Agency Perspective on Open Science for Data

Beth Plale

Office of Advanced Cyberinfrastructure (OAC),

Directorate for Computer & Information Science & Engineering (CISE)

National Science Foundation

NSF Public Access Repository (NSF PAR) for publications

- Centralized repository at NSF that is built on top of DOE PAGES system
- Accepts journal and juried conference papers
- Award recipients deposit author copy of publication as part of reporting process
- Recently extended on pilot basis to include workshop reports from NSF funded workshops

par.nsf.gov

Open science for data

- Dear Colleague Letter: Effective Practices for Data (May 2019)
- DCL encourages researchers to learn about effective practices for data, and to implement them in the proposals that they prepare for submission to NSF.
- Effective practices around:
  - Persistent IDs
  - Data Management Plans

Effective practices for data

- Assign **Globally persistent IDs** (DOIs) to research data
  - Cite research data in body of publication with corresponding reference in reference list
  - Include with research data a statement of data availability
- **Use DMP tools to create machine readable DMPs**
  - Data Management Plans (DMPs) required part of proposals to NSF
  - Machine readable DMPs can be read machines, shared with repositories, and changed over time (active DMPs)

Broader Landscape of Influencers

- AAU/APLU/ARL workshop series on Accelerating Public Access to research data, Oct 2018 and forthcoming
  - NSF and NLM sponsors
- NASEM *Open Science by Design: Realizing a Vision for 21st Century Research*, Washington, DC: The National Academies Press, 2018.
- NASEM *Reproducibility and Replicability in Science*, Washington, DC: The National Academies Press, 2018.
- Challenges in Irreproducible Research, Special Issue, Nature Oct 18, 2018, Springer Nature
- OSTP National Science and Technology Council (NSTC) Subcommittee on Open Science : reduce researcher burden

Community Workshop Activity

- Community workshop topics
  - Data reuse
  - Making data FAIR (findable, accessible, interoperable, reusable)
  - Publishing guidelines for chemical structures

Artificial Intelligence for Data Discovery and Reuse

An NSF-supported conference @ Carnegie Mellon University, in-cooperation with ACM

NSF WORKSHOP

MPS FAIR Hackathon

Findable Accessible Interoperable and Reusable (FAIR) Hackathon Workshop for Mathematical and Physical Sciences (MPS) Research Communities

February 27-28, 2019

Open Infrastructure to Reduce Burden on Researcher and Federal Agencies

Johns Hopkins University, in partnership with six other institutions, [...] will transform the deposit of manuscripts into federal agency repositories, thereby reducing burden on researchers, universities, and federal agencies.

Award Abstract #1939291

EAGER: Open Infrastructure to Reduce Burden on Researchers and Federal Agencies

goFAIR

SDSC's Research Data Services to Host First U.S. GO FAIR Office

Newly Rebranded Division to Offer FAIR Data Consulting, Other Services

Published February 28, 2019

Midwest Big Data Innovation Hub: ***midwestbigdatahub.org***

Northeast Big Data Innovation Hub: ***nebigdatahub.org***

South Big Data Hub: ***southbigdatahub.org***

West Big Data Hub: ***westbigdatahub.org***

NSF BIG DATA
INNOVATION HUBS

Leveraging Scientific Societies for Open and FAIR Scientific Data
Award Number:1838990; Principal Investigator:Shelley Stall; Co-Principal Investigator:; Organization:American Geophysical Union;NSF Organization:OAC Start Date:11/01/2018; Award Amount:$50,000.00; Relevance:45.83;

Type-Based Automation of Scientific Data Management
Award Number:1838981; Principal Investigator:Giridhar Manepalli; Co-Principal Investigator:; Organization:Corporation for National Research Initiatives (NRI);NSF Organization:OAC Start Date:10/01/2018; Award Amount:$297,796.00; Relevance:44.67;

Framework: Data: HDR: Extensible Geospatial Data Framework towards FAIR (Findable, Accessible, Interoperable, Reusable) Science
Award Number:1835822; Principal Investigator:Xiaohui Carol Song; Co-Principal Investigator:Jian Jin, Uris Lantz Baldos, Venkatesh Merwade, Jack Smith; Organization:Purdue University;NSF Organization:OAC Start Date:10/01/2018; Award Amount:$4,571,811.00; Relevance:37.56;

Creating FAIR Data in Lake Observatories of the Future
Award Number:1938743; Principal Investigator:Kathleen Weathers; Co-Principal Investigator:Renato Figueiredo, Paul Hanson, Cayelan Carey; Organization:Institute of Ecosystem Studies;NSF Organization:OAC Start Date:12/01/2019; Award Amount:$49,958.00; Relevance:47.62;

Workshop: Extending Open Access By Bridging the Data Steward Gap: GO FAIR Train-the-Trainer (T3)
Award Number:1937953; Principal Investigator:Christine Kirkpatrick; Co-Principal Investigator:Melissa Cragin; Organization:University of California-San Diego;NSF Organization:OAC Start Date:09/01/2019; Award Amount:$49,999.00; Relevance:47.62;

bplale@nsf.gov
703 292 7004

Always welcoming your views on open science for data and software