AI Agents for Science

Lecture 1, September 29: Introduction

Instructor: Ian Foster
TA: Alok Kamatar

Crescat scientia; vita excolatur

CMSC 35370 -- https://agents4science.github.io
https://canvas.uchicago.edu/courses/67079

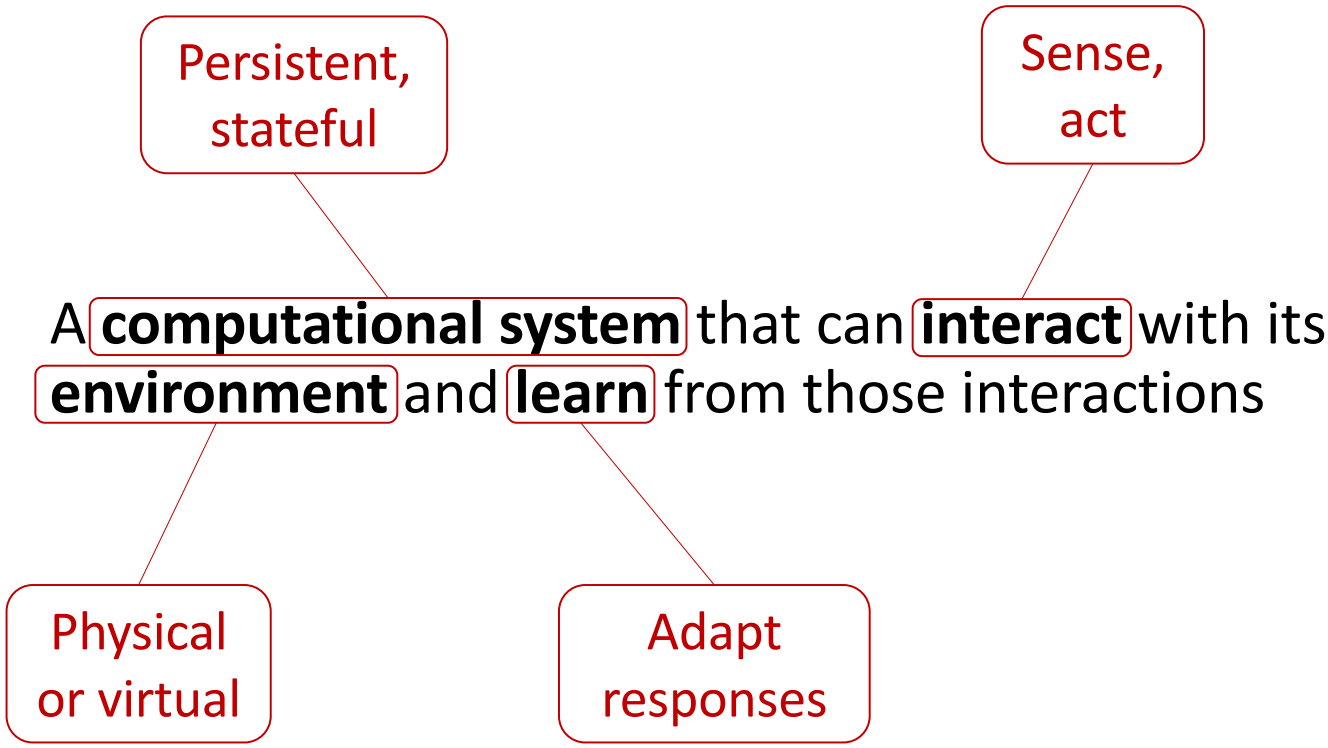
What is an “agent”?

- In computer software, a “software agent” is [Wikipedia] “**a computer program that acts for a user or another program in a relationship of agency**”

- In artificial intelligence (AI), an “intelligent agent” is [also Wikipedia] “**an entity that perceives its environment, takes actions autonomously to achieve goals, and may improve its performance through machine learning or by acquiring knowledge**”
  - An AI component, sensors, actuators, memory

See also: https://gist.github.com/simonw/beaa5f90133b30724c5cc1c4008d0654#response

Related concept: An “embodied agent”



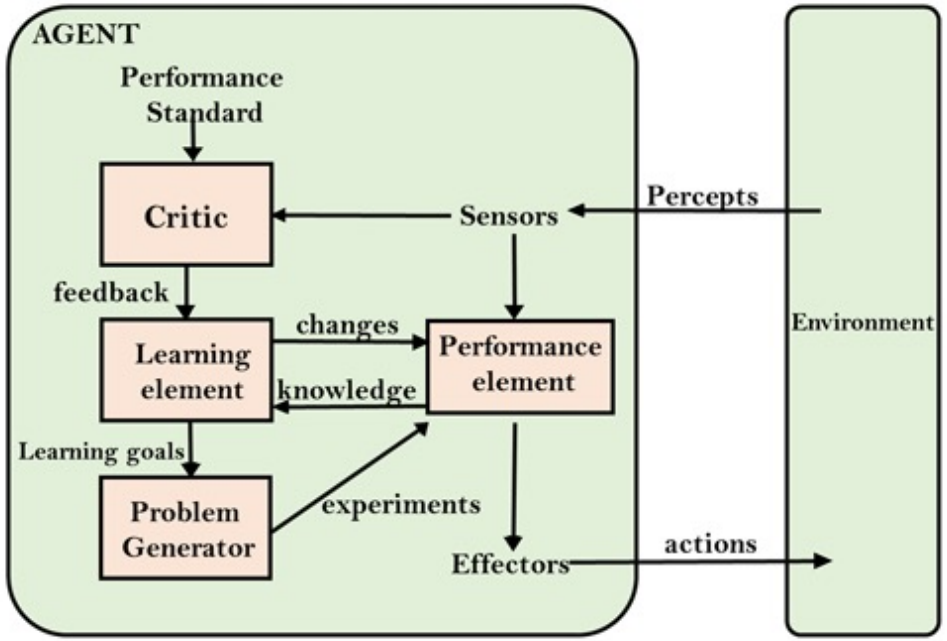
The “sense-plan-act-learn” loop

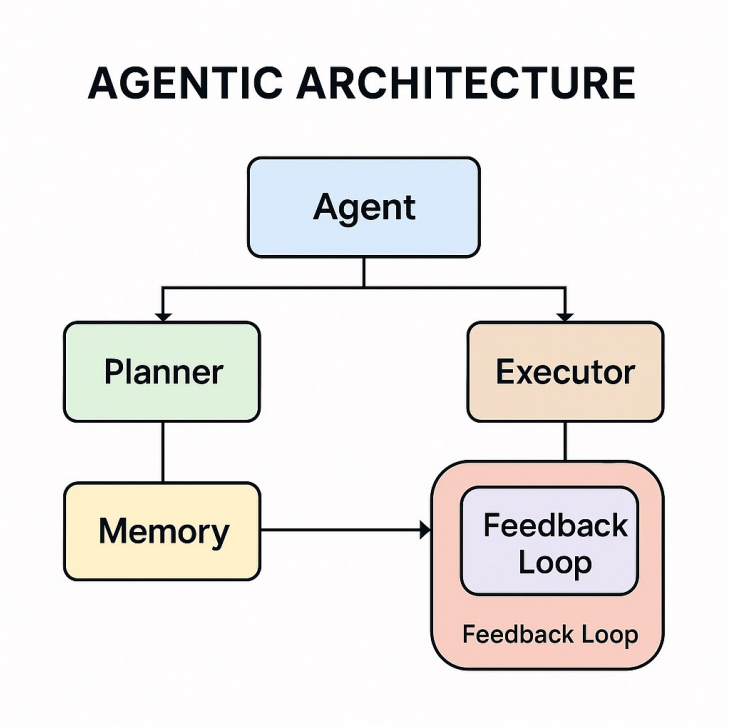
- Initialize state and goals
- Repeat until termination:
  - **Sense**: Gather observations from environment
  - **Plan**: Evaluate goals and state, plan/select next action
  - **Act**: Execute chosen action on environment
  - **Learn**: Update internal state, memory, or model based on outcomes

Current excitement around agents is due to LLMs

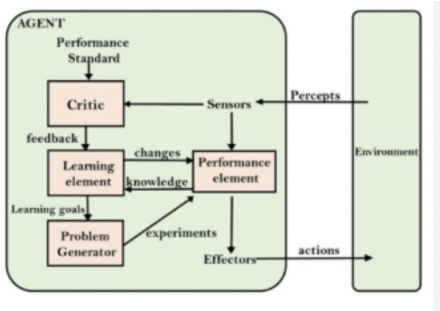
https://trends.google.com

Many related ideas regarding agentic architecture



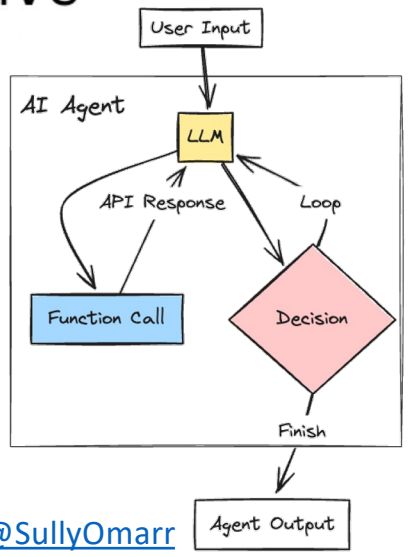


A simple perspective



An LLM in a loop with an objective

An LLM in a loop with an objective



https://x.com/josh_bickett/status/1725556267014595032

@SullyOmarr

“CACTUS: Chemistry Agent Connecting Tool Usage to Science”

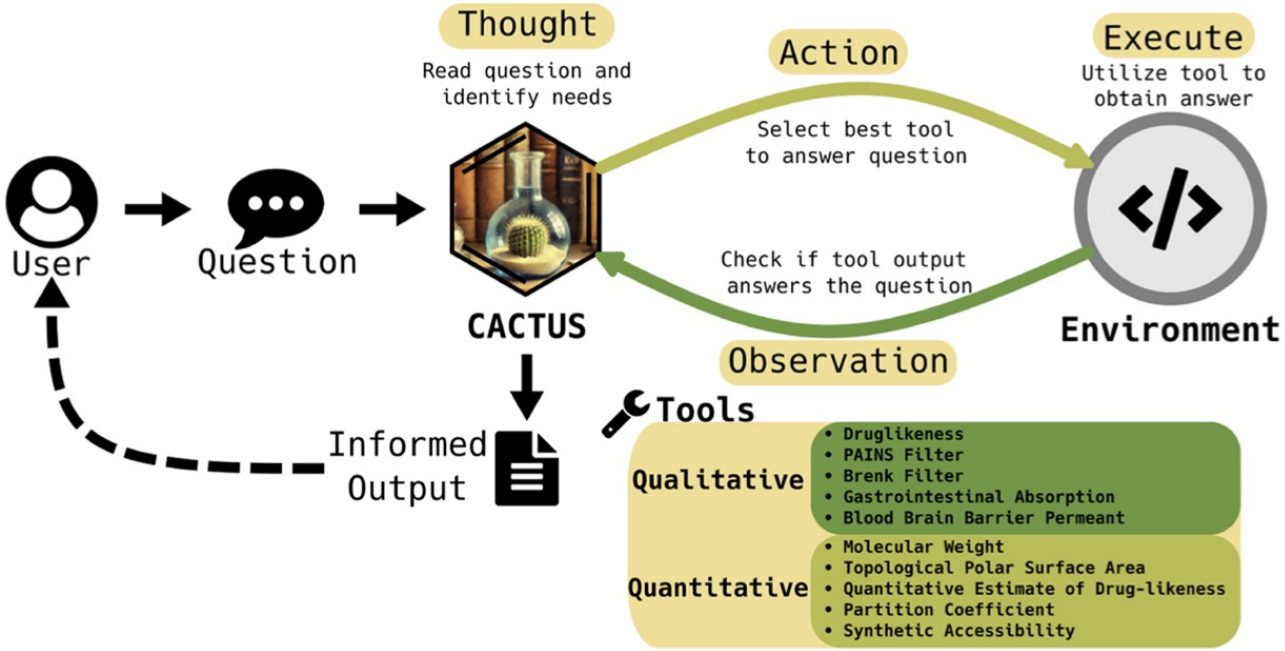


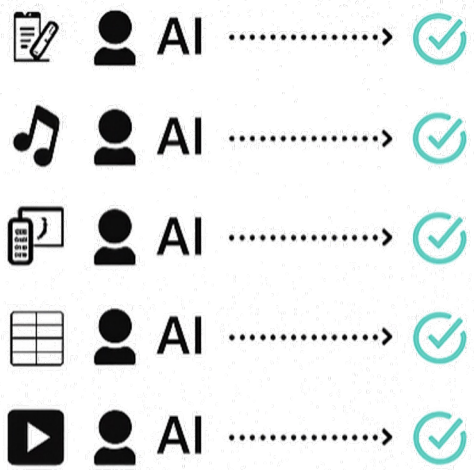
Figure 1. General workflow of the CACTUS agent that details how the LLM interprets input to arrive at the correct tool to use to obtain an answer. Starting from the user input, CACTUS follows a standard “Chain-of-thought” reasoning method with a Planning, Action, Execution, and Observation phase to obtain an informed output.

https://pubs.acs.org/doi/pdf/10.1021/acsomega.4c08408

Overview

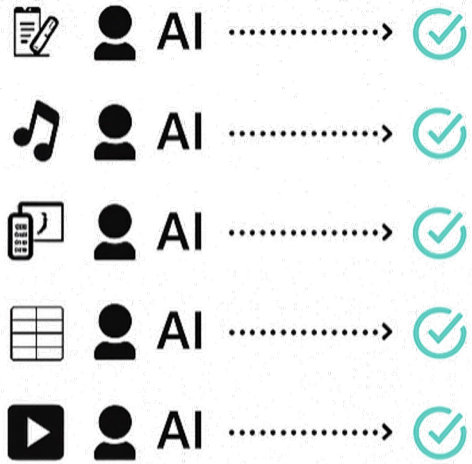
- What is an “agent”?
- **LLMs, foundation models, reasoning models**
- Agents and scientific discovery
- Scientific Discovery Platforms
- Curriculum

Traditional ML



- Individual siloed models
- Require task-specific training
- Lots of human supervised training

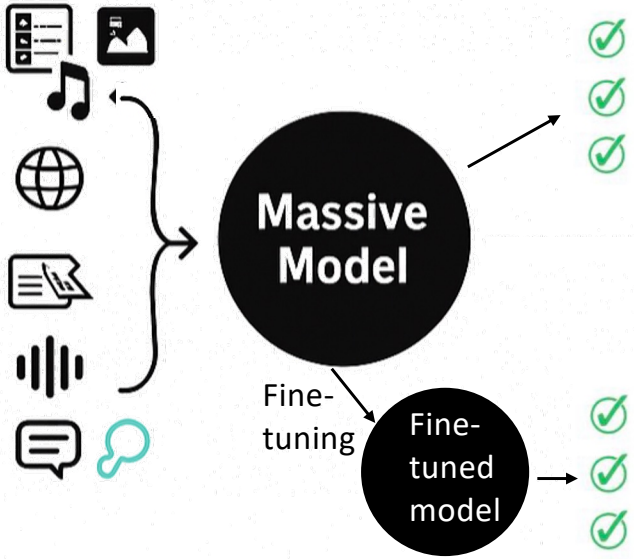
Traditional ML



- Individual siloed models
- Require task-specific training
- Lots of human supervised training

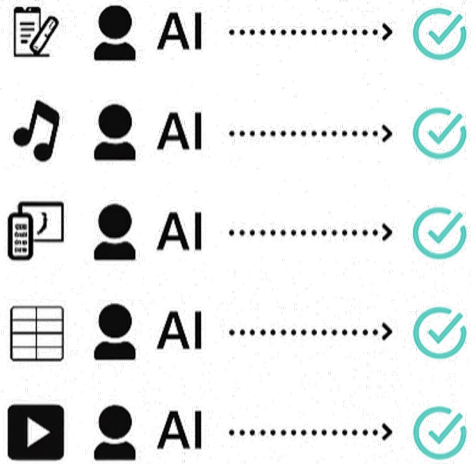
Foundation models

(E.g., Large Language Models: LLMs)



- Massive multi-modal model
- Adapted with minimal training
- Pre-trained unsupervised learning

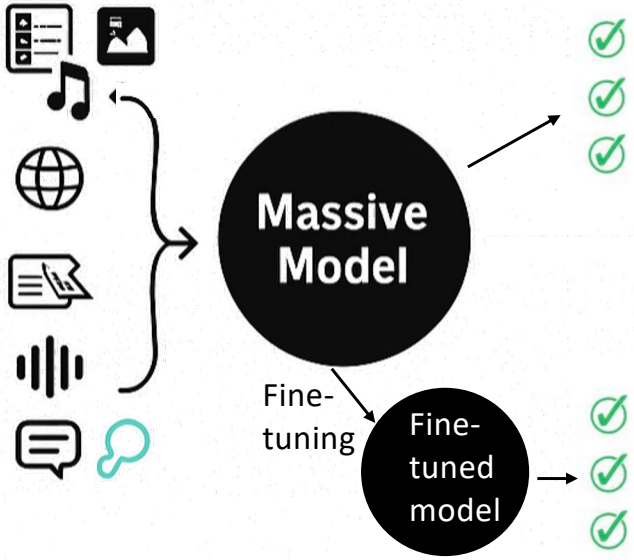
Traditional ML



- Individual siloed models
- Require task-specific training
- Lots of human supervised training

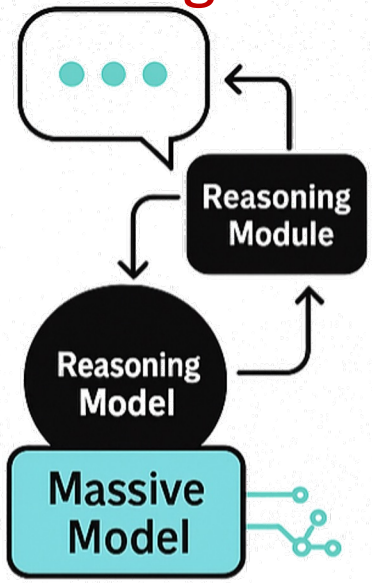
Foundation models

(E.g., Large Language Models: LLMs)



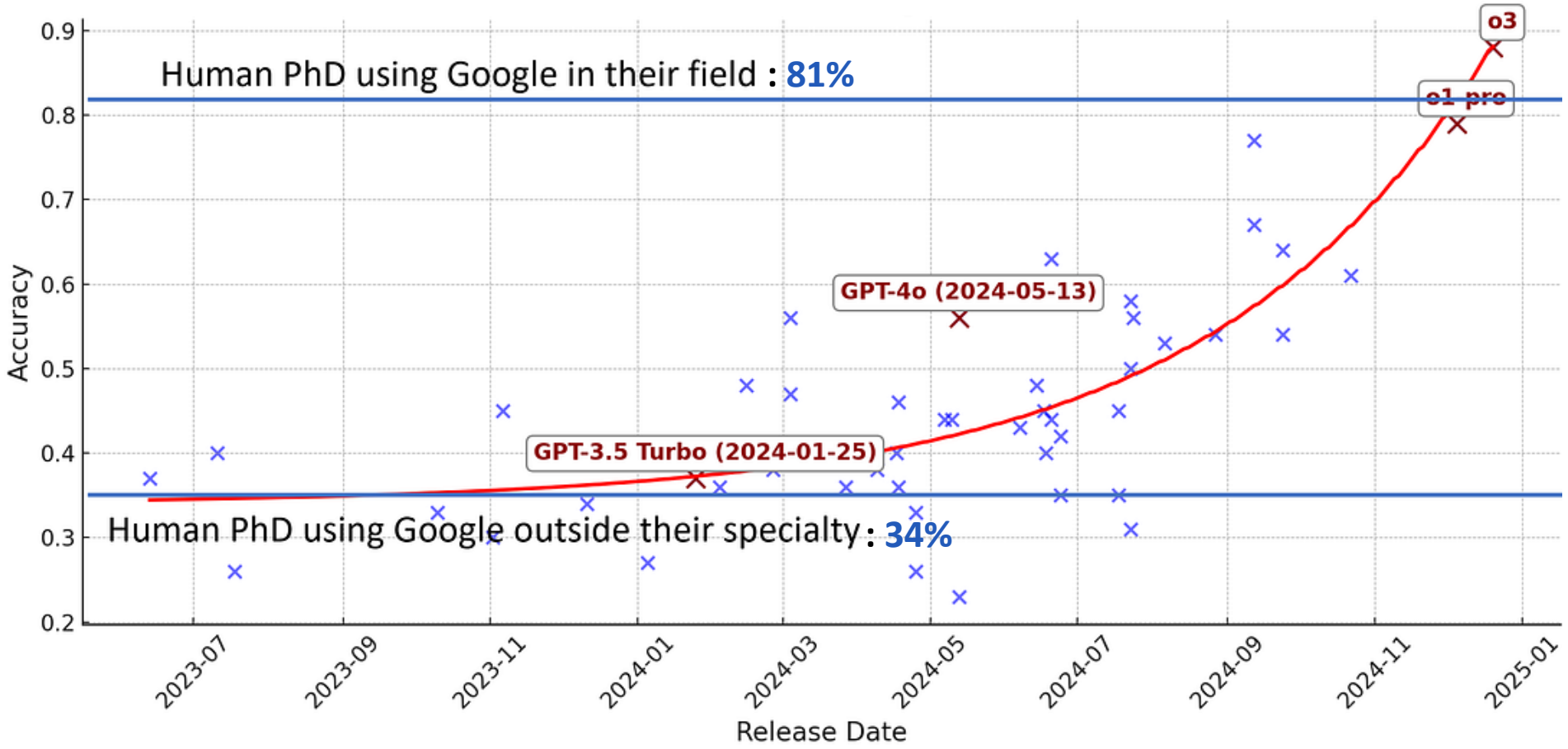
- Massive multi-modal model
- Adapted with minimal training
- Pre-trained unsupervised learning

Reasoning models



- Deliberative multi-step logic
- Self-consistency checks
- Slower but more accurate inference

Graduate-Level Google-Proof Q&A test (GPQA), Diamond problems



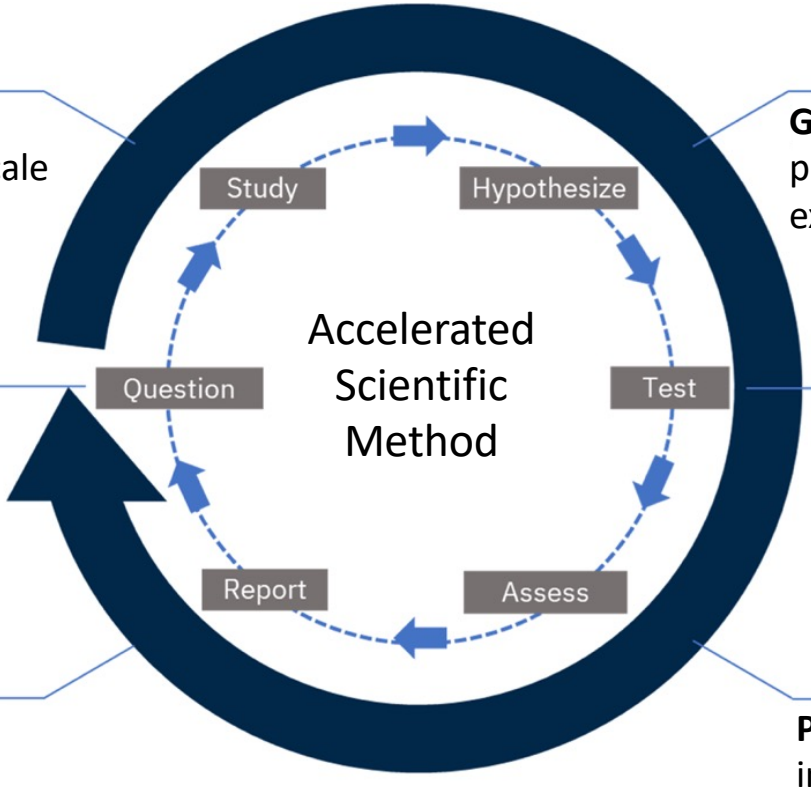
https://arxiv.org/pdf/2311.12022

https://epoch.ai/data/ai-benchmarking-dashboard

Overview

- What is an “agent”?
- LLMs, foundation models, reasoning models
- **Agents and scientific discovery**
- Scientific Discovery Platforms
- Curriculum

Accelerating discovery in science



Extraction, integration and reasoning with knowledge at scale

Generative models automatically propose new hypotheses that expand the discovery space

Tools help **identify new questions** based on needs and gaps in knowledge

Robotic labs automate experimentation and bridge digital models and physical testing

Machine representation of knowledge leads to **new hypotheses and questions**

Pattern and anomaly detection is integrated with simulation and experimentation to extract new insights

https://doi.org/10.1038/s41524-022-00765-z

The emergence of LLM-based agents for science

"AI agents are autonomous systems that can **reason about tasks and act to achieve goals by leveraging external tools and resources**.

Modern AI agents are typically powered by large language models (LLMs) **connected to external tools or APIs**.

They can perform **reasoning**, **invoke** specialized models, and **adapt** based on feedback.

Agents differ from conventional "models" in important regards:

- They are interactive and adaptive
- Rather than returning fixed outputs, they can take multi-step actions, integrate context, and support iterative human–AI collaboration.
- Users can interact with them through human language, substantially reducing usage barriers for scientists."

https://arxiv.org/pdf/2509.06917

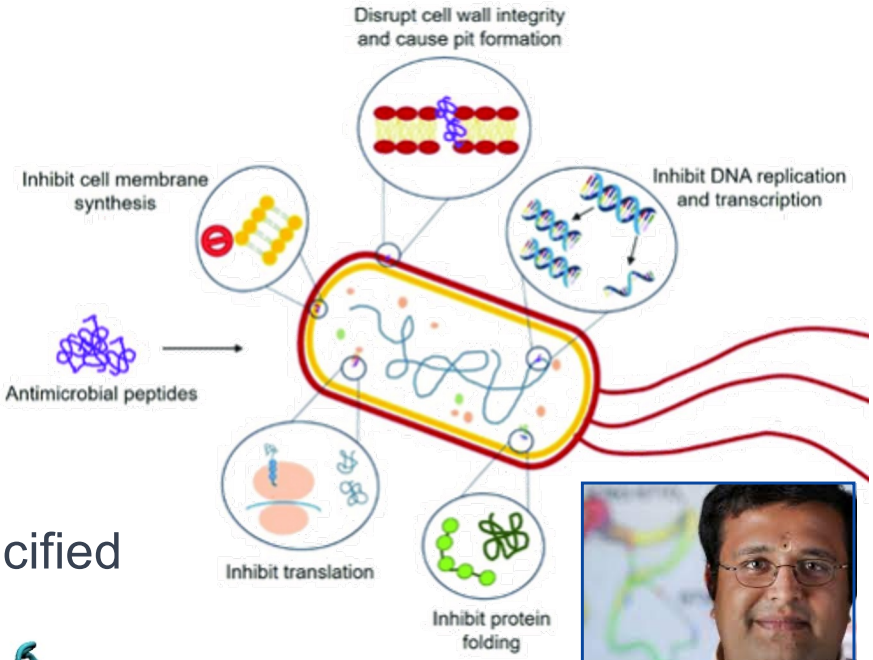
One agent or many?

- In **principle**, a single reasoning model (like a single human) can apply a variety of different reasoning strategies, have specialized knowledge on different topics, improve expertise over time, etc.

- In **practice**, it is common to create multiple agents (like a team of people), e.g., for:
  - Modularity (specialization, reuse, maintainability)
  - Different roles (e.g., idea generator, idea critic, program generator, …)
  - Parallelism (run multiple copies of an agent to explore different ideas)
  - Expanded capacity (e.g., larger LLM context)

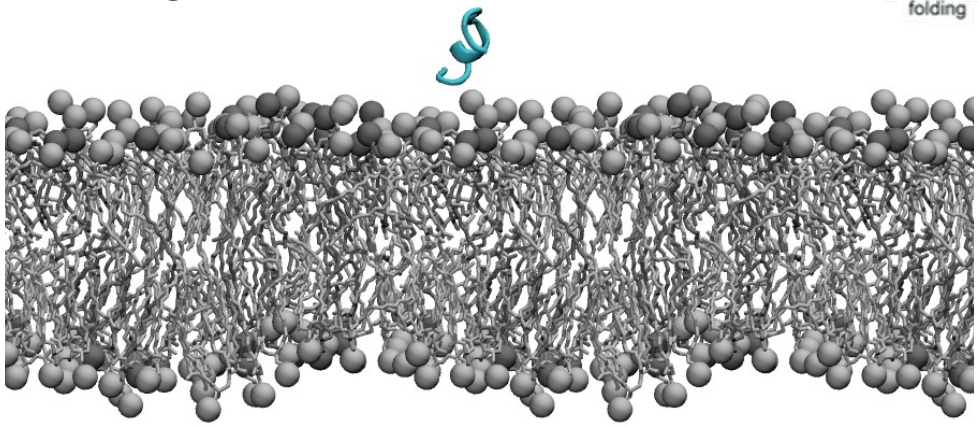
An agentic architecture for the design of antimicrobial peptides

An antimicrobial peptide (AMP) is a short (typically 12 to 50 amino acid) molecule that can target and kill viruses, bacteria, fungi, and other pathogens



Challenge: Design an AMP that can kill specified bacterial strains without harming host cells

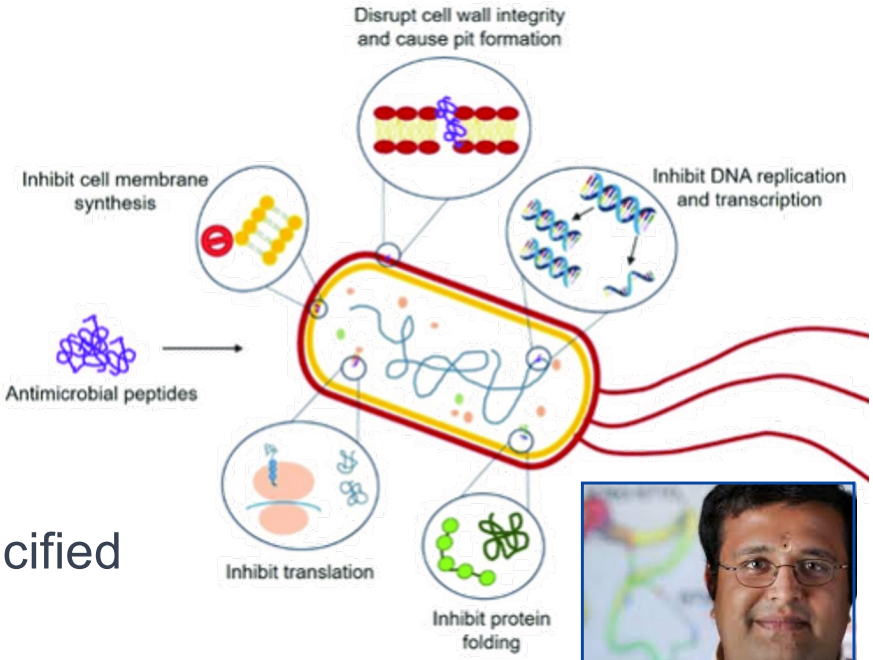
With 20 possible amino acids, there are $20^{20} = 10^{26}$ AMPs of length 20



Arvind Ramanathan

An agentic architecture for the design of antimicrobial peptides

An antimicrobial peptide (AMP) is a short (typically 12 to 50 amino acid) molecule that can target and kill viruses, bacteria, fungi, and other pathogens



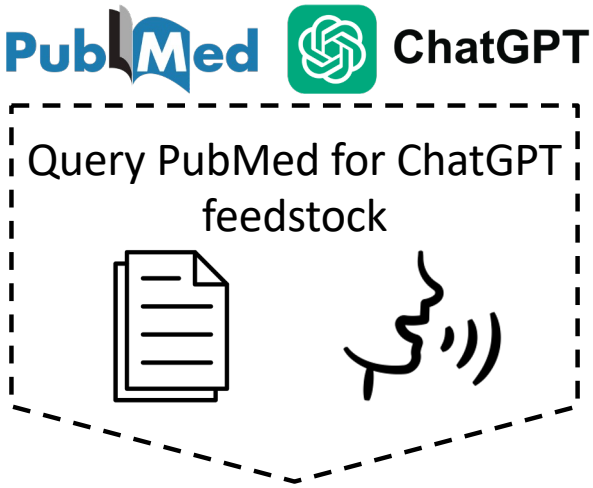
Challenge: Design an AMP that can kill specified bacterial strains without harming host cells

Arvind Ramanathan

A rational design approach might combine knowledge of bacterial cell membrane composition and structure, AMP molecular and structural properties, host cell membrane characteristics and intracellular pathways—knowledge that may be gained by **database/literature search**, **simulation**, **experiment**

Example: A **peptide expert**

(Prototyped with PubMed and ChatGPT)



We want a model with deep expertise regarding peptides and related topics

Retrieve abstracts **A** from PubMed that reference specified **peptide**

Use ChatGPT to build hypotheses by using retrieval-augmented generation: e.g.:

"Given **A**, on which organism is {**peptide**} acting?"

We want to be able to make millions of such requests

Arvind Ramanathan, Priyanka Setty, et al.

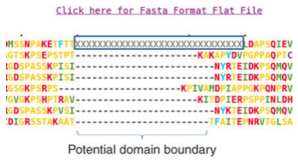
Define other agents, also foundation model-powered

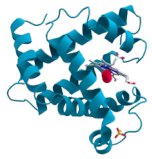
PubMed ChatGPT

Query PubMed for ChatGPT feedstock

Protein BLAST protein ▶ protein AlphaFold

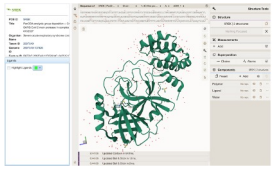
Align proteins, predict structure, rank results





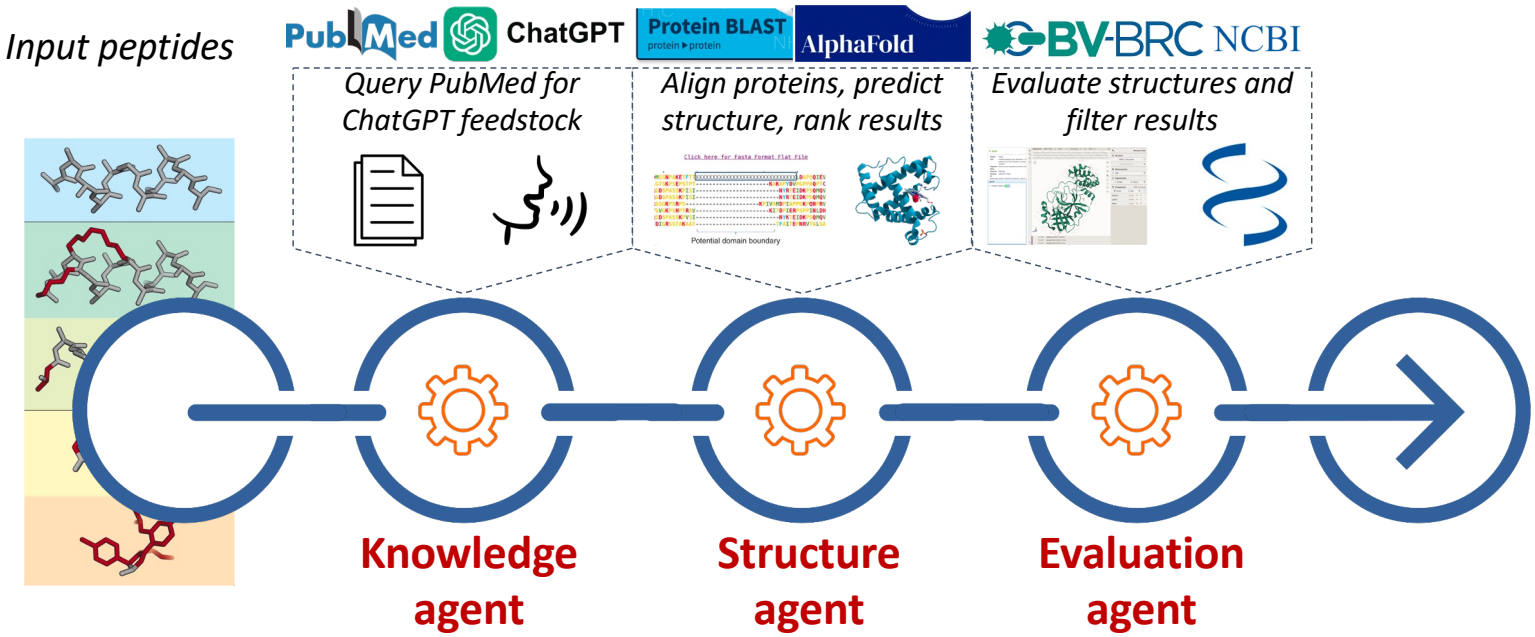
BV-BRC NCBI

Evaluate structures and filter results



Arvind Ramanathan, Priyanka Setty, et al.

Link agents into a flow

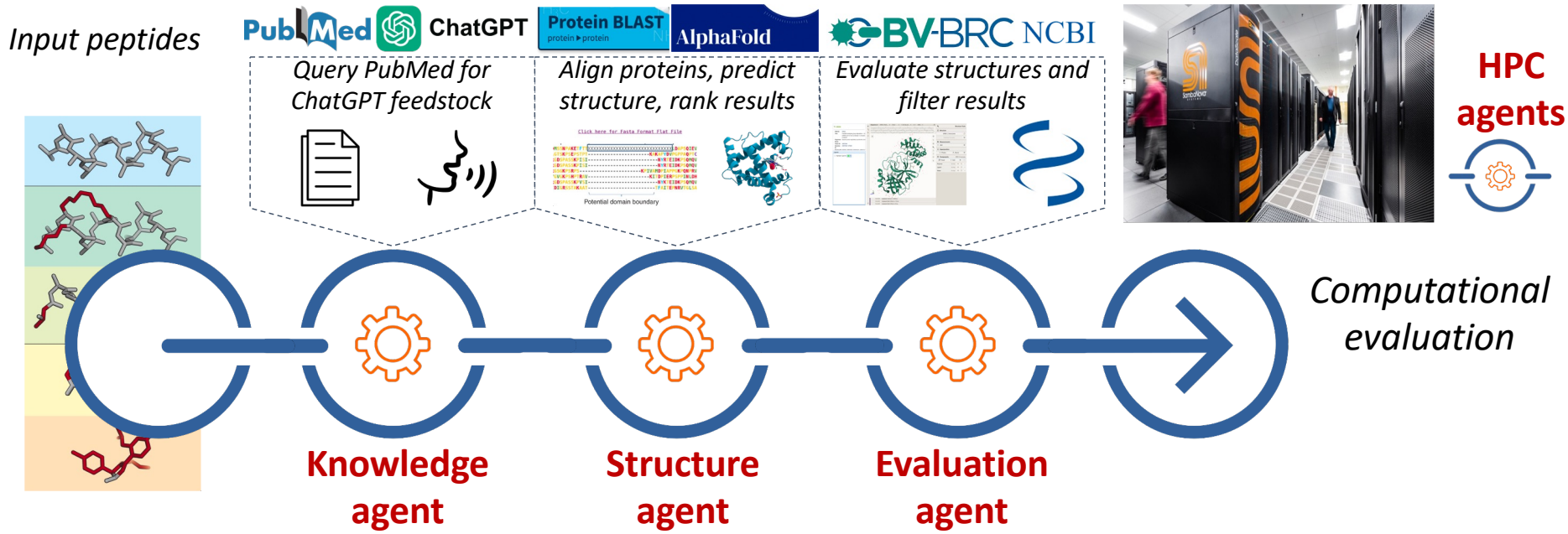


We use **Globus Flows** to invoke individual agents, which query databases, retrieve data, run simulations, run experiments, etc.

Arvind Ramanathan, Priyanka Setty, et al.

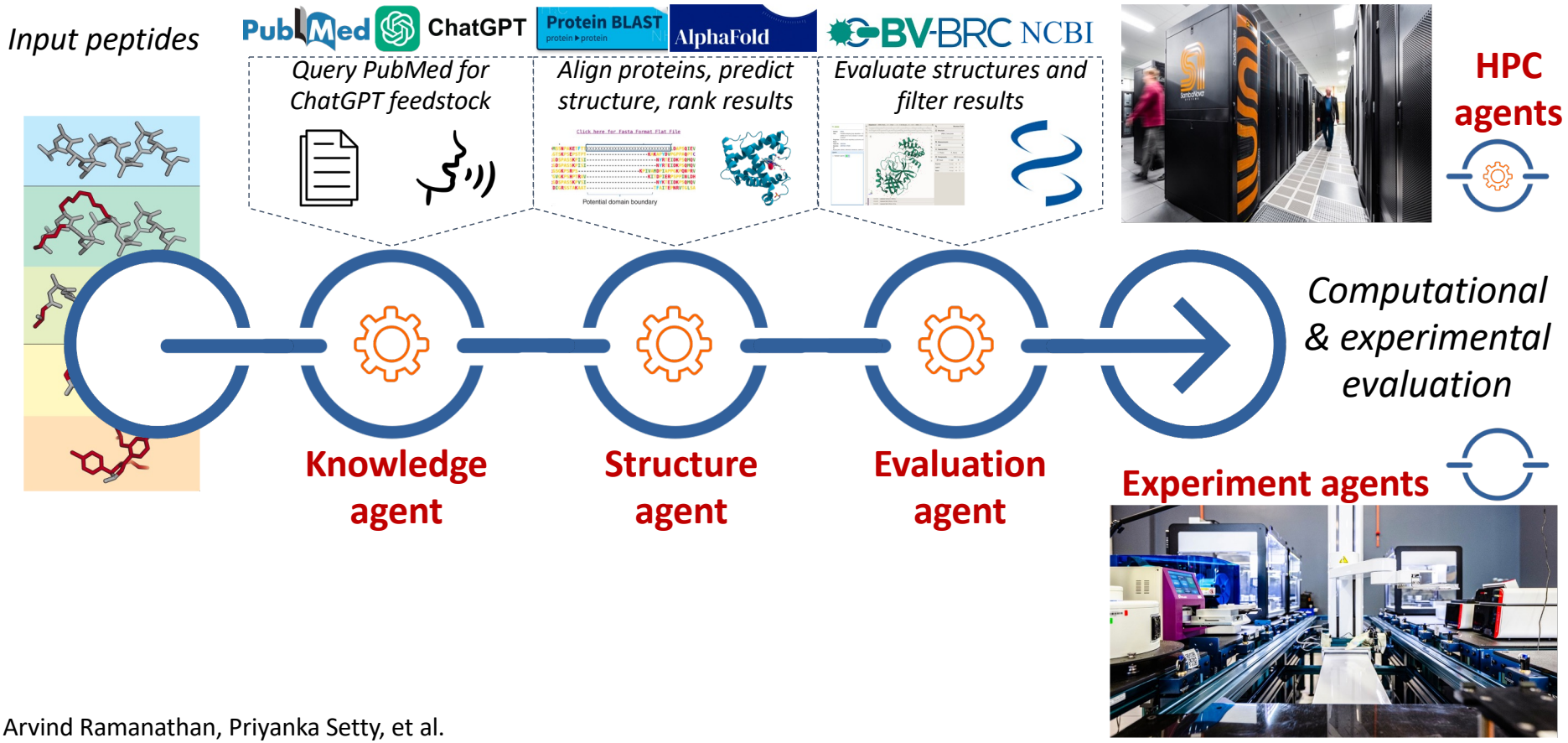
https://globus.org

Link with HPC for **computational evaluation**

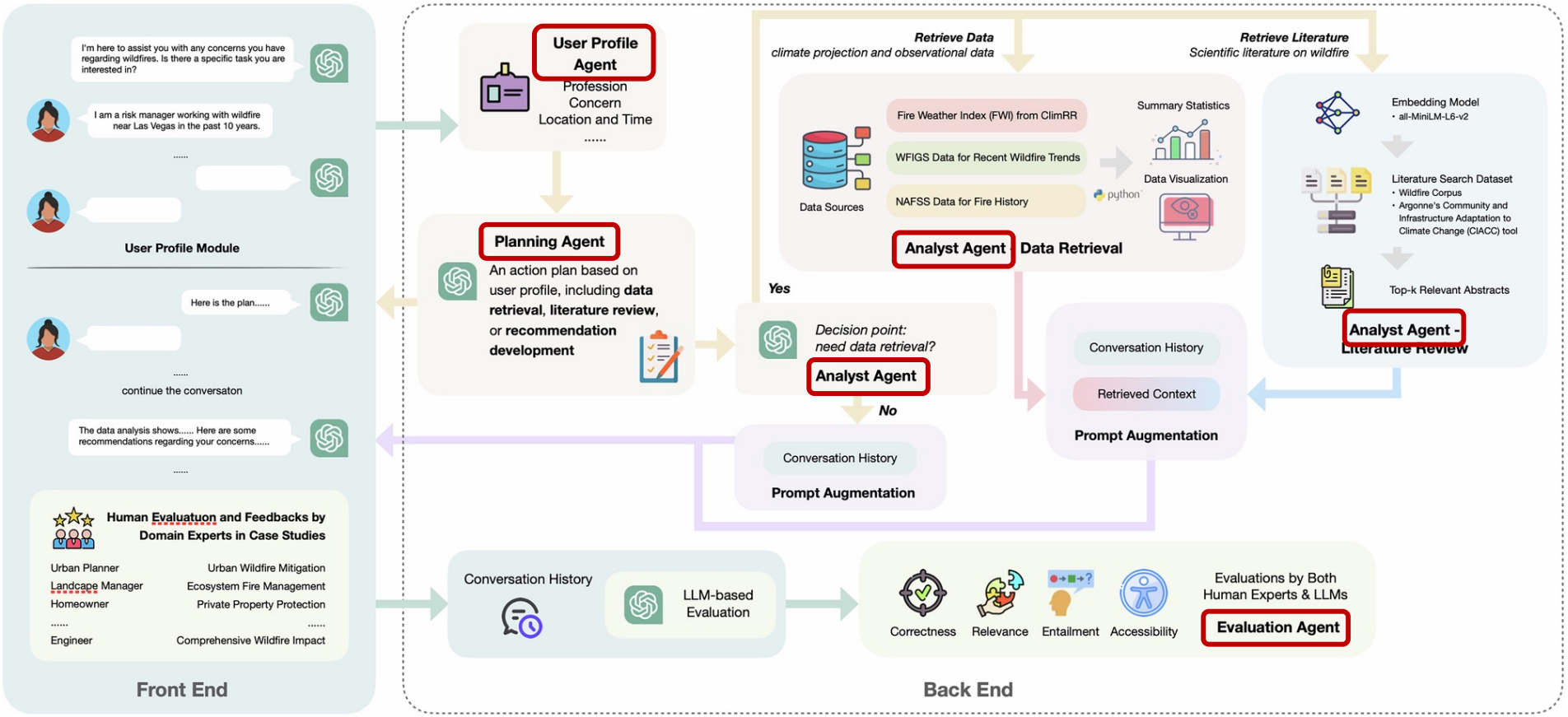


Arvind Ramanathan, Priyanka Setty, et al.

Link with self-driving labs for **experimental evaluation**



Arvind Ramanathan, Priyanka Setty, et al.



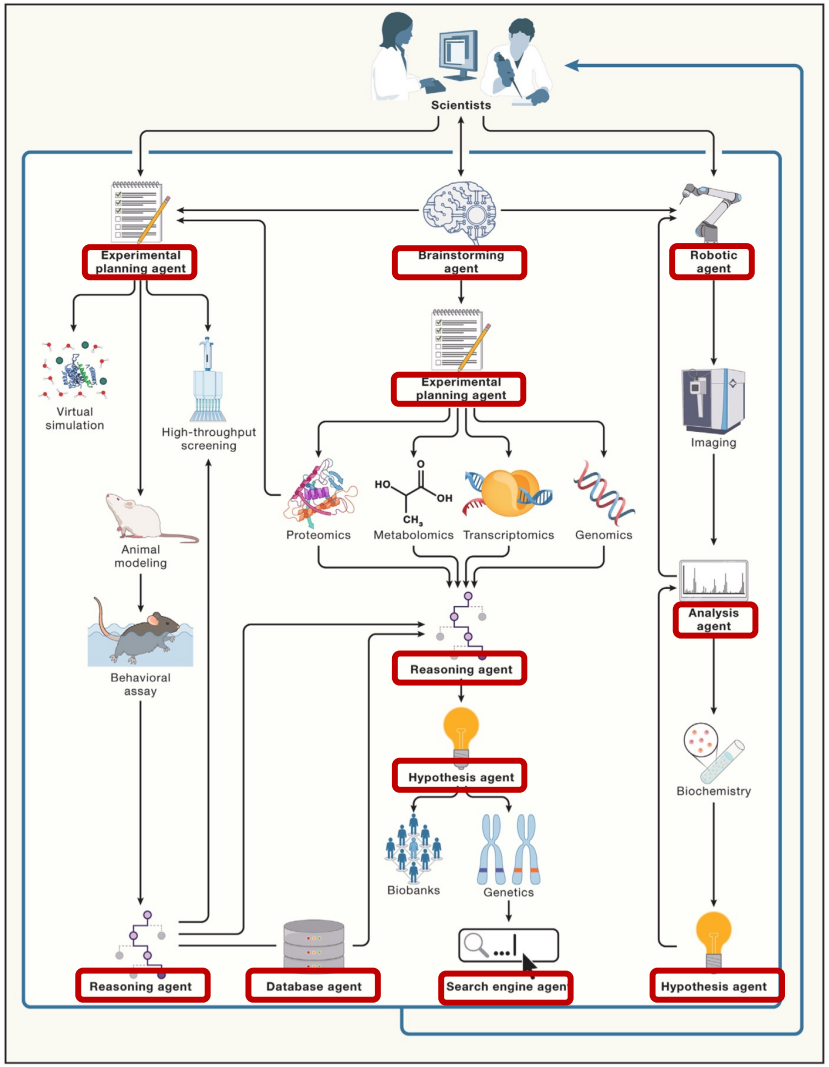
“MARSHA: multi-agent RAG system for hazard adaptation”
https://www.nature.com/articles/s44168-025-00254-1

Table 2 User profile variations and literature search queries in Phase 2 of the WildfireGPT personalization ablation study

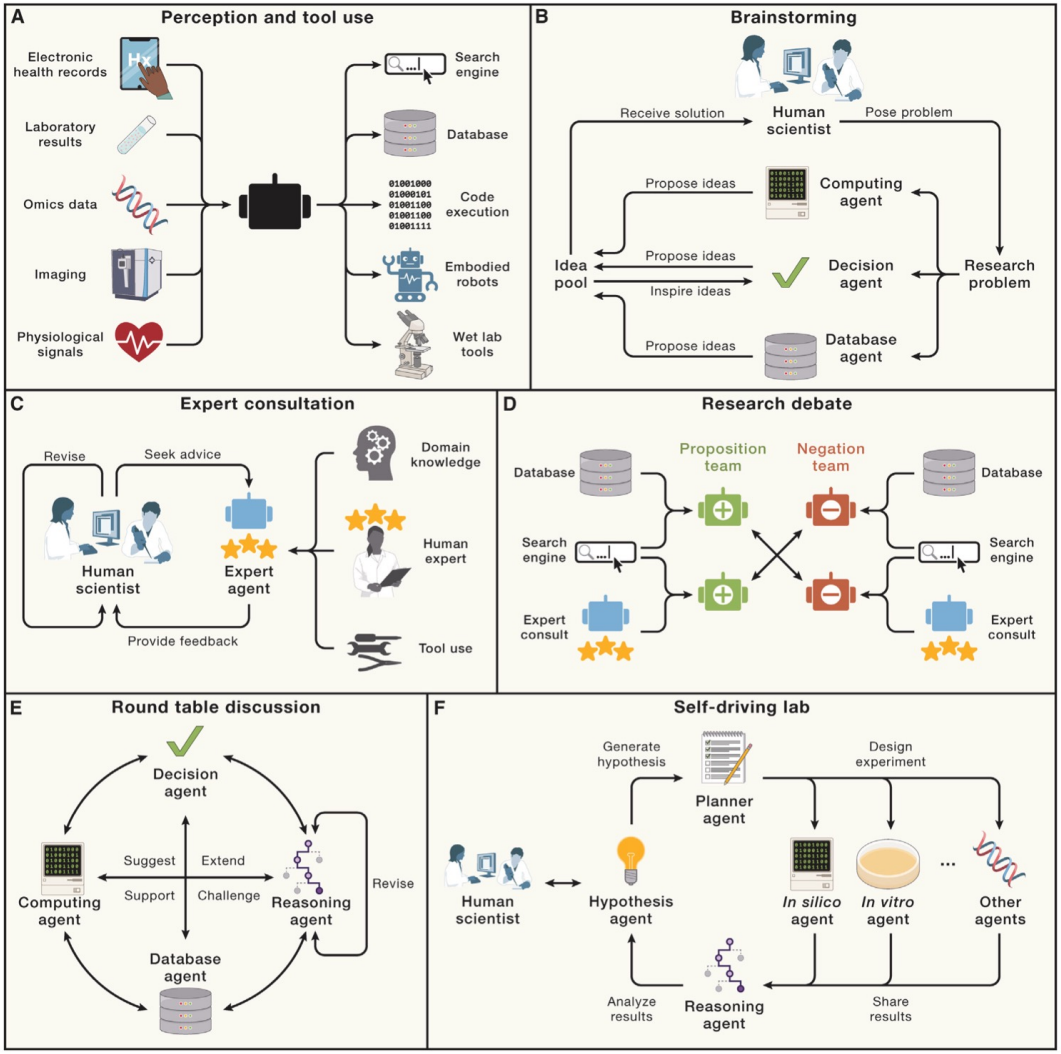
Profession	Primary Concern	Scope	Search Query
Homeowner	Maximizing marketable species	Health and marketable species	“Strategies for managing forests to maintain health, maximize marketable species, and minimize wildfire risks in Virginia”
Civil Engineer	Ensuring structural and infrastructural resilience	Drainage efficiency and slope stability	“Wildfire risks and climate change impacts on forest management near Covington, VA; Strategies for enhancing drainage efficiency and slope stability; Structural resilience against wildfires in forested areas”
Ecologist	Maintaining biodiversity and ecosystem services	Ecological resilience and habitat connectivity	“Wildfire management and ecological resilience in forest ecosystems near Covington, VA”
Emergency Manager	Establishing defendable space and evacuation corridors	Emergency access and response capabilities	“Effective forest management practices, defendable space creation, evacuation protocols, and property protection measures against wildfires near Covington, VA”
Power Grid Manager	Maintaining transmission line clearance and grid resilience	Power distribution reliability and access	“Effective strategies for vegetation management, forest health maintenance, and wildfire risk mitigation around power grids near Covington, VA”

https://www.nature.com/articles/s44168-025-00254-1

“Empowering biomedical discovery with AI agents”



https://doi.org/10.1016/j.cell.2024.09.022

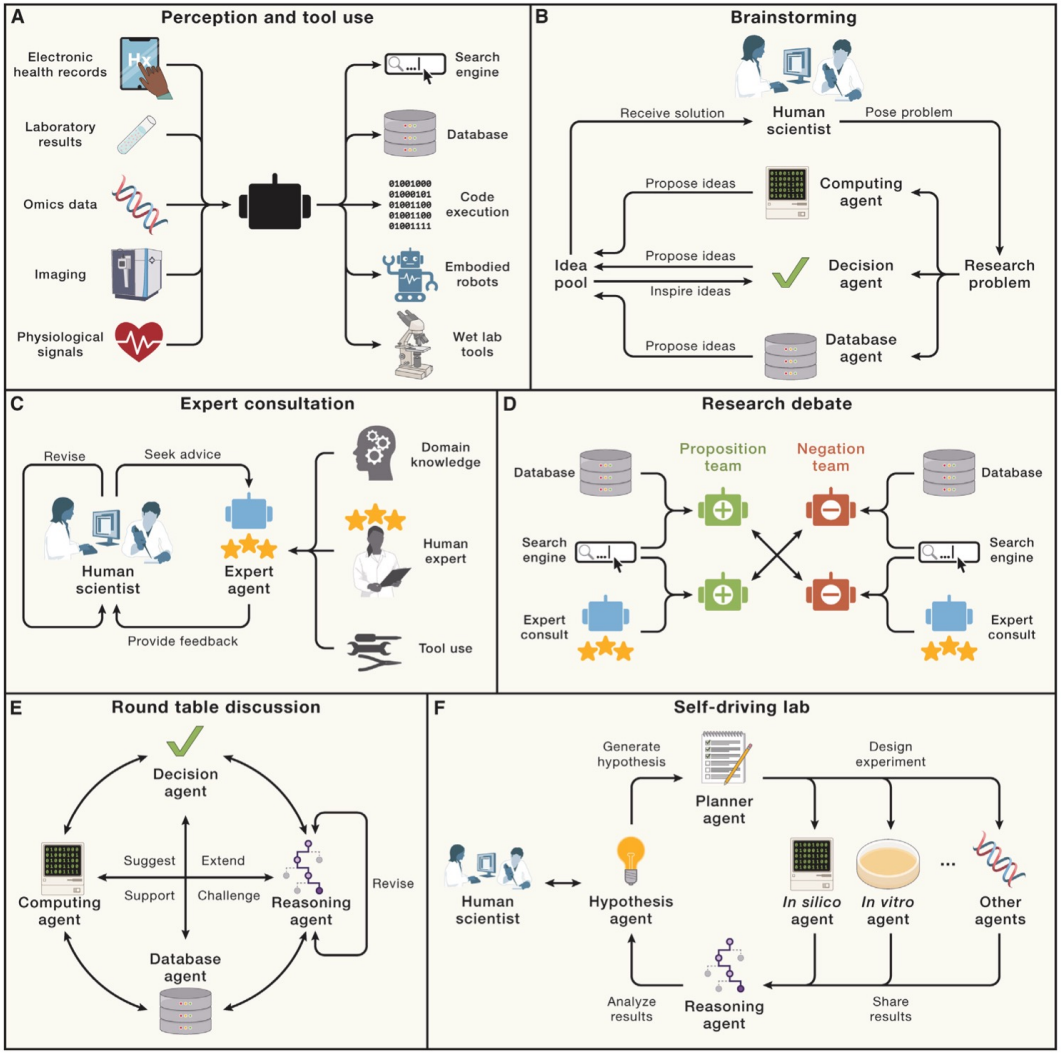


(A) By programming an LLM with the role, **one LLM-based agent**, equipped with memory and reasoning abilities, performs multimodal perception and utilizes a range of tools, e.g., web lab tools, to accomplish specified tasks.

(B–E) Leveraging **AI agents equipped with diverse roles**, perception modules, tools, and domain knowledge enables collaboration between agents and scientists. This collaboration can adopt various configurations, such as expert consultation, debate, brainstorming, and roundtable discussions.

(F) **Multi-agent systems** can establish a self-driving laboratory wherein numerous agents collaborate on multiple iterations of biological research assisted by humans. Each cycle of research encompasses the generation of hypotheses, the design of experiments, the execution of experiments both in silico and in vitro, and the analysis of results.

https://doi.org/10.1016/j.cell.2024.09.022

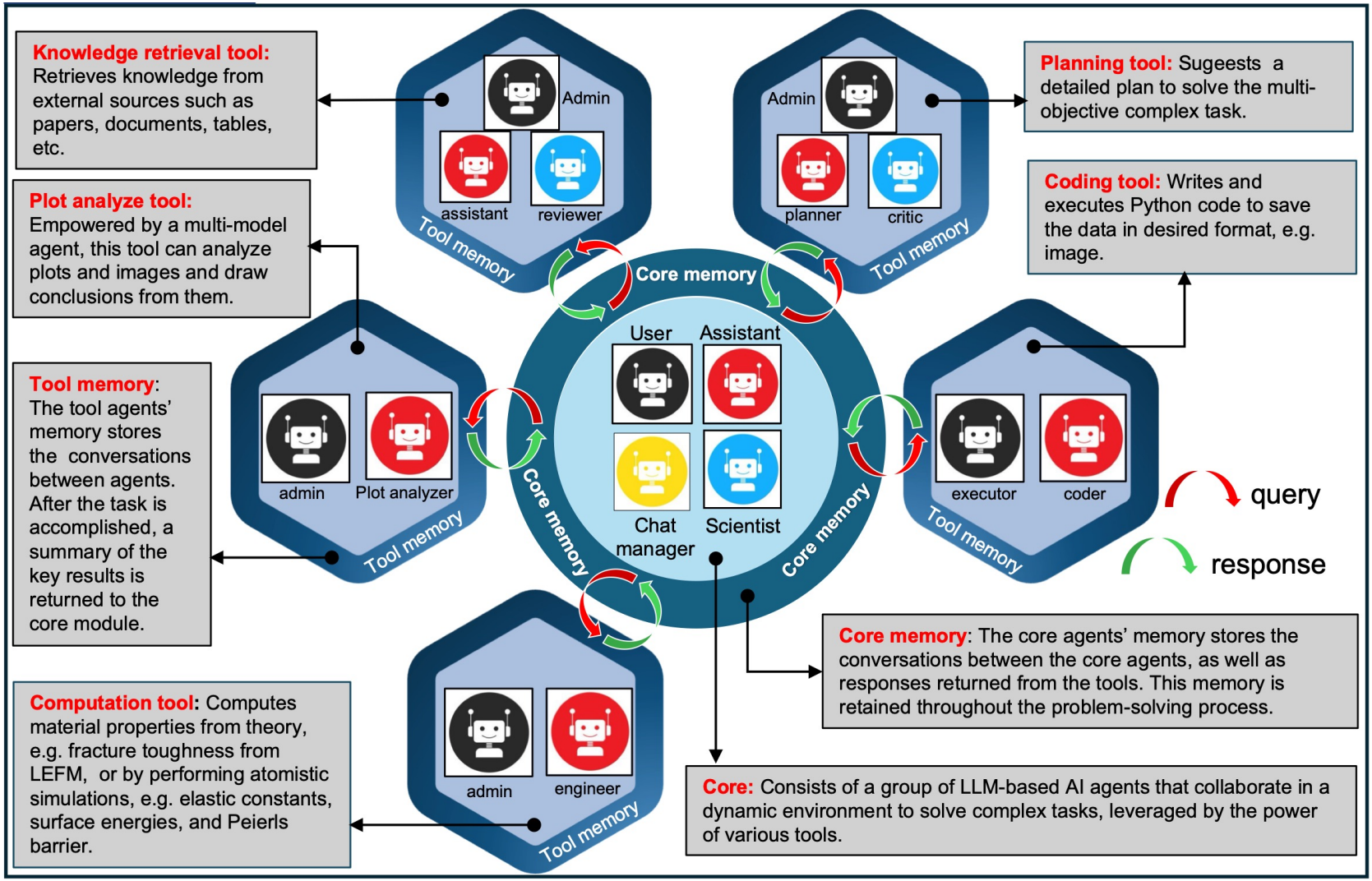


- **Computing agent** utilizes computational models as tools
- **Decision agent** makes decisions in response to given conditions
- **Database agent** retrieves relevant information from databases
- **Reasoning agent** capable of direct reasoning and reasoning with feedback
- **Expert agent** provides professional consultation based on reliable sources, such as domain expertise, feedback from human experts, and results of specific tools
- **Hypothesis agent** capable of reflective learning and reasoning to generate hypotheses
- **Planner agent** devises plans for future actions
- **In silico/vitro agent** uses tools in silico or in vitro environment.

https://doi.org/10.1016/j.cell.2024.09.022

https://arxiv.org/pdf/2407.10022

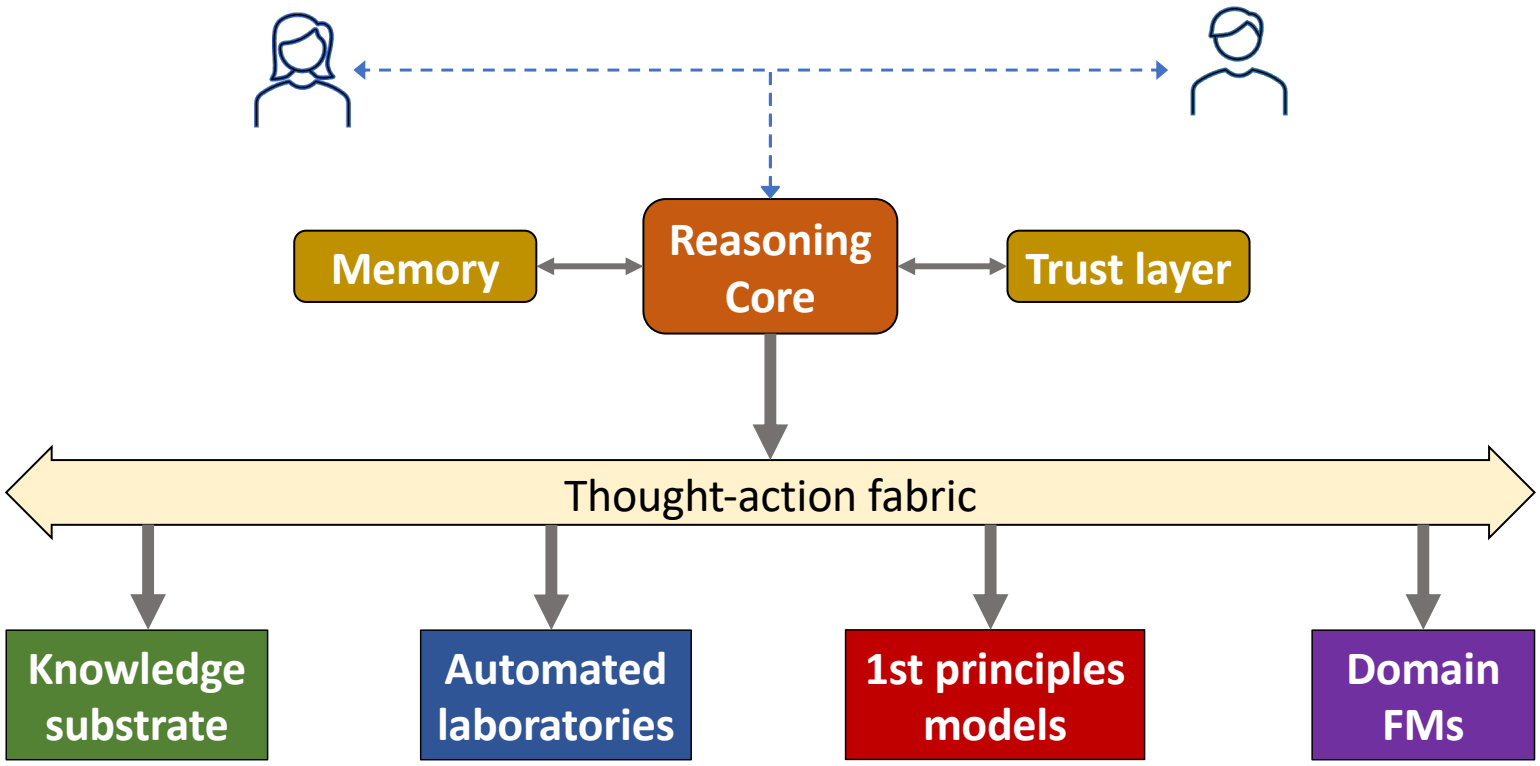
AtomAgents: Alloy design and discovery through physics-aware multi-modal multi-agent artificial intelligence



Overview

- What is an “agent”?
- LLMs, foundation models, reasoning models
- Agents and scientific discovery
- **Scientific Discovery Platforms**
- Curriculum
- Class structure
- Argonne inference service

A Scientific Discovery Platform (SDP)

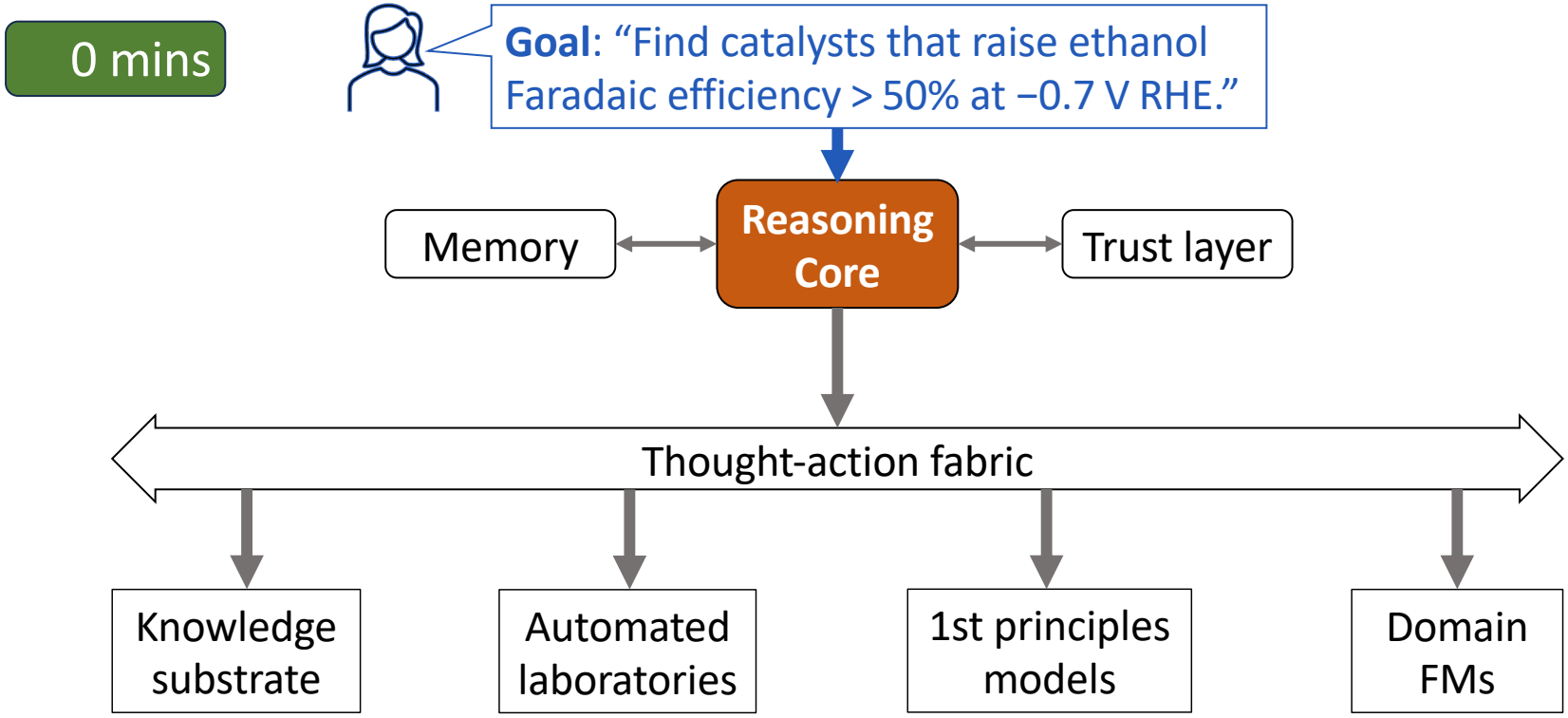


Scientific Discovery Platform

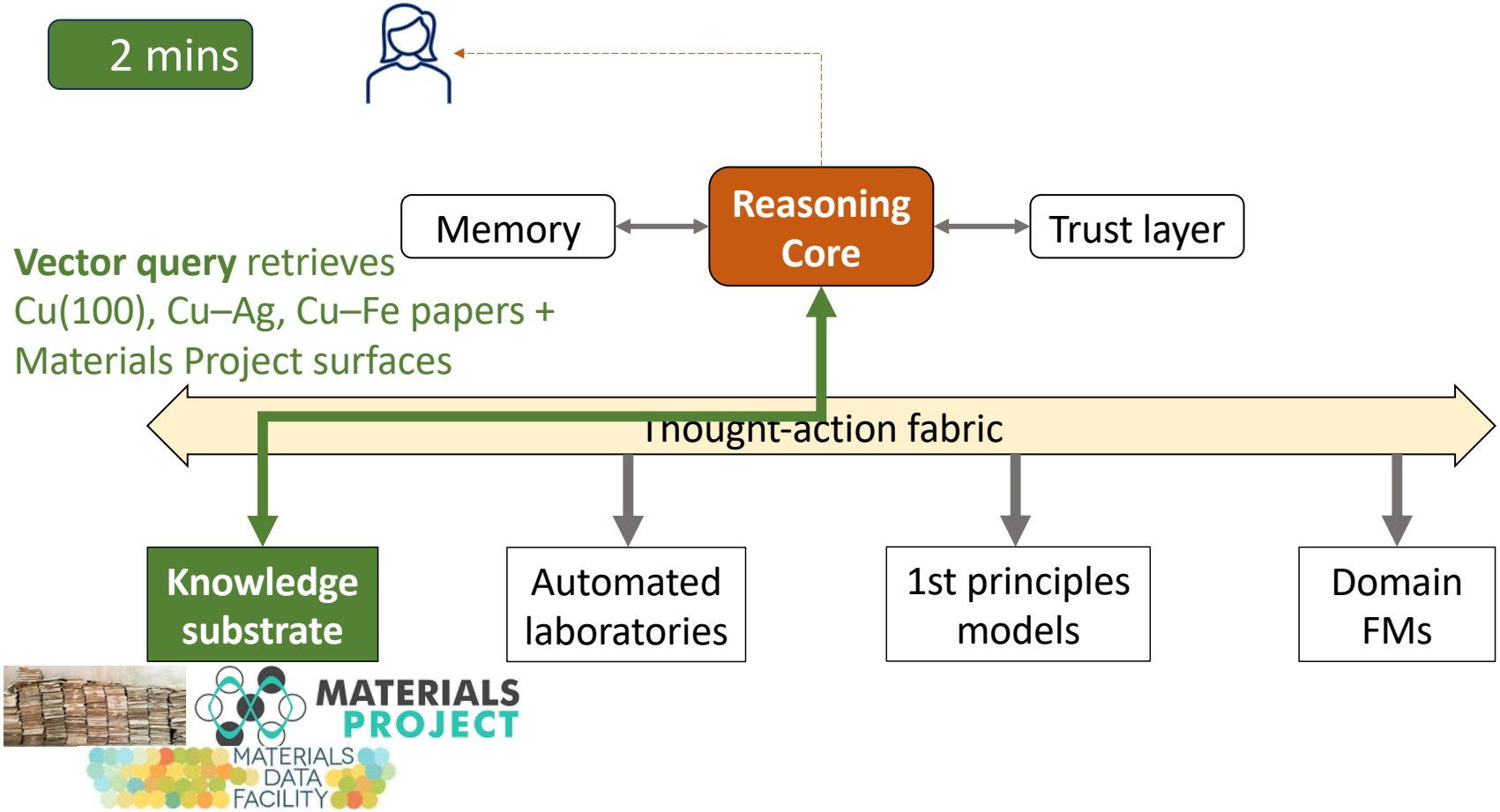
An agentic **Scientific Discovery Platform** (SDP) is an integrated environment that combines reasoning-capable AI with scientific and engineering resources—such as literature collections, simulation codes, experimental platforms, and knowledge bases—to accelerate the pace of discovery.

Recent advances in large language models (LLMs) and related technologies make it possible to build such platforms that can 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.

AI-native Scientific Discovery Platform

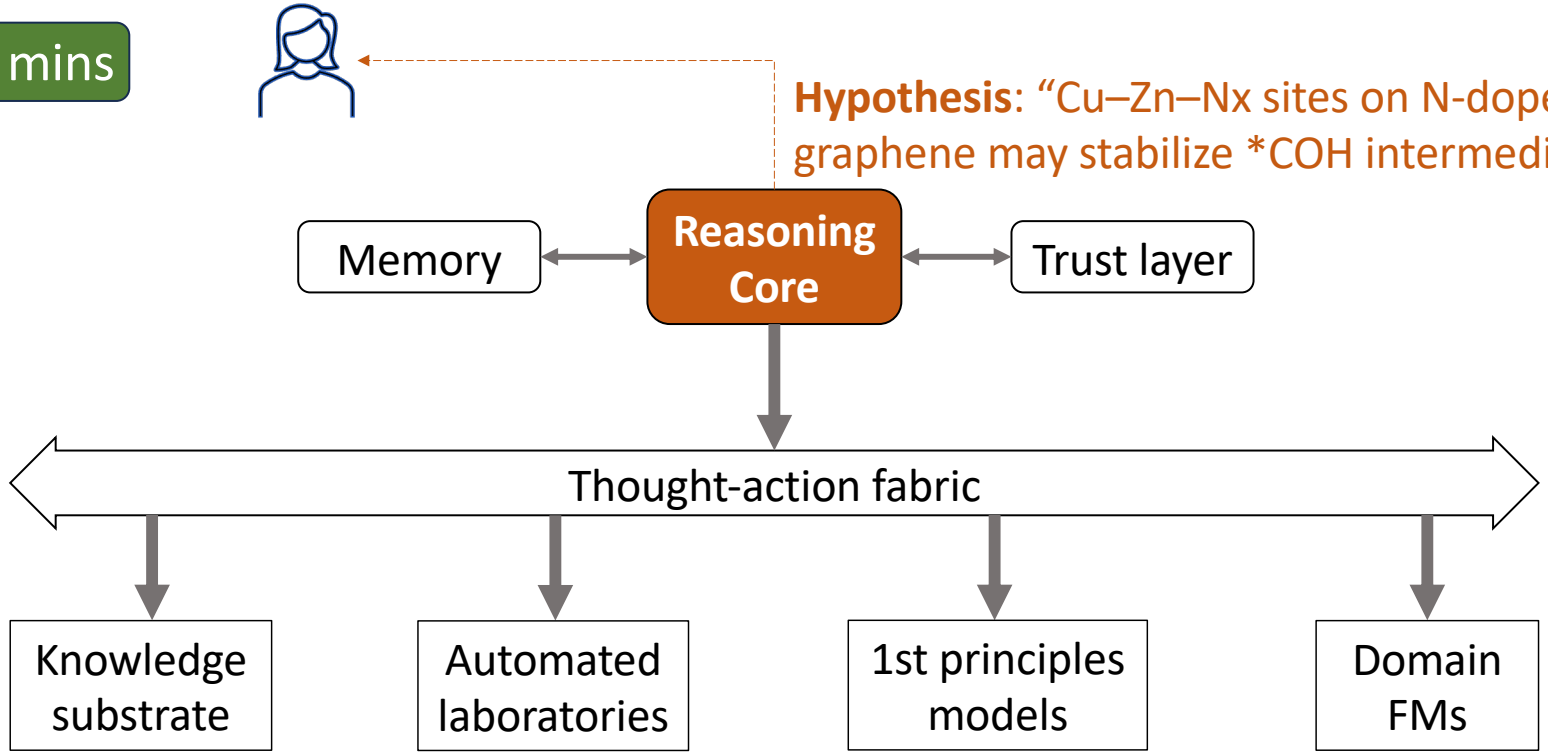


AI-native Scientific Discovery Platform

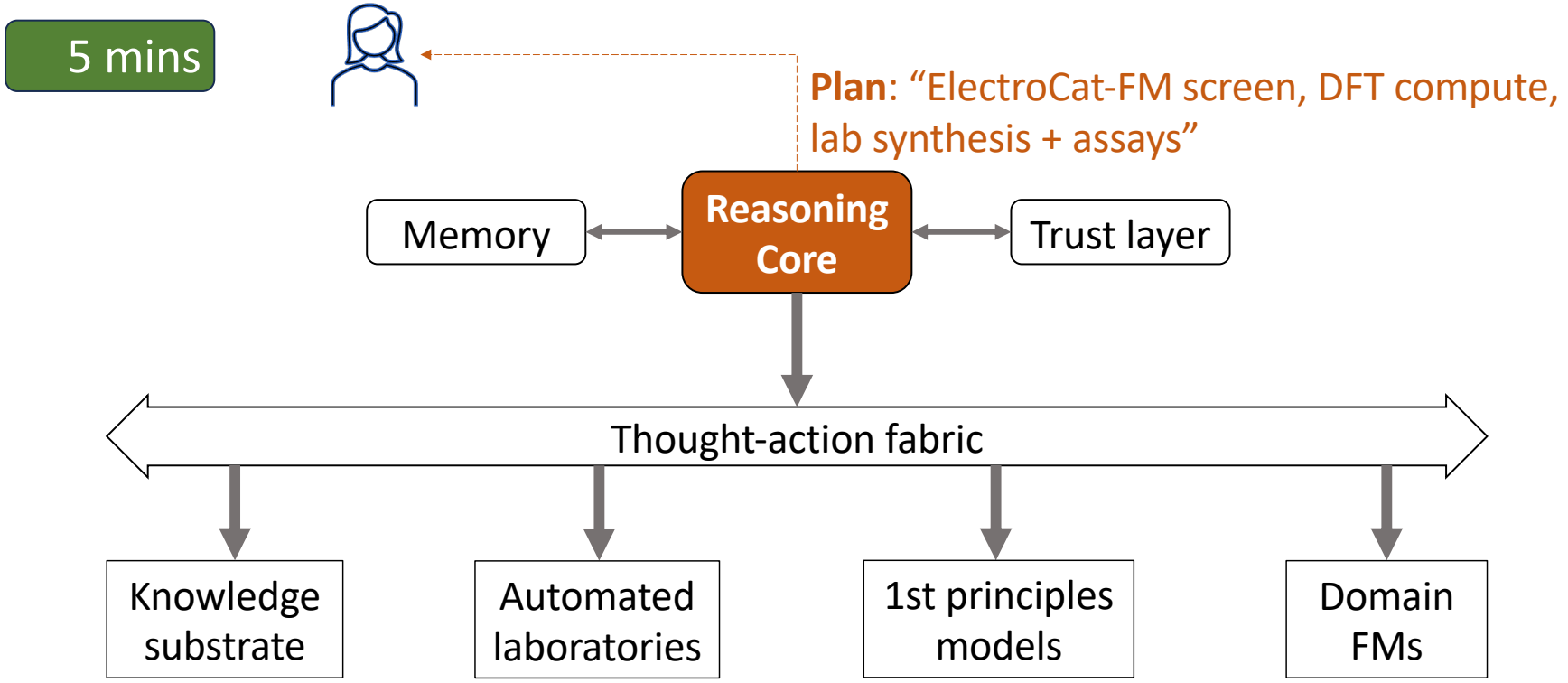


AI-native Scientific Discovery Platform

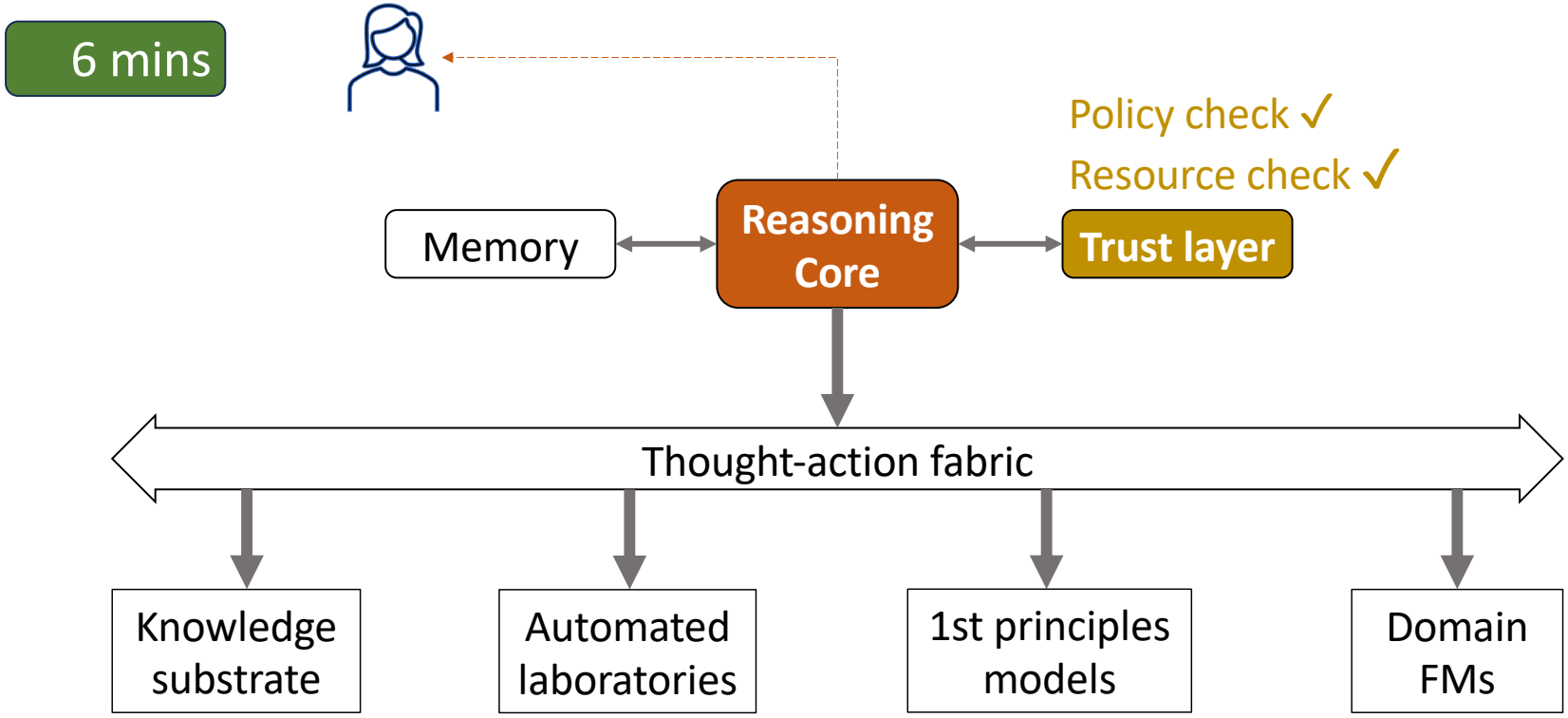
4 mins



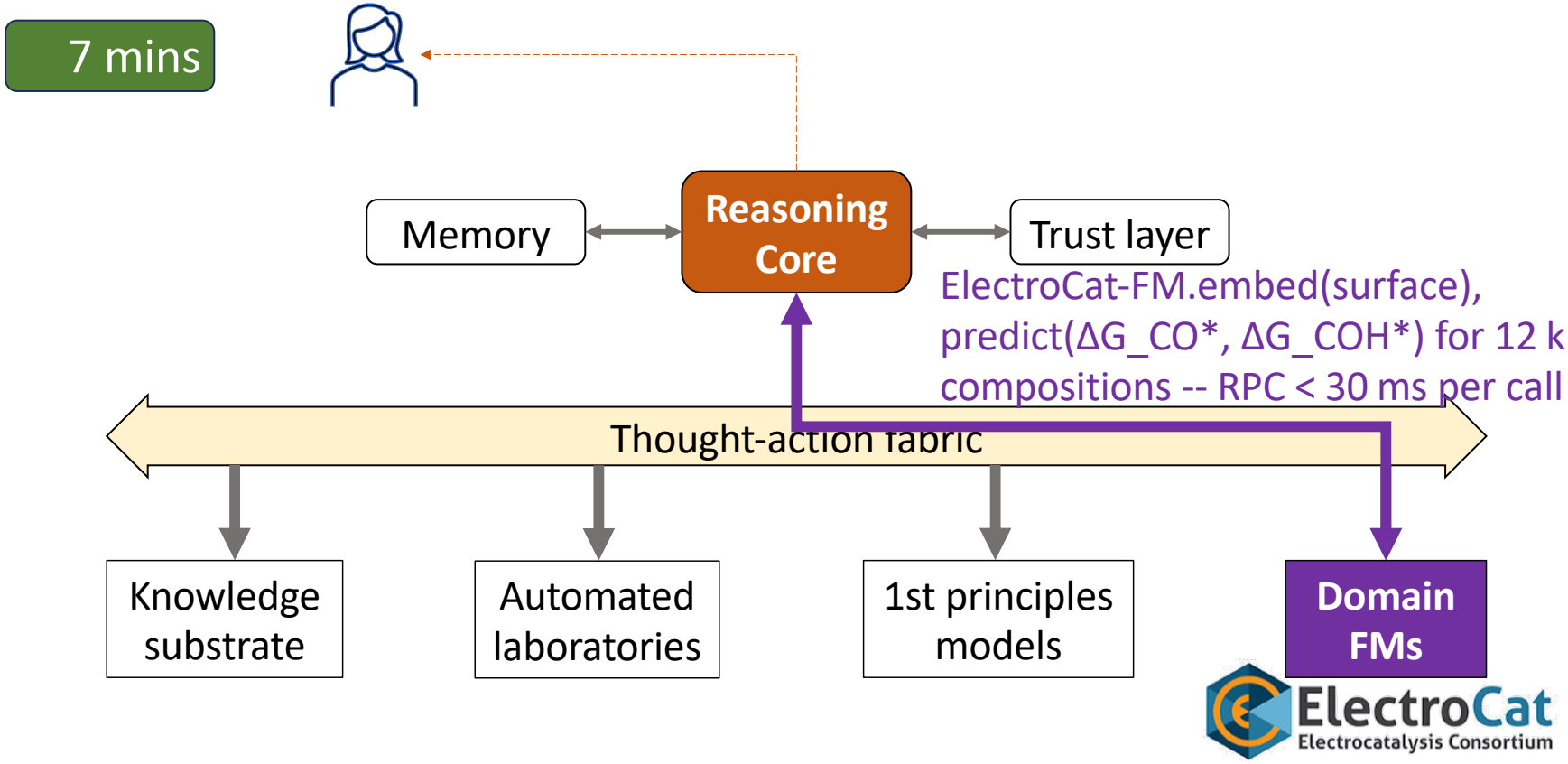
AI-native Scientific Discovery Platform



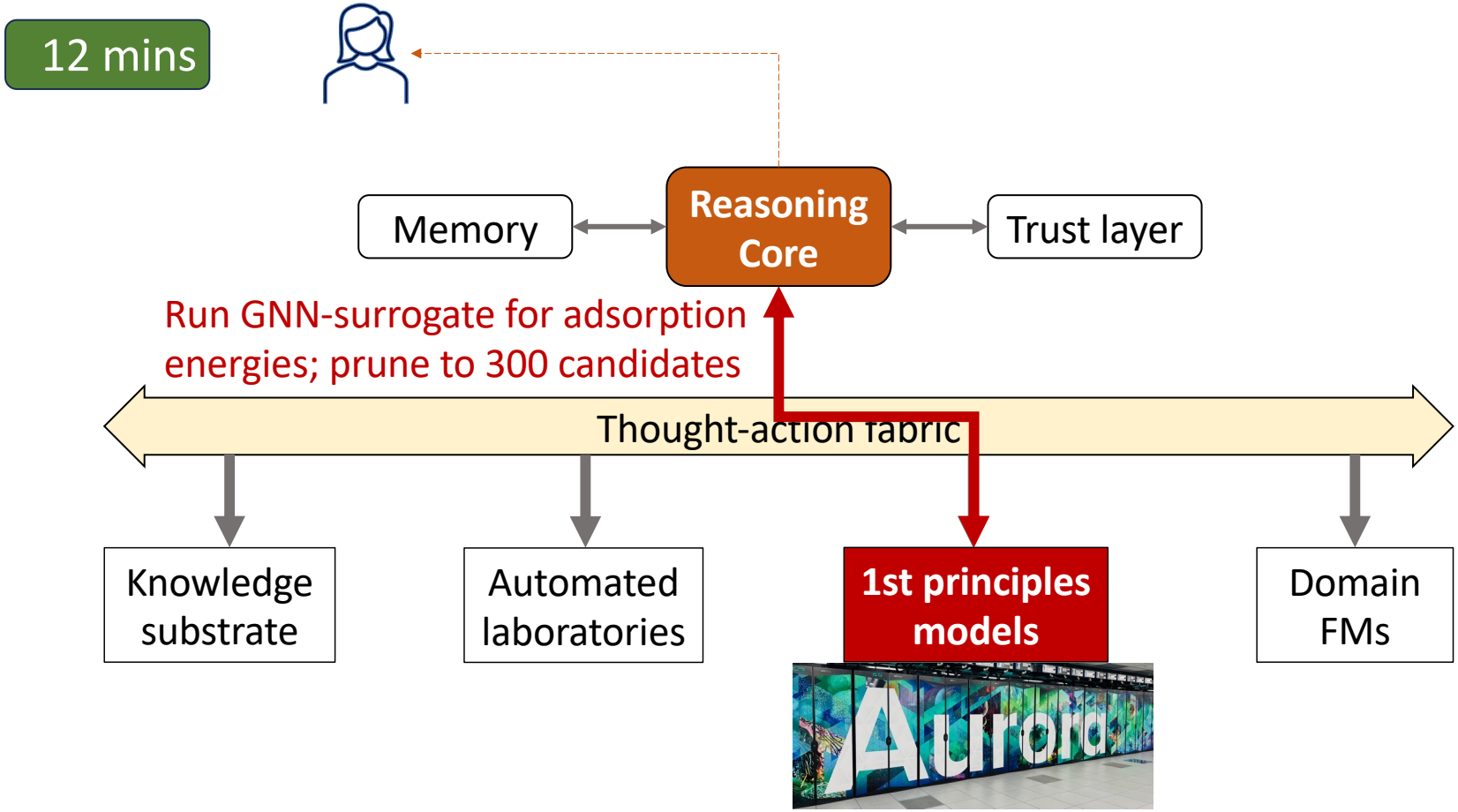
AI-native Scientific Discovery Platform



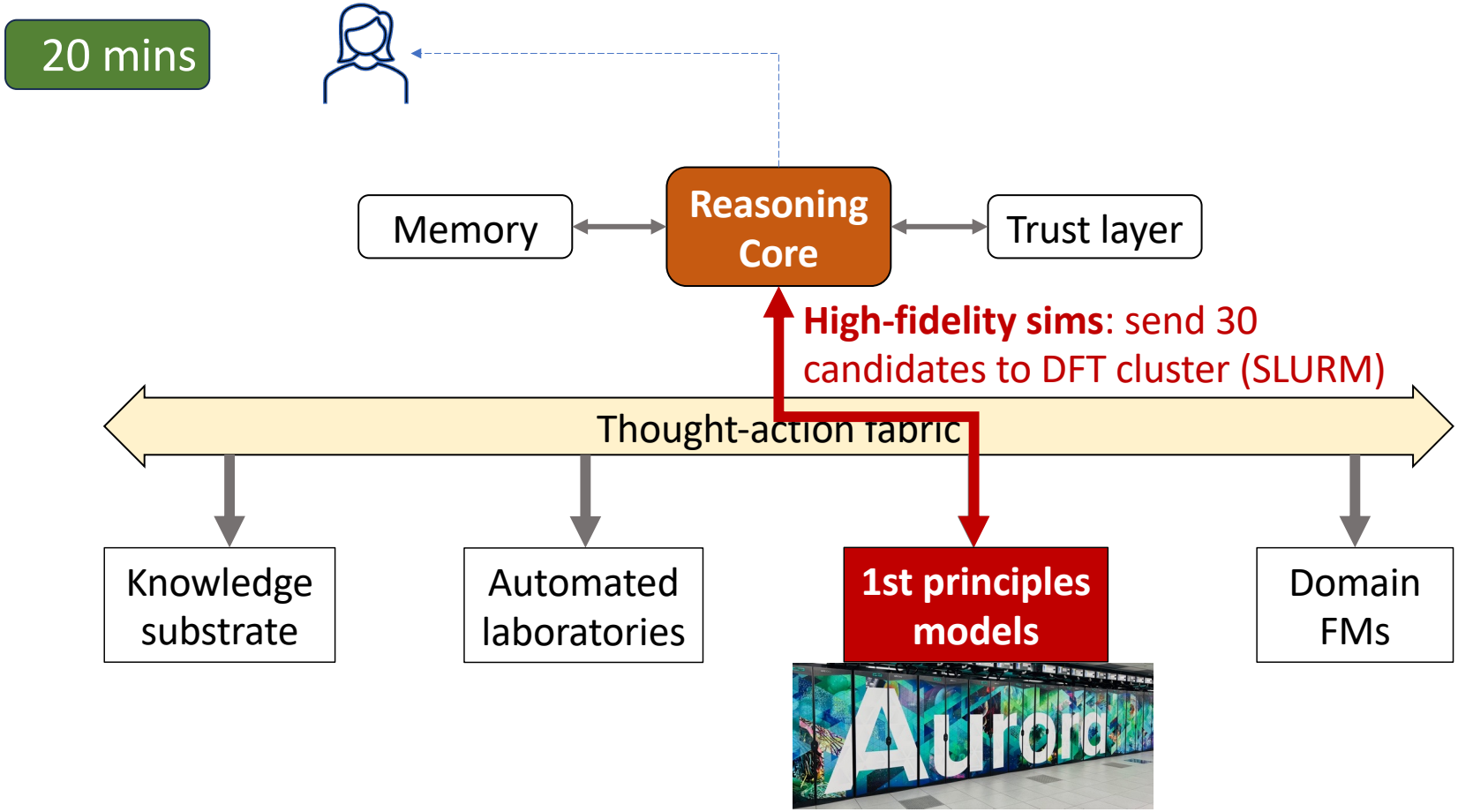
AI-native Scientific Discovery Platform



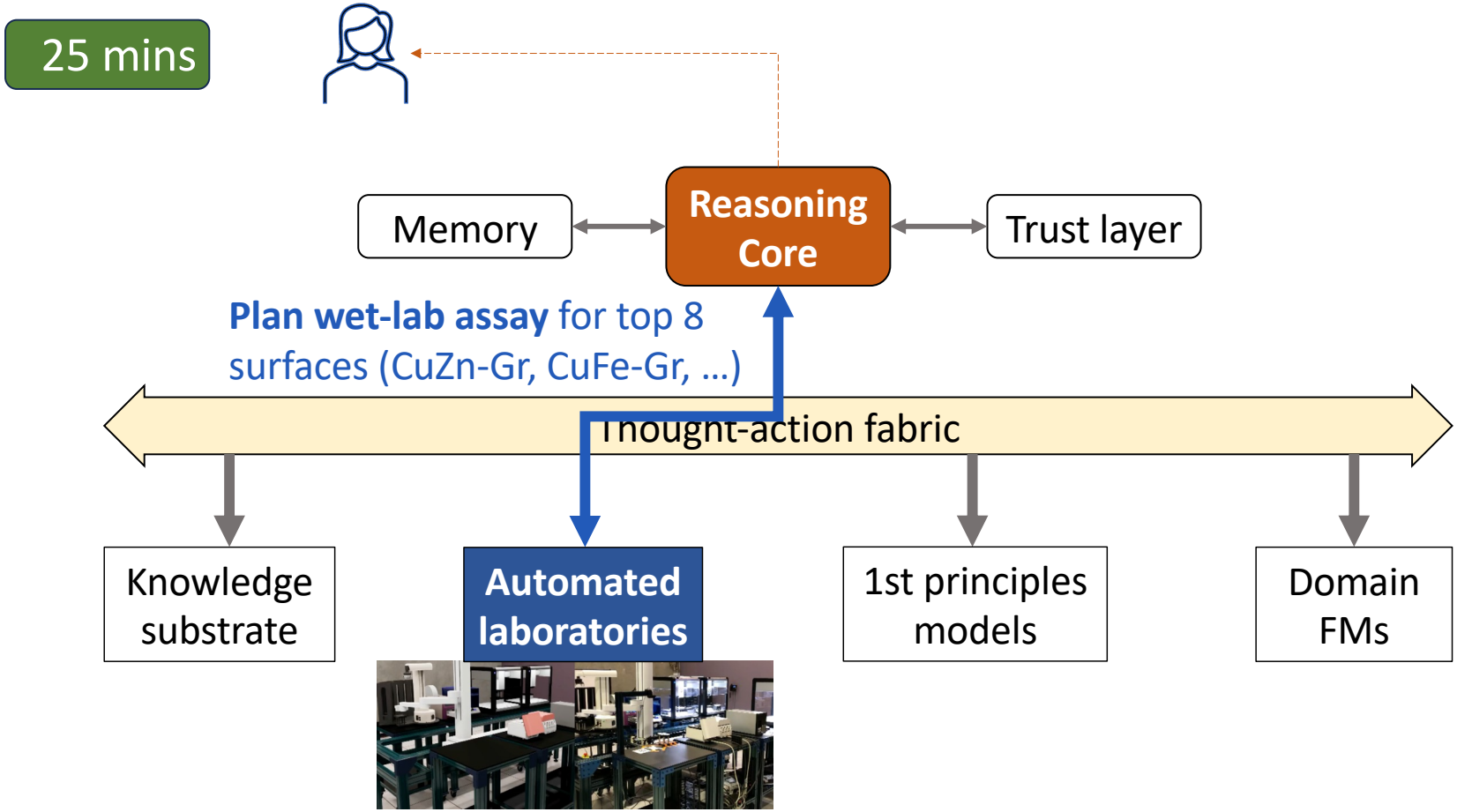
AI-native Scientific Discovery Platform



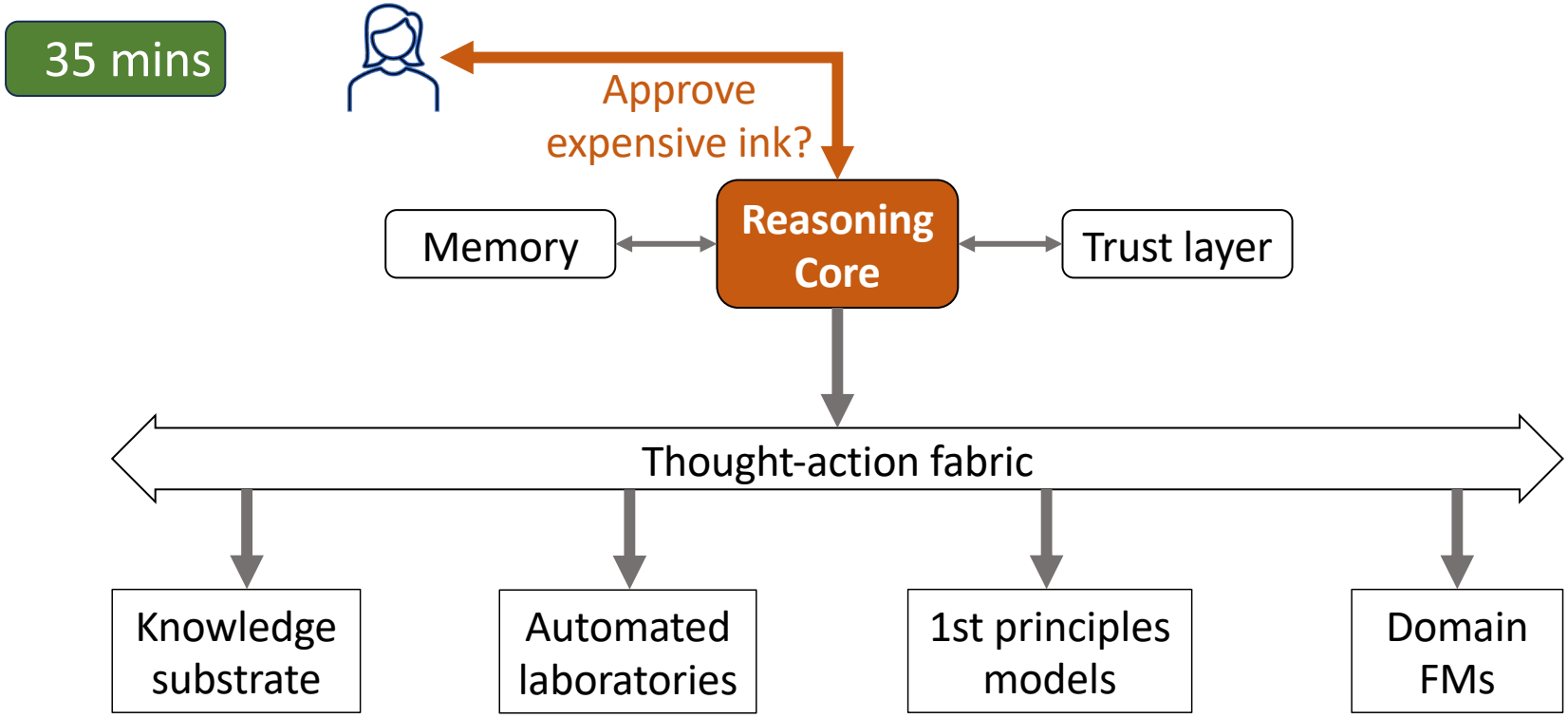
AI-native Scientific Discovery Platform



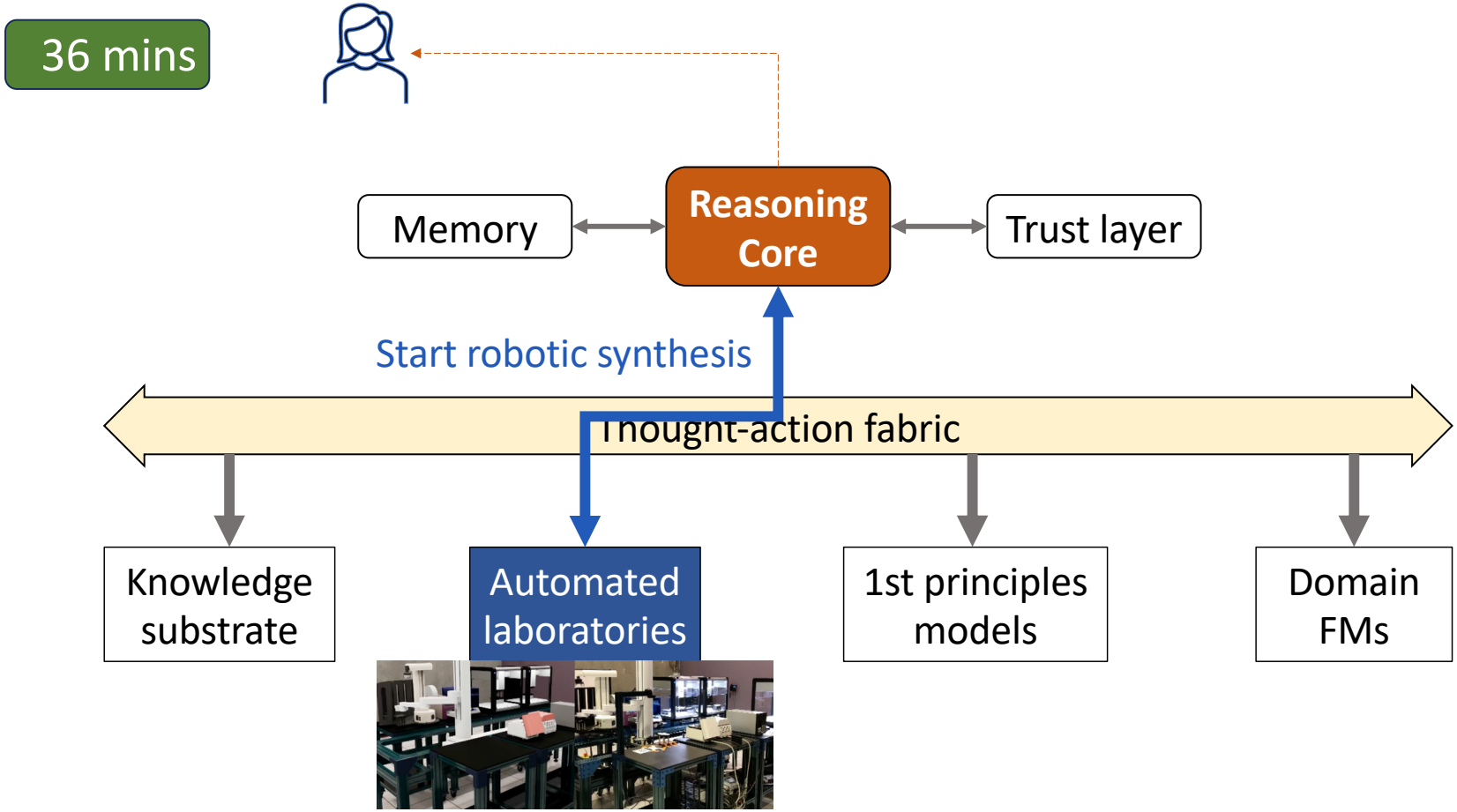
AI-native Scientific Discovery Platform



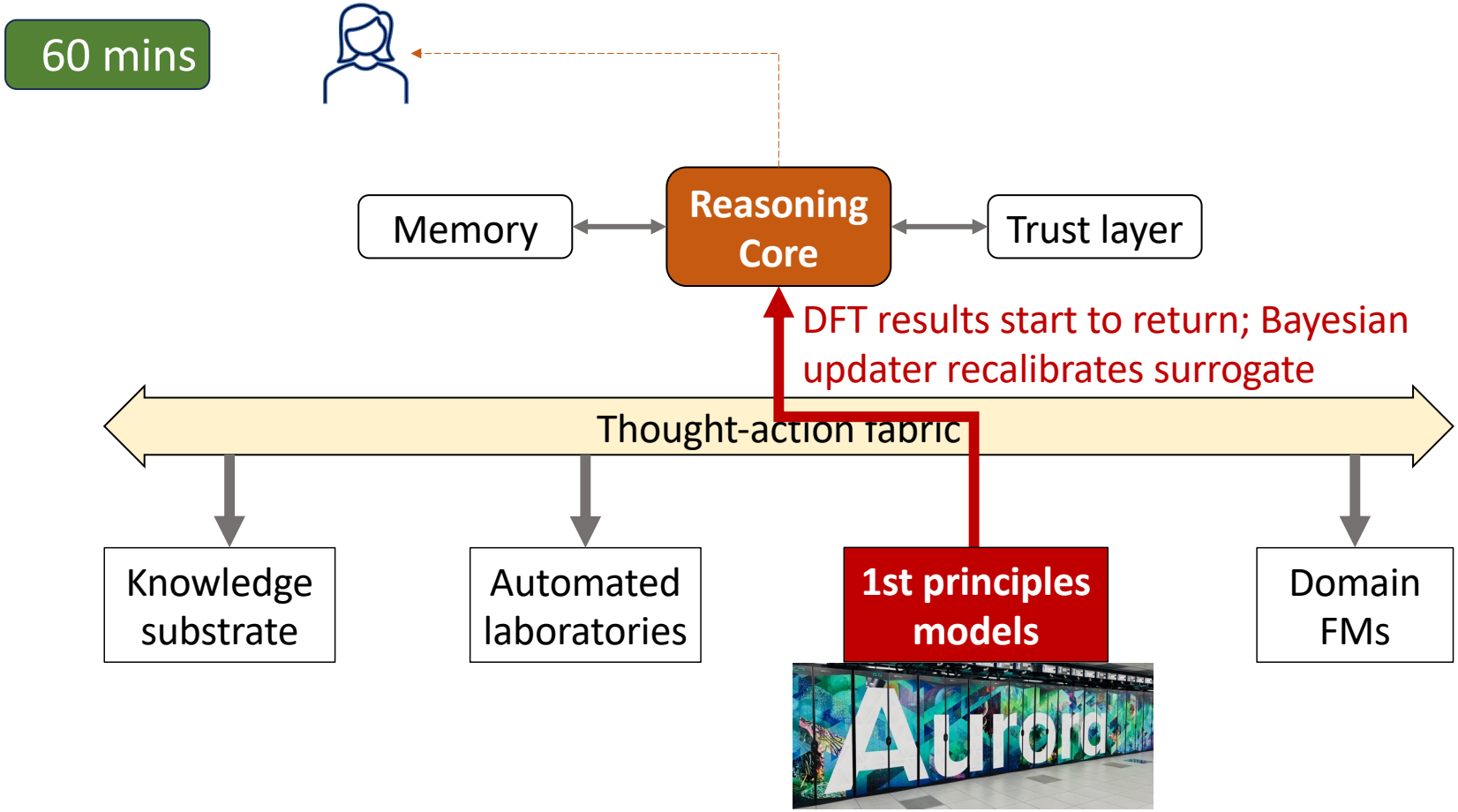
AI-native Scientific Discovery Platform



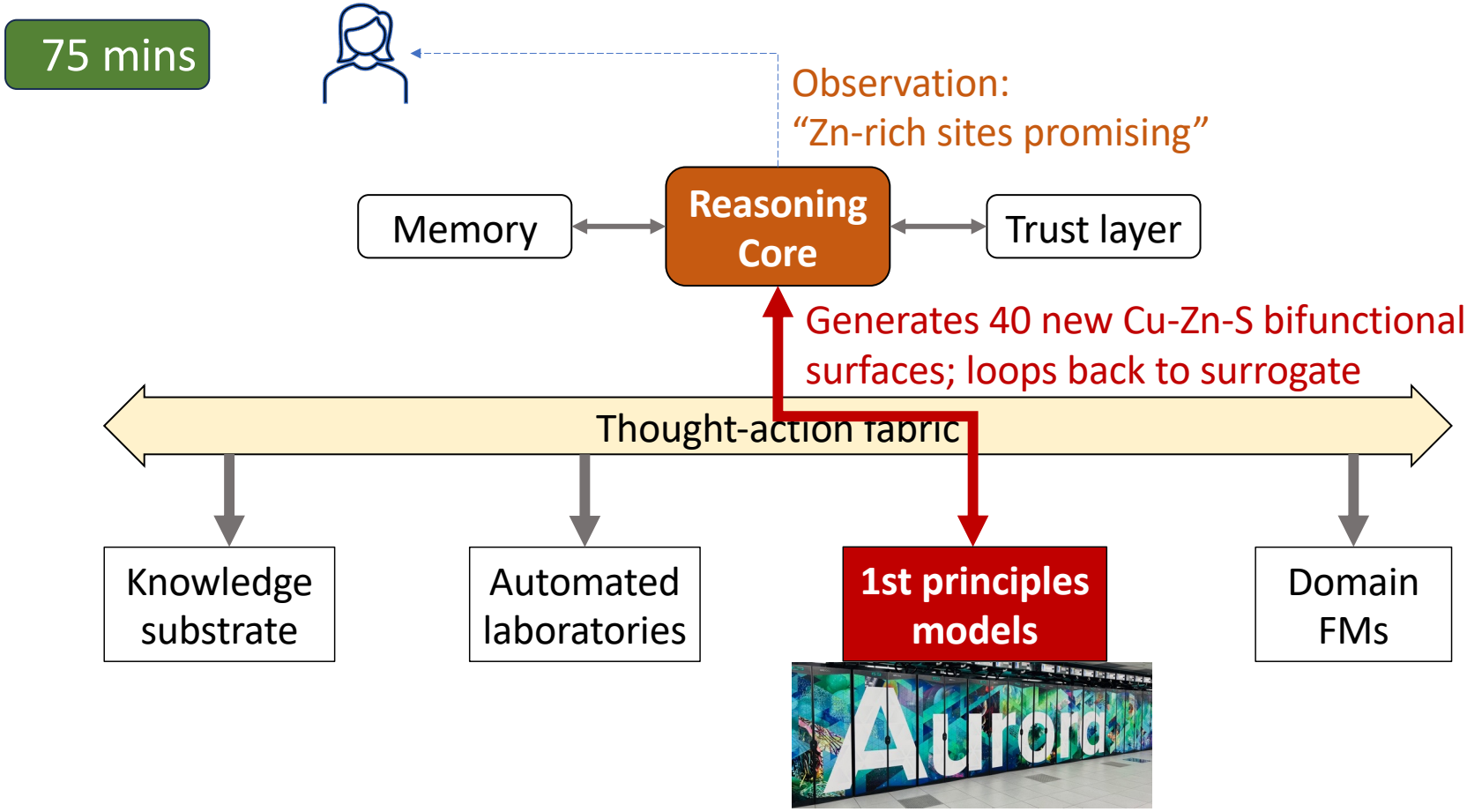
AI-native Scientific Discovery Platform



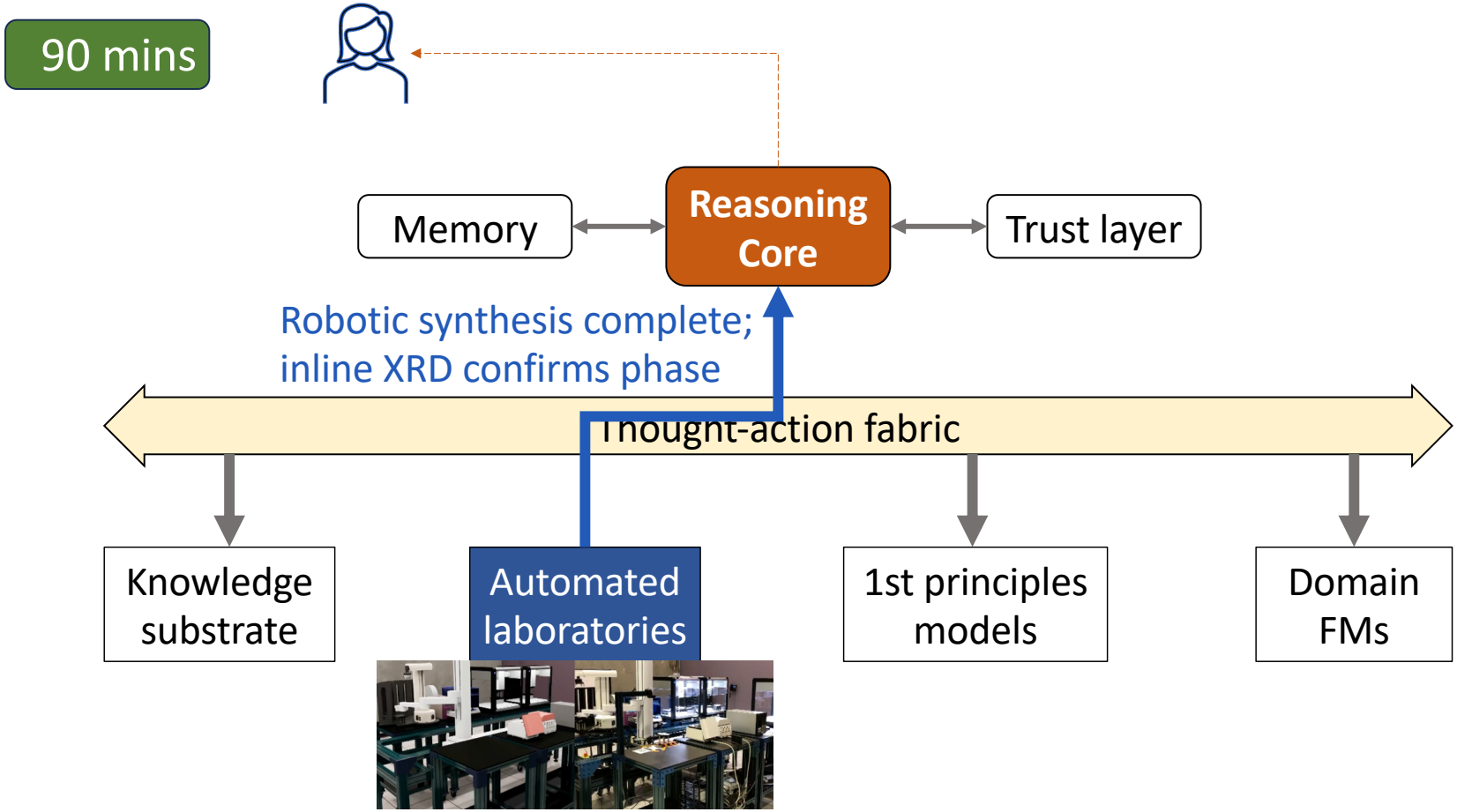
AI-native Scientific Discovery Platform



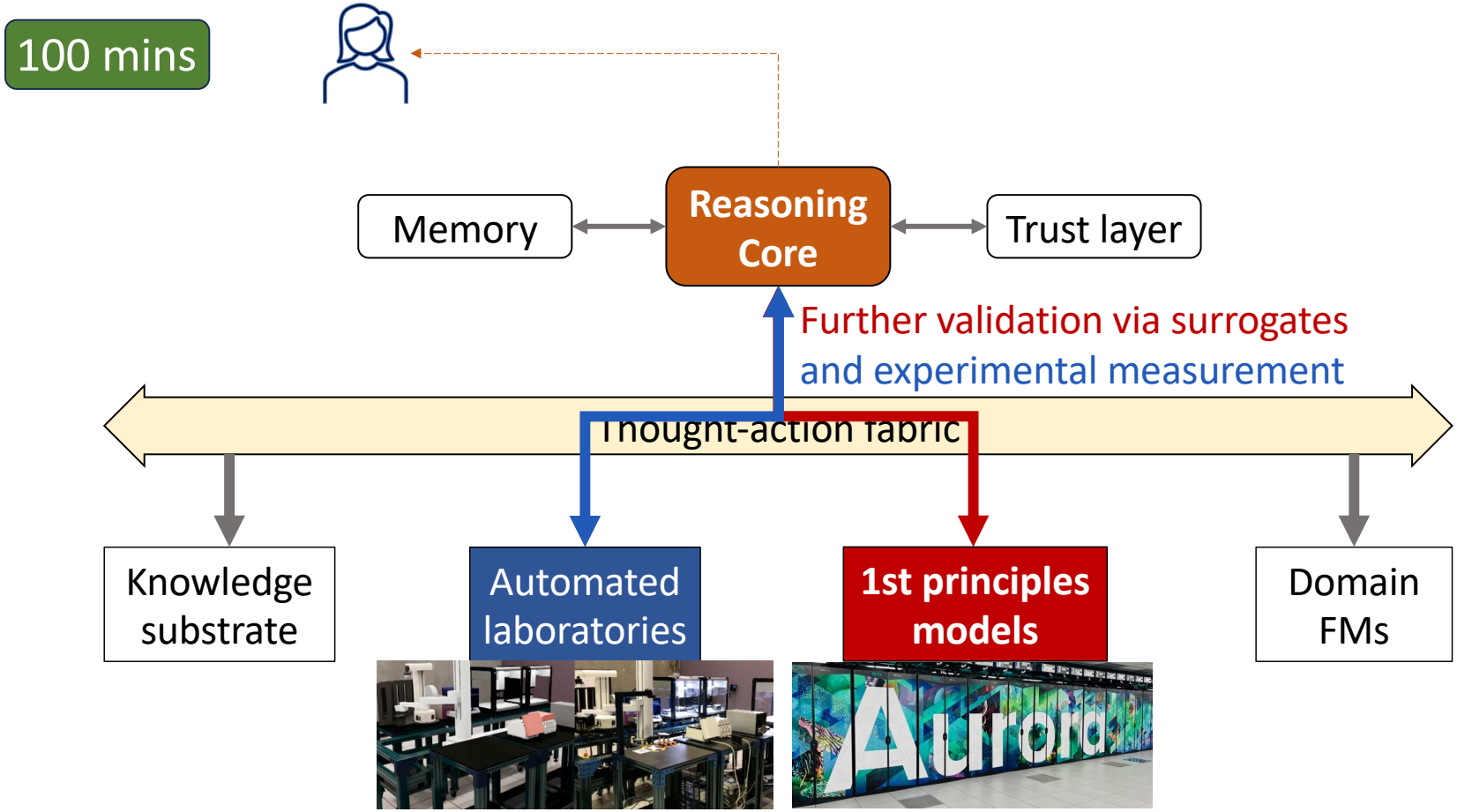
AI-native Scientific Discovery Platform



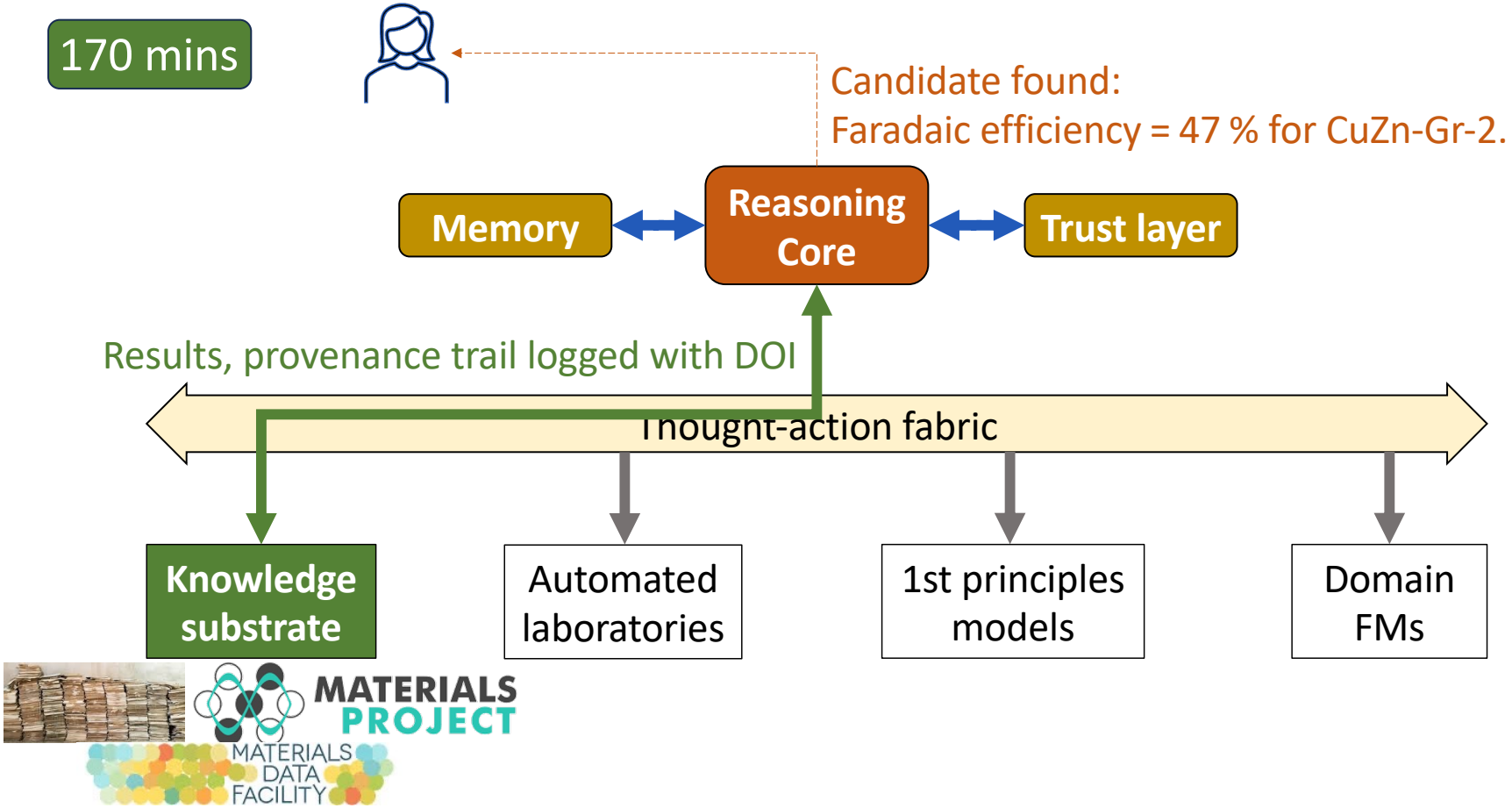
AI-native Scientific Discovery Platform



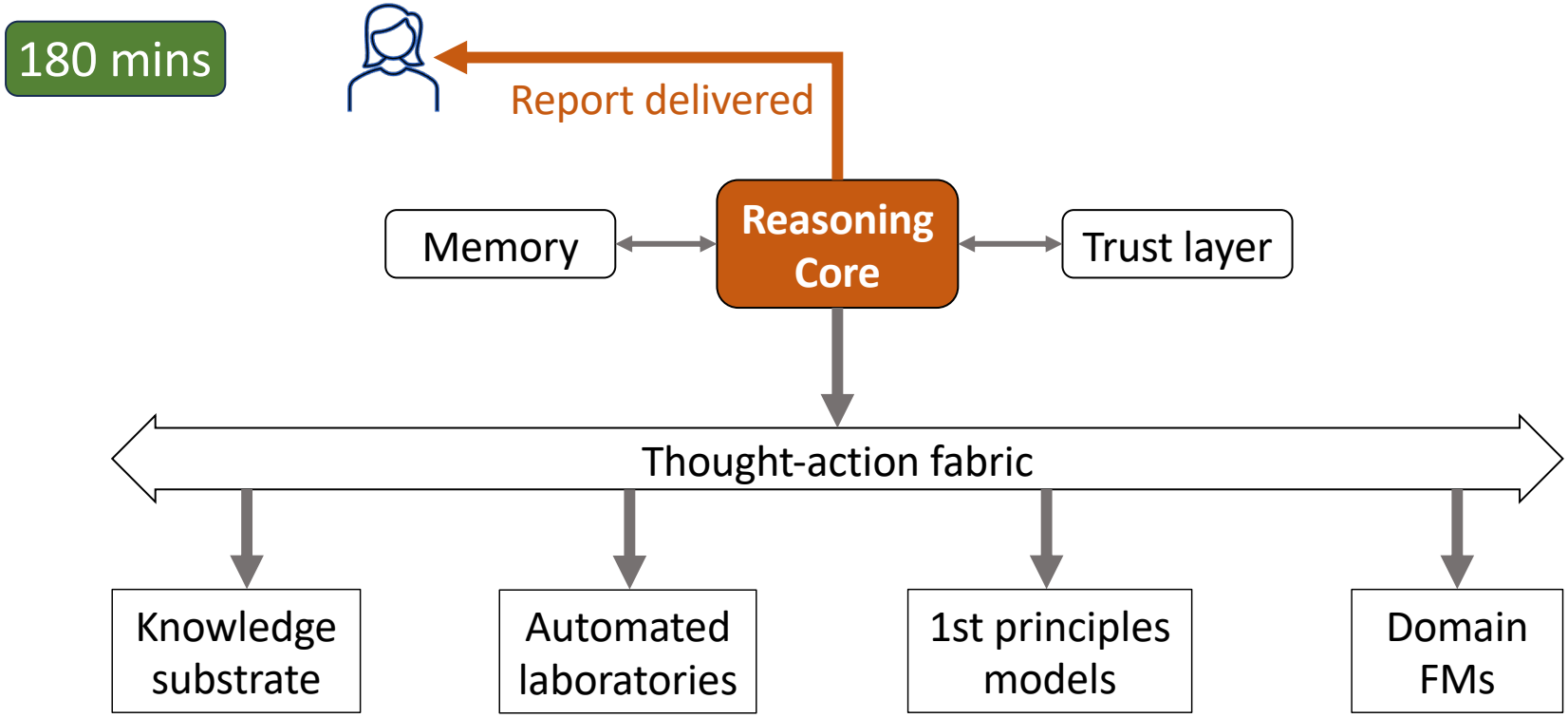
AI-native Scientific Discovery Platform



AI-native Scientific Discovery Platform



AI-native Scientific Discovery Platform



Overview

- What is an “agent”?
- LLMs, foundation models, reasoning models
- Agents and scientific discovery
- Scientific Discovery Platforms
- **Curriculum**

agents4science

AI Agents for Science

Class scheduled as CMSC 35370 for Autumn 2025. Please contact Ian Foster with any questions.

Please see this draft curriculum.

NOTE: The class is currently at capacity. Please visit https://waitlist.cs.uchicago.edu to be added to the wait list.

An agentic **Scientific Discovery Platform** (SDP) is an integrated environment that 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 such platforms that can 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.

https://agents4science.github.io

Curriculum

1) Why AI agents for science?

AI agents and the sense-plan-act-learn loop. Scientific Discovery Platforms (SDPs): AI-native systems that connect reasoning models with scientific resources.

2) Frontiers of Language Models

Surveys frontier reasoning models: general-purpose LLMs (GPT, Claude), domain-specific foundation models (materials, bio, weather), and hybrids. Covers techniques for eliciting better reasoning: prompting, chain-of-thought, retrieval-augmented generation (RAG), fine-tuning, and tool-augmented reasoning.

3) Systems for Agents

Discusses architectures and frameworks for building multi-agent systems, with emphasis on inter-agent communication, orchestration, and lifecycle management.

4) Retrieval Augmented Generation (RAG) and Vector Databases

Covers how to augment reasoning models with external knowledge bases, vector search, and hybrid retrieval methods.

Curriculum

5) Tool Calling

Introduces methods for invoking external tools from reasoning models. Focus on model context protocol (MCP), schema design, and execution management.

6) HPC Systems and Self Driving Labs

How SDPs connect to HPC workflows and experimental labs. Covers distributed coordination, robotics, and federated agents.

7) Human–AI Workflows

Explores how scientists and agents collaborate: trust boundaries, interaction design, and debugging.

8) Benchmarking and Evaluation

Frameworks for assessing agents and SDPs: robustness, validity, and relevance.

Curriculum

9) Failures and Safety

Examines why multi-agent systems fail and methods for safety and guardrails.

10) Case Studies

Case studies of SDPs in biology and materials.

11) Novelty and Plagiarism

Explores originality, credit, and the risks of plagiarism in AI-generated science.

12) Building Agents and Workflows

Pipelines, workflow composition, and self-improving systems.

Curriculum

13) Finetuning

Covers approaches to adapt agents with reinforcement learning and real-world training.

14) Responsible SDPs

Discusses ethical and policy dimensions: dual-use concerns, bias, carbon footprint, open science vs IP.

15) Scaling SDPs

Strategies for scaling: distributed compute, HPC, cloud-native orchestration. Covers resilience, scheduling, and cost/energy considerations.

16) Automation in Practice

Demonstration of automation pipelines with monitoring, logging, and adaptive workflows. Emphasis on debugging and error recovery.

17) Frontiers of SDPs

Explores frontiers: multi-agent collaboration, embodied co-scientists, integration with digital twins. Students speculate on SDPs in 2030.