

# Advanced Scientific Computing Research (ASCR)

Hal Finkel and Ben Brown

<https://science.osti.gov/ascr/officehours>



U.S. DEPARTMENT OF  
**ENERGY**

Office of  
Science

[Energy.gov/science](https://energy.gov/science)

# Office of Science Statement of Commitment & other Guidance

- ◆ **SC Statement of Commitment** – SC is fully and unconditionally committed to fostering safe, diverse, equitable, inclusive, and accessible work, research, and funding environments that value mutual respect and personal integrity. <https://science.osti.gov/SW-DEI/SC-Statement-of-Commitment>
- ◆ **Expectations for Professional Behaviors** – SC’s expectations of all participants to positively contribute to a professional, inclusive meeting that fosters a safe and welcoming environment for conducting scientific business, as well as outlines behaviors that are unacceptable and potential ramifications for unprofessional behavior. <https://science.osti.gov/SW-DEI/DOE-Diversity-Equity-and-Inclusion-Policies/Harassment>
- ◆ **How to Address or Report Behaviors of Concern**– Process on how and who to report issues, including the distinction between reporting on unprofessional, disrespectful, or disruptive behaviors, and behaviors that constitute a violation of Federal civil rights statutes. <https://science.osti.gov/SW-DEI/DOE-Diversity-Equity-and-Inclusion-Policies/How-to-Report-a-Complaint>
- ◆ **Implicit Bias** – Be aware of implicit bias, understand its nature – everyone has them – and implicit bias if not mitigated can negatively impact the quality and inclusiveness of scientific discussions that contribute to a successful meeting. <https://kirwaninstitute.osu.edu/article/understanding-implicit-bias>



U.S. DEPARTMENT OF  
**ENERGY**

Office of  
Science

## Our Mission:

Deliver scientific discoveries and major scientific tools to transform our understanding of nature and advance the energy, economic, and national security of the United States.



More than **34,000** researchers supported at more than **300** institutions and **17** DOE national laboratories



Steward **10** of the 17 DOE national laboratories

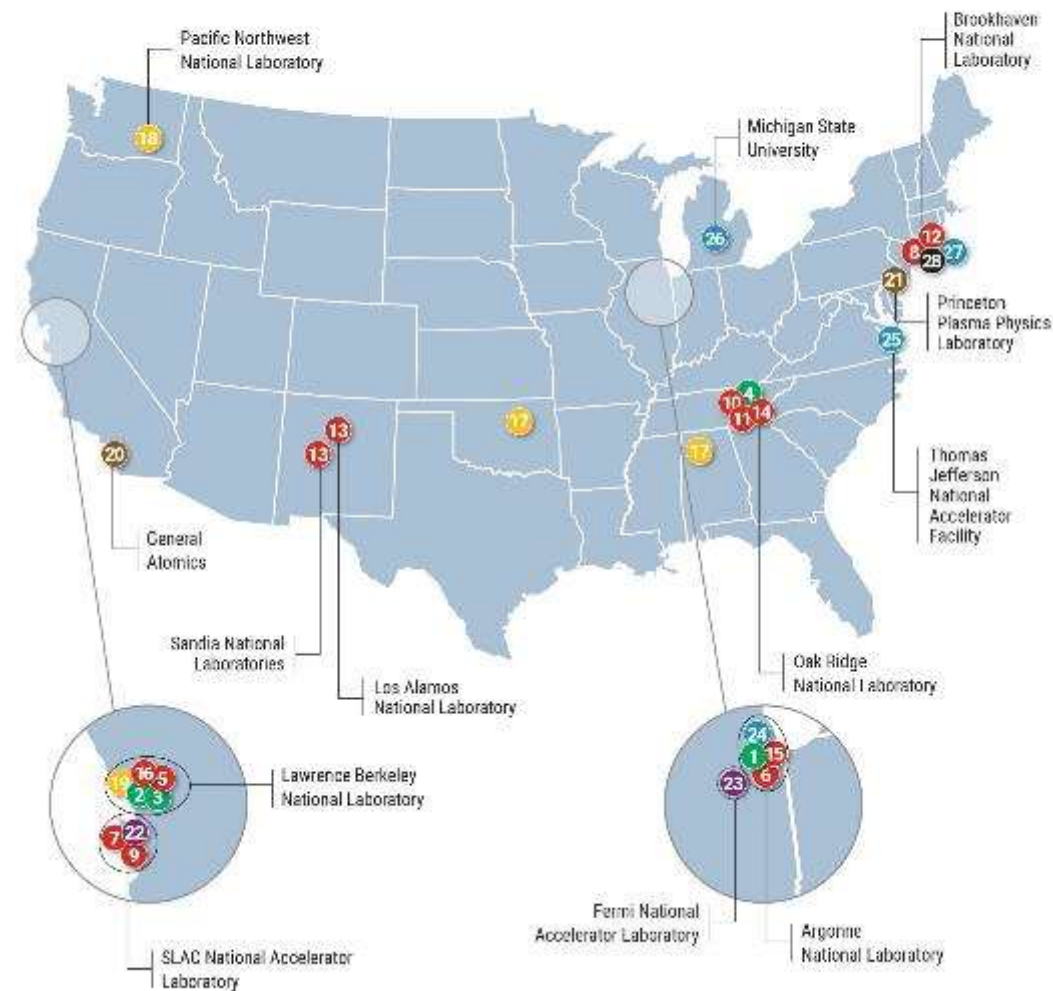


More than **37,000** users of **28** Office of Science scientific user facilities



**\$8.1B**  
(FY 23 enacted)

# U.S. Department of Energy Office of Science User Facilities



## Advanced Scientific Computing Research (ASCR)

- 1 Argonne Leadership Computing Facility (ALCF)  
Argonne National Laboratory
- 2 Energy Sciences Network (ESnet)  
Lawrence Berkeley National Laboratory
- 3 National Energy Research Scientific Computing Center (NERSC)  
Lawrence Berkeley National Laboratory
- 4 Oak Ridge Leadership Computing Facility (OLCF)  
Oak Ridge National Laboratory

## Biological and Environmental Research (BER)

- 17 Atmospheric Radiation Measurement (ARM) User Facility  
Fixed and Mobile Sites Across the Globe
- 18 Environmental Molecular Sciences Laboratory (EMSL)  
Pacific Northwest National Laboratory
- 19 Joint Genome Institute (JGI)  
Lawrence Berkeley National Laboratory

## Basic Energy Sciences (BES)

### LIGHT SOURCES

- 5 Advanced Light Source (ALS)  
Lawrence Berkeley National Laboratory
- 6 Advanced Photon Source (APS)  
Argonne National Laboratory
- 7 Linac Coherent Light Source (LCLS)  
SLAC National Accelerator Laboratory
- 8 National Synchrotron Light Source II (NSLS-II)  
Brookhaven National Laboratory
- 9 Stanford Synchrotron Radiation Lightsources (SSRL)  
SLAC National Accelerator Laboratory

### NEUTRON SOURCES

- 10 High Flux Isotope Reactor (HFIR)  
Oak Ridge National Laboratory
- 11 Spallation Neutron Source (SNS)  
Oak Ridge National Laboratory

### NANOSCALE SCIENCE RESEARCH CENTERS

- 12 Center for Functional Nanomaterials (CFN)  
Brookhaven National Laboratory
- 13 Center for Integrated Nanotechnologies (CINT)  
Sandia National Laboratories and  
Los Alamos National Laboratory
- 14 Center for Nanophase Materials Sciences (CNMS)  
Oak Ridge National Laboratory
- 15 Center for Nanoscale Materials (CNM)  
Argonne National Laboratory
- 16 The Molecular Foundry (TMF)  
Lawrence Berkeley National Laboratory

## Fusion Energy Sciences (FES)

- 20 DIII-D National Fusion Facility  
General Atomics
- 21 National Spherical Torus Experiment Upgrade (NSTX-U)  
Princeton Plasma Physics Laboratory

## High Energy Physics (HEP)

- 22 Facility for Advanced Accelerator Experimental Tests (FACET)  
SLAC National Accelerator Laboratory
- 23 Fermilab Accelerator Complex  
Fermi National Accelerator Laboratory

## Nuclear Physics (NP)

- 24 Argonne Tandem Linac Accelerator System (ATLAS)  
Argonne National Laboratory
- 25 Continuous Electron Beam Accelerator Facility (CEBAF)  
Thomas Jefferson National Accelerator Facility
- 26 Facility for Rare Isotope Beams (FRIB)  
Michigan State University
- 27 Relativistic Heavy Ion Collider (RHIC)  
Brookhaven National Laboratory

## Accelerator R&D and Production (ARDAP)

- 28 Accelerator Test Facility (ATF)  
Brookhaven National Laboratory

# OFFICE OF SCIENCE BY THE NUMBERS

Delivering scientific discoveries and major scientific tools to transform our understanding of nature and advance the energy, economic, and national security of the United States

FY23

## 6 CORE SCIENCE PROGRAMS

- Advanced Scientific Computing Research
- Basic Energy Sciences
- Biological and Environmental Research
- Fusion Energy Sciences
- High Energy Physics
- Nuclear Physics

## 3 ENGINEERING AND TECHNOLOGY OFFICES

- Accelerator Research and Development and Production
- Isotope Research and Development and Production
- Small Business Innovation Research (SBIR) and Small Business Technology Transfer (STTR)

## 5 NATIONAL QUANTUM INFORMATION SCIENCE RESEARCH CENTERS

ACROSS ITS 10 NATIONAL LABS, OFFICE OF SCIENCE MAINTAINS APPROXIMATELY

**24 MILLION**  
SQUARE FEET OF SPACE

**1,600**  
BUILDINGS

**38,000**  
ACRES OF  
LAND OWNED

SUPPORTS RESEARCH SPANNING

**16**  
DOE  
NATIONAL LABS

**50**  
STATES, GUAM,  
PUERTO RICO, AND  
WASHINGTON, D.C.

**>310**  
UNIVERSITIES AND  
HIGHER-LEARNING  
INSTITUTIONS

## 4

BIOENERGY  
RESEARCH  
CENTERS

## 2

ENERGY  
INNOVATION  
HUB  
PROGRAMS

STEWARDS

## 10

DOE NATIONAL  
LABORATORIES

ESTIMATED  
RESEARCHERS  
SUPPORTED

**11,100** Permanent PhDs

**3,400** Postdoctoral  
Associates

**5,200** Graduate Students

**9,700** Other Scientific  
Personnel

OVER

## 39,500

USERS AT

## 28

OFFICE OF SCIENCE  
FACILITIES

## 10

SITE OFFICES

## 1

CONSOLIDATED  
SERVICE CENTER

OVER

## 100

NOBEL  
PRIZES

**\$8.1 BILLION**

OVERALL  
OFFICE OF  
SCIENCE BUDGET

**\$918 MILLION**

USER  
FACILITY  
CONSTRUCTION

**\$281 MILLION**

SCIENCE  
LABORATORIES  
INFRASTRUCTURE

## 3

World-Leading  
Supercomputers

## 51

ENERGY  
FRONTIER  
RESEARCH  
CENTERS

# The Office of Science Research Portfolio



## Advanced Scientific Computing Research

- Delivering world leading computational and networking capabilities to extend the frontiers of science and technology

## Basic Energy Sciences

- Understanding, predicting, and ultimately controlling matter and energy flow at the electronic, atomic, and molecular levels

## Biological and Environmental Research

- Understanding complex biological, earth, and environmental systems

## Fusion Energy Sciences

- Supporting the development of a fusion energy source and supporting research in plasma science

## High Energy Physics

- Understanding how the universe works at its most fundamental level

## Nuclear Physics

- Discovering, exploring, and understanding all forms of nuclear matter

## Isotope R&D and Production

- Supporting isotope research, development, production, processing and distribution to meet the needs of the Nation

## Accelerator R&D and Production

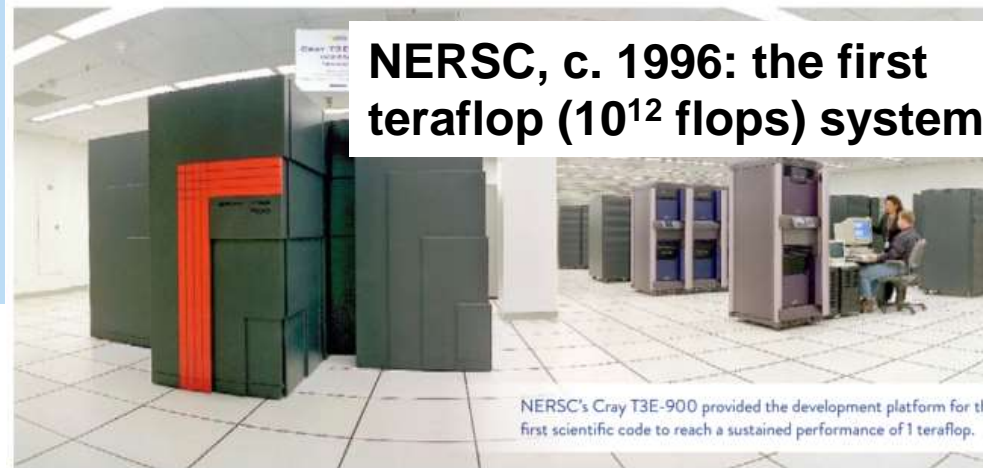
- Supporting new technologies for use in SC's scientific facilities and in commercial products

# ASCR – over 70 years of Advancing Computational Science

**Beginnings:** During the Manhattan Project, John Von Neumann advocated for the creation of a Mathematics program to support the continued development of applications of digital computing



Over 40+ years, ASCR has a rich history of investment in computational science and applied mathematics research, and revolutionary computational and network infrastructure.

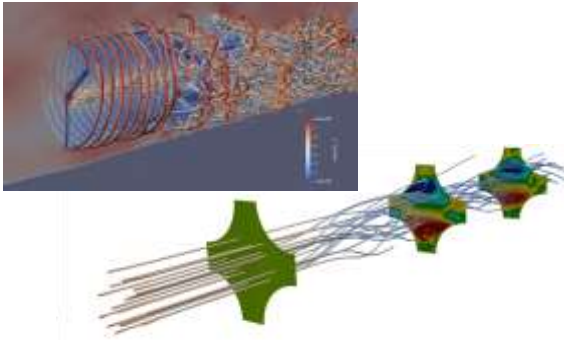


## WHY COMPUTATIONAL SCIENCE?

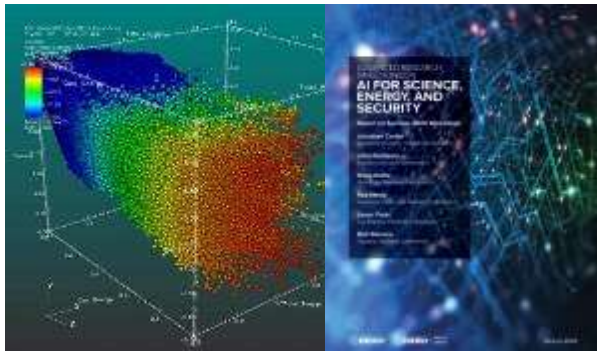
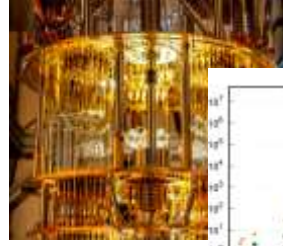
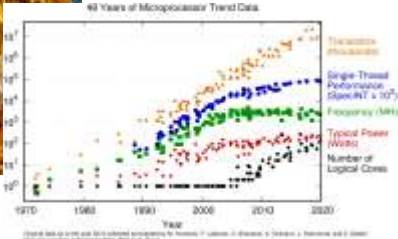
- Computational science adds a third pillar to researcher's toolkit along side theory and experiments
- Computational science is essential when experiments are too expensive, dangerous, time-consuming or impossible
- Computational science facilitates idea-to-discovery that leads from equations to algorithms
- Virtually every discipline in science and engineering has benefited from DOE's sustained investments in computational science

# Emerging Technology Trends for Scientific Computing

## Advanced Modeling, Simulation, and Visualization



## Trustworthy Artificial Intelligence and Data









Heterogeneous, Distributed, Co-Designed, Energy-Efficient Computing and Algorithms

## Software Complexity for Increased Versatility

HOW MANY LINES OF CODE MAKE UP THESE POPULAR TECHNOLOGIES

Technology	Lines of Code
Linux	10,000,000
Android	4,000,000
Windows	100,000,000
Python	230,000
OpenOffice	4,100,000
Field of View	620,000
Python 3 code	2,500,000
WebKit	11,000,000
Linux kernel	10,100,000
Python 2	10,000,000
Python	10,000,000
Python	10,000,000

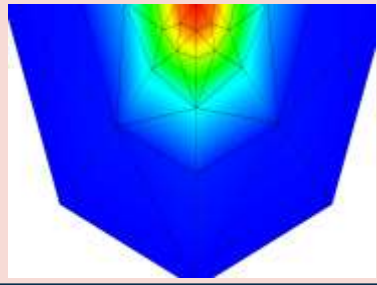





High-Performance Computing and Networking across Experiments, Exascale, and the Edge



# ASCR Research: Key To Enabling DOE and SC Scientific Enterprise

Simulation, modeling and data-driven discovery combined with testbeds and prototypes equip the ASCR community, big and small, to tackle scientific and societal crises.



## Discovery Science

ASCR's SciDAC partnership with Fusion Energy Sciences uses exascale-ready software to understand plasma motion.

## Lowering Energy Costs

Multi-scale mathematics algorithms and models led to insights to reduce energy in industrial coating by nearly a third.

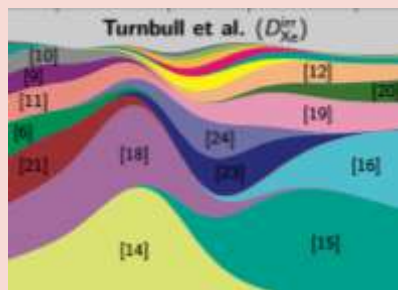
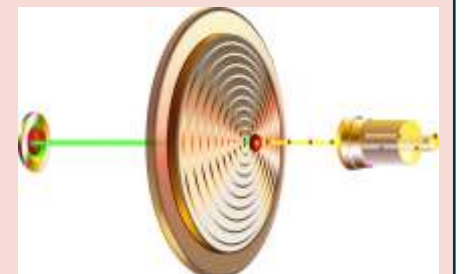


## Optimizing Experiments

Optimization and AI methods provided real-time experiment steering at beamlines and microscopes.

## Foundations For the Future

Design and demonstration of a deterministic single-photon source for quantum networking and computing.

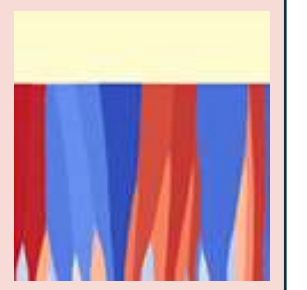


## Partnerships for Energy

ASCR's SciDAC partnership with Nuclear Energy predicts diffusion of xenon under irradiation conditions.

## Insights Unlocking Technologies

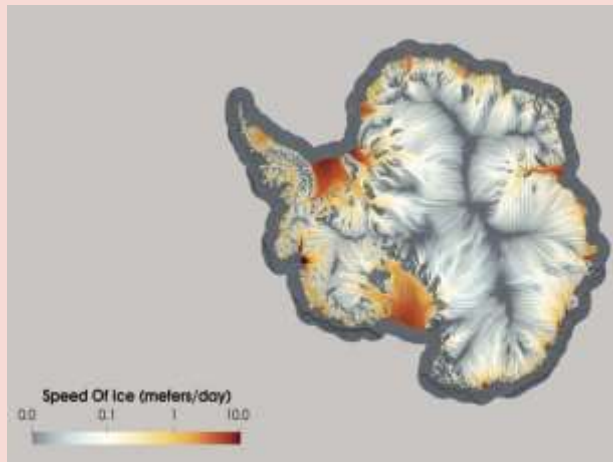
AI models predict the 3D grain structures of cooling metals to enable new advanced-manufacturing technologies.



# Understanding Changing Environmental Conditions: Sea & Fire

State-of-the-art research in simulation, modeling and data-driven discovery help us improve our understanding of fundamental processes and our projections for the changing global environment.

## Projected Land Ice Contribution to 21<sup>st</sup> Century Sea Level Rise



- The most comprehensive projections of sea-level rise from land ice to date.
- Antarctica remains a critical focus for reducing future sea level uncertainty.
- Limiting global warming to 1.5°C reduces 21st century land ice contribution to sea-level rise from 25 to 13 cm.

By simulating the flow of ice across Antarctica using an improved ice-sheet model, the researchers projected 2015-2100 land ice contribution to sea level for a range of emissions scenarios.

### An ASCR-BER SciDAC Partnership

## 5G Drones: Real Time Data Assimilation to Transform Wildfire Predictability



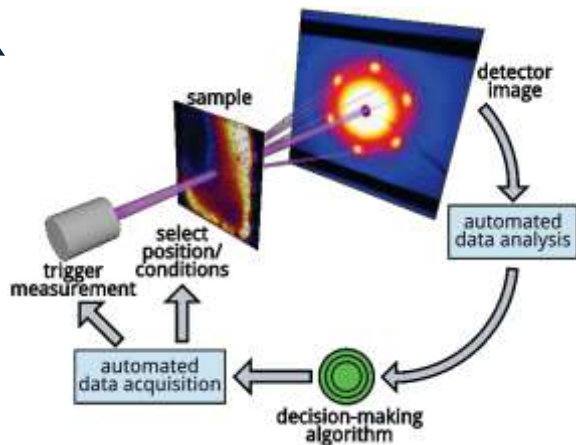
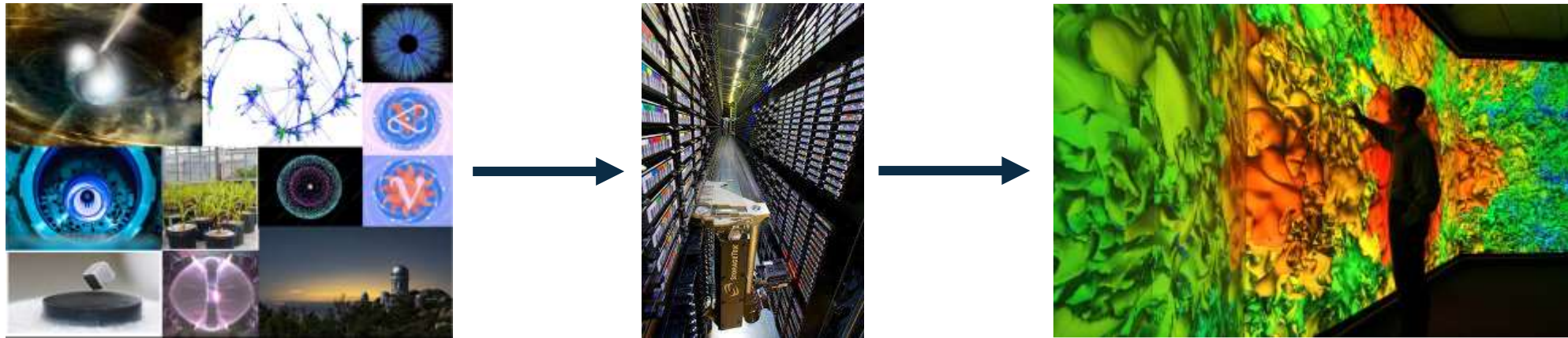
5G drone data will lead to better predictions of smoke and fire spread.

- Use 5G drones to assess changes in fire behavior and smoke characteristics.
- Leverage data gathered via various sources such as citizen scientists.
- Coordinate with partners to integrate fire modeling into fire master plans.



The time evolution of the Rio Medio (NM) fire was captured by citizen images and videos from multiple angles and distances. The researchers are harnessing this unique data set to inform their simulations and improve their models to enable better forecasts.

# Scientific Data at Extreme Scale



- Scientific computations and experiments produce terabytes or petabytes of data that must be efficiently stored.
- That data is stored on collections of disk drives and archive systems at ASCR computing facilities.
- As with ASCR's computing capabilities, high-performance data management requires performing many operations in parallel.
- ASCR invests in innovative ways to store, compress, search, and analyze data that maximizes parallelism and performance.
- ASCR also invests in advancements in streaming data and federated learning, allowing data in geographically-separated places to contribute to scientific modeling without needing to store all of the data in once place.

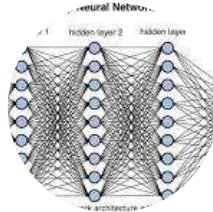
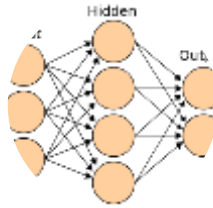
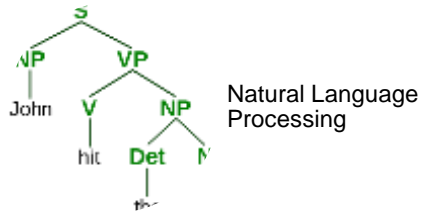
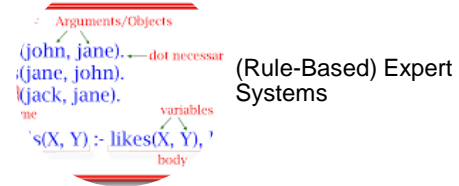
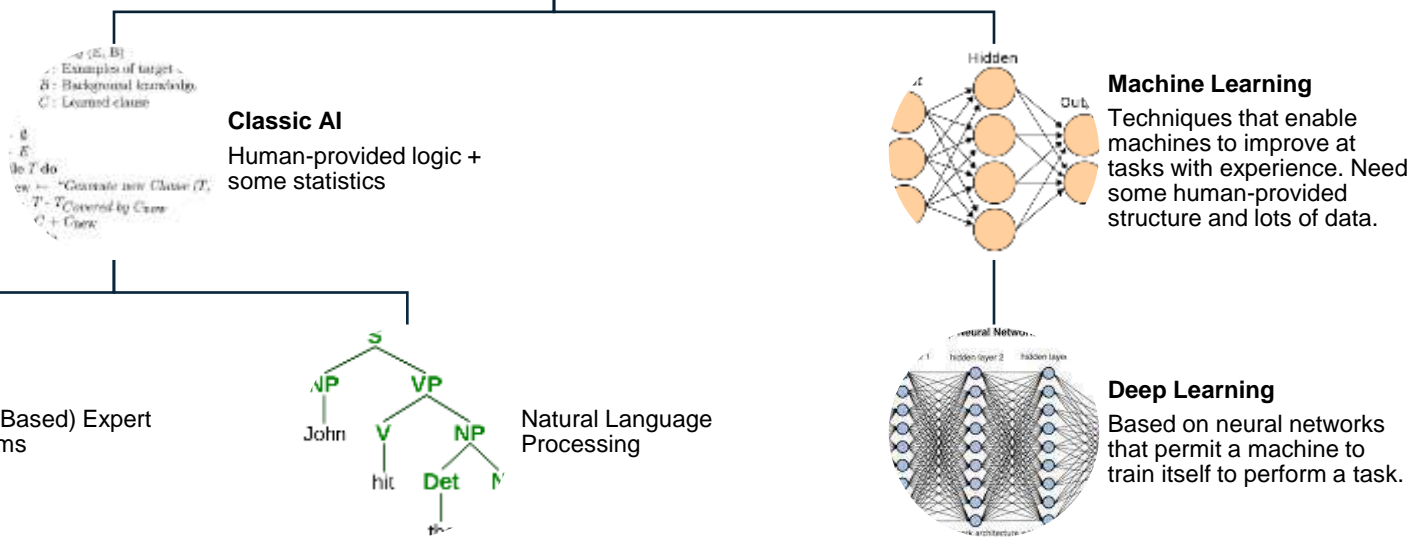
# A Rough Evolution of Artificial Intelligence



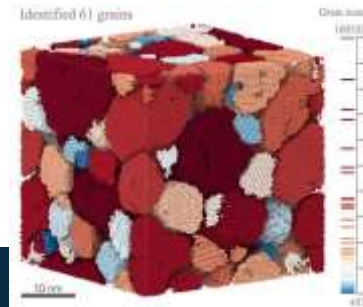
1960s



**AI**  
Any technique that enables computers to mimic human intelligence.



Today



## AI and Machine Learning R&D

**AI & Machine Learning** for predictive models from large-scale data

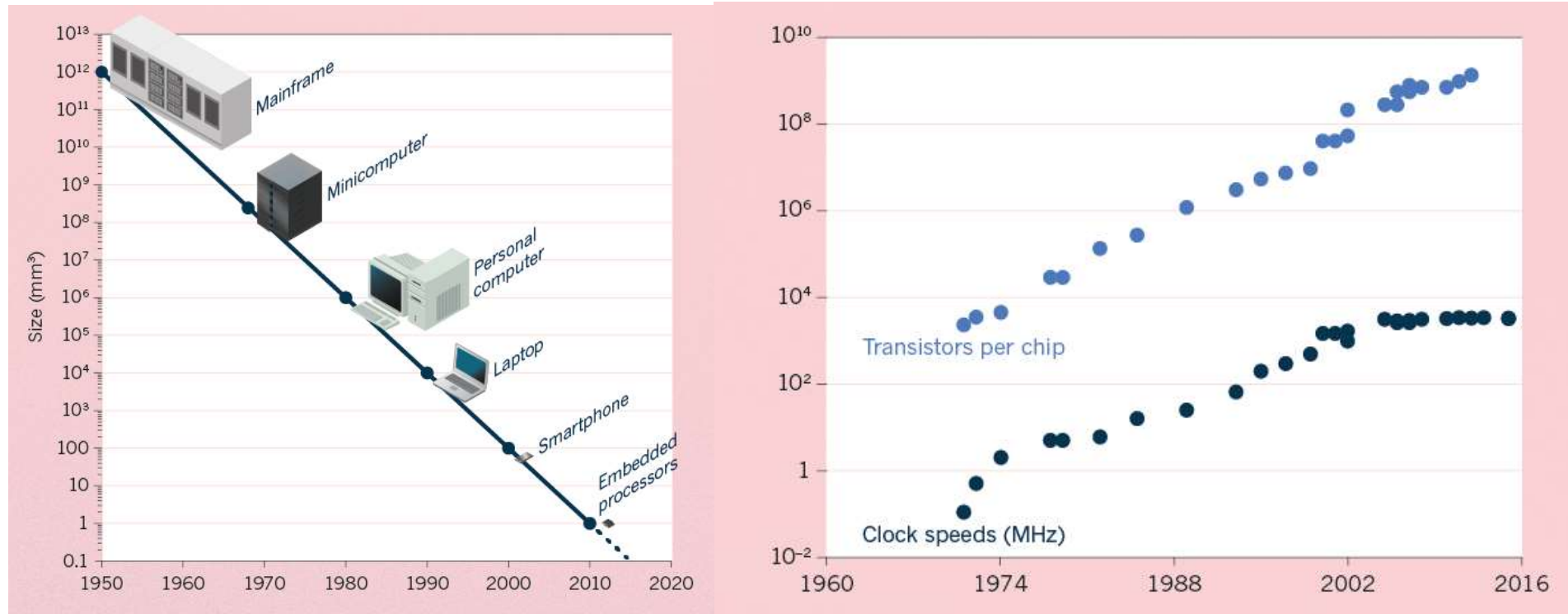
**Federated Learning** for broader insights via collectively shared datasets

**Privacy-preserving algorithms** for cybersecurity and AI at the edge

**AI hardware co-design** for energy-efficient **hybrid algorithms** & computing

Tools for ensuring **FAIR data for AI**  
AI/ML for **autonomous experiments**

# Moore's Law



<https://www.nature.com/news/the-chips-are-down-for-moore-s-law-1.19338>

- Moore's law is the observation that the number of transistors in an integrated circuit (IC) doubles about every two years.
- As Moore's law has continued computers have continued to shrink *and* become more capable.
- However, the clock speed of energy-efficient computers stopped increasing some time ago – this is why parallel computing, doing more simultaneously, is critical to modern computing including ASCR's supercomputers.

# Quantum Computer Simulation of Physical Systems

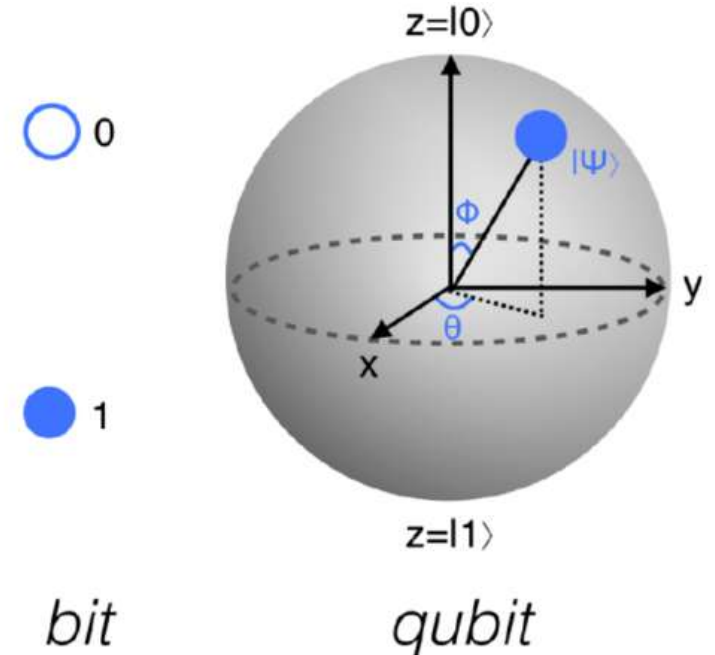


Richard P. Feynman  
*Simulating Physics with Computers,*  
*Int. J. Theor. Phys. (1982)*

## Power of Quantum Superposition

- |                                                                                                                                                                                                                                                           |                                                                                                                                                                                                                                                                     |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <ul style="list-style-type: none"><li>• Classical bit can represent 0 <b>or</b> 1</li><li>• Eight classical bits can represent one of 256 integers (<math>2^8</math>)</li><li>• N classical bits can represent one of <math>2^N</math> integers</li></ul> | <ul style="list-style-type: none"><li>• Quantum bit can be in a superposition of 0 <b>and</b> 1</li><li>• Eight quantum bits can represent all of 256 integers (<math>2^8</math>)</li><li>• N quantum bits can represent all of <math>2^N</math> integers</li></ul> |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

- Qubit states are fragile.
- Wiring qubits together into a functional architecture is hard.

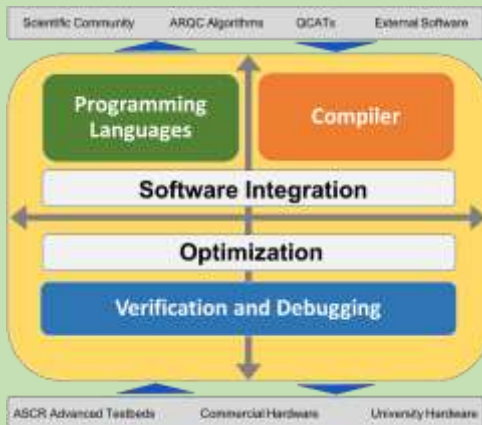


*Emani, P.S., Warrell, J., Anticevic, A. et al. Quantum computing at the frontiers of biological sciences, Nat. Methods (2021)*

# Quantum Computing in ASCR

## Fundamental Science

Programs support core basic research for quantum algorithms, quantum computer science and quantum networking.



AIDE-QC, an ARQC team, explores five thrusts to program emerging QC platforms and support the broader DOE quantum community.

## National QIS Research Centers

Support for the Centers, the first large-scale QIS effort that crosses the technical breadth of Office of Science.



Five National QIS Research Centers address major cross-cutting challenges in broad ranging topics in QIS including computing, communications and sensing.

## Quantum Internet Testbeds

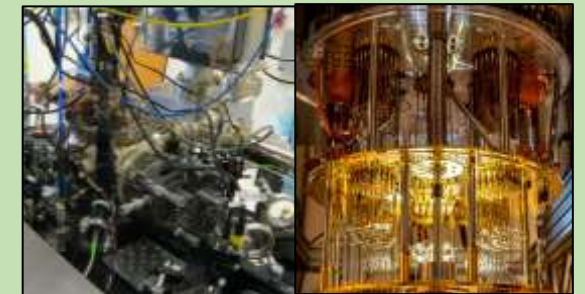
Research and development for the deployment of regional testbeds to provide early proof of concepts.



In FY21 ASCR awarded two projects, led by LBNL and ORNL to design, develop and demonstrate regional-scale quantum internet testbeds.

## Quantum Computing Testbeds

Provide the research community with fully transparent access to novel quantum computing hardware.



SNL's QSCOUT (left) is the world's first publicly-available trapped ion quantum computer. LBNL's AQT (right) offers access to a unique superconducting platform.

# ASCR R&D Funding (\*\*)

## Funding Opportunity Announcements (FOAs)

- <https://science.osti.gov/ascr/Funding-Opportunities>
- Announced on [grants.gov](https://www.grants.gov) (hint: sign up for email notifications for 'ASCR')
- Read each announcement carefully to understand who can apply and other restrictions/requirements
- Depending on the announcement, supports 2–5-year projects
- University researchers can apply directly (please coordinate with your organization's sponsored-research office)
- Subcontracting is often permitted, and sometimes collaborative applications are permitted

## Early Career Research Program

- <https://science.osti.gov/early-career>
- Research grants for five years
- Stays with PI if PI changes institutions
- Eligible within 10 years of Ph.D. (can apply up to three times)
- University-based researchers receive about \$175,000/year
- Topics released in the summer, pre-applications generally due in the fall

## DOE National Laboratory Announcements

- <https://science.osti.gov/ascr/Funding-Opportunities> (bottom of the page)
- Open only to DOE Laboratories
- Often allow subcontracts to support collaborators at other organizations

## Small Business Innovation Research (SBIR) and Small Business Technology Transfer (STTR)

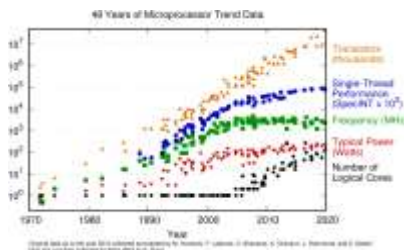
- <https://science.osti.gov/sbir>
- Grants to for-profit US businesses with 500 or fewer employees (including affiliates)
- Phase I: ~\$200k for 6-12 months, Phase II: ~\$1M for 2 years
- Subcontracting is permitted, STTR: requires collaboration with a research Institution
- Topics released in the summer, pre-applications generally due in the fall

## Computational Science Graduate Fellowship (CSGF)

<http://www.krellinst.org/csgf/>



# Transforming the Fundamentals of Computing



Heterogeneous, Distributed,  
Co-Designed, Energy-Efficient  
Computing and Algorithms



ASCR Workshop on Reimagining Codesign,  
March 2021: <https://doi.org/10.2172/1822199>



Quantum Computing for Biomedical  
Computational and Data Sciences  
Joint DOE NIH Quantum Roundtable  
March 2023:  
<https://doi.org/10.2172/2228574>



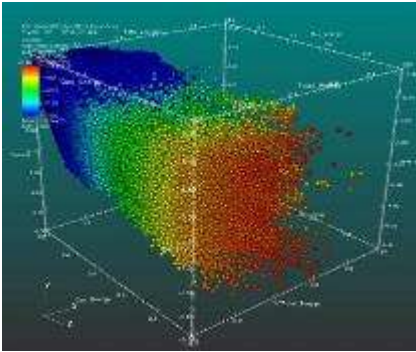
ASCR Basic Research Needs in Quantum  
Computing and Networking, July 2023:  
<https://doi.org/10.2172/2001044> (brochure;  
report forthcoming)

FY 2023

FY 2023

# Empowering Science Through Data Innovations

Data, Privacy, and  
Scientific Integrity



ASCR Workshop on Basic Research Needs for  
Management and Storage of Scientific Data,  
January 2022:  
<https://doi.org/10.2172/1845707>

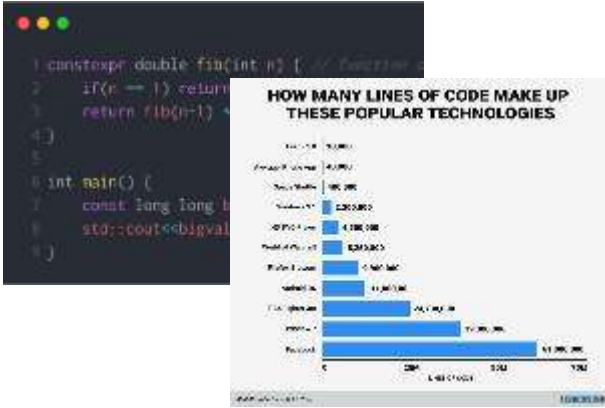


ASCR Basic Research Needs Visualization for  
Scientific Discovery, Decision-Making, and  
Communication, January 2022:  
<https://doi.org/10.2172/1845708> (brochure;  
report forthcoming)



# Enhancing Scientific Programming

## Exploding Software Complexity



ASCR Workshop on Basic Research Needs in The Science of Scientific Software Development and Use, December 2021: <https://doi.org/10.2172/1846009>



DOE/NSF Workshop on Correctness in Scientific Computing, June 2023: <https://arxiv.org/abs/2312.15640>

FY 2023

# Accelerating Science from Exascale to the Edge



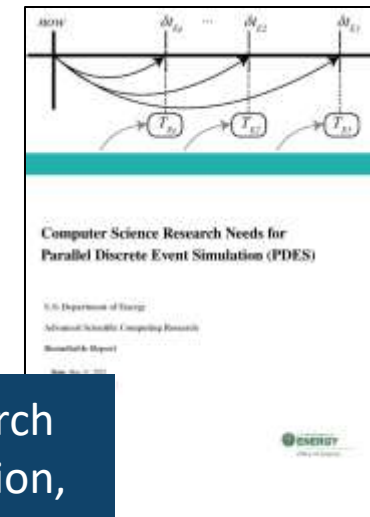
Scientific Computing and Networking: from Exascale to the Edge

Integrated Research Infrastructure Architecture  
Blueprint Activity, 2023:  
<https://doi.org/10.2172/1984466>

FY 2023



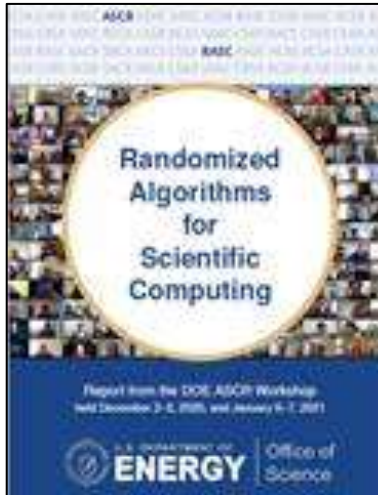
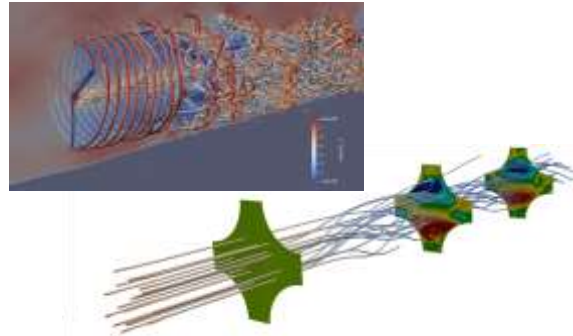
Roundtable on Foundational Science for  
Biopreparedness and Response, March 2022:  
Report available from  
<https://science.osti.gov/ascr/Community-Resources/Program-Documents>



Roundtable on Computer Science Research  
Needs for Parallel Discrete Event Simulation,  
2022: <https://doi.org/10.2172/1855247>

# Innovating in Algorithms and Mathematics

Advanced Modeling,  
Simulation, and Visualization



ASCR Workshop on Randomized Algorithms for  
Scientific Computing, January 2021:  
<https://doi.org/10.2172/1807223>



Data Reduction for Science, January 2021:  
<https://doi.org/10.2172/1770192>

# Additional Information on ASCR's Website

<https://science.osti.gov/ascr/Community-Resources/Program-Documents>

<https://science.osti.gov/ascr/Funding-Opportunities>

About

Research

Facilities

Science Highlights

Benefits of ASCR

**Funding Opportunities**

Closed Funding Opportunity Announcements (FOAs)

Closed Lab Announcements

Award Search / Public Abstracts

Additional Requirements and

## Funding Opportunities

Look at past opportunity announcements

Other non-profit organizations as well as those germane to the mission of DOE, and solicitations for each research program. The selection of researchers to fund is based on the solicitation. For the most current information, the original posting dates, check the Office of Science Guidance on ASCR's website.

Office of Science Guidance on ASCR's website

Look at abstracts for current awards

Look at recent reports from ASCR-sponsored workshops. These discuss priority research directions, as identified by the research community, along with relevant background information, in various areas.

## ASCR Program Documents

Provided below is a listing of relevant articles, plans and ASCR-sponsored workshop reports.

Select the link to view the ASCR Program Document Archive.

- ASCR@40: Four Decades of Department Of Energy Leadership in Advanced Scientific Computing Research**  
In December 2017, the Advisory Committee for DOE's Office of Advanced Scientific Computing Research (ASCR) was asked to document some of the major impacts of ASCR and its predecessor organizations. The summary report includes a list of multi-year processes of information gathering, drafting, reviewing, and releasing input was provided by over 100 scientists.  
Full Report
- A Quantum Path Forward**  
Today, many scientific experts recognize that building and scaling quantum-powered and enhanced communication networks are among the most important technological frontiers of the 21st century. The international research community perceives the construction of a first prototype global quantum network—the Quantum Internet—as the within reach over the next decade.  
In February 2021, the U.S. Department of Energy (DOE)'s Office of Advanced Scientific Computing Research hosted the Quantum Internet Strategic workshop to define a potential roadmap toward building the first reconfigurable quantum internet. The workshop participants included representatives from DOE national laboratories, universities, industry, and other U.S. agencies with various interests in quantum networking. The goal was to produce an outline of the essential research needed, critical engineering and design barriers, and suggest a path forward to review from today's limited social network experiments to a viable, secure quantum internet.  
Workshop Report
- 5G Enabled Energy Innovation Workshop (5GEEIW)**  
On March 10-12, 2020, the Office of Science (OS) organized a three-day workshop to deliver a consensus-based report highlighting 5G and beyond 5G research, development, applications, technology transition, infrastructure, and dissemination opportunities in support of the U.S. DOE mission. The literature and report will help the OS/OS Office of Science understand both the challenges and the opportunities offered by 5G and emerging advanced wireless technologies in the areas of basic research, development, and integration into scientific user facility operations.  
Cover | Abstracts | Workshop Report
- Data and Models: A Framework for Advancing AI in Science**  
On June 5, 2019, the Office of Science (OS) organized a meeting to establish a focus on enhancing access to high-quality and fully traceable research data, models, and computing resources to increase the value of such resources for artificial intelligence (AI) research and development and the OS mission. In this report, we consider AI to be inclusive of, for example, machine learning (ML), deep learning (DL), neural networks (NN), computer vision, and natural language processing (NLP). The computer "data for AI" means the digital artifacts used to generate AI results either employed in combination with AI results during inference. In sum, this report was motivated by the recognition that a large portion of advanced data currently are not well suited for AI.  
View Technical Report
- Storage Systems and I/O: Organizing, Storing, and Accessing Data for Scientific Discovery**  
In September, 2018, the Department of Energy, Office of Science, Advanced Scientific Computing Research Program convened a workshop to identify key challenges and define research directions that will advance the field of storage systems and I/O over the next 5-7 years. The workshop concluded that addressing these current challenges and opportunities requires tools and techniques that greatly extend traditional approaches and require new research directions. Key research opportunities were identified.  
View Technical Report
- ASCR Workshop on In Situ Data Management**  
In January 2018, ASCR convened a workshop on In Situ Data Management (ISDM). The goal was to identify priority research directions (PRDs) to support current and future scientific computing needs, which will increasingly incorporate a number of different tasks that need to be managed along with the main simulation or data analysis tasks. The

# AI4SES Report

- AI for Science, Energy, and Security Report, released May 2023:  
<https://www.anl.gov/ai-for-science-report>
- Created by a confederation of laboratories, informed by a series of workshops held in 2022.
- Covers AI approaches:
  - AI and Surrogate Models for Scientific Computing
  - AI Foundation Models for Scientific Knowledge Discovery, Integration, and Synthesis
  - AI for Advanced Property Inference and Inverse Design
  - AI-Based Design, Prediction, and Control of Complex Engineered Systems
  - AI and Robotics for Autonomous Discovery
  - AI for Programming and Software Engineering
- Also covers crosscuts, including workflows, data, AI hardware, computing infrastructure, and workforce



# Exascale Today Enables the AI of Tomorrow

Long-term investments in applied mathematics and computer science enabled exascale.



TOP500  
# 1

GREEN500  
# 2

HPL-MxP  
# 1

Frontier, #1 on the Top500, **leads the world in computational capability**, and is also **#2 in the world in energy efficiency**, and is **#1 in the world for AI capability**.

The exascale and AI-enabled science era will lead to dramatic capabilities to predict extreme events and their impacts on the electric grid across weather and climate time scales...



and will accelerate the design and deployment of clean-energy technologies to create a better future.





# Exascale Computing Project (ECP)

*DOE's Exascale Computing Initiative: A partnership between SC and NNSA/ASC to accelerate R&D, acquisition, and deployment to deliver exascale computing capability to DOE national labs by the early- to mid-2020s*

6 Core DOE Labs  
100 R&D Teams  
1000 Researchers

Exascale System  
deployment  
Frontier, Aurora,  
El Capitan

## APPLICATION DEVELOPMENT

*Develop and enhance the predictive capability of applications critical to DOE*

## SOFTWARE TECHNOLOGY

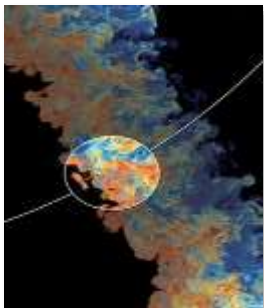
*Expanded & vertically integrated software stack for capable exascale computing*

## HARDWARE AND INTEGRATION

*Integrated delivery of ECP products on targeted systems at leading DOE HPC facilities*

### National security

Stockpile Stewardship  
Reentry-vehicles  
High-energy density physics



### Energy security

Wind farms  
Small Modular Reactors  
Nuclear materials  
Subsurface Science  
Combustion  
Clean fossil fuels  
Biofuel catalysts

### Economic security

Additive manufacturing  
Power grid  
Seismic risk



### Scientific discovery

Astrophysics  
Lattice QCD  
Accelerators  
Materials  
Chemistry  
Fusion  
Standard Model

### Earth system

Earth system models  
Biomass  
Metagenomics (DOE applications)

### Health care

Cancer



**On track for CD-4 in FY24**

# Then (2016) and Now (2023): WarpX

Modeling of charged particle beams and accelerators, lab & astro plasmas, fusion devices

## WarpX & Spinoffs History & Roadmap

**Legend:**  
 PIC = Particle-In-Cell  
 EM = Electromagnetic  
 ES = Electrostatic  
 QS = Quasistatic

## Overview of Warp/WarpX

Warp and WarpX are multiphysics codes/frameworks for the modeling of charged particle beams and accelerators, lab & astro plasmas, fusion devices & more.

Codes are constructed around the Particle-In-Cell (PIC) algorithm:

**Eulerian**  
electromagnetic fields  
on structured grid

**Lagrangian**  
charged macroparticles

**Challenge ECP problem:** the modeling of chains of plasma-based particle accelerators for future high-energy physics colliders

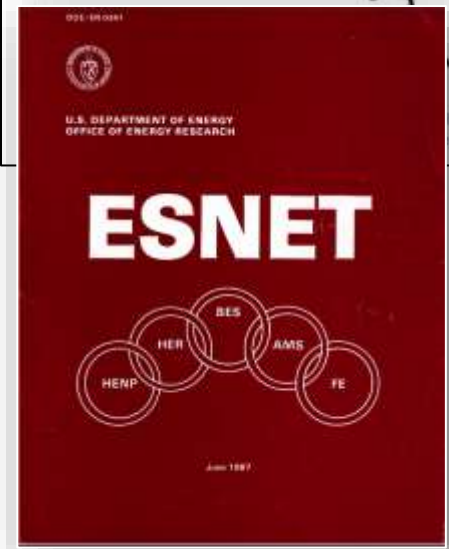
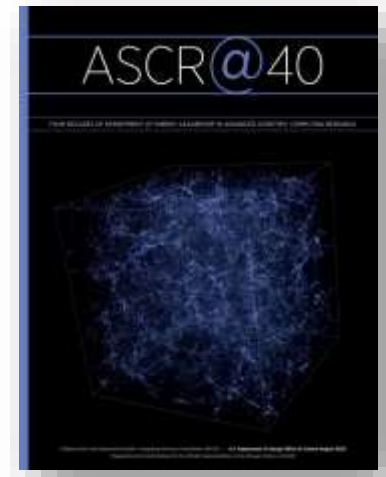
Warp as of 2016		WarpX as of 2023
large set of advanced, novel algorithms	<b>Algorithms</b>	Warp advanced algos + new algorithms introduced during ECP
50% Fortran + 50% Python (including programmable frontend) had grown to large >1M lines of codes w/ varying programming styles	<b>Source code</b>	Source code: C++17 & optional Python programmable frontend compact thanks to C++ templating
CPUs, MPI-parallel	<b>Supported hardware</b>	CPUs, 3 flavors of GPUs, MPI-parallel
limited	<b>Performance optimization</b>	extensive
limited support, independently	<b>Load balancing &amp; AMR</b>	combined native support
compilation from source some support for binaries	<b>Installation</b>	standard (CMake) compilation from source one step with Spack/Conda/PyPI, multi-platform
small team (2+) of computational physicists + individual contributions over several decades	<b>Development team</b>	tightly integrated team of computational physicists + applied mathematicians + computer scientists + software engineers
manual runs of test suite partial online documentation, outdated in part informal code reviews for critical changes	<b>Development policies/practices</b>	extensive test coverage with continuous integration extensive online documentation formal code reviews for all changes
could perform 3-D modeling of single plasma accelerator stage at moderate resolution	<b>ECP science case</b>	can perform 3-D modeling of chain of tens of plasma accelerator stages at twice the resolution in each direction

# The ASCR Facilities ecosystem began with the National Magnetic Fusion Energy Computing Center (later renamed NERSC) (1974) and ESnet (1985).



**CDC 7600**  
S/N 1 - delivered 01/69 retired 10/88  
Cost: \$5.1 million

- 36 million operations per second
- 4,000,000 bytes(chars) magnetic core memory
- Small core memory 65,000 sixty-bit words
- Large core memory 512,000 sixty-bit words
- 3,360 modules
- 120 miles of wire



**A legacy of project excellence**

ESnet through 6  
NERSC through 9  
OLCF through 5  
ALCF through 3  
Exascale Computing Project

# ASCR Facilities: History



Ewing "Rusty" Lusk at ANL's Advanced Computing Research Facility which fielded an array of early parallel systems.



NERSC's Cray T3E-900 provided the development platform for the first scientific code to reach a sustained performance of 1 teraflop.



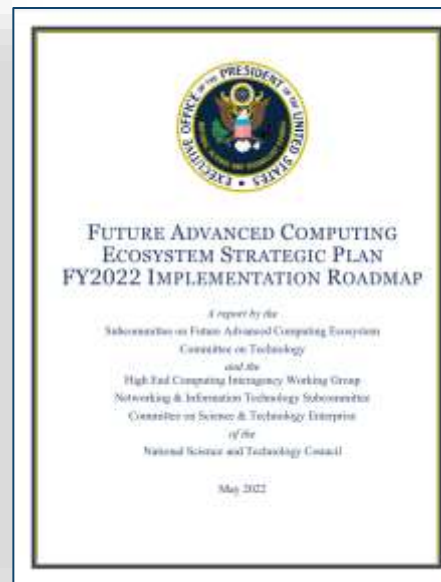
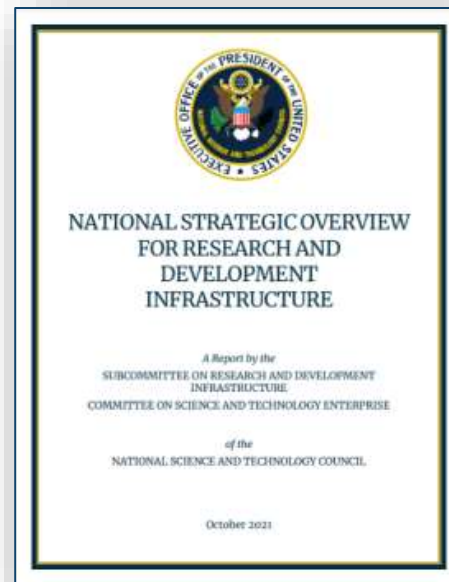
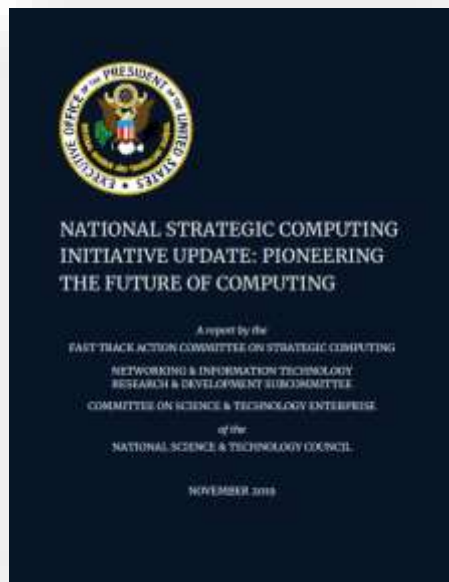
to address the grand challenge applications listed ORNL's Intel Paragon system was installed in 1995, comprising 3,072 processors to support research into Grand Challenge problems.

# The ASCR Facilities mission: Research infrastructure for the nation

Our mission is to **achieve the greatest impact for science and the nation** by delivering **first-of-a-kind high-uptime** high performance computing, data, and networking infrastructure capable of meeting the requirements of extreme scale science.

We seek to **influence the trajectory of computing, data, and networking technology to benefit U.S. competitiveness and the national research enterprise**, and

We seek to influence **how researchers use computing, data, and networking to benefit the practice of science**.



DOE is an apex provider of national research infrastructure (RI). Other USG agencies and industry rely on DOE RI.

DOE's extreme scale RI is unique in the national advanced computing ecosystem.

# The ASCR Facilities are Scientific User Facilities

FY 2023  
28 scientific  
user facilities  
>37,000 users



# ASCR Facilities provide world-leading computing, data, and networking infrastructure for extreme-scale science while advancing U.S. competitiveness

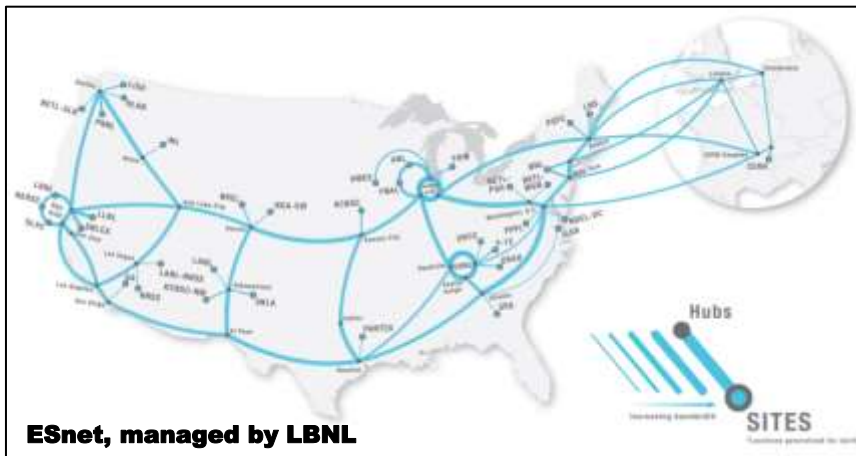
## High Performance Computing Facilities: ALCF, OLCF, NERSC



**Leadership Computing Facilities (ALCF, OLCF):**  
Unique national HPC resources for extreme-scale applications, delivering the exascale ( $10^{18}$ ) era of supercomputing

**High Performance Production Computing Facility (NERSC):**  
Dedicated HPC resource for the Office of Science research community, serving many thousands of users annually

## High Performance Network Facility: ESnet



### Energy Sciences Network (ESnet):

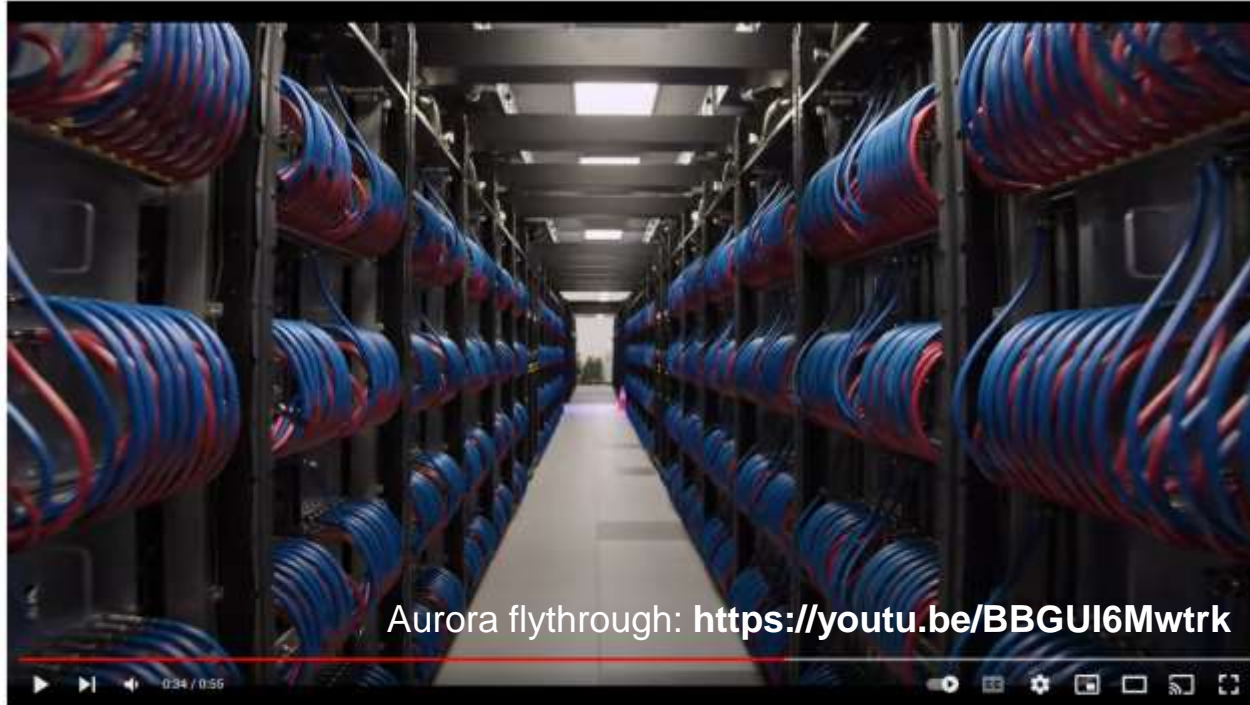
Connects all DOE national labs and dozens of other DOE sites to 150+ global research networks, commercial cloud providers, and the internet

Engineered for lossless transmission of huge data flows

# Today: Exascale systems

**NERSC > 10,000 users!**  
**ESnet6 deployed**

Aurora at Argonne



Aurora Installation Flythrough  
Argonne Leadership Computing Facility  
Subscribe  
Like Share Download Save

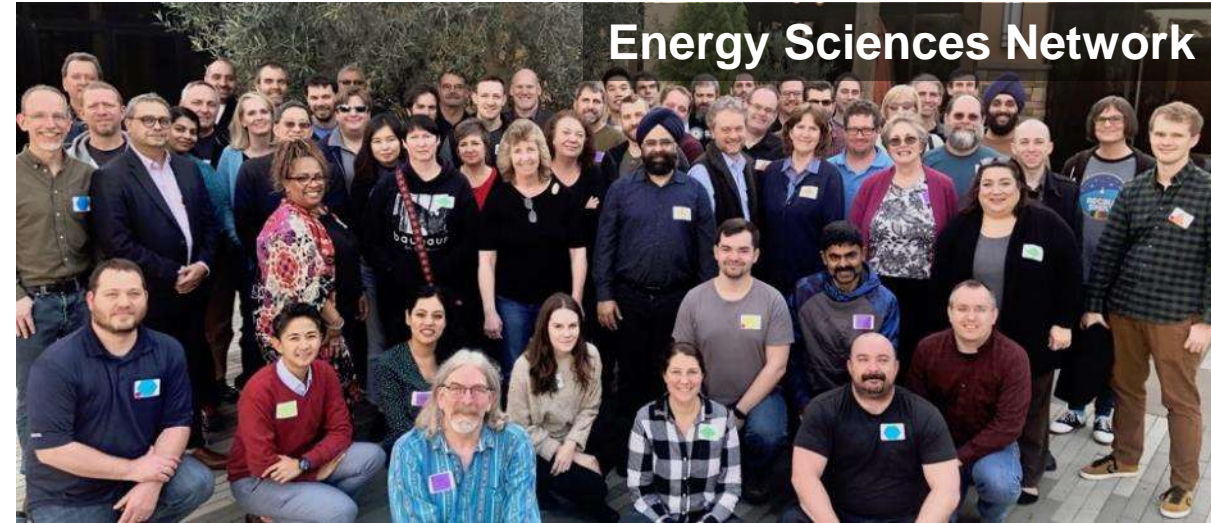


Frontier at Oak Ridge





# The people of the ASCR Facilities



# High Performance Computing Allocation Programs

	INCITE	ALCC	ERCAP	Director's Discretionary
<b>Allocation Program Mission</b>	Advance science and engineering	Advance DOE mission priorities; respond to national emergencies	Advance DOE Office of Science and SBIR/STTR research	Advance science and engineering
<b>Allocatable Time</b>	ALCF, OLCF: 60% NERSC: N/A	ALCF, OLCF: 30% NERSC: 10%	ALCF, OLCF: N/A NERSC: 80%	ALCF, OLCF: 10% NERSC: 10%
<b>Managing Office</b>	ALCF/OLCF	ASCR	DOE Office of Science Programs, SBIR/STTR	Each Facility
<b>Award Duration</b>	One year	One year (offset 6 months relative to INCITE)	One year	One year

For more information, see: <https://science.osti.gov/ascr/Facilities/Accessing-ASCR-Facilities>

# ASCR HPC system lifecycle timeline 2022-2035

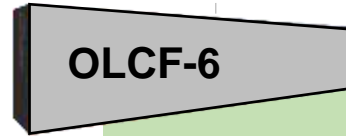
When “accepted,” a system enters a five-year operations window (green bar); the red bar indicates a possible 6<sup>th</sup> year life extension.

2022 2023 2024 2025 2026 2027 2028 2029 2030 2031 2032 2033 2034 2035

## ORNL OLCF



HPE/AMD



OLCF-6

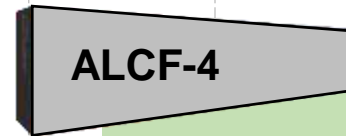


OLCF-7

## ANL ALCF



HPE/Intel



ALCF-4

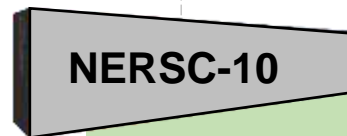


ALCF-5

## LBNL NERSC



HPE/AMD/NVIDIA



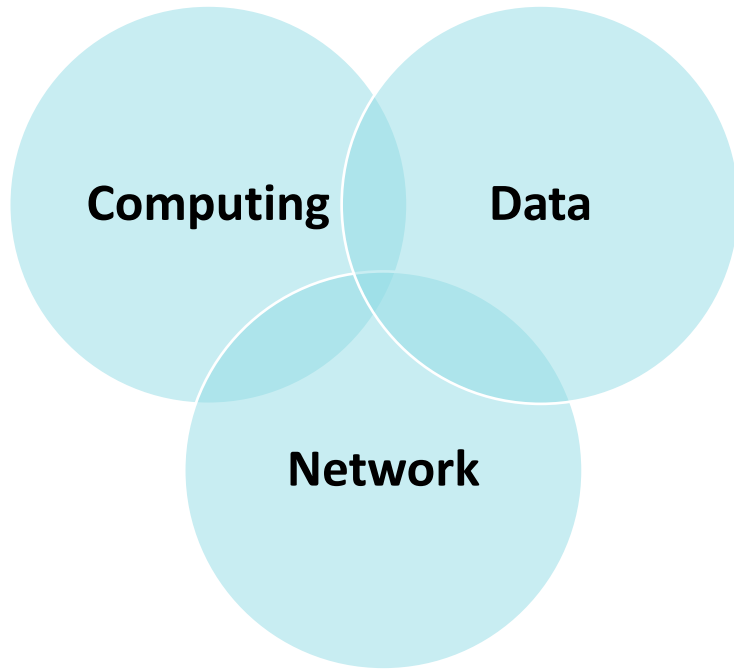
NERSC-10



NERSC-11

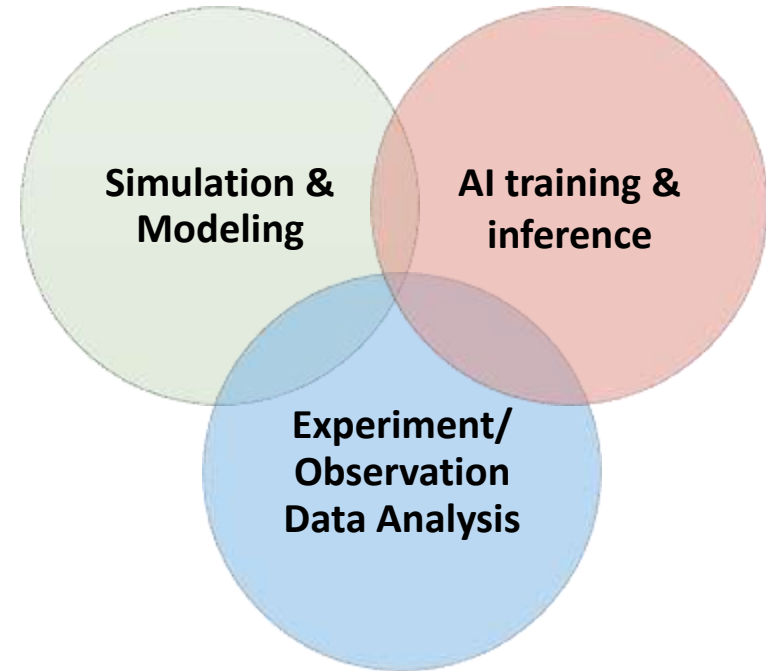
# Major triads in the ASCR Facilities ecosystem

ASCR Facilities  
Research Infrastructure Ecosystem



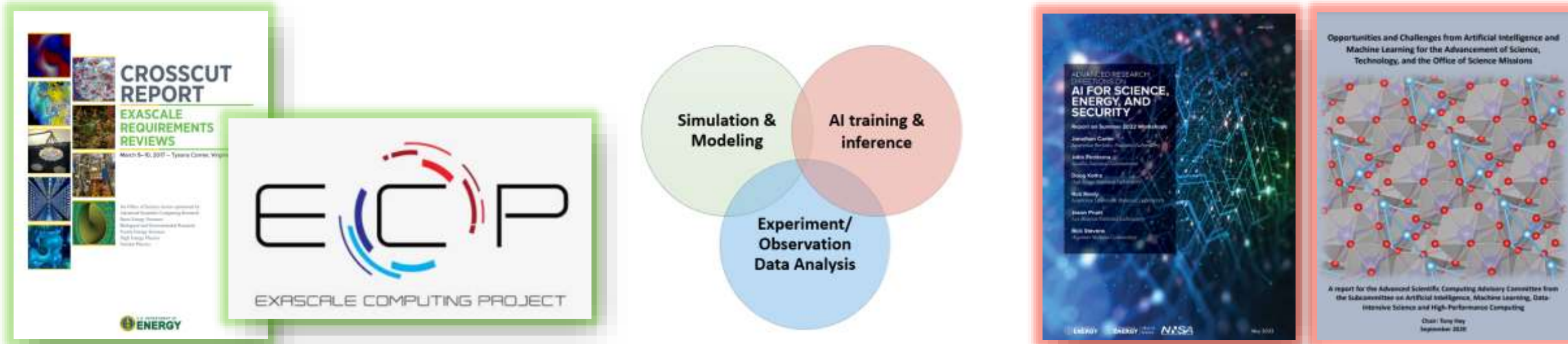
**Workflows!**

How researchers use  
the ASCR Facilities Ecosystem

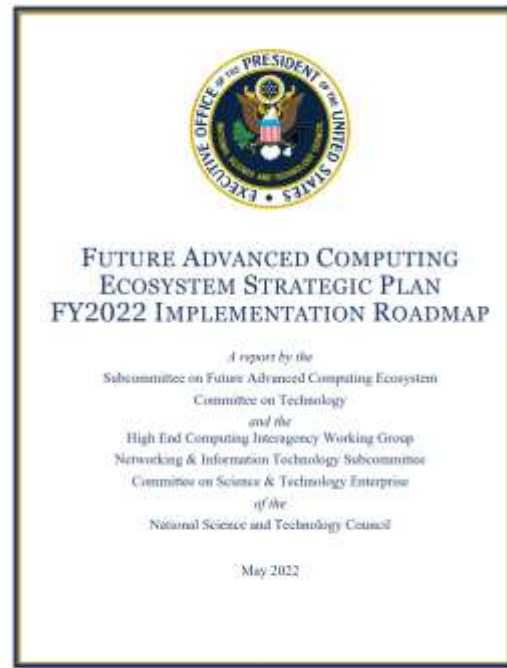
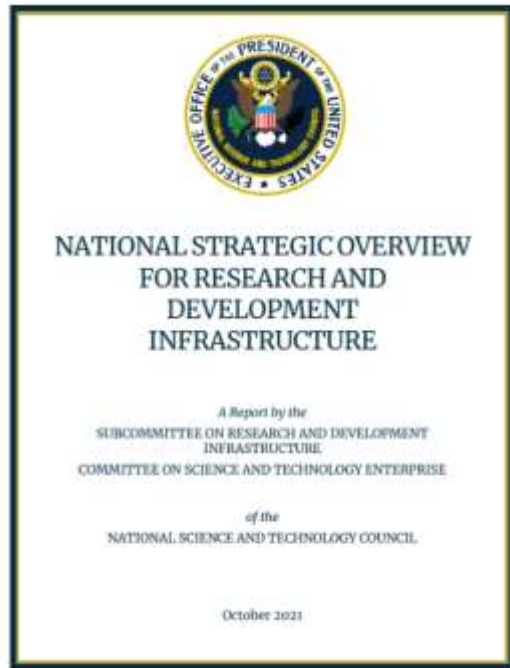


**Core Principle: Provide the right resources for the research project or workflow at the right time. Allocate resources appropriately.**

# The ASCR Facilities work together to synthesize infrastructure requirements from formal requirements reviews, stakeholder and user engagements, and other key reports.



# Interconnectivity and integration of instrumentation, data and computing are essential requirements for national R&D objectives



***"R&D continues to shift from smaller to bigger science, driven in large part by advances in computing and other research cyberinfrastructure, which interlink[s] research data, analytics, ... and experimental instrumentation."***

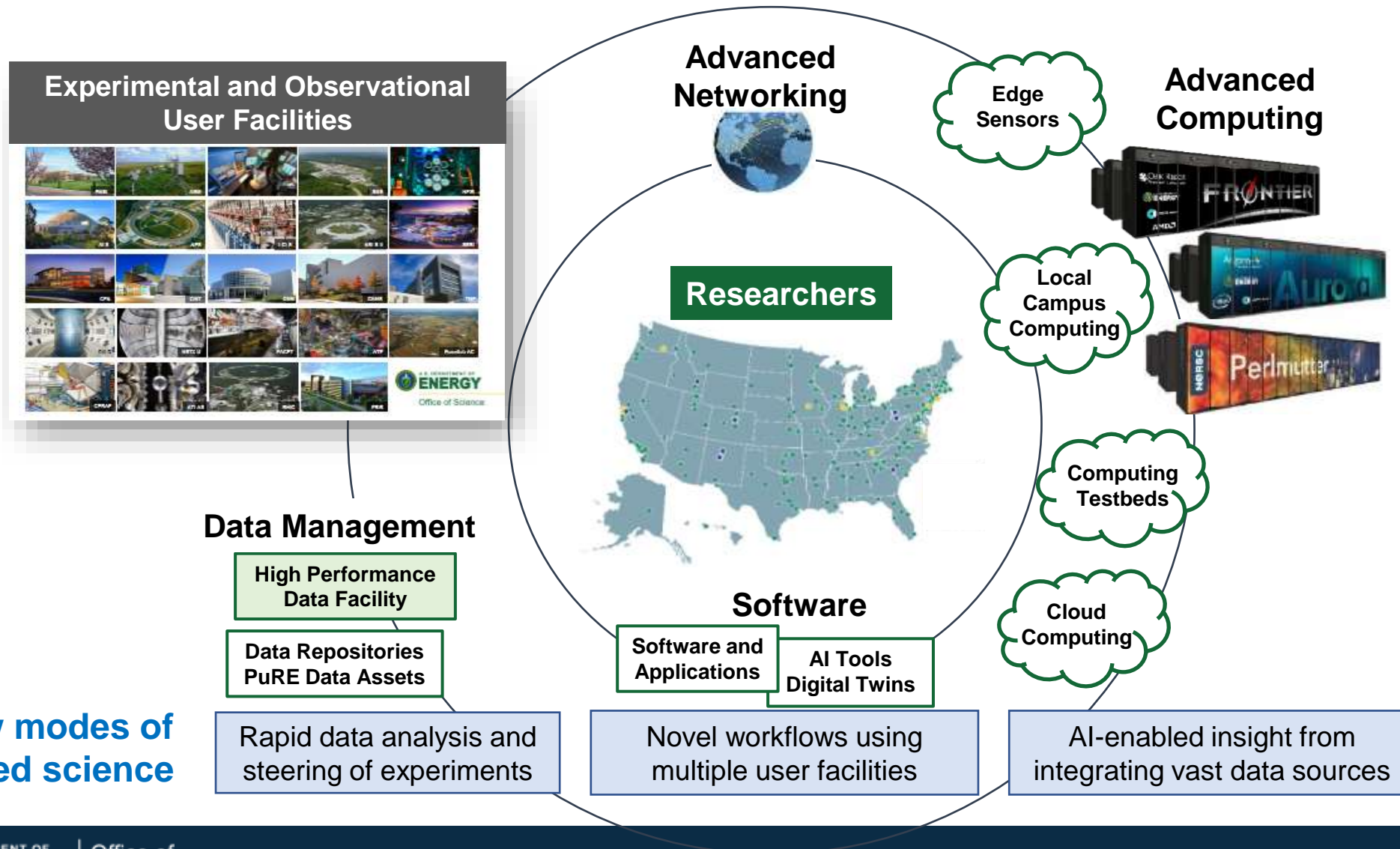
**~ 2021 National Strategic Overview of R&D Infrastructure**

## International Context

China Science and Technology Cloud  
European Open Science Cloud  
IRIS UKRI SFTC initiative

# DOE's Integrated Research Infrastructure (IRI) Vision:

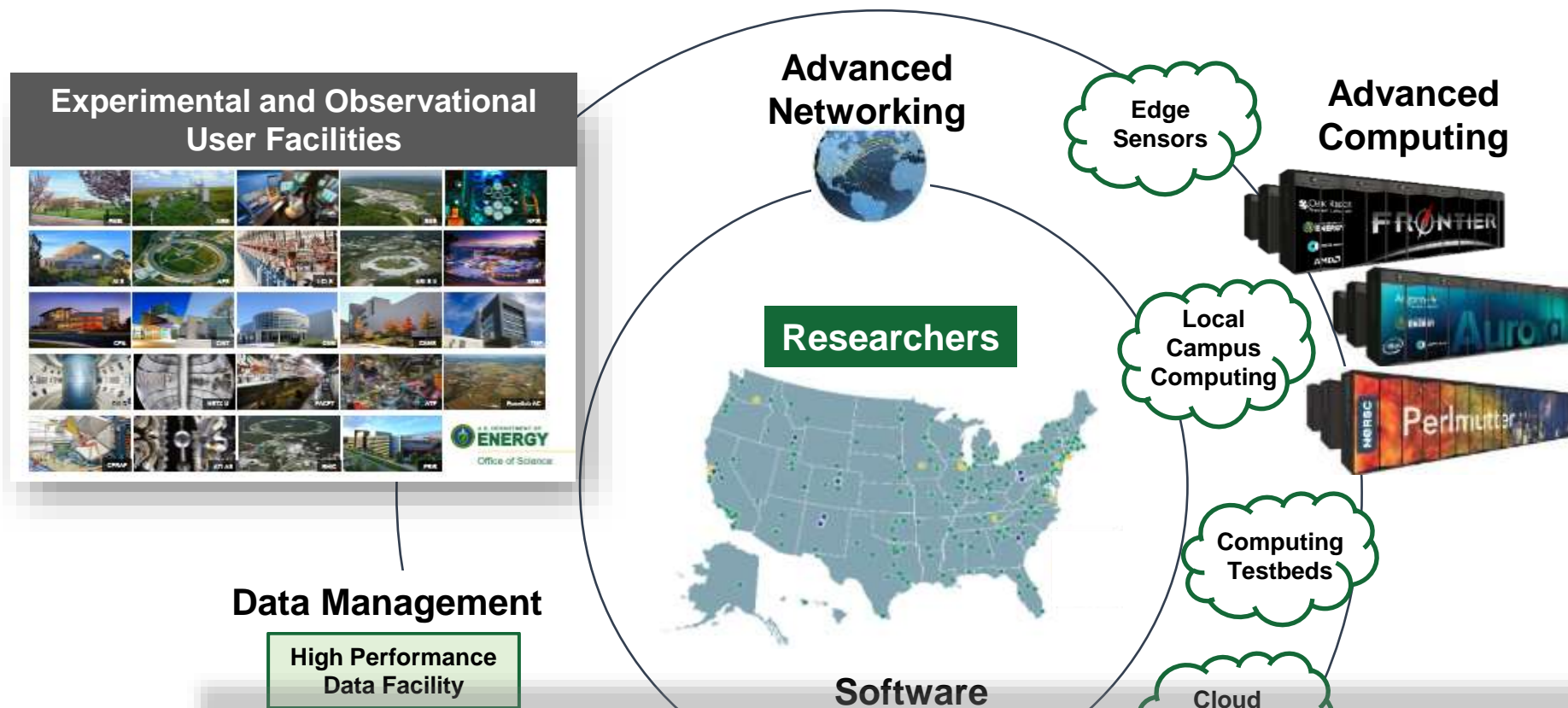
*To empower researchers to meld DOE's world-class research tools, infrastructure, and user facilities seamlessly and securely in novel ways to radically accelerate discovery and innovation*



**New modes of integrated science**

# DOE's Integrated Research Infrastructure (IRI) Vision:

*To empower researchers to meld DOE's world-class research tools, infrastructure, and user facilities seamlessly and securely in novel ways to radically accelerate discovery and innovation*



New modes of  
integrated science

**The IRI Vision:**  
**It's about empowering people.**  
**It's about data. It's about software.**



# Partnerships to Deliver Future Leaders

## *DOE Computational Science Graduate Fellowship (CSGF)*

- Started in 1991 to broadly train advanced computational scientists
- Funded by both DOE-SC/ASCR and NNSA/ASC
  - Currently, CSGF supports 99 students at 41 universities in 22 states.
  - More than 500 students at 65 U.S. universities have trained as fellows.
- Requires that fellows
  - plan and follow a plan of study that transcends the bounds of traditional academic disciplines
  - participate in 12-week research experience at DOE lab
- Benefits
  - Up to four years of support, including full tuition/required fees paid
  - Yearly stipend of \$38,000 plus an Academic allowance
  - Annual program review with peers, Alumni and DOE/Lab scientists

<https://www.krellinst.org/csgf/>

### *2019 incoming class of Computational Science Graduate Fellows*



*CSGF alumni work in DOE laboratories, industry and educational institutions*



# Finding Out More About ASCR – ASCAC

science.osti.gov/ascr/ascac/Meetings

## Meetings

- September 2022
- July 2022
- March 2022
- September 2021
- July 2021
- September 2020
- April 2020
- January 2020
- September 2019
- March 2019
- December 2018
- September 2018

## Meetings

### ASCR Advisory Committee Meetings

### ASCR ASCAC [YouTube Channel](#)

Like and subscribe all ASCAC meetings

### Next ASCAC Meeting

Public participants must identify themselves and their organizational affiliation to be admitted to

Friday, September

- [Agenda](#)
- [Presentations](#)

9:05 AM- 10:45 AM

View from GERMANTOWN, PA  
[Barbara Helland](#), Associate Director, Advanced Scientific Computing Research (ASCR)

10:30 AM- 11:15 AM

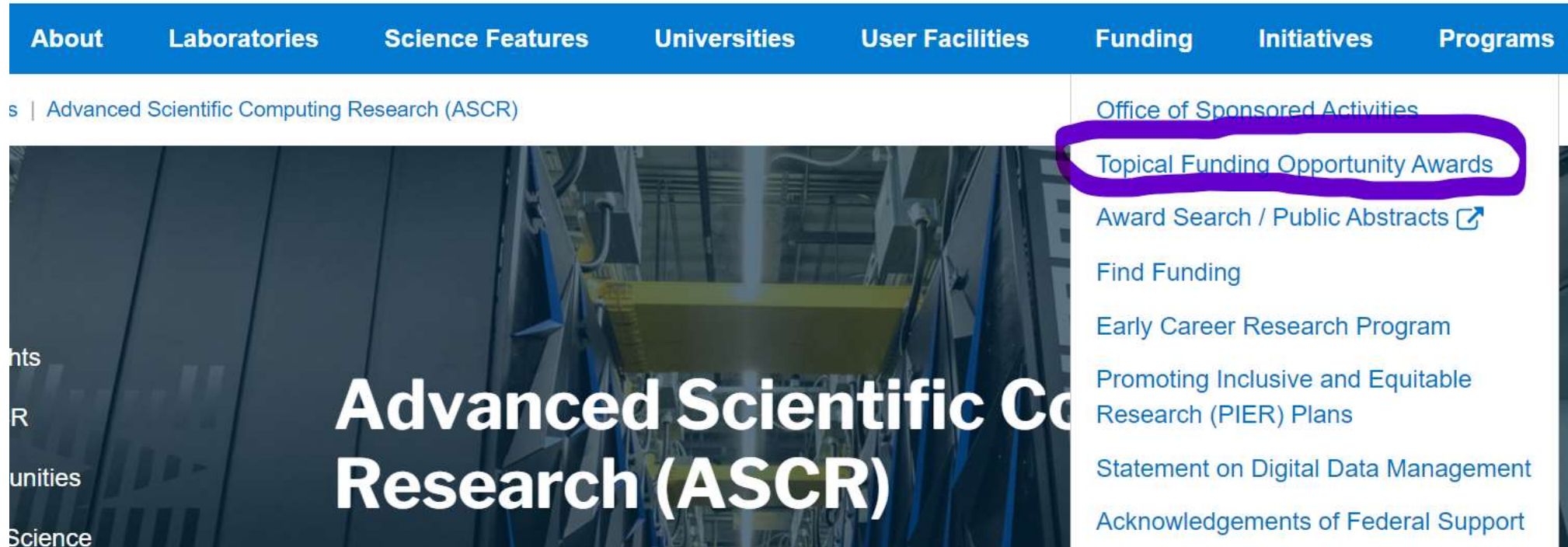
ASCR Research Priorities  
[Ceren Susut](#), Research Division Director, Advanced Scientific Computing Research

Presentation videos are available.

The presentations for each meeting are posted.

Look for presentations by program leadership for information on future priorities.

# Award Lists – A New Website Location



Award lists are now posted to <https://science.osti.gov/Funding-Opportunities/Award> along with other awards from the Office of Science. To receive award and solicitation announcements, and other ASCR-related news, signup for the Office of Science's GovDelivery email service, and check the box for the Advanced Scientific Computing Research Program in your subscriber preferences:

**Join Mailing List**

Signup for the Office of Science's GovDelivery email service, and check the box for the *Advanced Scientific Computing Research Program* in your subscriber preferences.

# ASCR Office Hours

- ◆ Starting in March, ASCR will hold virtual office hours on the second Tuesday of the month, 2 PM ET
- ◆ Researchers, educators, and leaders within research administration from all institutional types are encouraged to join
- ◆ A primary goal of the virtual office hours is to broaden awareness of our programs; no prior history of funding from DOE is required to join
- ◆ Program managers will be available to answer questions
- ◆ Upcoming topics include:
  - Tuesday, March 12, 2024 at 2pm ET - *Introduction to ASCR and its program mission and history*
  - Tuesday, April 9, 2024 at 2pm ET - *Introduction to ASCR's Computer Science research program*
  - Tuesday, May 14, 2024 at 2pm ET - *Introduction to ASCR's Applied Mathematics research program*

Check the ASCR website (<https://science.osti.gov/ascr/>) for Zoom registration links.

# A Selection of Highlights and Backup Slides

# Submodular Matchings for Balancing Data and Computations

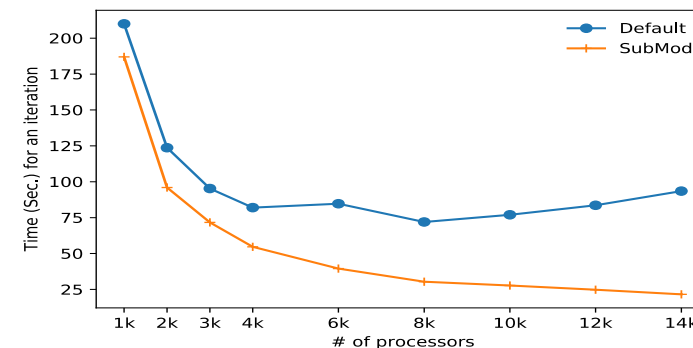
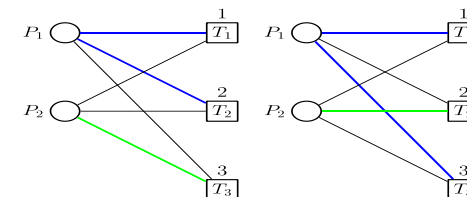
## A scalable parallel algorithm and a case study in Chemistry

### Scalable Quantum Chemistry via Submodular Matching

- Computing electronic properties of molecules via density functional theory involves the data intensive and compute intensive Fock matrix, whose elements consist of multidimensional integrals. The computation scales as  $O(n^4)$ , where  $n$  is the number of basis functions.
- We provide a scalable parallel algorithm for computing the Fock matrix within the NWChemEx software from Pacific Northwest National Lab.
- The algorithm assigns blocks of Fock submatrix computations to processors in order to balance the data and work load among the processors, and also the number of messages each processor is involved in.
- This is accomplished by computing a  $b$ -matching in the block-processor graph, with a nonlinear (submodular) objective function, to satisfy both objectives mentioned above.
- A submodular function balances the load on the processors, whereas a linear function cannot distinguish between unbalanced and balanced task assignments.
- Although the submodular  $b$ -matching problem is computationally intractable, we design fast approximation algorithms that provide constant-factor approximations to the optimal matching.

### Performance of NWChemEx on Summit

- We designed a submodular matching algorithm and incorporated it with the NWChemEx library.
- The code speeded up the Fock matrix computation for the ubiquitin protein molecule by a factor of four over the current task assignment.
- It also scaled the NWChemEx code to 14000 processors on Summit, from 4000 processors.
- More work could be done to reduce the size of the data even further by means of matrix factorizations.
- We collaborated with colleagues at PNNL from the ExaGraph and NWChemEX projects.



Top Fig. : A submodular matching balances the work in assigning tasks  $T$  to processors  $P$  (left), while a linear matching does not (right).

Bottom Fig.: Submodular assignment balances the load in computing the energy levels of the Ubiquitin protein, reducing the time on 14K Summit processors four-fold over the default.

PI: Alex Pothen

Collaborating Institutions: Purdue University, PNNL

ASCR Program: Computer Science

ASCR PM: Hal Finkel

Publication(s) for this work: S M Ferdous et al., "A parallel approximation algorithm for submodular  $b$ -matching," Proceedings SIAM Applied Computational Discrete Algorithms, (2021): pp. 45-56, . Doi:

[10.1137/1.9781611976830.5](https://doi.org/10.1137/1.9781611976830.5)

# Stochastic Learning for Binary Optimal Design of Experiments

## Scientific Achievement

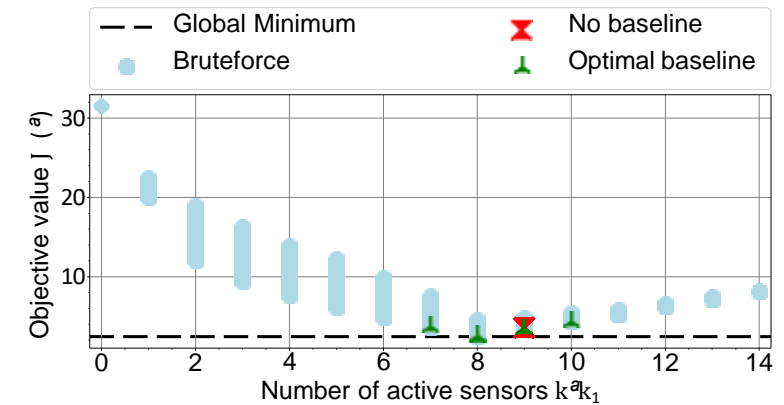
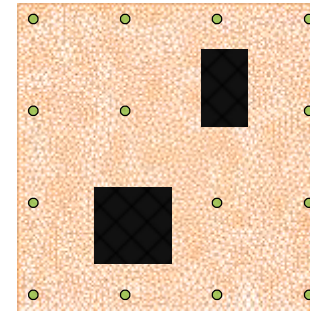
New stochastic approach to binary optimization for optimal experimental design (OED) for Bayesian inverse problems governed by mathematical models such as partial differential equations.

## Significance and Impact

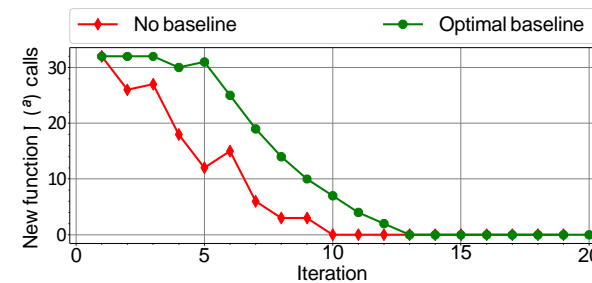
Binary OED problems are crucial for designing optimal data acquisition schemes, such as sensor placement or spatiotemporal data collection, in inverse problems in order to improve inversion accuracy (e.g., identifying the source of a contaminant) and long-term predictability of data assimilation systems.

## Research Details

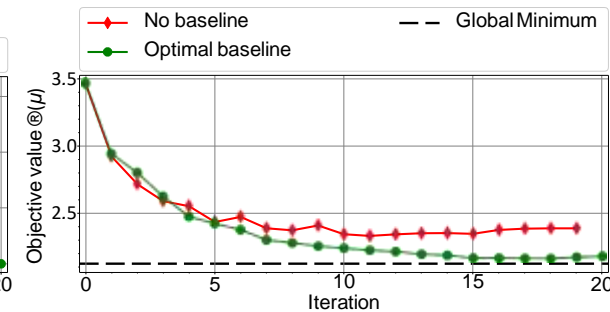
- We have developed a new probabilistic approach to efficiently solving binary OED optimization problems, without needing to relax the design, or carry out heuristic rounding techniques.
- The stochastic approach does not require differentiability of the utility function with respect to the design, and is directly interpretable:
  - Enable employment of sparsity-enforcing penalty functions such as  $l_1$ ,
  - Massively reduce the computational cost compared with traditional OED,
  - Sample efficient observational policies in a small number of optimization steps.
- Computationally efficient policy gradient (reinforcement learning) optimization algorithms, with convergence guarantees.



Sensor placement for parameter identification in an Advection-Diffusion experiment to locate contaminant source. Results of the policy gradient procedures, compared with the brute-force search of all candidate binary designs.



The value of the objective function at each iteration of the optimization procedures.



Number of new function evaluations at each optimization step

## References:

- [1] Attia, Ahmed, Sven Leyffer, Todd Munson. Stochastic Learning Approach for Binary Optimization: Application to Bayesian Optimal Design of Experiments. SIAM Journal on Scientific Computing, 2022.

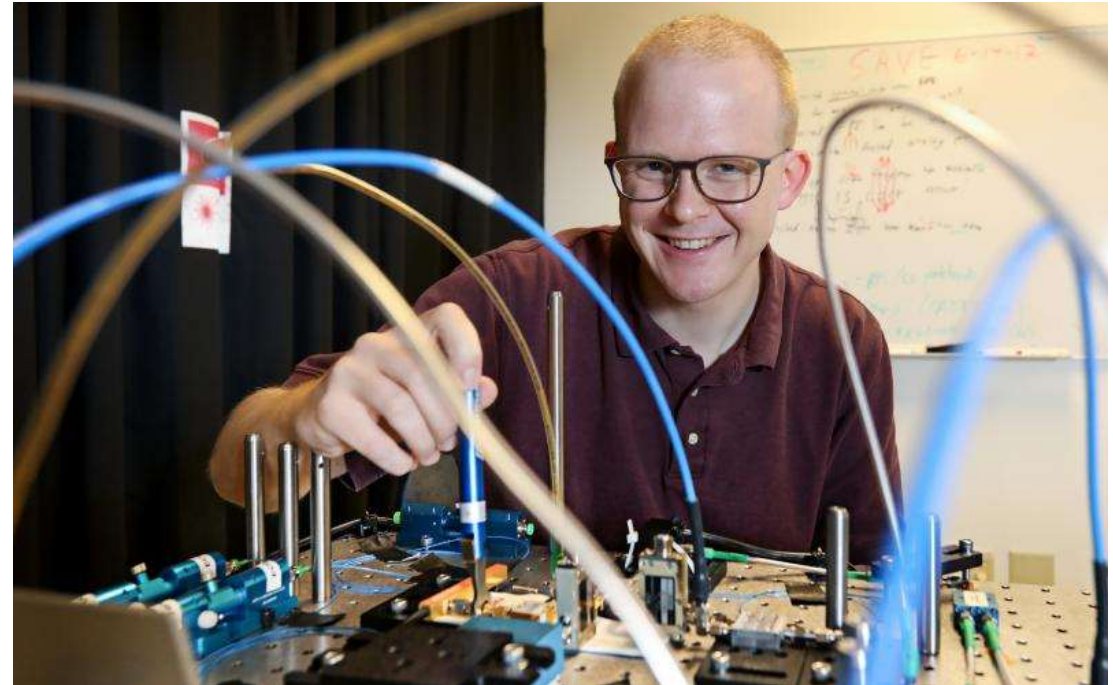
# Giant leap toward quantum internet realized with Bell state analyzer

## The Science

A multi-institutional team featuring ORNL's Joe Lukens has made strides toward a fully quantum internet by designing and demonstrating the first ever Bell state analyzer for frequency bin coding. Measuring Bell states is critical to performing many of the protocols necessary to perform quantum communication and distribute entanglement across a quantum network. The team's method represents the first Bell state analyzer developed specifically for frequency bin coding, a quantum communications method that harnesses single photons residing in two different frequencies simultaneously.

## The Impact

The analyzer was designed with simulations and has experimentally demonstrated 98% fidelity for distinguishing between two distinct frequency bin Bell states. This incredible accuracy is expected to enable new fundamental communication protocols necessary for frequency bins.



ORNL's Joseph Lukens runs experiments in an optics lab. Credit: Jason Richards/ORNL, U.S. Dept. of Energy

PI(s)/Facility Lead(s): Joe Lukens (ORNL)

ASCR Program/Facility: N/A

ASCR PM: Lali Chatterjee

Funding: Office of Science through the Early Career Research Program

Publication for this work: Navin B. Lingaraju, Hsuan-Hao Lu, Daniel E. Leaird, Steven Estrella, Joseph M. Lukens, and Andrew M. Weiner. "Bell state analyzer for spectrally distinct photons," *Optica* Vol. 9, Issue 3, pp. 280-283 (2022).

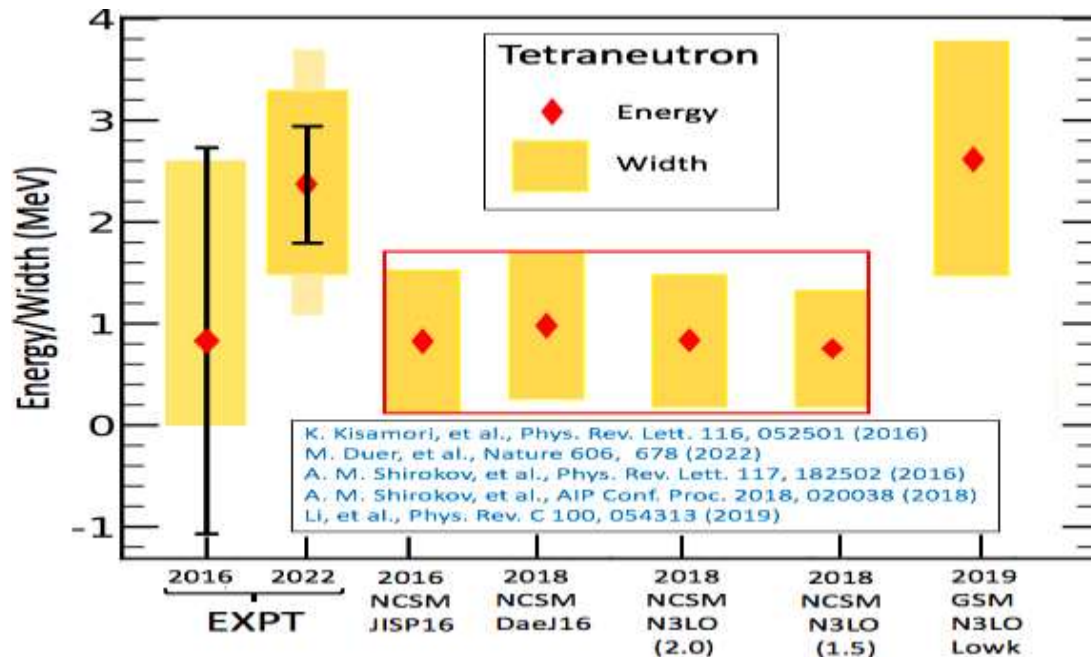
Date submitted to ASCR: Spring 2022



# Tetraneutron Discovery Confirms Prediction

## Objectives

- *Ab initio* nuclear theory aims for parameter-free predictions of nuclear properties with controlled uncertainties using supercomputer simulations.
- Specific goal is to predict if the tetraneutron (4-neutron system) has a bound state, a low-lying resonance or neither



Experiment and theory for the tetraneutron's resonance energy and width. *Ab initio* No-Core Shell Model (NCSM) and Gamow Shell Model (GSM) predictions use different neutron-neutron interactions and different basis function techniques.

## Impact

- Discovery announced in Nature [1] confirms *ab initio* theory predictions from 2016 [2] of a short-lived tetraneutron resonance at low energy and the absence of a tetraneutron bound state
- Demonstrates the predictive power of *ab initio* nuclear theory since theory and experiment are within their combined uncertainties
- Sets stage for further experimental and theoretical research on new states of matter formed only of neutrons
- Shows need to anticipate a long wait time for experimental confirmation of such an exotic phenomena, ~ 6 years in this case
- Emphasizes the value of DOE supercomputer allocations (NERSC) and support for multi-disciplinary teamwork (SciDAC/NUCLEI)

## Publications

- [1] M. Duer, et al., Nature 606, 678 (2022)
- [2] A.M. Shirokov, G. Papadimitriou, A.I. Mazur, I.A. Mazur, R. Roth and J.P. Vary, "Prediction for a four-neutron resonance," Phys. Rev. Lett. 117, 182502 (2016)

# Scalable Transformers on Frontier for Real-Time Experiment Steering

## Scientific Achievement

ORNL developed a scalable transformer on OLCF Frontier for real-time decision-making in neutron diffraction experiments at the TOPAZ beamline of SNS. This work:

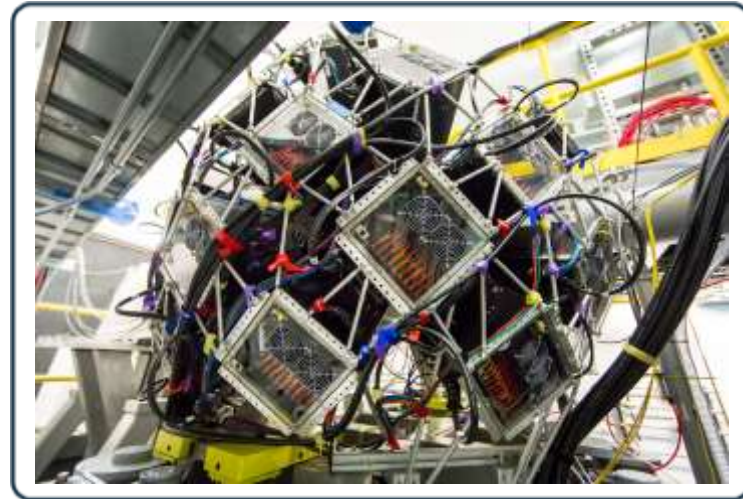
- Develops a stochastic process model for the time-of-flight neutron scattering data and exploits a temporal fusion transformer to **help reduce the experiment time**.
- Demonstrates outstanding scalability of the ML model on Frontier, which is necessary to synchronize neutron diffraction experiments, data analysis, and decision making.

## Significance and Impact

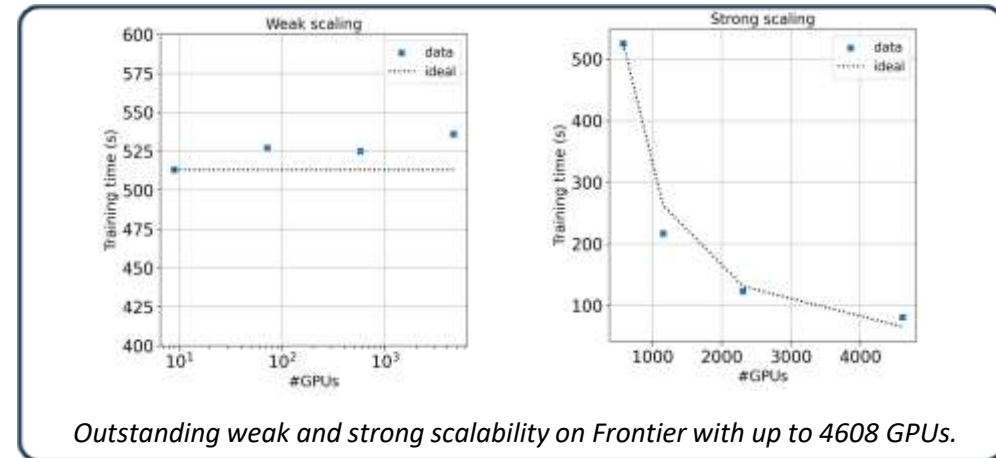
- The ML algorithm could help neutron scientists to **reduce the over-counting beamtime by around 30%** at TOPAZ, while achieving the similar data quality.
- This effort proves the concept of connecting BES's neutron facilities and ASCR's HPC facilities through AI/ML, **forming an integrated research infrastructure**.

## Technical Approach

- The developed stochastic process model provides a novel and effective approach to describe the time-of-flight neutron scattering data.
- The hierarchical parallelization approach effectively uses ~60% of Frontier's computing power to keep up with the neutron experiment speed.



*A single-crystal diffractometer on the TOPAZ beamline at SNS*



PI : Guannan Zhang (ORNL); ASCR Program: Data-Intensive Scientific Machine Learning and Analysis; ASCR PM: Steve Lee  
Publication: J. Yin, S. Liu, V. Reshniak, X. Wang, and G. Zhang, *A scalable transformer model for real-time decision making in neutron scattering experiments*, *Journal of Machine Learning for Modelling and Computing*, Vol 4 (1), pp. 95-107, 2023

# SuperNeuro: An Accelerated Neuromorphic Computing Simulator

## Scientific Achievement

ORNL scientists have developed SuperNeuro, the world's fastest simulator for neuromorphic computing. It was designed for speed and scalability, and is capable of running **300 times faster** than its competitors, garnering the team the **2023 R&D 100 Award in the Software/Services Category**.

## Significance and Impact

Neuromorphic architectures have the potential to increase computing power and efficiency, as well as advance AI applications. SuperNeuro provides an indispensable capability for this effort via the leveraging of GPU computing to provide superior performance for neuroscience, increased adaptability, spiking neural networks (SNNs), and general-purpose computing workloads.

## Technical Approach

Two novel approaches used: matrix computation (MAT) and agent-based modeling (ABM).

- MAT Mode: Homogeneous simulations, built-in learning, CPU execution
- ABM Mode: Heterogeneous simulations, GPU acceleration

PI(s): Prasanna Date, Chathika Gunaratne, Shruti Kulkarni, Robert Patton, Mark Coletti, and Thomas Potok

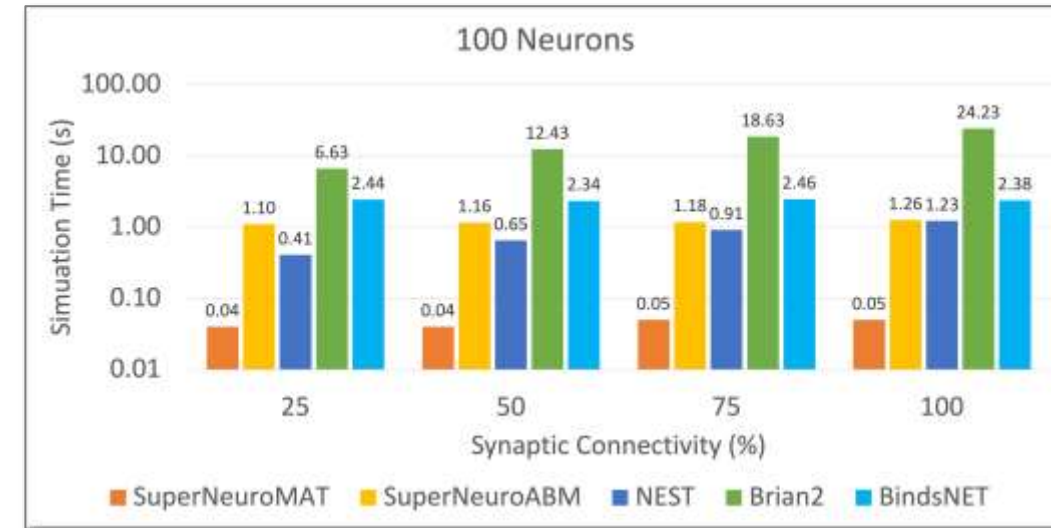
Collaborating Institutions: Oak Ridge National Laboratory

ASCR Program: Neuromorphic Computing for Accelerating Scientific Discovery

ASCR PM: Robinson Pino

Publication(s) for this work: Date, Prasanna, Chathika Gunaratne, Shruti R. Kulkarni, Robert Patton, Mark Coletti, and Thomas Potok.

"SuperNeuro: A Fast and Scalable Simulator for Neuromorphic Computing." In Proceedings of the 2023 International Conference on Neuromorphic Systems, pp. 1-4. 2023.



*Simulating 100 neurons on 5 neuromorphic simulators with 4 different synaptic connectivities. SuperNeuroMAT performs 300 times faster than other neuromorphic simulators.*

<https://github.com/ORNL/superneuromat>



# Privacy-Preserving Federated Learning as a Service using APPFL

## Scientific Achievement

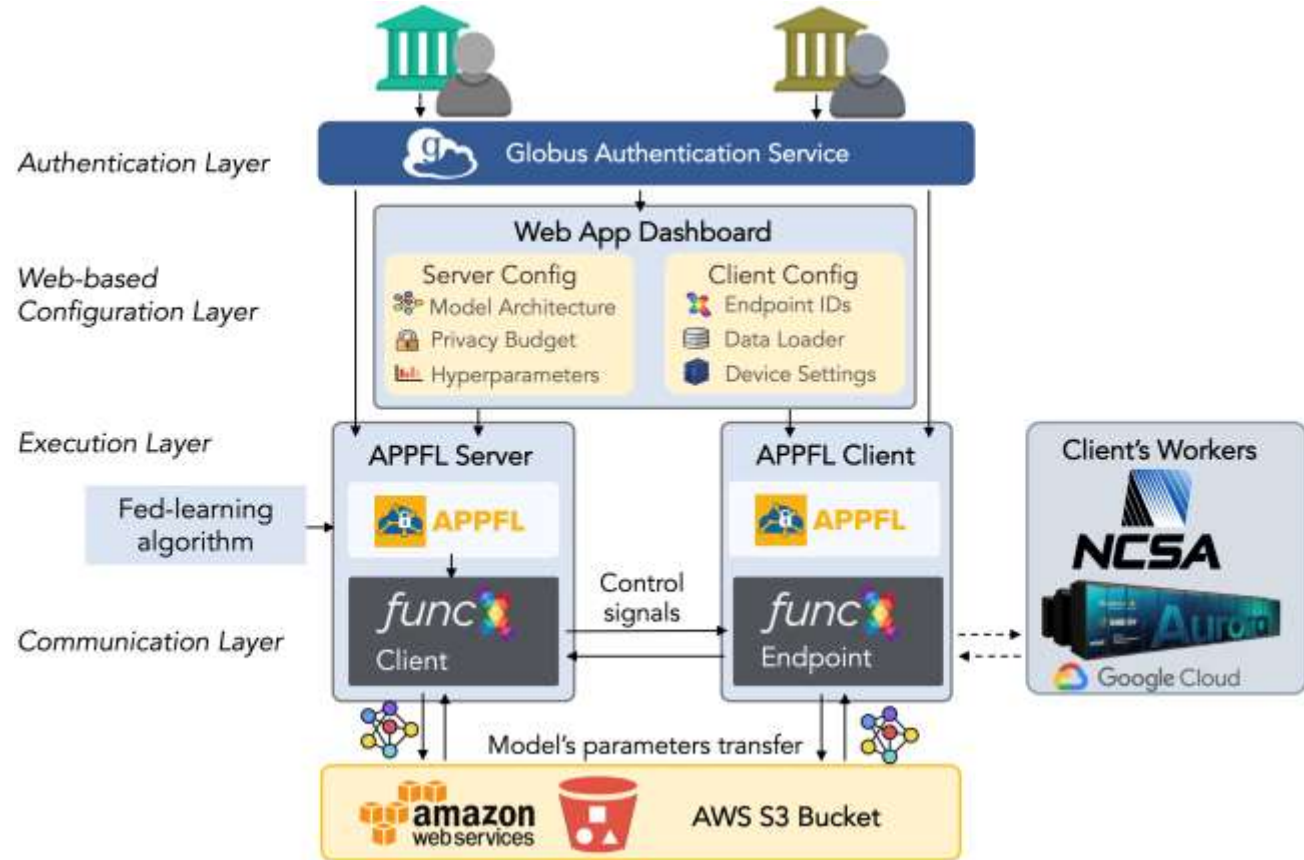
APPFL as a Service (APPFLaaS) enables end-to-end secure and privacy-preserving federated learning. Using APPFL and Globus services, the service provides supervised learning of a model on distributed sensitive datasets while preserving data privacy.

## Significance and Impact

APPFLaaS will enable secure collaborations across countries and institutions while addressing the privacy and data shift challenges in many DOE applications (e.g., scientific machine learning, critical infrastructure) leading to fair and trust-worthy AI models

## Research Details

- Integration with Globus Auth and Compute enables secure access controls and integration with heterogenous compute resources
- Novel distributed optimization algorithms with differential privacy result in better convergence and learning performance
- In collaboration with medical institutions, APPFL is used to train various ML models for disease prognosis, diagnosis and treatment planning
- APPFL used for federated control of power system operations maintaining data privacy against an adversary
- APPFLaaS provides comprehensive report for each federation learning experiment including training logs, hyperparameters, validation results, training metrics and Tensorboard visualization



PI(s): Ravi Madduri and Kibaek Kim; Argonne National Laboratory  
ASCR Program: Bridge2AI And Privacy-Preserving Artificial Intelligence Research  
ASCR PM: Steven Lee  
Publication(s) for this work: Ryu, Kim, Kim, Madduri. "APPFL: Open-Source Software Framework for Privacy-Preserving Federated Learning" 2022 IEEE IPDPS Workshop



<https://github.com/APPFL/APPFL>

# GenSLM: Genome-scale Language Models for predicting evolutionary dynamics of SARS-CoV-2



## Scientific Achievement

Developed foundation models for genome-scale datasets to predict variants of concern for SARS-CoV-2; awarded the IEEE/ACM Gordon Bell Prize for HPC in COVID-19 research (2022)

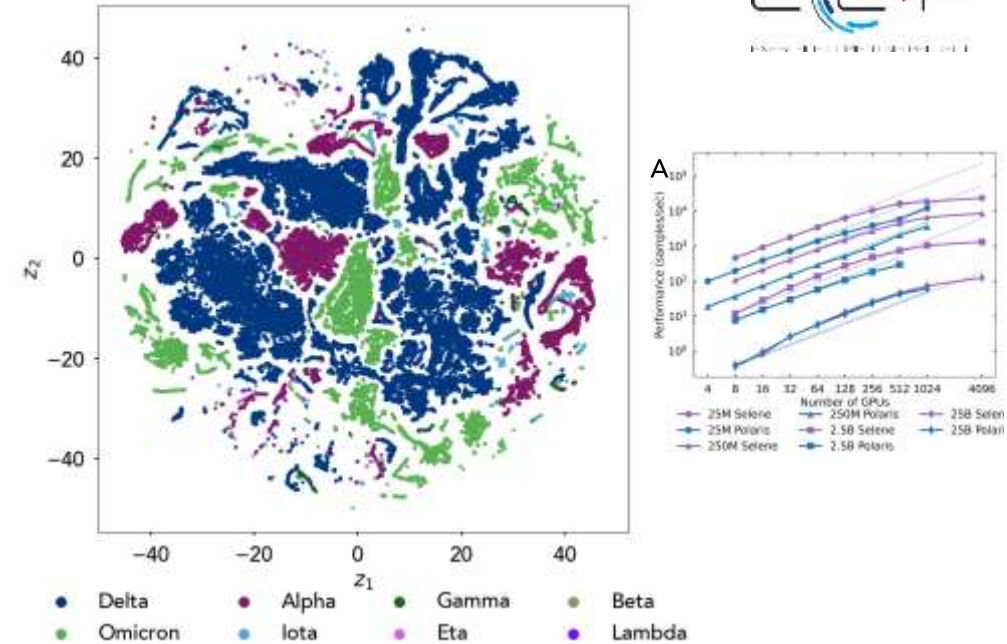
## Significance and Impact

Demonstrated ability to scale generalized pre-trained transformer architectures for nucleotide data using both ALCF Polaris and Cerebras architectures, achieving near perfect scaling for large-scale AI models

## Research Details

- GenSLM models build on codon-level representations to enable biological interpretation and include a hierarchical diffusion model to infer longer range context from genomic data
- An end-to-end workflow that integrates both genomic and biophysical data to identify SARS-CoV-2 variants of concern
- GenSLM models include models with 25 million to 25 billion parameters trained on >110 million gene sequences (leveraging openly available data such as K-Base)
- On Cerebras architecture, we achieve linear speedups with model training convergence < 0.5 days
- Performed a total of 1.64 Zettaflops for training, with sustained computational rates of 121 PFLOPS and a peak of 850 PFLOPS in mixed precision
- Generated sequences from GenSLMs achieve similarity to the recently found BQ.1 variant, implying deployability of GenSLMs for real use-cases

<https://www.acm.org/media-center/2022/november/gordon-bell-special-prize-covid-research-2022>



Despite being trained only on one year's worth of data (Alpha and Beta variants), GenSLM can correctly identify all variants in the subsequent data, demonstrating its ability to correctly classify VOCs.

	GenSLM 123M		GenSLM 1.3B	
	1 CS-2	4 CS-2	1 CS-2	4CS-2
Training steps	5,000	3,000	4,500	3,000
Training samples	165,000	396,000	49,500	132,000
<b>Time to train (h)</b>	<b>4.1</b>	<b>2.4</b>	<b>15.6</b>	<b>10.4</b>
Validation accuracy	0.9615	0.9625	0.9622	0.9947
Validation perplexity	1.031	1.029	1.031	1.025

Contact: Arvind Ramanathan ([ramanathana@anl.gov](mailto:ramanathana@anl.gov)), Venkatram Vishwanath ([Venkat@anl.gov](mailto:Venkat@anl.gov))

# Dehallucination of LLMs for High-Level Planning

## Scientific Achievement

- Large language models can generate plans for solving high-level planning problems, such as the operation of robots in DOE national laboratories.
- While the plans may appear to be of high quality, it is not uncommon for the produced plans to contain actions that cannot be executed in reality.
- We have developed a framework that mitigates hallucinations (generated errors) in LLM generated plans.

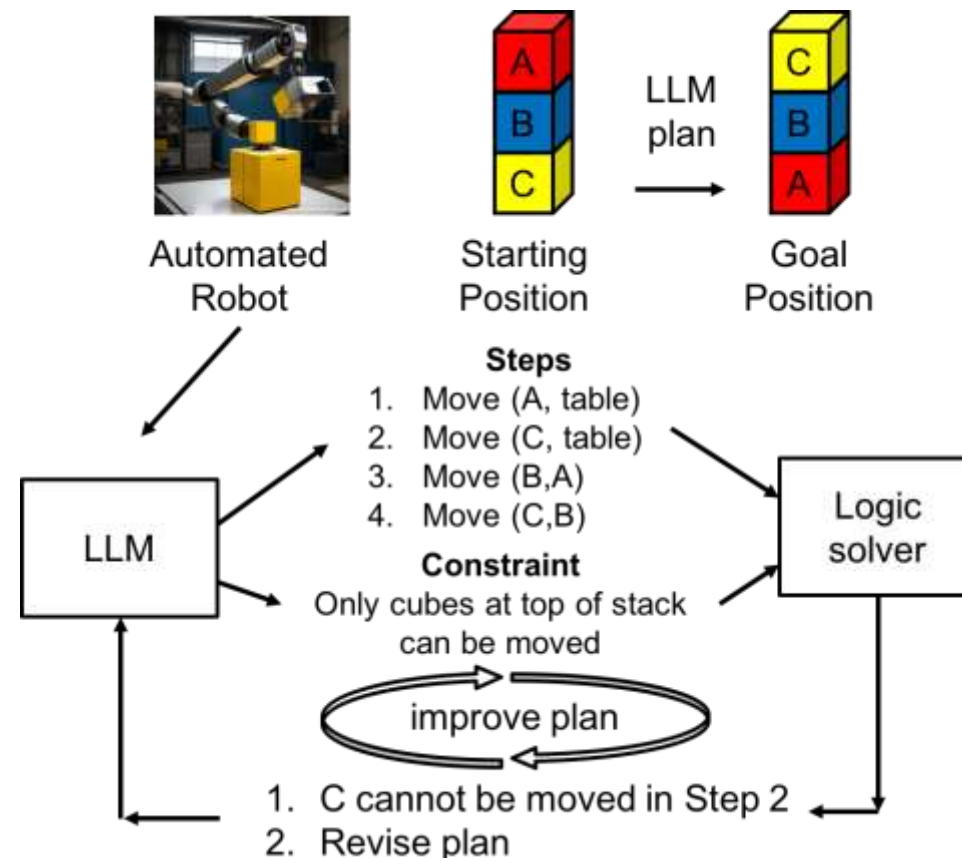
## Significance and Impact

The project provides a solution to specifying scientific problems in natural languages (or text) while solving them using neuro symbolic methods. This is a step towards lowering technical barriers for future engineers and scientists.

## Technical Approach

- The code generation capabilities of the LLM is used to specify logical constraints that every generated plan must satisfy.
- A solver is used to automatically check the adherence to the constraints and provide feedback to the AI model regarding unsatisfied constraints.
- The feedback allows the LLM to generate a new provably correct plan.

PI(s)/Facility Lead(s): Rickard Ewetz, Sumit Kumar Jha; University of Central Florida  
Collaborating Institutions: Florida International University  
ASCR Program: EXPRESS, Explainable AI  
ASCR PM: Margaret Lentz  
Publication(s) for this work: S. Jha, et al., "Counterexample Guided Inductive Synthesis Using Large Language Models and Satisfiability Solving," MILCOM, November, (2023). (to appear).



The LLM generates a high-level plan for moving the starting position to the goal position. The LLM also generates mathematical constraints describing how cubes are allowed to be moved. The plan and the constraints are fed into a logic solver, which determines that a constraint is violated in step 2. The C cube is attempted to be moved while cube B is on top. The solver provides feedback to the LLM why the plan is infeasible such that a new legal plan can be generated.

# ECP Industry and Agency Council Members



# The breadth of exascale-ready applications is remarkable; indicative of a sea change in computing abilities for DOE and the nation

