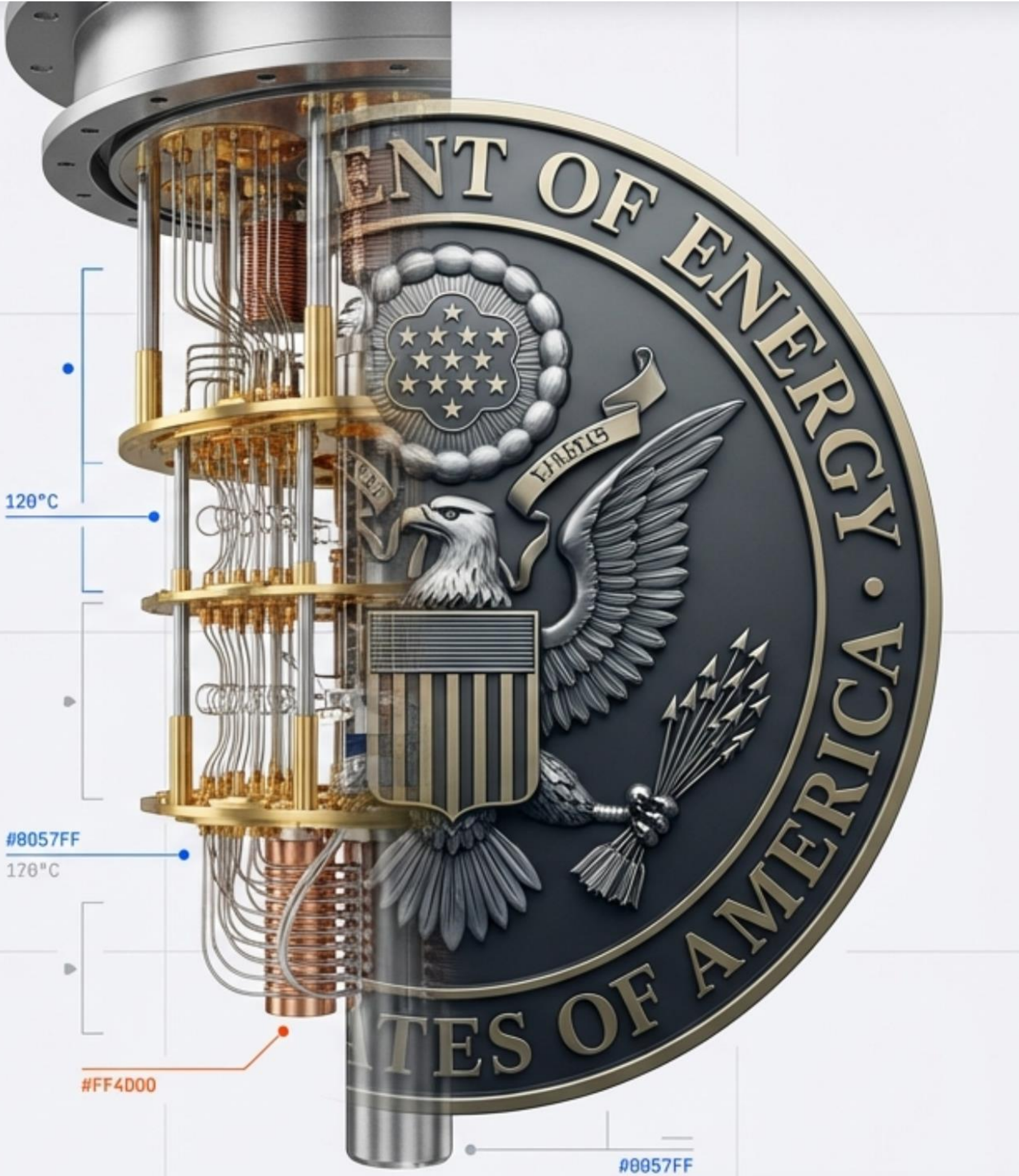
The Genesis Mission

A National Strategy to Accelerate Science Through Artificial Intelligence.



AUTHORITY: Executive Order | November 24, 2025
LEAD AGENCY: Department of Energy (DOE) + White House OSTP
LEADERSHIP: Dr. Darío Gil, Under Secretary for Science

Genesis Mission

National Initiative for Science, National Security, and Innovation

Mission Goal: Use artificial intelligence to dramatically accelerate American scientific discovery and **strengthen** U.S. economic and national security leadership.

Genesis is not a tool or program it is a national capability that compounds scientific advantage over time

WHAT IS HAPPENING RIGHT NOW

The Largest Marshaling of Federal Scientific Resources Since Apollo

> *"We are going to act with an urgency that will feel deeply uncomfortable... Let's act like our lives depend on our execution (because they do)."*
>
> — Darío Gil, Under Secretary for Science & Genesis Mission Director

Executive Order

Signed Nov 24, 2025
Presidential mandate

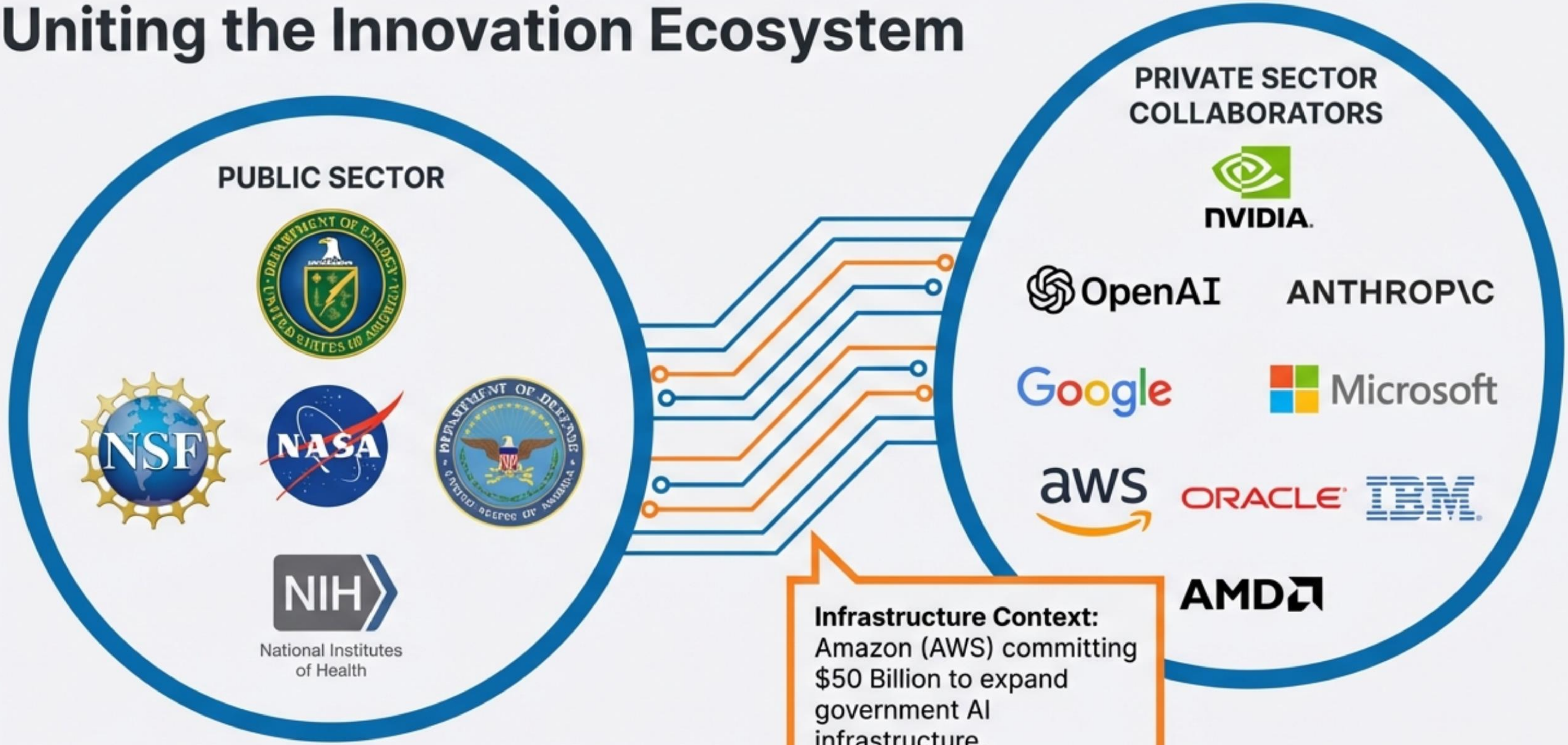
National Scale

17 National Labs + industry
+ academia unified

Urgency

Lighthouse challenges
being defined now

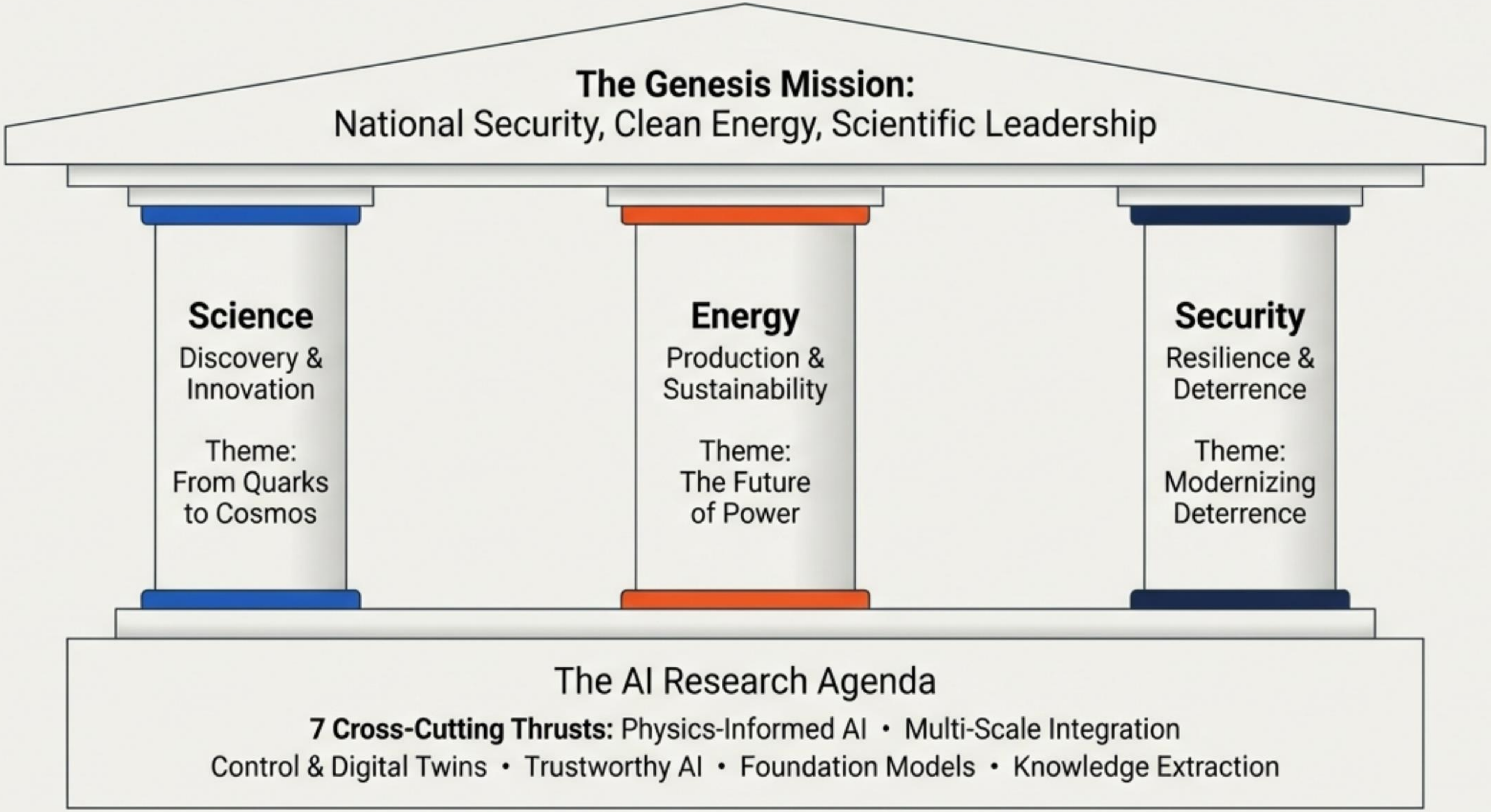
Uniting the Innovation Ecosystem



Infrastructure Context: Amazon (AWS) committing $50 Billion to expand government AI infrastructure.

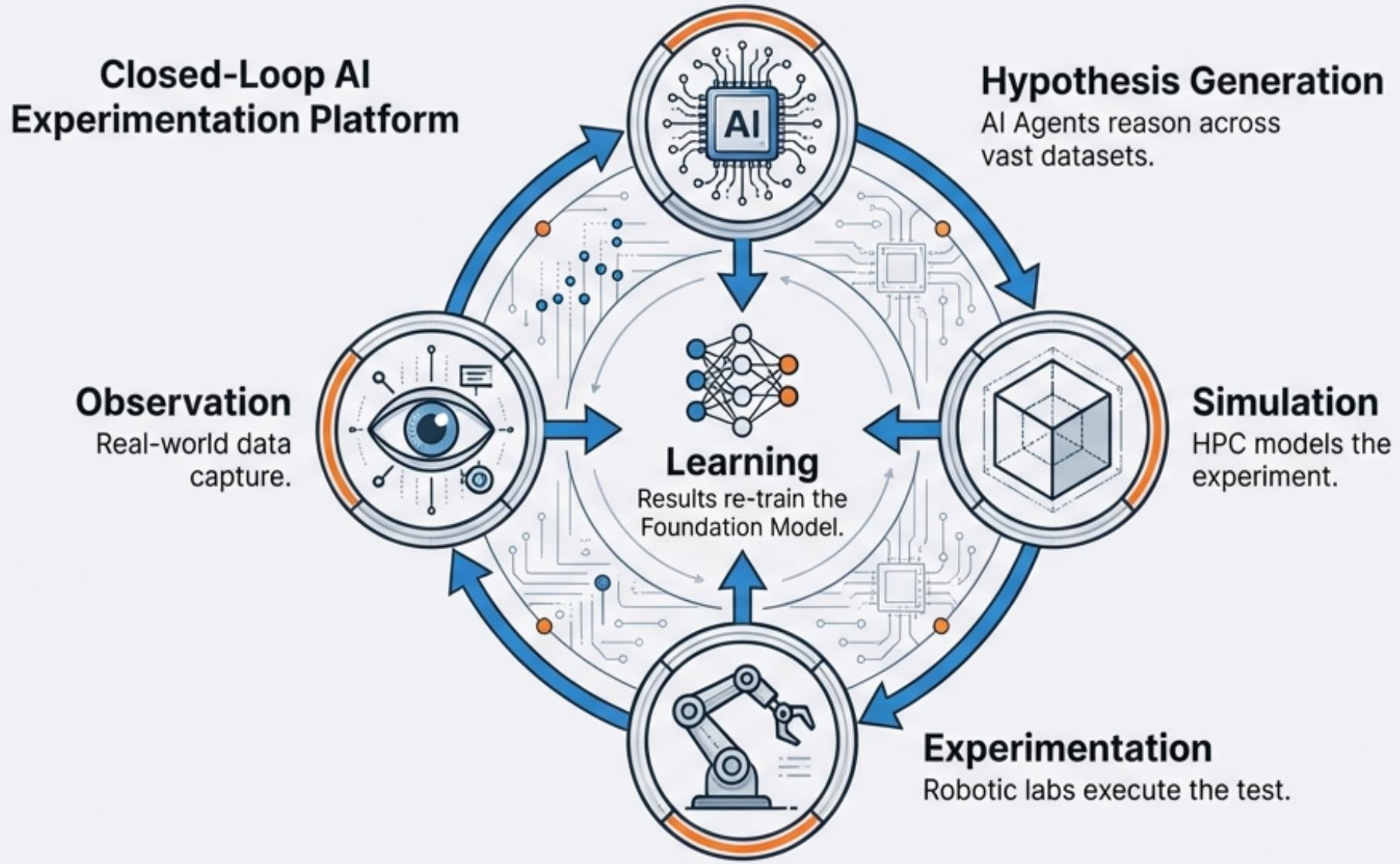
17 National Laboratories | 40,000 Scientists & Engineers

The Strategic Ecosystem Architecture



A Paradigm Shift: Closed-Loop Discovery

"It's AI for discovery, not automation." — Department of Energy



> "An intelligent network capable of sensing, simulating, and understanding nature at every scale."

THE QUESTION GENESIS ASKS

If you had 200 dedicated people,
unlimited frontier AI and exascale computing,
and robotic experimentation at national scale —

What 10-year problem could you solve in 2?

This is the bar. Not a pilot. Not a study. A paradigm shift.

200+ Dedicated researchers	**10^{18}** FLOPs of compute	**T+** Tokens of scientific data	**24/7** Robotic laboratories

Over the past five months, OpenAI has been running an experiment: building and shipping an internal beta of a software product with **0 lines of manually-written code**.

The product has internal daily users and external alpha testers. It ships, deploys, breaks, and gets fixed. What's different is that every line of code—application logic, tests, CI configuration, documentation, observability, and internal tooling—has been written by Codex. We estimate that we built this in about 1/10th the time it would have taken to write the code by hand.

Humans steer. Agents execute.

We intentionally chose this constraint so we would build what was necessary to increase engineering velocity by orders of magnitude. We had weeks to ship what ended up being a million lines of code. To do that, we needed to understand what changes when a software engineering team's primary job is no longer to write code, but to design environments, specify intent, and build feedback loops that allow Codex agents to do reliable work

The Strategic Imperative

Why DOE? Why Now?

The Opportunity

Current commercial AI excels at pattern matching and language. The new frontier is **Agentic and Physics-Informed AI**—systems capable of causal reasoning, respecting conservation laws, and navigating multi-scale dynamic systems.

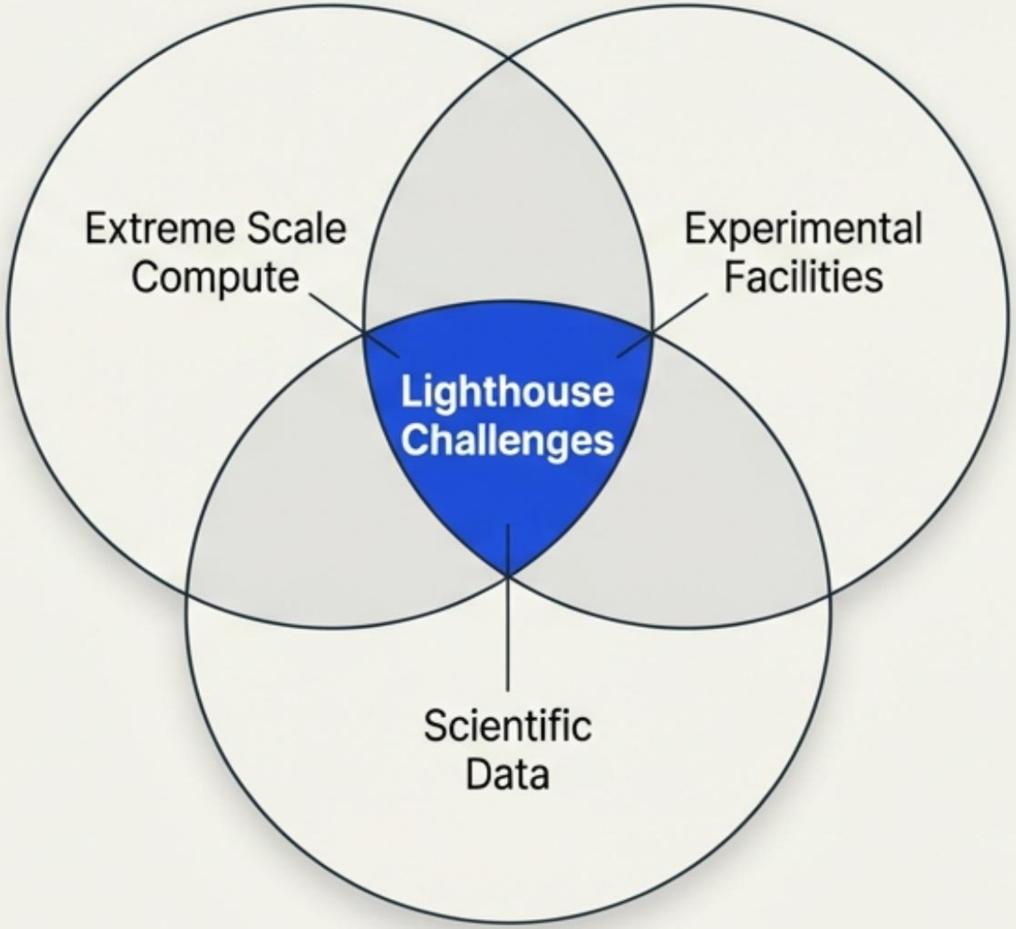
The Unique DOE Asset

We possess the unique convergence required for this leap:

- World-class User Facilities (Equinox, Synchrotrons)
- Massive Proprietary Datasets (80+ years of nuclear data)
- Deep Domain Expertise

The Mission

To guide innovation through the 'Valley of Death'—bridging the gap between scientific discovery and commercially viable deployment in manufacturing, energy, and defense.



The Engine: Why Commercial AI Isn't Enough

The Gap Between Generative AI and Scientific AI

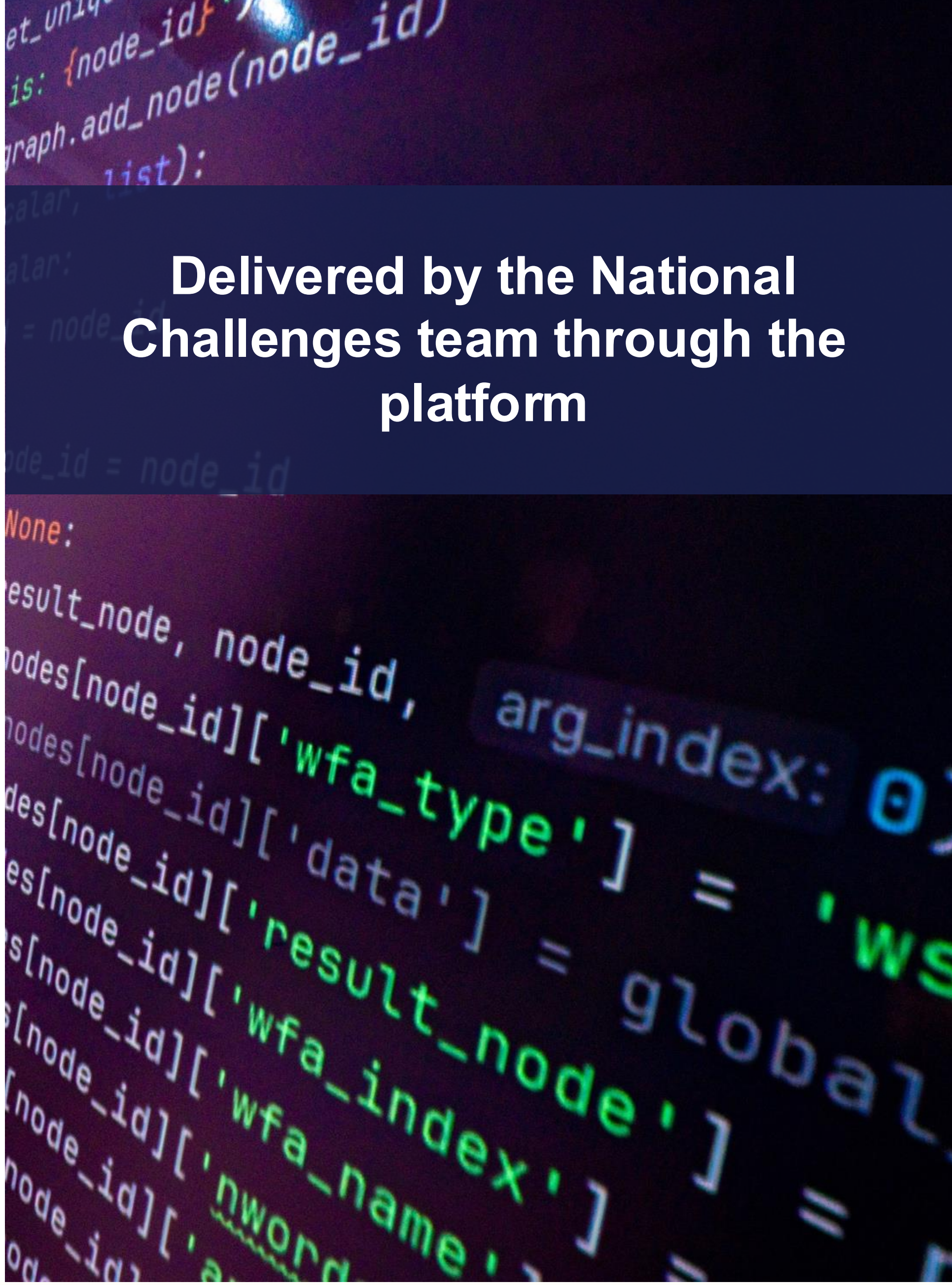
Commercial AI (Generative)	DOE Scientific AI
• Optimized for **Plausibility**	• Optimized for **Truth & Physical Consistency**
• Data-Rich (Internet Scale)	• Data-Sparse (High-Cost Experiments)
• Pattern Matching (Statistical)	• Causal Reasoning & Conservation Laws
• Black Box Operation	• Explainable & Auditable
• Hallucinations are Common	• Must Handle Rare/Extreme Events

The Goal: Machines capable of Scientific Reasoning, Multi-Scale Integration, and Trustworthy Operation in high-consequence environments.

Genesis Mission

The National Science and Technology Challenges

The **National Challenges** are high-impact technical problems aligned to urgent national priorities – where AI can dramatically accelerate progress. They demand AI model, data, computing, and automation capabilities unified in a single solution platform. Expert teams develop innovative solutions while results strengthen and extend the platform itself to tackle ever more ambitious challenges.



Delivered by the National Challenges team through the platform

Pillar I - Science: Unifying Physics & Discovery

From Human-Speed Hypothesis to AI-Speed Reasoning

Unifying Physics

Moving from pattern matching to physics reasoning. AI that internalizes the Standard Model to integrate results from particle collisions, nuclear decays, and cosmological surveys.

Key Focus: **Foundation Model for the Cosmic Frontier (HEP) >**

Accelerators

Predicting chaotic beam dynamics where small perturbations cascade into major problems.

AI-driven digital twins enabling real-time optimization of the 6-dimensional state of particle beams.

Solution: **Real-time optimization & self-updating facilities. >**

Quantum Systems

Controlling the delicate nature of quantum systems (computing, sensing, communication).

Key Focus: **AI for Quantum Systems Design (BES) to uncover causal relationships. >**

Autonomous Labs

Shifting to self-driving laboratories by integrating robotics and edge AI.

Impact: **Automating hypothesis generation and experimentation. >**

Pillar I - Science: Designing Reality

Materials, Biology, and The Shift to Inverse Design

The Core Shift: Inverse Design

Instead of trial-and-error, we tell the AI the properties we want, and it identifies the material or molecule to build.

Materials Design

Developing physics-aware AI frameworks to bridge atomic scales to macroscopic properties.

- *Focus:* "AI-Driven Materials Processing (BES)"

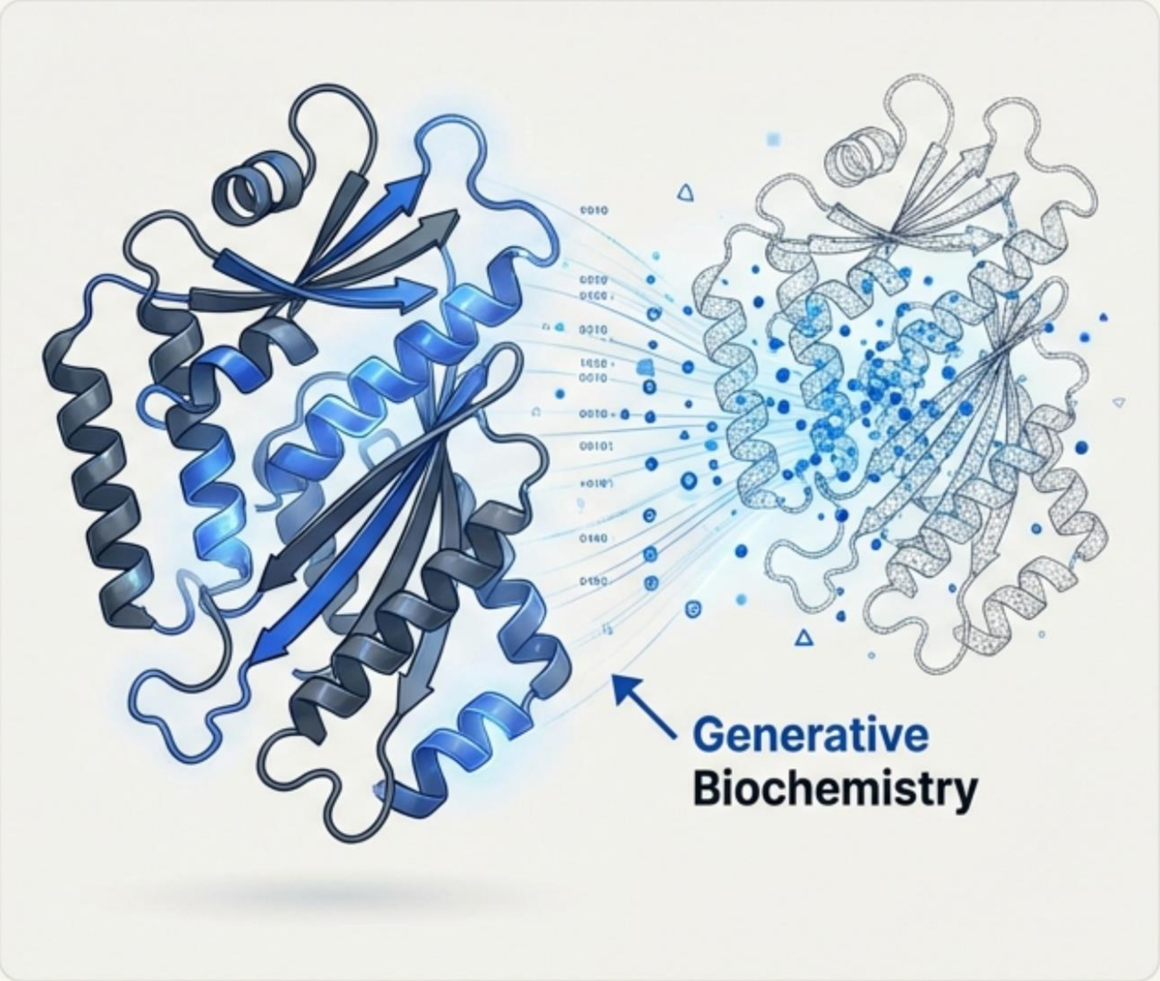
The Bio-Revolution

Establishing genotype-phenotype relationships to design biology on demand.

- *Spotlight:* "**Biochemistry in Five Dimensions** — Incorporating energy and time into protein folding models."
- *Spotlight:* "**Agentic AI-Driven Chemical Manufacturing**"

Water Security

Predicting water cycles by coupling fast (cloud) and slow (groundwater) processes using AI surrogates.



Pillar II: Discovery Science

Understanding Nature from Quarks to Cosmos

MOLECULAR DYNAMICS

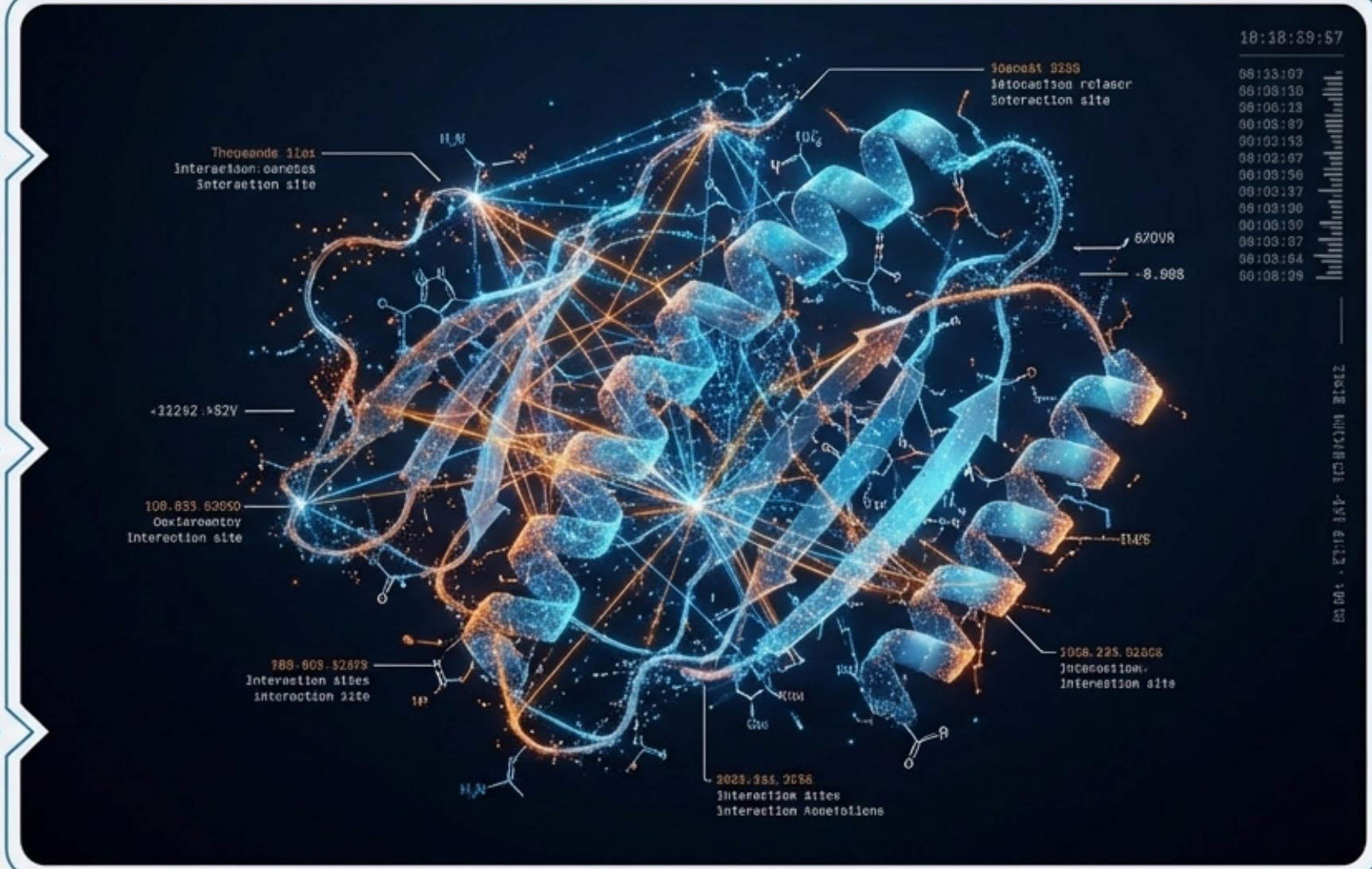
- Seeing chemical and biological processes in real-time. Accelerating breakthroughs in medicine and polymers.

PARTICLE PHYSICS

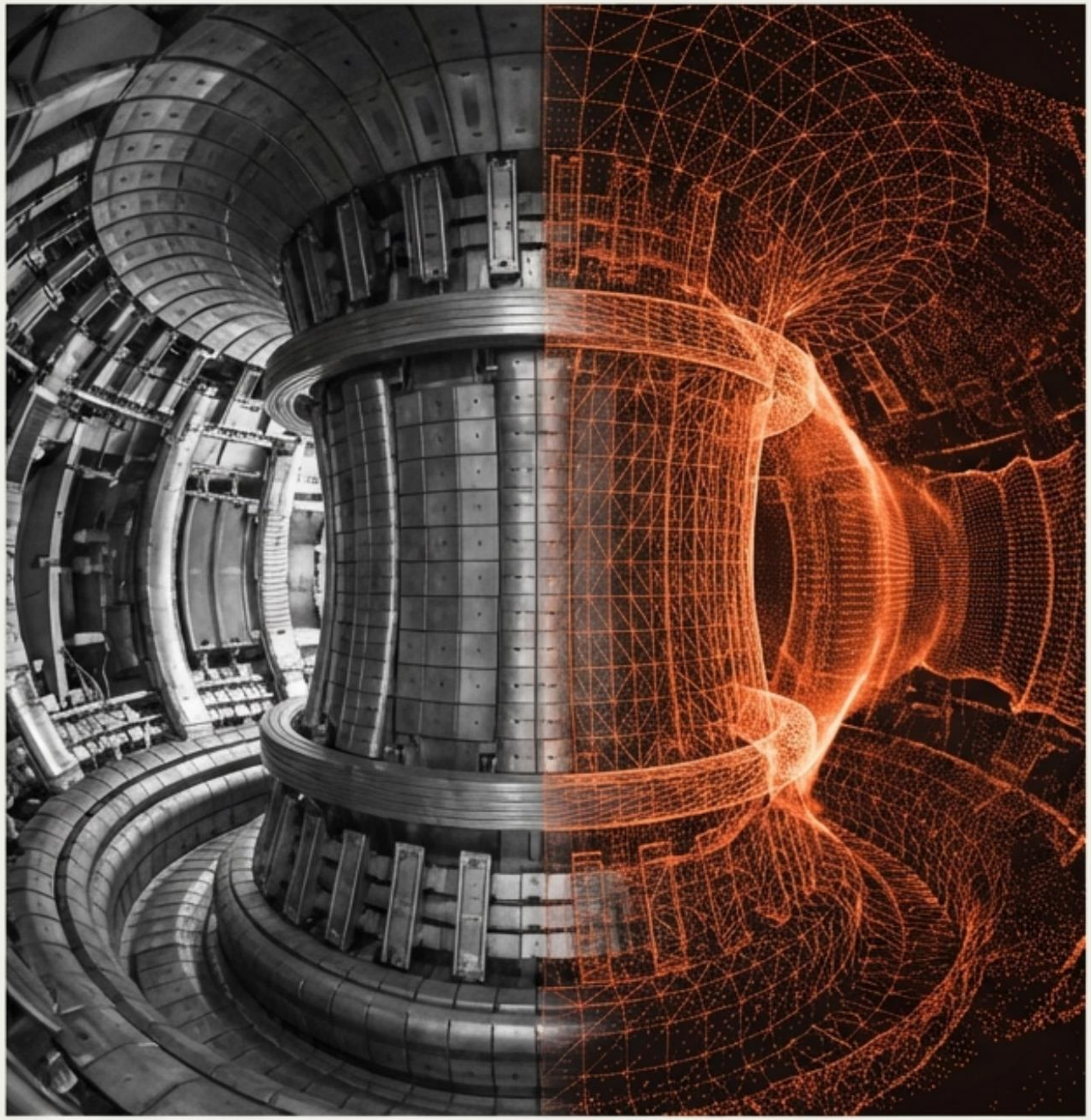
- AI as a reasoning partner to probe dark matter, dark energy, and fundamental laws.

QUANTUM COMPUTING

- Using AI to generate and verify new classes of quantum algorithms.



Pillar II - Energy: The Future of Power Generation



Digital Twin Visualization: Tokamak Reactor Core.

Fusion Energy

Accelerating delivery via the **AI-Fusion Digital Convergence Platform (DCP)**. Integrating plasma, nuclear, and materials behavior into a unified predictive framework.

Nuclear Energy

Goal: **2x schedule acceleration** and **>50% cost reduction**.

Method: Autonomous design, licensing, and operation. Using AI digital twins to interpret complex operational data in real-time.

Focus: Autonomous Research and Development to condense 50 years of irradiation data.

Subsurface Assets

Unleashing geothermal and hydrocarbon value by reasoning under extreme uncertainty. Connecting molecular-scale mechanisms to field-scale resource availability.

Pillar II - Energy: Infrastructure & Manufacturing

Resilience, Competitiveness, and The Grid

The Grid: Scaling to handle the AI data center load. Enabling **20-100x faster decision-making** with >10% reliability improvement. *Focus:* **Grid Foundational Model**.

Manufacturing: Reimagining industrial productivity. Solving the ‘Valley of Death’ between discovery and deployment. *Tool:* ATHENA (AI-Toolkit for Holistic Economic Network Analysis).

Critical Minerals: Securing the supply chain via AI-driven resource mapping. *Focus:* **Nature’s Path to Alternative Chemistries** — Designing industrial catalysts that eliminate the need for critical minerals entirely.

Buildings & Data Centers: Reimagining construction and operation. De-risking advanced cooling and energy integration for extreme compute power.

Genesis Mission

The American Science and Security Platform

The **platform** supports AI-driven experimentation, analysis, discovery, design, and manufacturing. It unifies access to AI, computing resources, scientific data, and automated facilities, allowing application of unprecedented capabilities. It delivers scientific self-improving feedback that will accelerate R&D in the energy dominance, discovery science, and national security pillars.

Delivered by the infrastructure, data, and models teams

The mission technical efforts are organized into five teams

National Challenges
Engages with projects addressing high-impact national challenges—such as energy security and advanced materials—that are accelerated by the platform. Serves as a living portfolio that evolves as goals and technologies change.

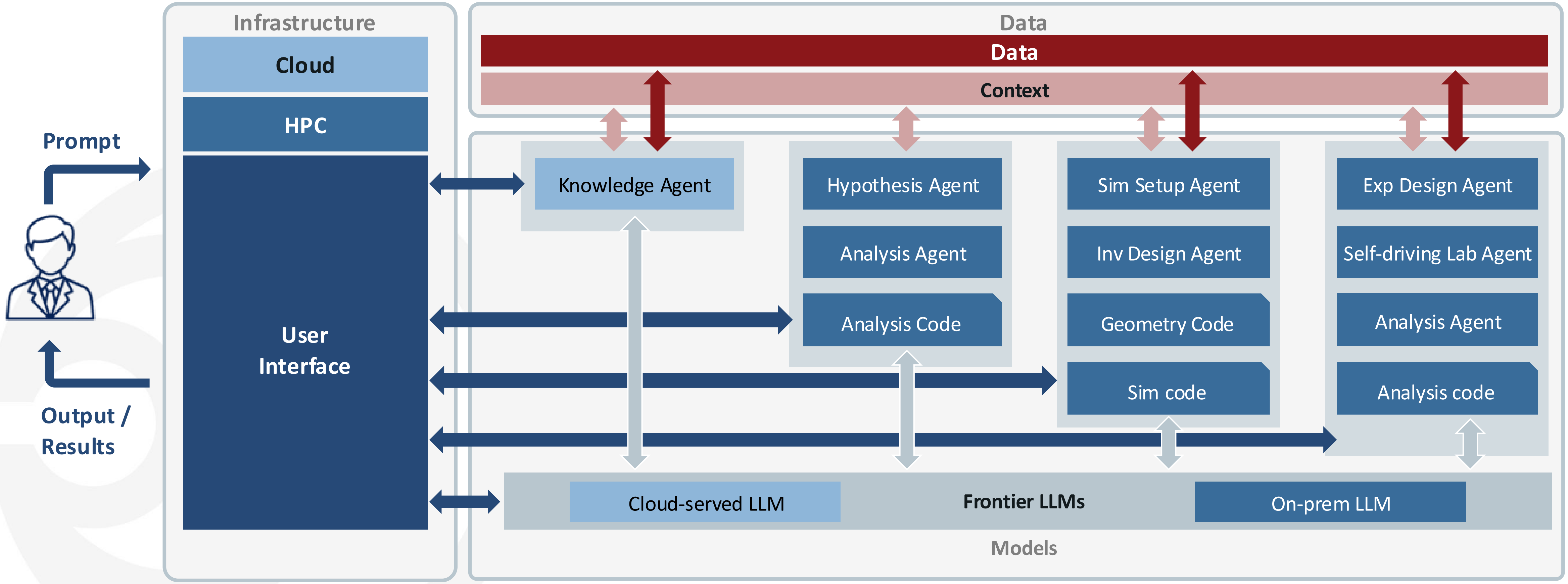
Infrastructure
Provides hardware and software foundations to support large-scale AI training and inference, data management, and scalable agentic workflows. Ensures security of models, data, and systems and interoperability across DOE facilities and production agencies.

Data
Organizes, curates, and prepares scientific and engineering data from experiments, simulations, and observations for AI use. Ensures high-quality, reliable data to accelerate discovery and problem-solving.

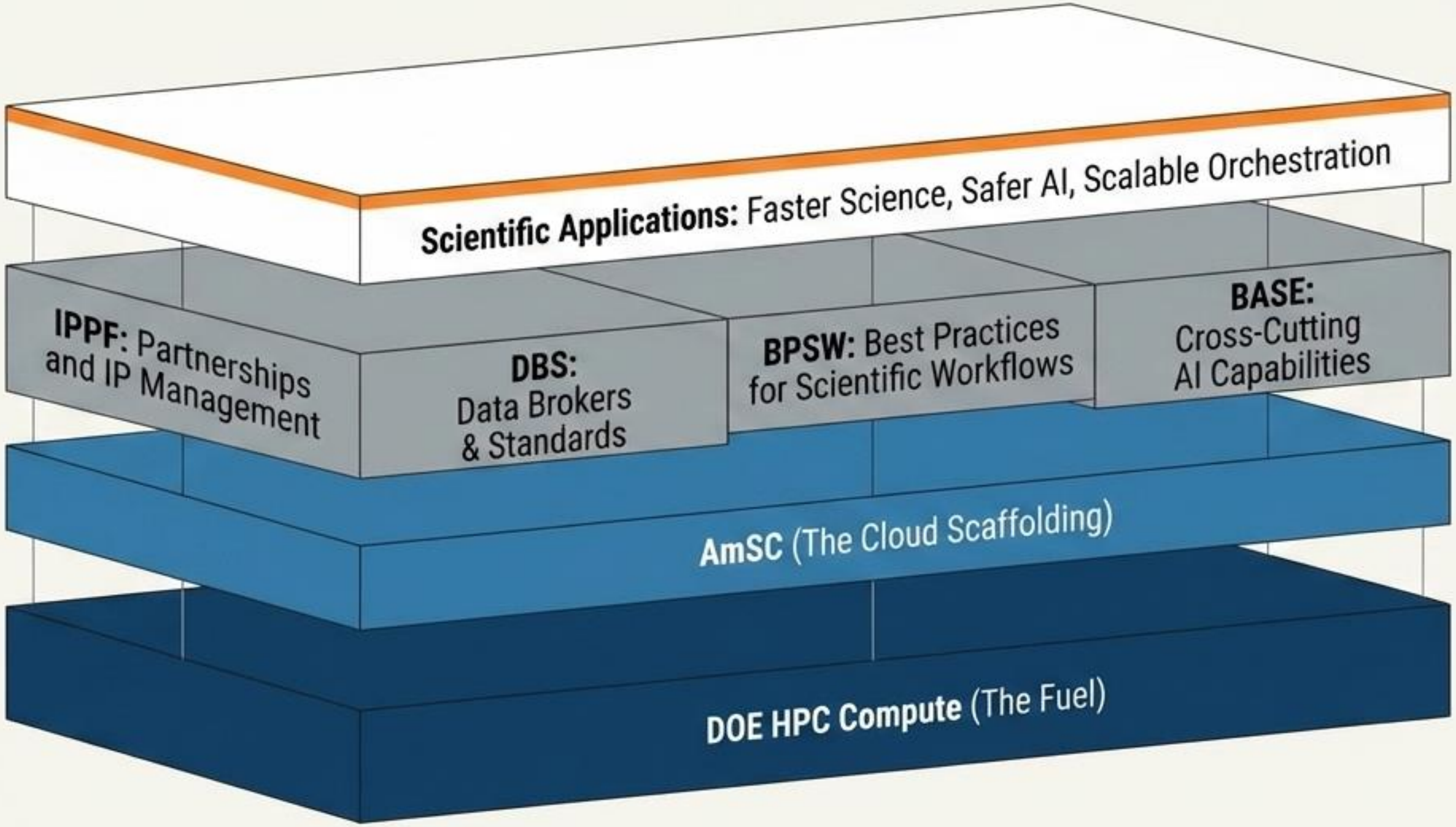
Models
Develops an AI ecosystem of agents and models that combines frontier industry capability with DOE-specific expertise to discover and implement novel solutions to the nation's open science and technology challenges.

Partnerships
Builds collaborations across government, industry, and academia to advance AI innovation. Leverages shared expertise and resources to deliver real-world impact at national scale.

Each team provides a collection of subcomponents that interoperate to enable new workflows

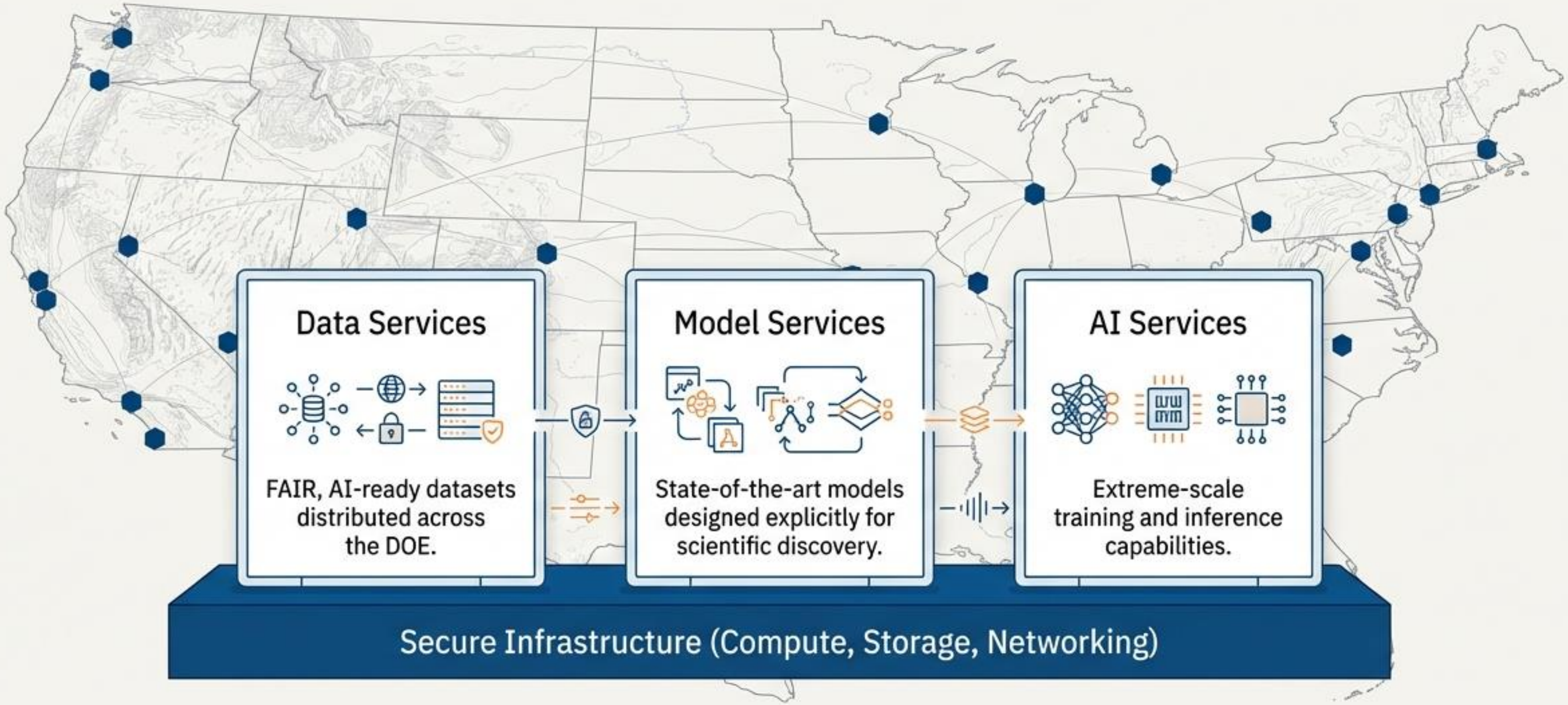


The Architecture of AI-Driven Discovery



The American Science Cloud (AmSC) Scaffolds the Genesis Mission

AmSC is a secure, federated, and science-optimized environment integrating DOE's world-leading computing, experimental facilities, and HPC networks.



The Cross-Cutting Services Scaffolding

Data Brokers & Standards (DBS)

Managing the Genesis Central Catalog. Ensuring petabytes of data are FAIR and AI-ready using standardized data cards and automated curation tools.

Best Practices for Scientific Workflows (BPSW)

Providing a catalog of AI integration patterns (e.g., RAG, surrogate optimization) and maintaining a Genesis AI Resource Hub of 330+ curated model and agent cards.

IP & Partnership Formation (IPPF)

Streamlining collaboration. Offering pre-approved agreement templates (SOWs, NDAs) and a Virtual Lab Teaming Framework to rapidly integrate industry and academic partners.

Baseline AI R&D Capabilities (BASE)

Diagnostic Dashboard for Model Teams

Core Agentic Framework (CAF)

Scalable orchestration supporting 1000s of concurrent agents and tool calls.

AI-Ready Data (DATA)

Autonomous pipelines transforming raw scientific data into training-ready form.

Model Evaluation (EVAL)

Comprehensive suites assessing robustness, generalizability, and scientific validity.

Multimodal Reasoning (MM)

PRISM architecture for cross-modal reasoning (text, images, graphs, geometry).

Self-Improving Models (SIM)

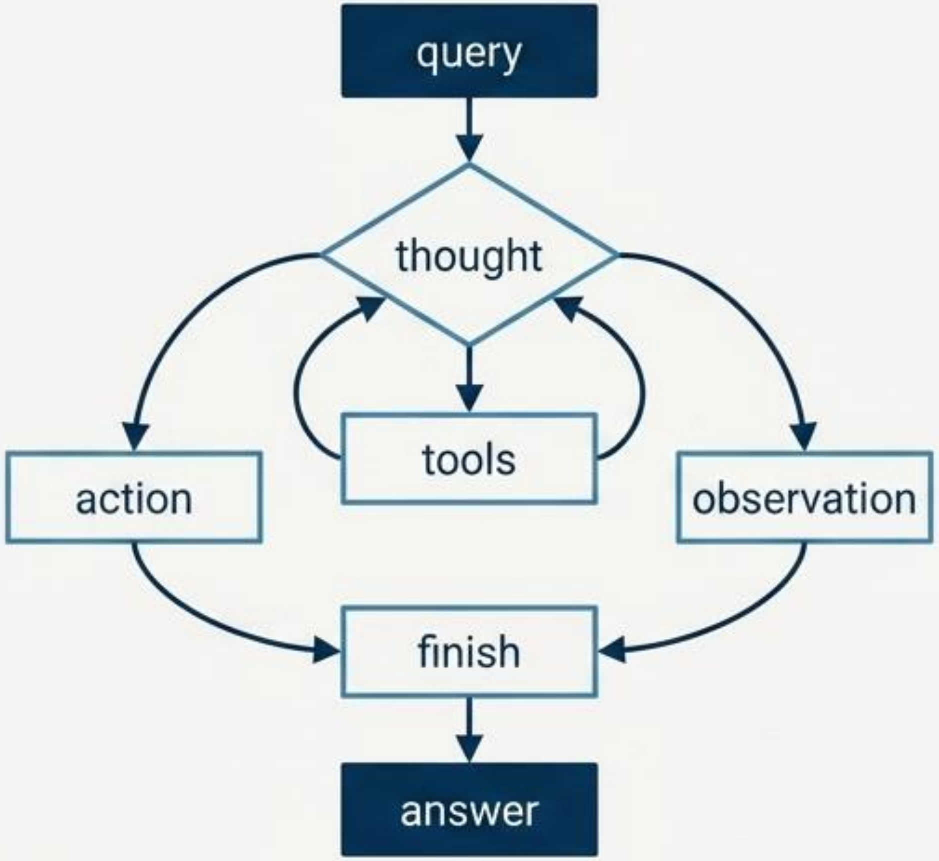
APEIRON framework for continuous learning as new experimental data emerges.

Safety & Security (SAFE)

Rigorous frameworks for red/blue teaming, uncertainty quantification, and CI pipeline safeguarding.

Scaling Closed-Loop Scientific Workflows

Agent Logic Enforces Autonomous Execution



CAF Deployment Scale

1. Local Execution:
Persistent agents on a workstation.

2. Federated Execution:
Cross-institutional execution under shared governance.

3. HPC Parallel Inference: Fanning out thousands of LLM requests simultaneously. (e.g., 1.4B tokens processed in 35 minutes on 2048 Aurora nodes).

National Science and Technology Challenges vs Projects

What is a National Challenge?

- It addresses a big scientific and/or technological problem
- It has ambitious goals
- It is achievable (does not break the laws of physics)
- It is resource intensive
- It has potential for high impact

What is a National Project?

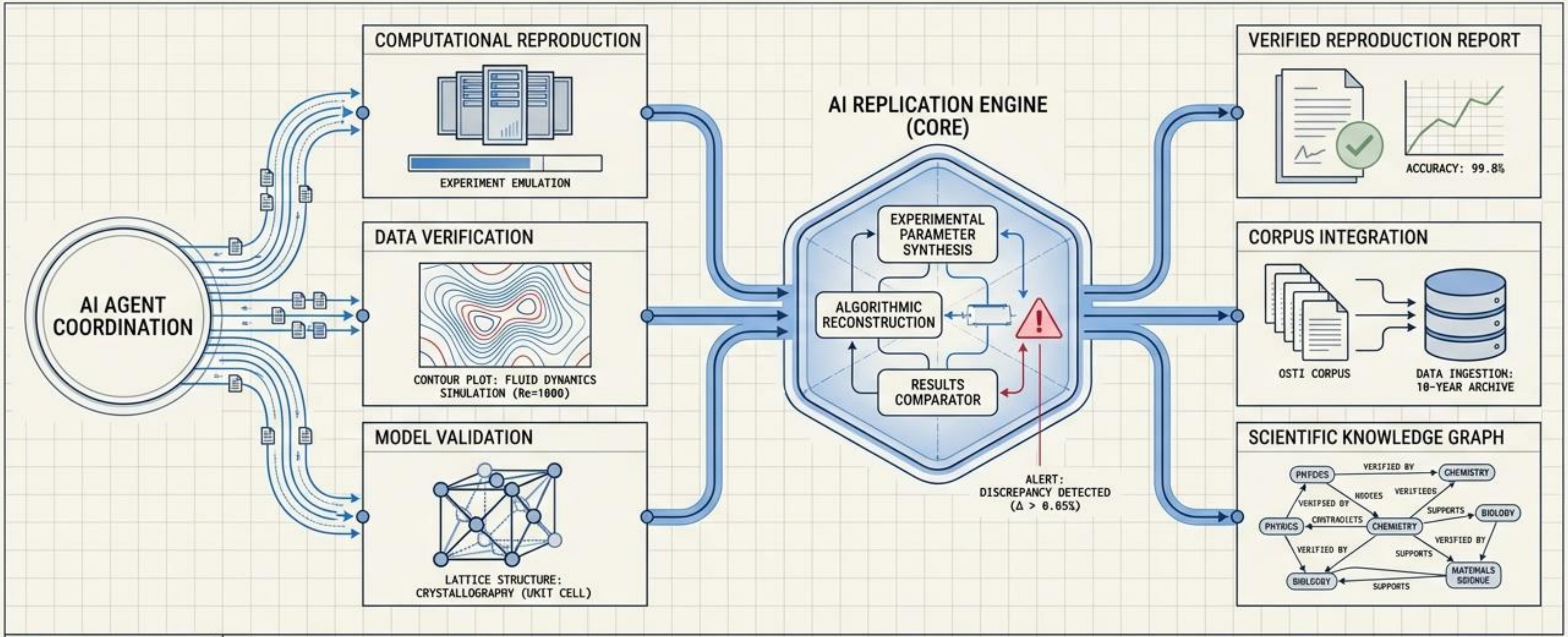
- It addresses a part (or all) of the overall lighthouse challenge
- It defines the research frontier and “illuminates” future research directions
- It is a highly visible, ambitious, and focused multidisciplinary research endeavor
- It involves multi-institutional partnerships, including public-private partnerships

National priorities define National Challenges and inform investments in National Projects

PROJECT:
Systematic reproduction of computational science via autonomous AI agents.

THE AI REPLICATION ENGINE

Systematic reproduction of computational science using AI and autonomous agents.

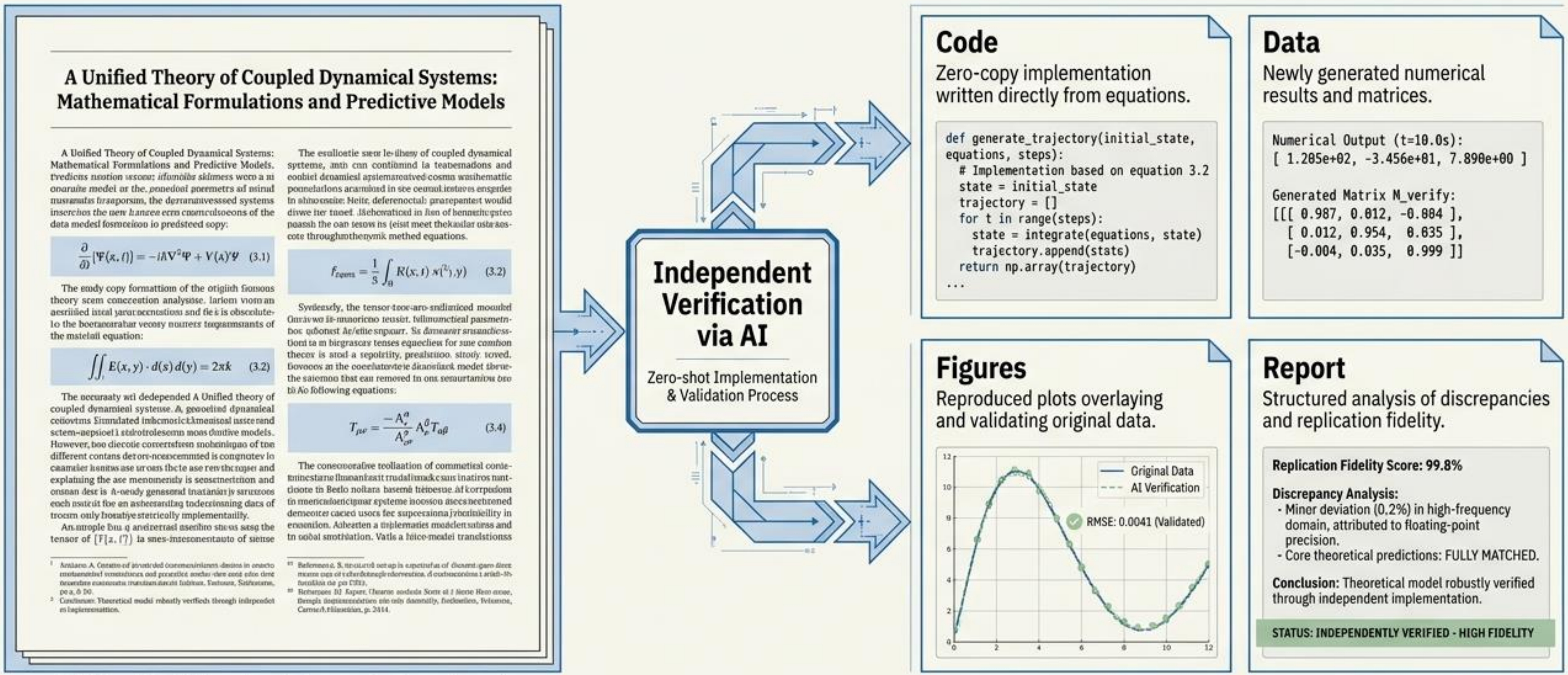


STATUS: ACTIVE

ENVIRONMENT: Rick Stevens, Argonne National Laboratory

OBJECTIVE: 10-Year OSTI Corpus Verification

True replication requires independent implementation from theory.



Original Publication (Theory & Equations)

The autonomous workflow from raw PDF to verified Git commit.

Independent Verification via AI: Five-Stage Computational Pipeline.

1. Ingest & Analyze

Parse the original paper and isolate key equations and parameters.

```
[PDF] -> [Equations, Parameters]
```

2. Implement

Code the methods entirely independently, avoiding source code copying.

```
def model(params):
… # Independent Code
```

3. Execute & Reproduce

Generate key results, numerical values, and output figures.

```
Output: y=1.285e+02,
Plot: [Figure 1]
```

4. Validate

Document discrepancies and objectively assess reproducibility metrics.

```
RMSE: 0.0041,
Discrepancy: <0.2%
```

5. Compile

Generate a structured LaTeX report and output final repository artifacts.

```
[Report.tex, Git Commit] -> Repository
```

SYSTEM STATUS: Iterative execution active across 15 OSTI subject categories.

Generating comprehensive, publishable computational artifacts.

```
replication-project/
├── README.md (Systematic framework documentation)
├── COMPUTATIONAL_TOOLS_INVENTORY.md
│   (Hardware and software environment tracking)
└── 1275503-COSMIC-REIONIZATION/
    ├── replication_plan.tex
    │   (42.7% of corpus output is native TeX source)
    ├── aggregated_datasets.pdf
    │   (Reproduced figures and verification tables)
    └── main_model.py
        (Independent Python implementation)
```

ARTIFACT STANDARD:

Every successful pipeline execution produces native LaTeX reports, aggregated datasets, and zero-copy code implementations ready for immediate community review.

THE QUESTION

What fraction of DOE-funded computational papers can an AI agent replicate end-to-end?

ARGONNE PUBLICATION LANDSCAPE

- **~3,500–4,300** papers/year published by Argonne (OSTI)
- **~400/year (9.3%)** contain computational results potentially replicable without wet-lab
- **~3,500 papers** in the 10-year computational backlog
- **15+ disciplines** from nuclear physics to ML to astrophysics

THIS STUDY

- **30 papers** selected from OSTI database
- **15 scientific domains** covered
- **17 days** of agent operation (Apr 5–22, 2026)
- **1 AI agent** with HPC access (Polaris, Aurora, DGX)

SCIENTIFIC BREADTH

12 Distinct Domains in 13 Papers

- ✓ Materials Science / ML
- ✓ Computational Chemistry
- ✓ Computer Science
- ✓ Astrophysics (× 2)
- ✓ ML / Dynamical Systems
- ✓ Pure Mathematics
- ✓ Optimization / ML
- ✓ Quantum Optics
- ✓ Quantum Computing
- ✓ Applied Mathematics
- ✓ Nuclear Engineering
- ✓ Particle Physics

The same agent, same approach — no domain-specific fine-tuning or prompting

TIME & EFFICIENCY

The Bottleneck Is Compute, Not Comprehension

27 sec
Fastest (CPW Resonator)

54 sec
Photo-z PDFs

~2 days
MSM (with MD)

~10 days
PVMol-Gen

WHERE TIME IS SPENT

- **Understanding the paper:** minutes (fast — this is the AI's strength)
- **Writing code:** minutes to hours (depends on domain complexity)
- **Running simulations:** seconds to days (GPU training, MD, Monte Carlo)
- **Waiting for queues:** minutes to hours (HPC scheduling)
- **Debugging infra:** variable (driver mismatches, disk quotas, network issues)

SCALING ANALYSIS

From 1 Agent to 1,000: Clearing the Backlog

Scale	Papers/Day	Time to Clear 3,500	Resources
1× (current)	~1.1/day	8.8 years	1 machine, 1 agent
10×	11/day	10.5 months	~10 agents, shared GPU cluster
100×	**110/day**	**5 weeks**	**500–1,000 GPUs, 75–100 agents**
1,000×	1,096/day	3.5 days	Aurora-scale system

At 100 × — achievable with current technology — the entire annual computational publication output of a national lab can be audited in about 5 weeks.

FROM REPLICATION TO DISCOVERY

The Next Step: AI Agents Doing Original Science

REPLICATION (proven)

- Follow published methodology
- Reproduce known results
- 1 × compute per paper
- Quality audit function

→ ~30 ×

DISCOVERY (next)

- Formulate new hypotheses
- Design novel experiments
- ~30 × compute per paper
- Original scientific contributions

RESOURCE REQUIREMENTS AT 100 × DISCOVERY RATE

- **14,000–20,000 GPUs** needed (200–280 Oberon racks)
- **≈ 1 Aurora-class system** fully utilized by AI agents doing original science

IMPLICATIONS

What This Means

- **1.** AI agents can replicate real scientific papers across diverse domains with high fidelity
- **2.** The bottleneck is compute, not understanding — the agent consistently grasped the methodology
- **3.** Reproducibility problems in published papers are real — even 'confirmed' replications required detective work
- **4.** This scales — with more compute and parallel agents, a lab's annual output becomes auditable in weeks
- **5.** Quality matters more than quantity — each replication serves as a reproducibility audit

Genesis Mission

U.S. DEPARTMENT of ENERGY