Michael E. Papka
Division Director, Argonne Leadership Computing Facility
Deputy Associate Laboratory Director, Computing, Environment and Life Sciences, Argonne National Laboratory

Research Professor of Computer Science, Illinois Institute of Technology

Argonne National Laboratory

The U.S. Department of Energy's Argonne National Laboratory delivers world-class research, technologies, and new knowledge that aim to make an impact — from the atomic to the human to the global scale.

About Argonne

Argonne is a multidisciplinary science and engineering research center located outside Chicago.

- Born out of the University of Chicago's work on the Manhattan Project in the 1940s.
- Managed by UChicago Argonne, LLC, for the U.S. Department of Energy's Office of Science.
- Works with universities, industry, and other national labs on questions and experiments too large for any one institution to do by itself.

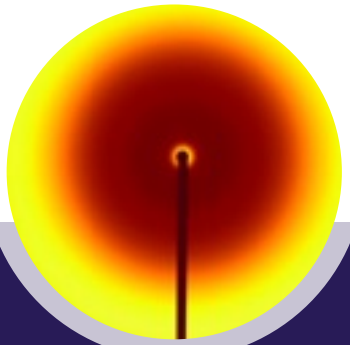
Argonne's Research Directorates

Computing, Environment and Life Sciences

Couples computing-related activities with science domains whose futures are closely tied to progress in computing

Energy and Global Security

Conducts applied R&D, creates tools that enable scientific and technological breakthroughs, and translates discoveries through engineering to the marketplace



Photon Sciences

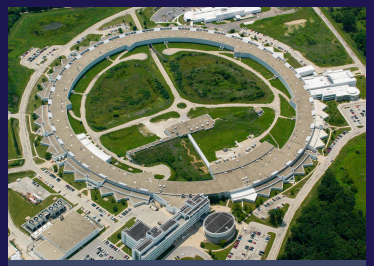
Provides the brightest X-ray beams in the Western Hemisphere and provides discoveries in nearly every scientific discipline

Physical Sciences and Engineering

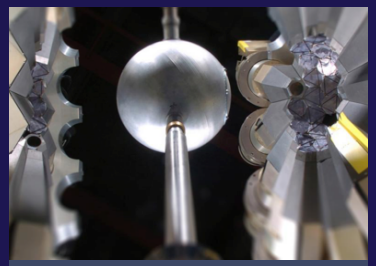
Creates new materials and chemistries and advances accelerator physics

Our one-of-a-kind facilities enable science from the nanoscale to the exascale

Argonne's five flagship facilities support one of the largest user communities in the U.S. Department of Energy complex.



Advanced Photon Source



Argonne Tandem Linear Accelerator System

Argonne Leadership Computing Facility

Center for Nanoscale Materials

Atmospheric Radiation Measurement – The Southern Great Plains

DOE Leadership Computing Facility

- Established in 2004 as a collaborative, multi-lab initiative funded by DOE's ***Advanced Scientific Computing Research*** program
- Operates as **one facility** with two centers, at Argonne and at Oak Ridge National Laboratory
- Deploys and operates at least two advanced architectures that are **10-100 times more powerful** than systems typically available for open scientific research
- **Fully dedicated** to open science to address the ever-growing needs of the scientific community

Broad Engagement in HPC

We enable and support science campaigns

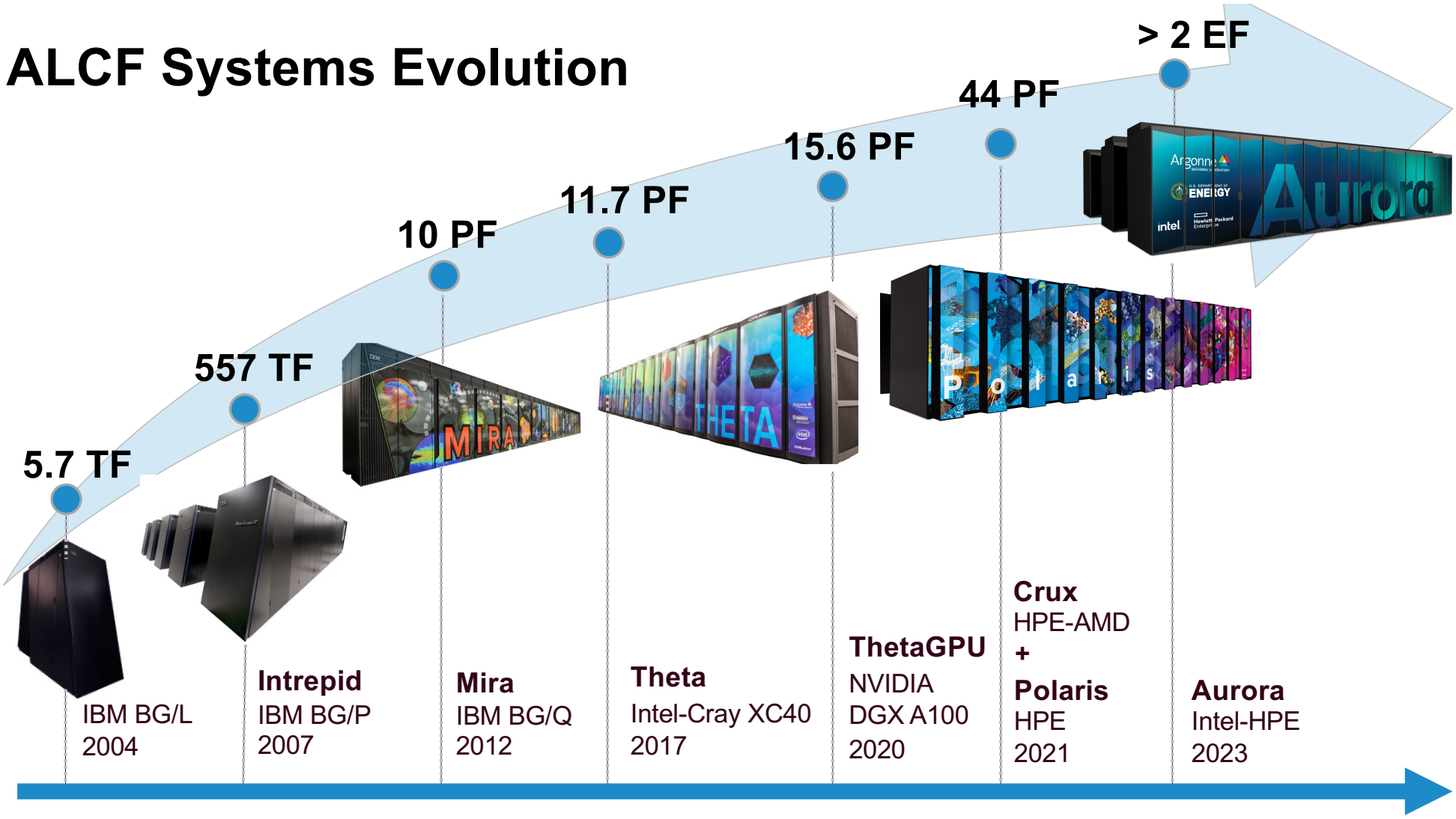
- ALCF computational scientists assist science teams to ready their codes to efficiently use our resources
- Researchers are supported by performance engineers, user support staff, and data analysis and visualization services
- Each year, ALCF-supported research results in hundreds of refereed publications, in journals such as Proceedings of the National Academy of Sciences, Nature, and Physical Review Letters

We deliver cycles to computational scientists

- Delivers millions of node hours of compute time
- Scheduled availability for the resource exceeds 99%

We partner with community on R&D in hardware and software

ALCF Systems Evolution



Argonne NATIONAL LABORATORY

Aurora

Argonne's upcoming exascale supercomputer will leverage several technological innovations to support machine learning and data science workloads alongside traditional modeling and simulation runs.

SUSTAINED PERFORMANCE
≥1 Exaflop DP

Xᵉ ARCHITECTURE-BASED GPU
Ponte Vecchio

INTEL XEON SCALABLE PROCESSOR
Sapphire Rapids

PLATFORM
HPE Cray EX

Compute Node
2 Intel Xeon scalable "Sapphire Rapids" processors; 6 Xᵉ arch-based GPUs; Unified Memory Architecture; 8 fabric endpoints; RAMBO

GPU Architecture
Xᵉ arch-based "Ponte Vecchio" GPU; Tile-based chiplets, HBM stack, Foveros 3D integration, 7nm

CPU-GPU Interconnect
CPU-GPU: PCIe
GPU-GPU: Xᵉ Link

System Interconnect
HPE Slingshot 11; Dragonfly topology with adaptive routing

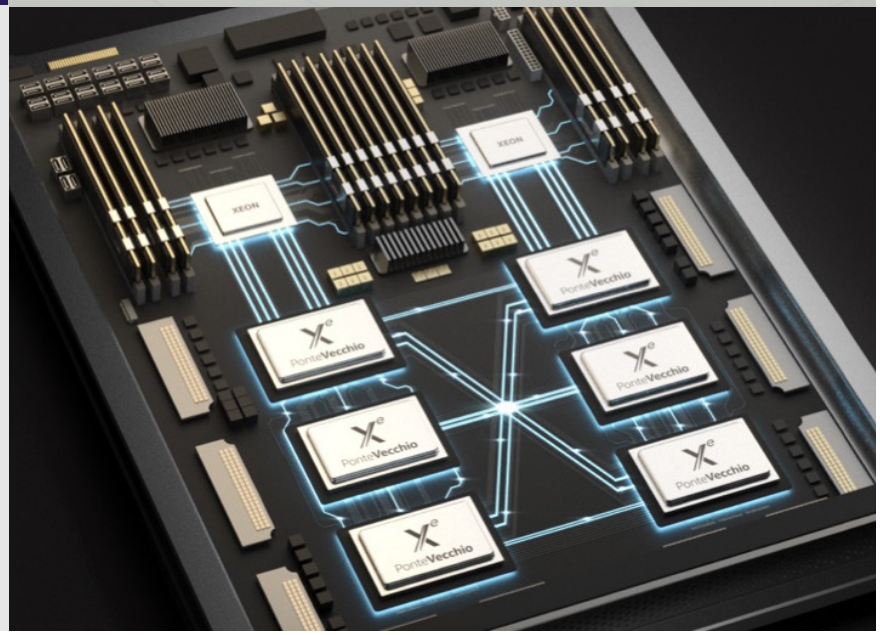
Network Switch
25.6 Tb/s per switch, from 64–200 Gbs ports (25 GB/s per direction)

High-Performance Storage
≥230 PB, ≥25 TB/s (DAOS)

Programming Models
Intel oneAPI, MPI, OpenMP, C/C++, Fortran, SYCL/DPC++

Node Performance
>130 TF

System Size
>9,000 nodes



Computing Resources

Polaris
- HPE Apollo Gen10+
- AMD processors/NVIDIA GPUs
- 44 petaflops (double precision)
- NVIDIA GPU A100; HBM stack
- AMD EPYC Processor Milan
- 560 nodes

Theta

KNL NODES
- Intel-Cray XC40
- 11.7 petaflops
- 4,392 nodes
- 281,088 cores
- 843 TB (DDR4); 70 TB (HBM) of memory

GPU NODES
- NVIDIA DGX A100
- 3.9 petaflops
- AMD EPYC 7742
- 24 nodes
- 24 TB of DDR4; 7.7 TB (HBM) of memory

Cooley
- Cray/NVIDIA
  126 nodes
- 1512 Intel Haswell CPU cores
- 126 NVIDIA Tesla K80 GPUs
- 48 TB RAM / 3 TB GPU

Iota
- Intel/Cray XC40 architecture
- 117 teraflops
- 44 nodes
- 2,816 cores
- 12.3 TB of memory

JLSE Experimental Testbeds
- 150 nodes
- Intel/AMD/IBM/Marvell/GPGPU
- EDR/100GbE/OPA
- Lustre/GPFS/DAOS

Grand and Eagle (Storage)
Each system has:
- HPE ClusterStor E1000
- 100 petabytes of usable capacity
- 8,480 disk drives
- Lustre filesystem
  - 160 Object Storage Targets
  - 40 Metadata Targets
- HDR InfiniBand network
- 650 GB/s rate on data transfers

ALCF AI-TESTBED

Next-Generation AI-Accelerator Systems

- Infrastructure of next-generation machines with hardware accelerators customized for artificial intelligence (AI) applications with a goal to integrate AI accelerators in existing and upcoming supercomputers
- Provides a platform to evaluate usability and performance of machine learning-based HPC science applications running on these accelerators.
- Promising results for diverse spectrum of science ranging from cancer , covid19, high-energy physics, biosciences, climate, among others.
- Close collaboration with AI accelerator vendors on their product developments and roadmaps

Cerebras (CS-2)

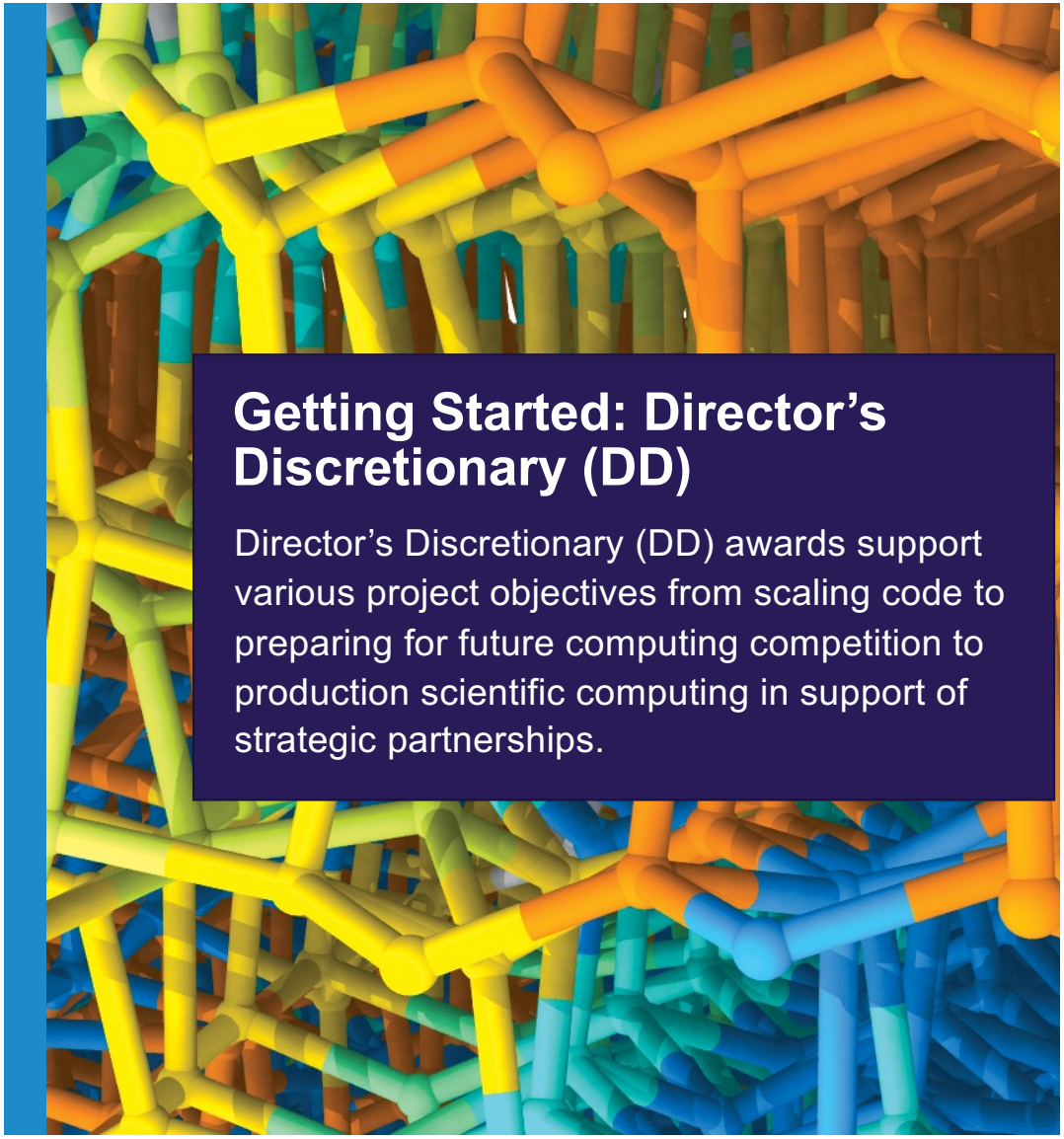
SambaNova

Graphcore

Habana

Groq

https://www.alcf.anl.gov/alcf-ai-testbed



Getting Started: Director's Discretionary (DD)

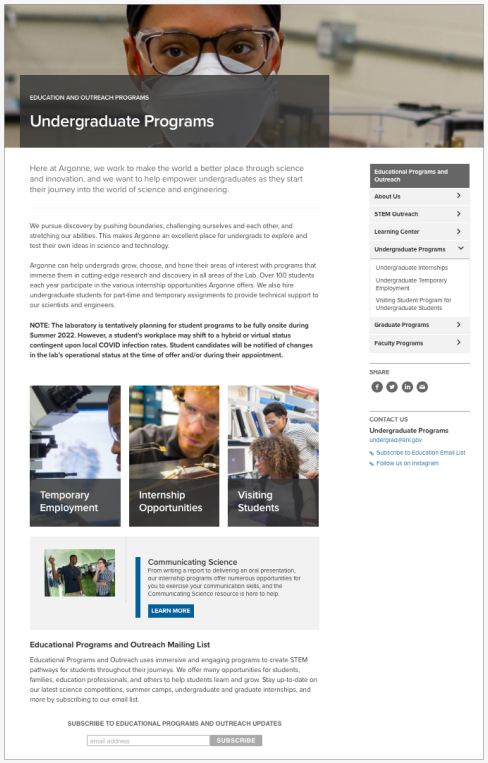
Director's Discretionary (DD) awards support various project objectives from scaling code to preparing for future computing competition to production scientific computing in support of strategic partnerships.

- **Award Cycle:** Ongoing (available year round)
- **Award size:** Up to several million compute-hours
- **Duration:** 3–6 months; renewable
- **Total percent of ALCF** resources allocated for DD projects: 20%

https://www.alcf.anl.gov/science/directors-discretionary-allocation-program

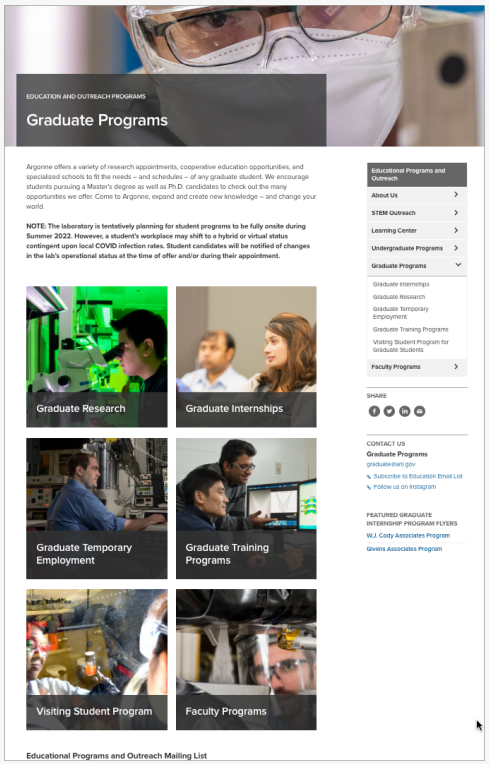
Opportunities at Argonne

Undergraduate



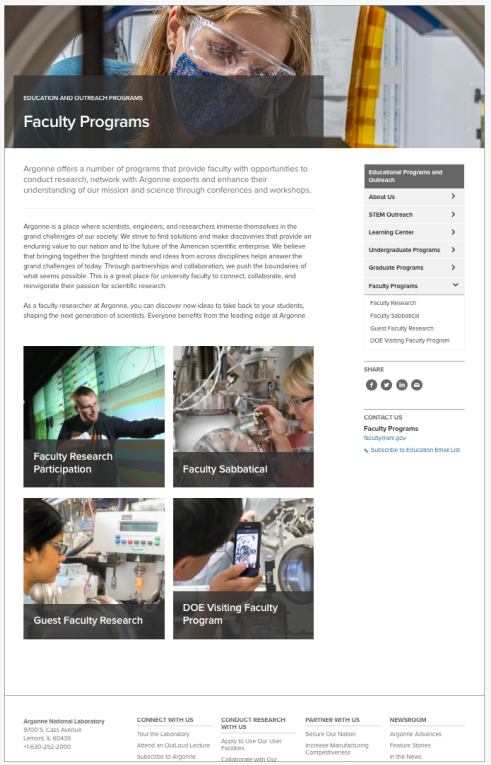
https://www.anl.gov/education/undergraduate-programs

Graduate



https://www.anl.gov/education/graduate-programs

Faculty



https://www.anl.gov/education/faculty-programs

Argonne NATIONAL LABORATORY

Computing@Argonne: Research

ALCF + CPS + DSL + MCS

Valerie Taylor
Division Director, Mathematics and Computer Science Division
Argonne Distinguished Fellow
Argonne National Laboratory

Computing@Argonne

Valerie Taylor
Director, Mathematics and Computer Science Division (**MCS**)

Ian Foster
Director, Data Science and Learning Division (**DSL**)

Salman Habib
Director, Computational Science Division (**CPS**)

Mike Papka
Director, Argonne Leadership Computing Facility Division (**ALCF**)

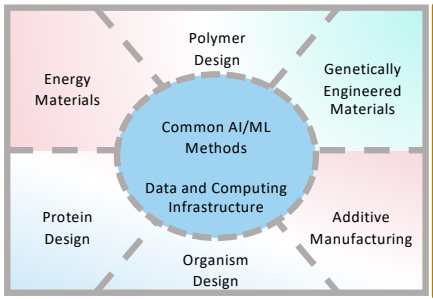
Future Drivers

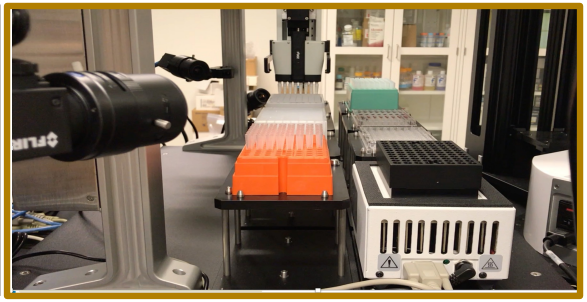
AI-Driven Modes of Discovery

- **Learned Models Begin to Replace Data**
  - — Queryable, portable, pluggable, chainable, secure
- **Experimental Discovery Processes Dramatically Refactored**
  - — ML models drive processes; experiments improve ML models
- **Many Questions Pursued Semi-Autonomously at Scale**
  - — Searching for materials, molecules and pathways
- **Simulation and AI Approaches Merge**
  - — Deep integration of ML, numerical simulation and UQ
- **Theory Becomes Data for Next Generation AI**
  - — AI begins to contribute to explaining the how and why of phenomena
- **AI Becomes Common Part of DOE Laboratory Activities**
  - — Infuses scientific, engineering and operations

Self-Driving Laboratories

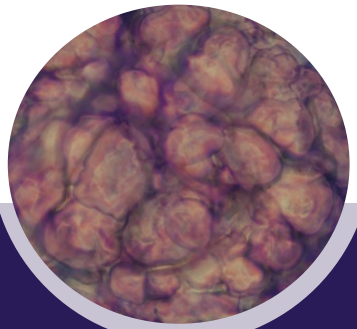
- Are emerging as the **next-generation facilities** for an accelerated scientific discovery process.
- They augment automated experimentation platforms with artificial intelligence to enable autonomous experimentation.





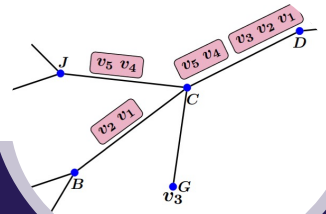
Computing Continuum

Research Capabilities: Applied Math



PDEs and PDE Solvers

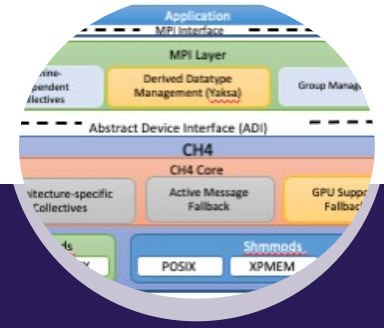
Develops, analyzes, and implements algorithms and software for obtaining numerical solutions to science problems that can be modeled as PDEs



Mathematical Optimization

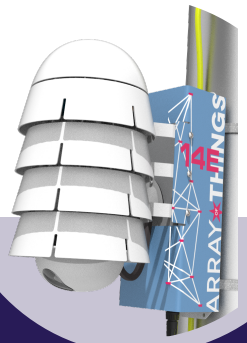
Develops the theory, algorithms, and software for minimizing some function subject to a set of constraints.

Research Capabilities: CS, Visualization and Data



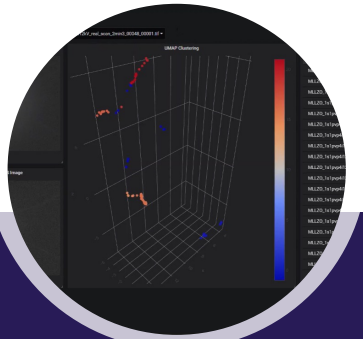
Operating Systems, Programming Models, and Compilers

Develop software that makes computational approaches to DOE science possible



Advanced Architectures

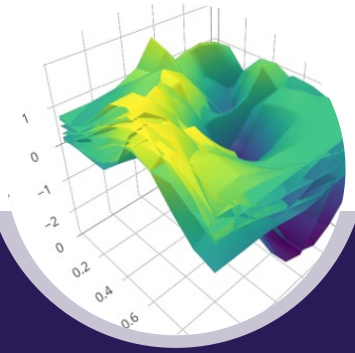
Accelerate the adoption of advanced architectures and lay foundations for Post-Moore opportunities, include edge computing, clouds, quantum information science



Data Management, Storage Systems, Visualization

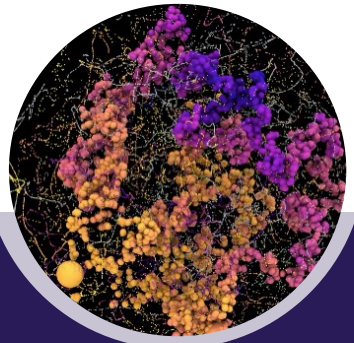
Enable the storing, organizing, transport, and transformation of scientific research data to be readily available and actionable

Research Capabilities: AI/ML



Foundations of ML, Data Analysis & Statistics

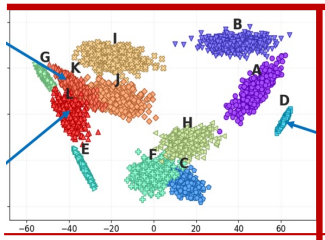
Leverage expertise in automatic differentiation, modeling, optimization and uncertainty quantification



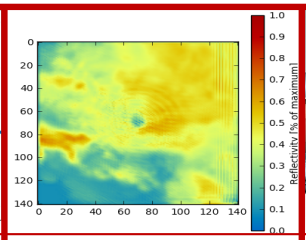
Application of AI/ML

Explore the application of AI/ML to problems in science and engineering; application to many different areas including design of new materials, additive manufacturing, climate models, weather prediction, COVID therapeutics

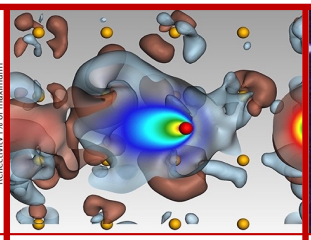
Applications of AI/ML



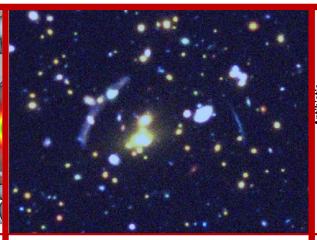
Reduced order modeling of laser sintering



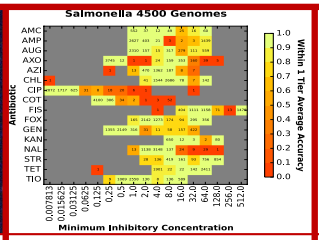
Nowcasting with convolutional LSTMs



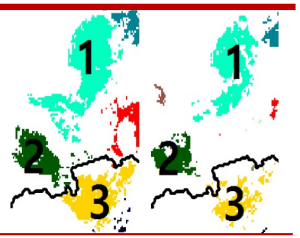
Prediction of radiation stopping power



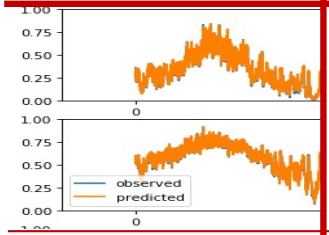
Strong and weak lensing in sky survey data



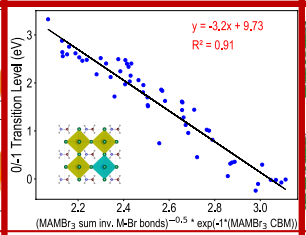
Prediction of antimicrobial resistance phenotypes



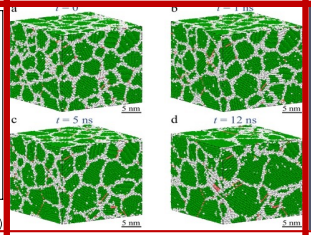
Identification and tracking of storms



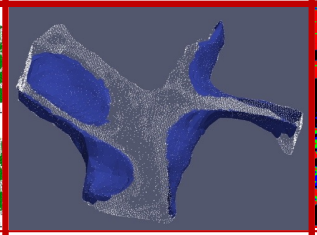
Efficient climate model emulators



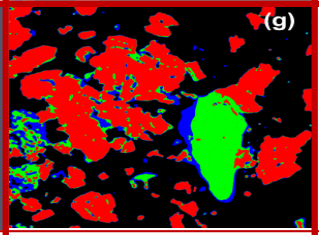
Defect-level prediction in seminconductors



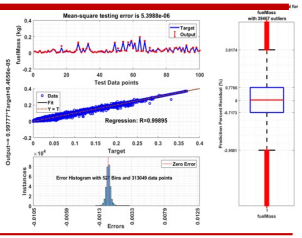
Structure-property-process triangle in additive manufact.



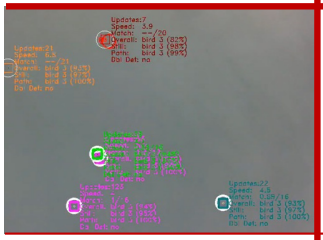
Parameter extraction in atom probe tomography



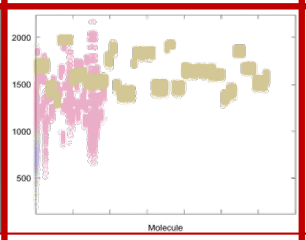
Learning for dynamic sampling in spectroscopy



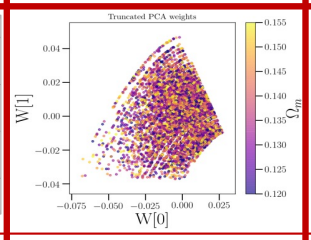
Vehicle energy consumption prediction



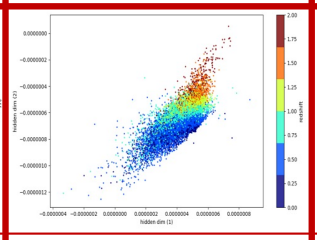
Flying object detector for edge deployment



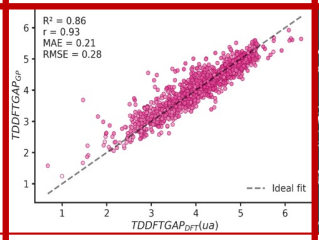
Discovery of new energy storage materials



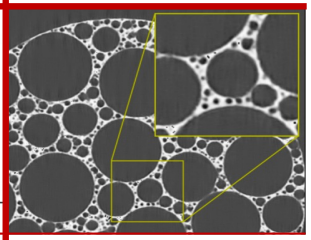
Cosmic Microwave Background emulation



Photometric red shift estimation



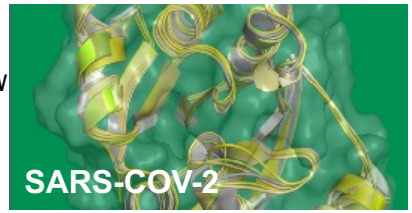
New materials for efficient solar cells



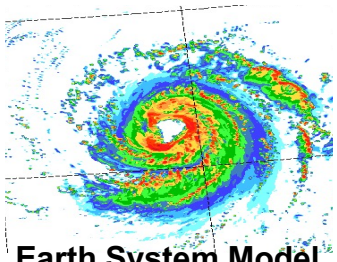
Enhancement of noisy tomographic images

Research Capabilities: Computational Science

Accelerate understanding of how the SARS-COV-2 virus binds to human cells

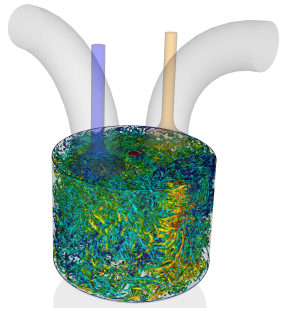


Predicting hurricane tracks to mitigate risks, hindcasting with earth system model data to gauge impact of global change

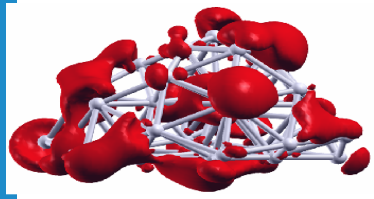


Earth System Model

Predictive simulation tools for nozzle flow, spray, combustion, and turbulence to understand cyclic variability

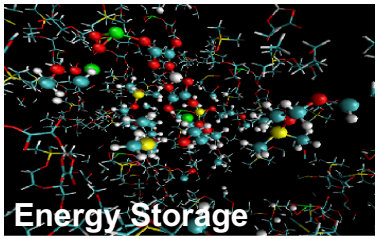


Internal Combustion

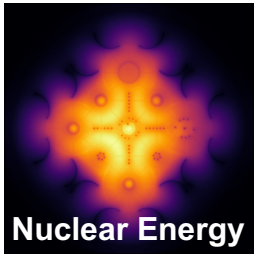


Nano Catalysts

Mapping out properties of nanoparticles to design catalysts for fuel cells and methane conversion

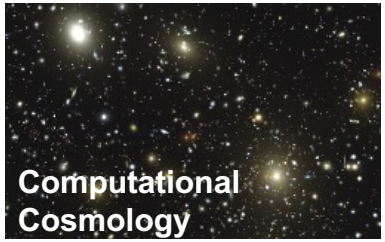


Next generation battery technologies. Here, optimizing electrolyte stability in metal-air batteries

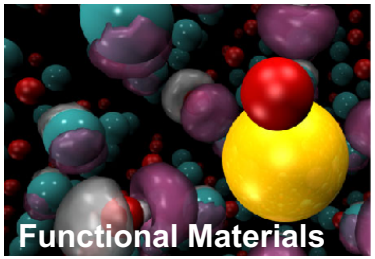


High-fidelity, Multiphysics simulations combining particle transport and fluid dynamics for modeling next-generation fission and fusion reactors

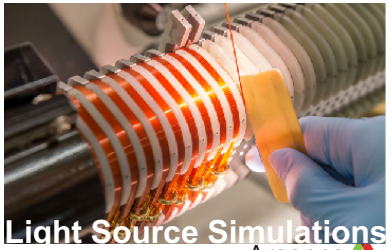
Extreme-scale multi-physics simulations for next-generation sky surveys investigating the "Dark Universe"



Enhanced predictive capability for functional materials with ab initio quantum Monte Carlo methods



Advanced accelerator simulation methods for next-generation light source design, and improved beam quality, as applied to APS-U



Initial IIT- CELS Seminar Series

Goal: Facilitate collaborations

- *March 9: AI/ML*
  - IIT: Mustafa Bilgic (AI/ML), Yan Yan (computer vision)
  - ANL: Prasanna Balaprakash (AI/ML), Venkat Vishwanath (AI Testbed)
- *March 23: HPC*
  - IIT: Stefan Muller (programming languages), Ruja Wang (computer architecture)
  - ANL: Yanfei Guo (MPI), Katherine Riley (science of ALCF)
- *March 30th: Edge Computing*
  - IIT: Kyle Hale (virtualization), Nik Sultana (networking)
  - ANL: Nicola Ferrier (AI@edge, computer vision), Tekin Bicer (scientific workflows)
- *April 6th: Data Science*
  - IIT: Lulu Kang (data science), Maggie Cheng (data analytics and graph theory)
  - ANL: Julie Bessac (multidimensional statistics), Ryan Chard(on-demand data science)