



U.S. DEPARTMENT OF
ENERGY

Office of
Science

Office of Science Data for AI Round Table

Implementing FAIR Data for People and Machines: Impacts and Implications
September 11, 2019

Laura Biven, PhD
Program Manager, Computer Science,
Advanced Scientific Computing Research
U.S. Department of Energy
Laura.Biven@science.doe.gov

FY19 Appropriation:
\$6,585,000,000

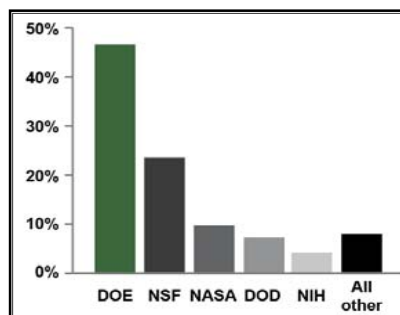
Office of Science

By the numbers

SC delivers scientific discoveries and tools to transform our understanding of nature and advance the energy, economic, and national security of the U.S.

Research

- Provides over 40% of the U.S. Federal support for basic research in the physical sciences;
- Supports over 23,000 Ph.D. scientists, graduate students, engineers, and support staff at over 300 institutions and all 10 DOE national laboratories;
- Maintains U.S. and world leadership in high-performance computing and computational sciences;
- Continues to be the major U.S. supporter of physics, chemistry, materials sciences, and biology for discovery and for energy sciences.



Support for basic research in the physical sciences by agency.

Source: NSF Science and Engineering Indicators 2012

Scientific User Facilities

- SC maintains the world's largest collection of scientific user facilities (aka research infrastructure) operated by a single organization in the world, used by more than 35,000 researchers each year.

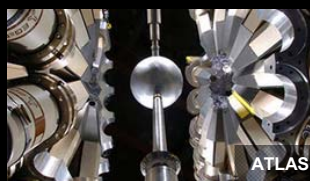
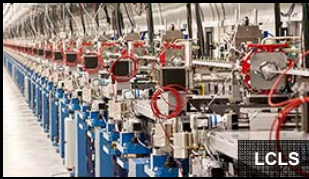


Shown is a portion of SLAC's two-mile-long linear accelerator (or linac), which provides the electron beam for the new Linac Coherent Light Source – the world's first hard x-ray, free-electron laser. For nearly 50 years, SLAC's linac had produced high-energy electrons for physics experiments. Now researchers use the very intense X-ray pulses (more than a billion times brighter than the most powerful existing sources) much like a high-speed camera to take stop-motion pictures of atoms and molecules in motion, examining fundamental processes on femtosecond timescales.

DOE Office of Science Research Portfolio^{*}

Basic Energy Sciences	• Understanding, predicting, and ultimately controlling matter and energy at the electronic, atomic, and molecular levels
Advanced Scientific Computing Research	• Extending the frontiers of science through world leading computational science, supercomputers, and networking
Biological and Environmental Research	• Understanding complex biological and environmental systems
Fusion Energy Sciences	• Studying matter at very high temperatures and densities and the scientific foundations for fusion
High Energy Physics	• Exploring the elementary constituents of matter and energy, the interactions between them, and the nature of space and time
Nuclear Physics	• Discovering, exploring, and understanding all forms of nuclear matter

**FY 2018
27 user facilities
35,000 users**



**U.S. DEPARTMENT OF
ENERGY**
Office of Science

17 DOE National Laboratories

Office of Science Laboratories

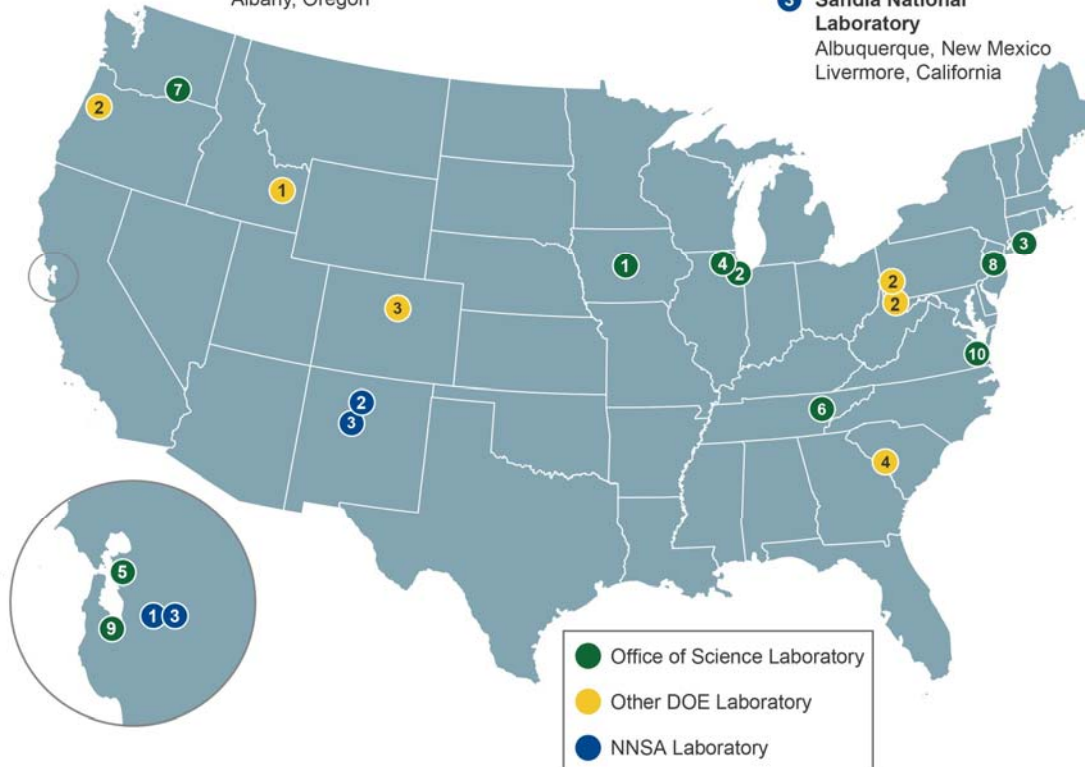
- 1 Ames Laboratory
Ames, Iowa
- 2 Argonne National Laboratory
Argonne, Illinois
- 3 Brookhaven National Laboratory
Upton, New York
- 4 Fermi National Accelerator Laboratory
Batavia, Illinois
- 5 Lawrence Berkeley National Laboratory
Berkeley, California
- 6 Oak Ridge National Laboratory
Oak Ridge, Tennessee
- 7 Pacific Northwest National Laboratory
Richland, Washington
- 8 Princeton Plasma Physics Laboratory
Princeton, New Jersey
- 9 SLAC National Accelerator Laboratory
Menlo Park, California
- 10 Thomas Jefferson National Accelerator Facility
Newport News, Virginia

Other DOE Laboratories

- 1 Idaho National Laboratory
Idaho Falls, Idaho
- 2 National Energy Technology Laboratory
Morgantown, West Virginia
Pittsburgh, Pennsylvania
Albany, Oregon
- 3 National Renewable Energy Laboratory
Golden, Colorado
- 4 Savannah River National Laboratory
Aiken, South Carolina

NNSA Laboratories

- 1 Lawrence Livermore National Laboratory
Livermore, California
- 2 Los Alamos National Laboratory
Los Alamos, New Mexico
- 3 Sandia National Laboratory
Albuquerque, New Mexico
Livermore, California



DOE is uniquely positioned to lead in AI/ML R&D

DOE has a unique combination of capabilities and infrastructure to lead the Nation in Artificial Intelligence (AI) and Machine Learning (ML) research and development:

- A broad **mission** that presents new and unique research problems on national and global scales to attract new talent.
- Sources of massive and/or complex science and engineering **data** from sensors, instruments, and from large-scale simulations.
- World-class high performance **computing** infrastructure capable of world-leading AI research.
- World-class high performance **network** infrastructure capable of integrating computing resources and data assets .
- An exceptional **workforce** with large numbers of scientists, computer scientists, and mathematicians currently engaged in AI and related fields.

AI Initiative



Artificial Intelligence for the **American People**

Enhance access to high-quality and fully traceable Federal data, models, and computing resources to increase the value of such resources for AI R&D, while maintaining safety, security, privacy, and confidentiality protections consistent with applicable laws and policies.

[Executive Order 13859](#) of February 11, 2019
**Maintaining American Leadership in
Artificial Intelligence**

Office of Science (SC) Round Table on Data for AI

Rockville Hilton

June 5, 2019

<https://www.ornl.gov/RTD-AI2019/>

Supported by the **Office of Science Working Group on Digital Data** (SCWGDD), which includes representatives from all six SC program offices and OSTI

POC: Laura Biven (ASCR)

Participants came from 13 DOE National Labs, NIH, & NSF with expertise in AI/ML, data management, data curation, metadata, storage systems and I/O, open data, big data, data analytics, and edge computing; with ties to SC research, facilities, and community data repositories

Round Table Focus, Goals, & Context

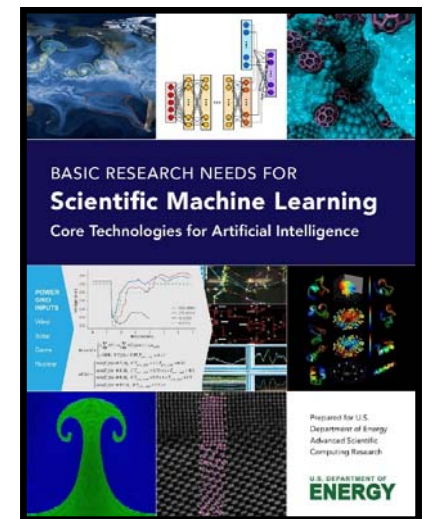
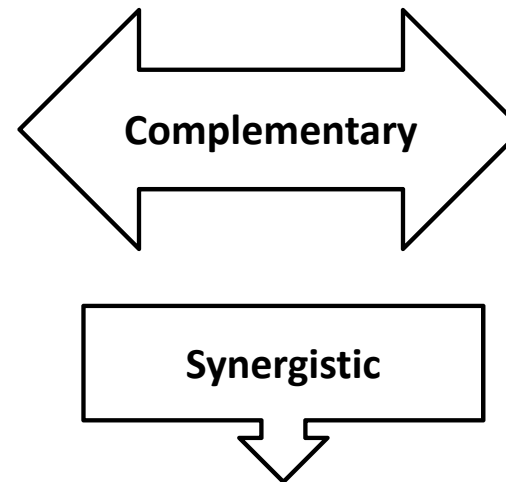
SC Round Table on Data for AI

Focus:

Enhancing and enabling access to high-quality and fully traceable research data, models, and computing resources to increase the value of such resources for AI R&D and the SC mission.

Goal:

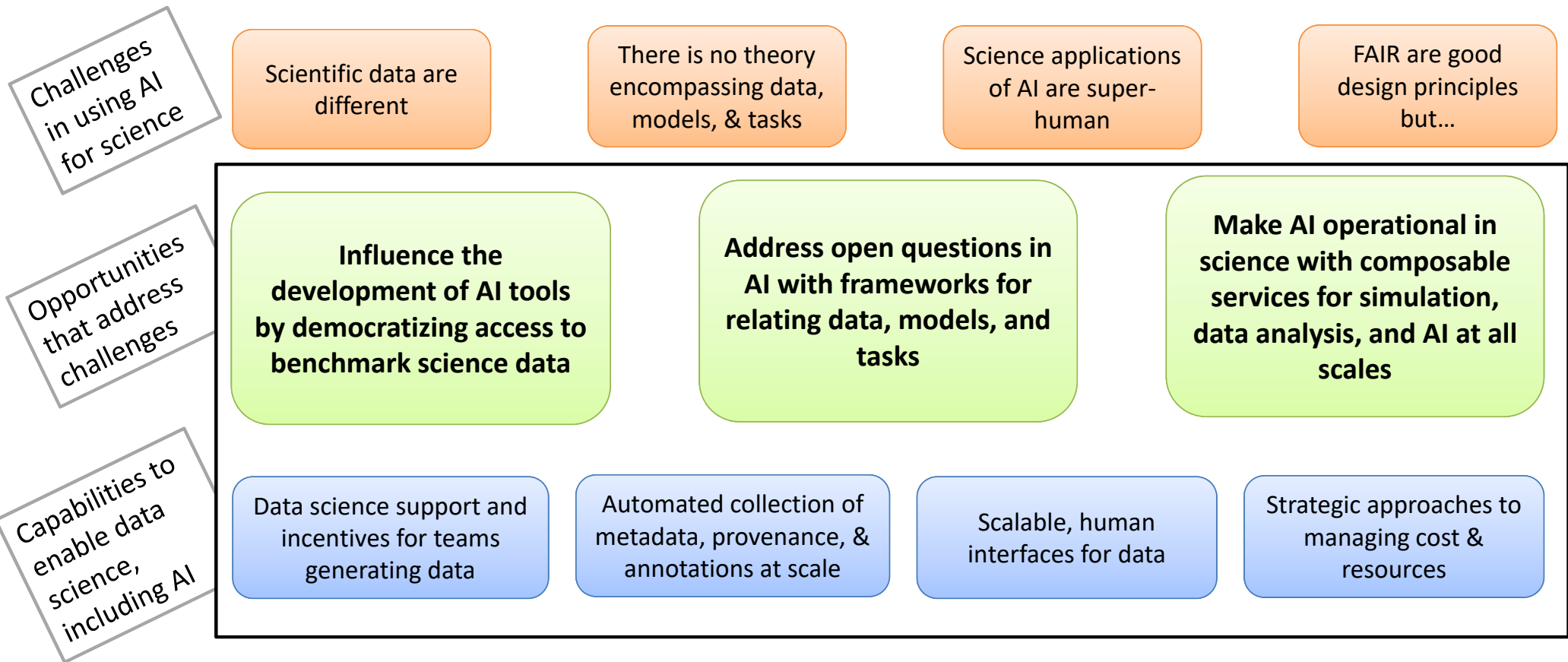
Identify key challenges/opportunities and potential next steps for the Office of Science.



AI FOR SCIENCE TOWN HALL

AI-focused workshops and round tables of sister SC program offices

Office of Science Data for AI Round Table: Challenges, Opportunities, & Enabling Capabilities



Thank you!!

SC Organizing Team

SC Working Group on Digital Data

POC: Laura Biven (ASCR)

Ben Brown (ASCR)

Michael Cooke (HEP)

Mariam Elsayed (SC-2)

Sujata Emani (AAAS - BER)

Jay Hnilo (BER)

Carolyn Lauzon (ASCR / AI Office)

Joanna Martin (OSTI)

Jessica Moerman (AAAS - BER)

Lab Writing Team



Kjersten
Fagnan
(LBNL/JGI)



Youssef Nashed
(ANL)



Gabriel Perdue
(FNAL)



Daniel Ratner
(SLAC)



Arjun
Shankar
(ORNL)



Shinjae Yoo
(BNL)