF_agent_of_the_week

## AI Agents for Scientific Discovery

**Agentic discovery** combines reasoning-capable AI with scientific and engineering resources—such as literature collections, simulation codes, experimental platforms, and knowledge bases—to accelerate the pace of discovery. Recent advances in large language models (LLMs) and related technologies make it possible to build agentic systems that automate key aspects of scientific work: synthesizing information from the literature, generating and prioritizing hypotheses, designing and executing protocols, running simulations or experiments, and interpreting results.

An agent is a persistent, stateful process that acts on behalf of a user or system. An agent may:

- Observe inputs or events
- Plan (decide on) actions using a policy (rules or LLM)
- Act: Execute tools or interact with other agents
- Learn: Update state to adapt over time




We can think of an agent as a *scientific assistant that can reason, act, and coordinate on our behalf.*


**Teaching**
AI Agents for Science curriculum (University of Chicago, Autumn 2025).


**Capabilities**
Deployment patterns from local prototypes to massively parallel HPC inference.


**Applications**
Example applications of agentic systems for scientific discovery.

https://agents4science.github.io/

# Summary

- **We've entered the 5th paradigm of science** — frontier models have crossed the expert threshold and agents are capable of driving the full scientific cycle
- **Agents will be the primary consumers of scientific infrastructure** — compute, data, instruments, and APIs will be orchestrated by AI rather than humans
- **DOE science demands specialized solutions** — scale, federation, security, provenance, and compliance require infrastructure designed for scientific agents
- **CAF delivers robust, autonomous, scalable framework** — a unified agentic framework that eliminates fragmentation, enables federation, and provides built-in coordination, oversight, and resource governance

Genesis Mission

# Thank you

Kyle Chard (and the CAF team)
chard@anl.gov

U.S. DEPARTMENT of ENERGY