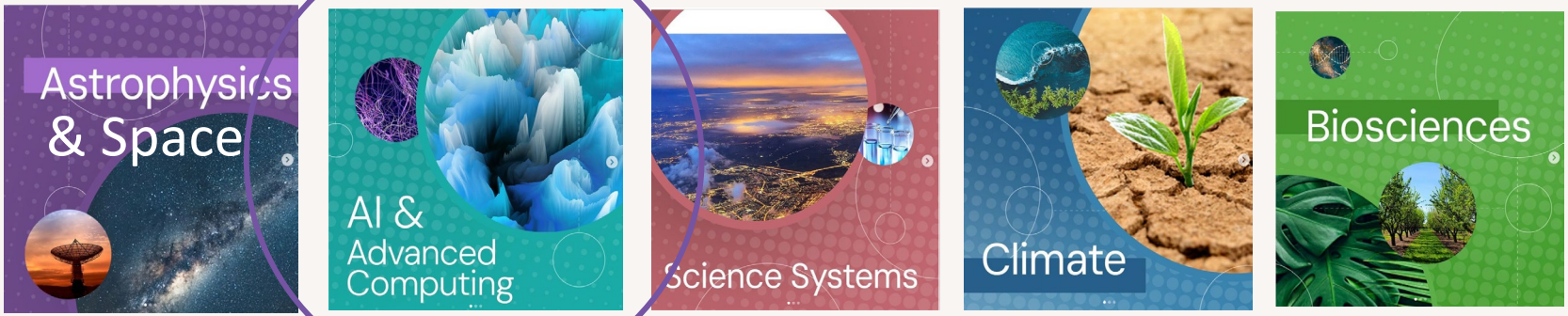
Schmidt Sciences

Braver, Newer Science

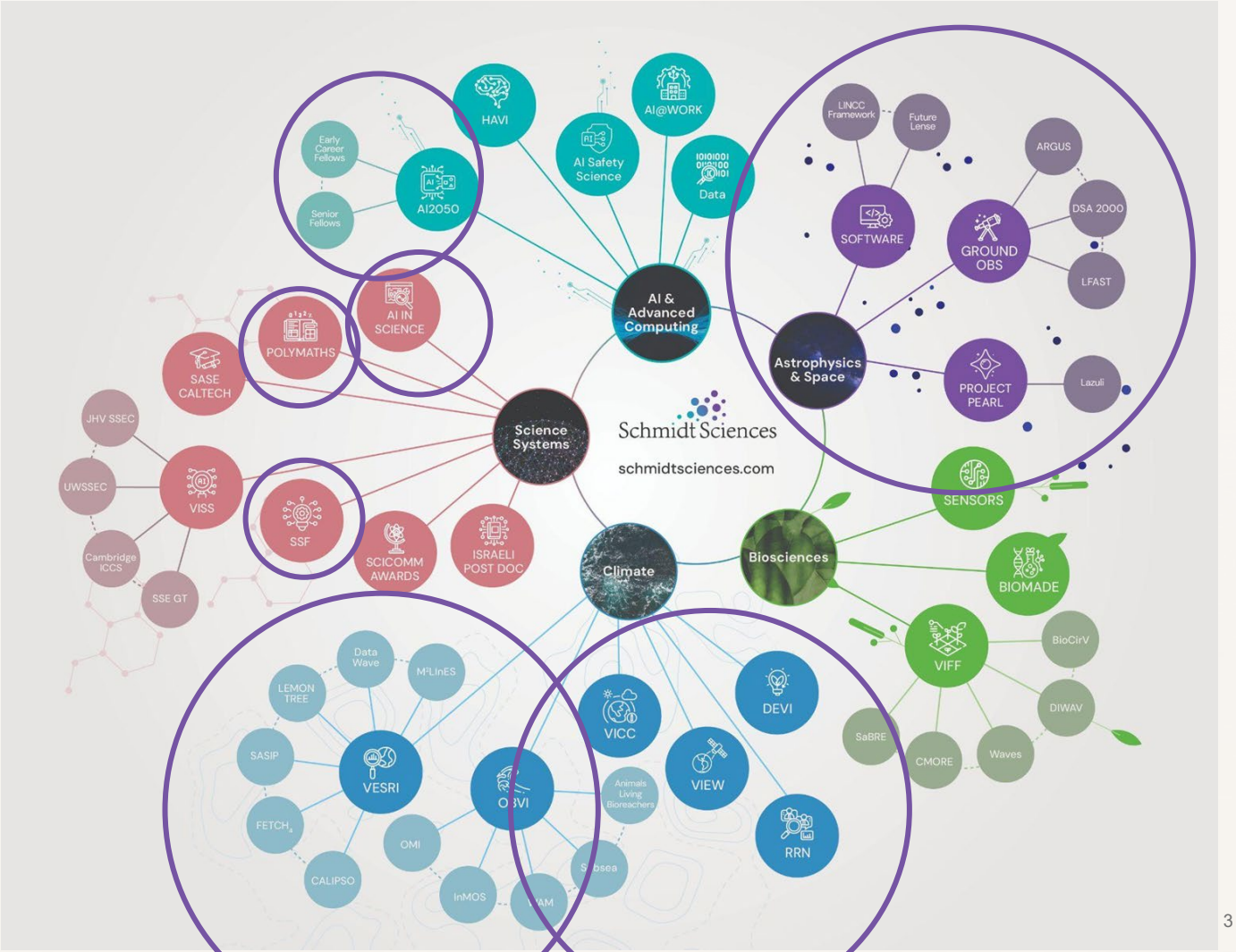
Suhas Mahesh

November 19 2025

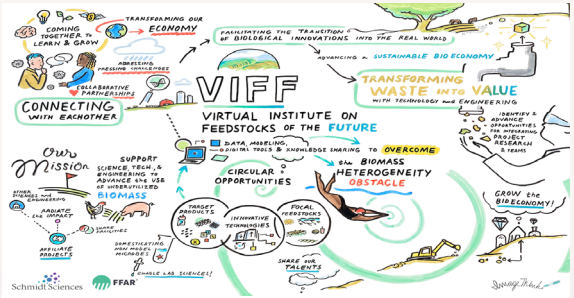
Schmidt Sciences accelerates scientific knowledge and breakthroughs with the most promising, advanced tools to support a thriving planet



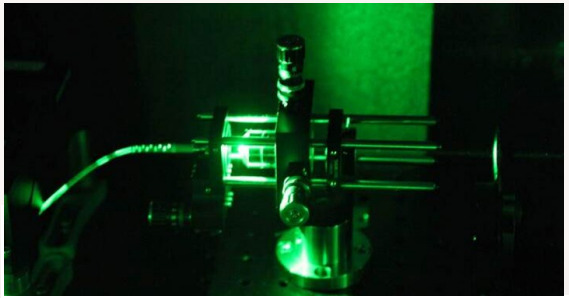
STRATEGIC APPROACH



Biosciences

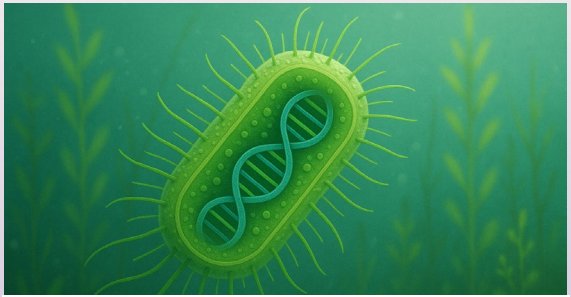


Virtual Institute on Feedstocks for the Future (VIFF)



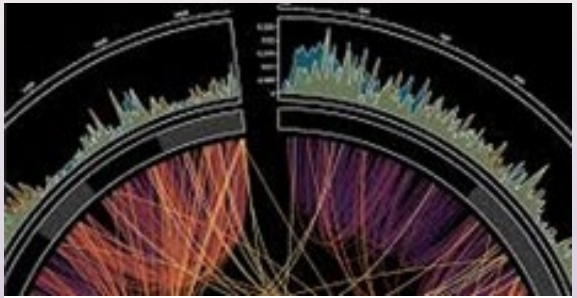
Biological Sensing

Innovations for Bioreactors



Microbes and Healthy Ecosystems

Climate and Bio



Biological Data and Advanced Computation

Reach out to: Andrea Hodgson (ahodgson@schmidtsciences.org)

Astrophysics & Space

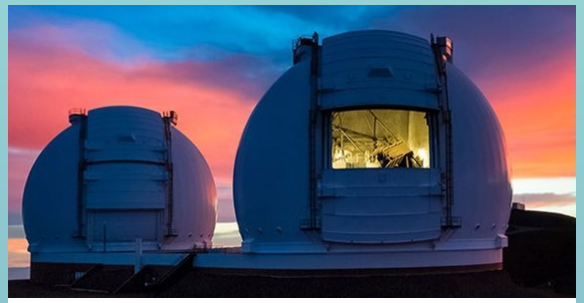
Ground - Based Observatories

PROJECT PEARL

Space- Based Observatories

Software Platforms for Astronomy

Cross - Observatory Research Groups



Data Stewardship for Astro

AI/ML in Astronomy

Reach out to: Arpita Roy (aroy@schmidtsciences.org)

Climate

Decarbonization, Water, Carbon Cycle, Ocean Biogeochemistry, Earth Systems

Critical Earth Observations

Regional Research Networks

Reach out to: Balaji (balaji@schmidtsciences.org)

GOALS

AI & Advanced Computing

Advance the science of AI to develop more robust, reliable, and safe AI-enabled systems, capable of realizing the beneficial promise of AI

Enable AI to make significant contributions to humanity's greatest scientific challenges and opportunities

Foster the creation of research platforms like datasets, experiments, evaluations and models that can both accelerate and measure the impacts of AI

Contribute to the growth of a vibrant, diverse, and global scientific community focused on advancing AI

Science is undergoing a phase change

Can the Academies be our crystal ball?

Science

+ AI = ?

Science is undergoing a phase change

Academic AI research faces risk of being irrelevant?

How to reassess risks and opportunities in face of inflow of enormous private capital?

Philanthropy

Science is undergoing a phase change

- Does AI improve scientist productivity? How about diversity of ideas? Quality of scientific code?
- As new AI-based tools are built for science, how can we prevent private capture?
- What Protein Data Banks are possible in other fields?
- What routine tasks could be automated away? (e.g. crystallographic analysis)

Our programs

AI at Work

Advance scholarship in the quantitative social sciences into the impacts on the workplace of new AI tools.

Trustworthy AI

Deepen our understanding of the safety properties of systems built with LLMs and to develop well - founded, concrete, implementable technical methods for testing and evaluating LLMs.

AI2050

Support exceptional people working on the key opportunities and hard problems that are critical to get right for society to benefit from artificial intelligence.

Humanities and AI Virtual Institute

Advance scholarship in the humanities through the use of AI - based technology for research.

2026 explorations

Post - Transistor Hardware

AI for Scientific Simulation

Interpretability

AI Benchmarking

And more…

How do we create the AlphaFold moment for weather, earth systems, retrosynthesis, materials simulation, drug discovery...?

Credits: Google DeepMind

Barriers to the next AlphaFold Moment

Investment Value

Investment value includes RCSB-PDB's operational costs, costs of data creation and deposition (from contributors), costs of annotating and adding value to the data (from collaborators), etc. While the costs of data creation and deposition are unknown, a reasonable **estimate to replicate the RCSB-PDB data archive is $12 billion** (assuming $100,000 average cost to replicate each entry). The annual operating costs of the RCSB-PDB are approximately $6.9 million per year ($5.1 million spent on operations in New Jersey and $1.8 million spent in California). These operating costs include all operating expenses (e.g., equipment, maintenance, supplies, salary and fringe, etc.). The operating costs can be considered the minimum annual investment value needed to keep RCSB-PDB archive available for public use. The annual operating costs in New Jersey of $5.1 million are used to estimate the economic impact to the New Jersey economy (presented below).

Rutgers Report on PDB

1. Sheer cost of ~100k experiments
2. Good experimental data cannot be gathered for many systems. E.g. catalyst surfaces

Other problems are not so simply expressed

1. Spans many length, time scales.
2. Theory not well developed (e.g. excited states)
3. Representations, architectures not clear.

More Concretely

WETLAB/BITLAB DIVIDE

How do we build tools that will make their way to the lab bench?

Can we build tools that will tempt the pipetting grad student or the postdoc spectroscopist to put down their tools and spin up AWS instead?

BETTER, NOT JUST FASTER

How do we use AI to push down the error bars on predictions to experimental levels?

Can AI be the load-bearer where theory is still immature?

Can AI help us fuse diverse data, length and time scales?

PREPARING FOR A WORLD OF AGENTS

How do engender trust in science as we move into a world of swarms of agents?

What do evaluations look like for agentic science?

How do we prevent private capture of this emerging ecosystem?

Philanthropy Needs Data

- Who will build an Epoch.ai for Science?
- Stanford HAI-style AIxScience index?
- Reports on usage and failure modes of AIxScience tools
- Workshops where Science and AI intersect less frequently: interpretability, RL, CV, Benchmarking…

Suhas Mahesh
smahesh@schmidtsciences.org