

Research Data Alliance: Introduction

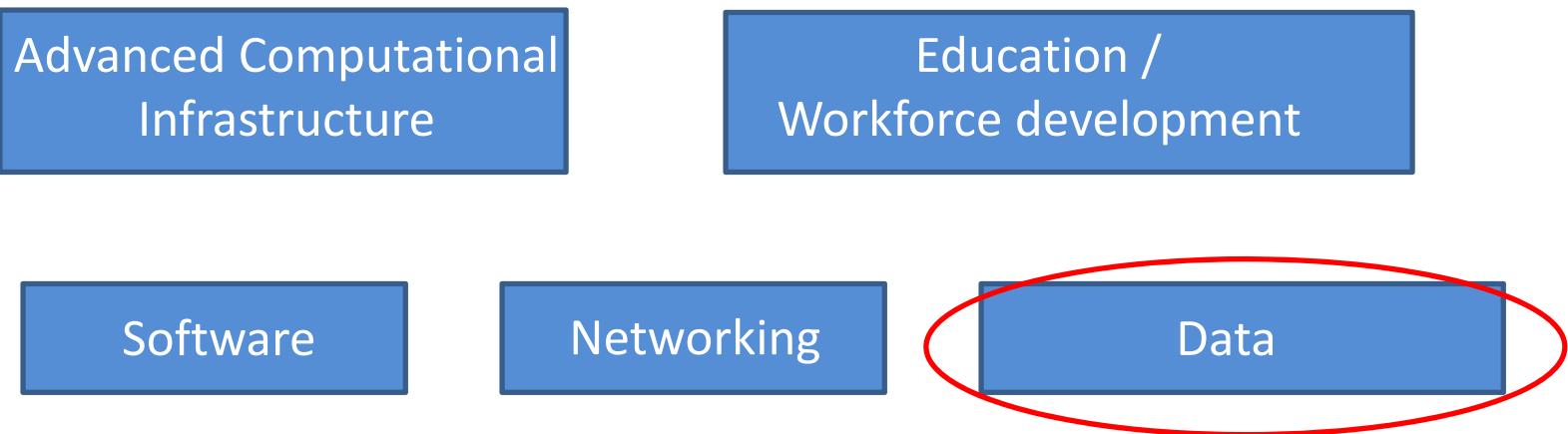
Emily Grumbling

AAAS Science and Technology Policy Fellow
placed in the Office of Cyberinfrastructure, NSF

CENDI Meeting, 1-9-13

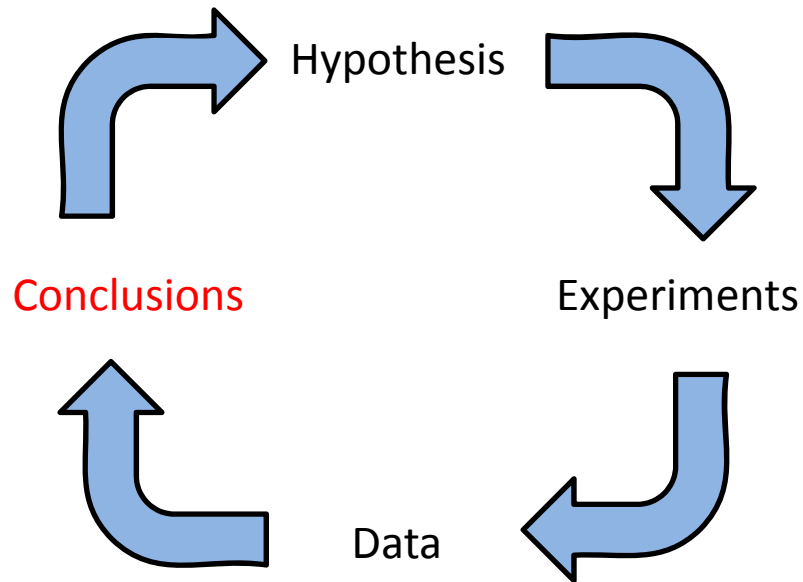
The Office of Cyberinfrastructure at the National Science Foundation

The Office of Cyberinfrastructure coordinates and supports the acquisition, development and provision of state-of-the-art cyberinfrastructure resources, tools and services essential to the conduct of 21st century science and engineering research and education.



Scientific Discovery

Scientific Method



Scientific Paradigms

Empirical

Theoretical

Computational

Data-centered

Data Explosion: a catalyst for new discovery!

Big Data

large, diverse, complex, longitudinal, and/or distributed data sets generated from instruments, sensors, Internet transactions, email, video, click streams, and/or all other digital sources available today and in the future (NSF 12-499)

... but don't forget

non-digitized / small data sets

Much data resides in the “long-tail” of science!

Data-Driven Science: Opportunities and Challenges

- data is abundant
- multiple use of data
- enhanced collaboration
- crowdsourcing
- access / interoperability
- community acceptance of data-sharing
- attribution / citation of data
- credit for contributions to infrastructure

Strategies and tools for managing, sharing and using data in the short- and long-term, within and between disciplines, are crucial to the scientific enterprise

OCI Data Efforts

Sustainable Digital Data Preservation and Access Network Partners (DataNet)

to catalyze the development of a system of science and engineering data collections that is open, extensible and evolvable... and support the creation of a set of exemplar national and global data research infrastructure organizations, dubbed DataNet partners (NSF 07-601)

DataONE, DataConservancy, SEAD, others
DataWeb Forum → Research Data Alliance

Research Coordination Networks (RCN)

to advance a field or create new directions in research or education by supporting groups of investigators to communicate and coordinate their research, training and educational activities across disciplinary, organizational, geographic and international boundaries (NSF 13-520)

Research Data Alliance (RDA)

*to facilitate the sharing of research data world-wide
and help advance data-driven scientific discovery*

Now underway!

- Award active: September 1, 2012
- US efforts supported by NSF and NIST
- International partners in EU, Australia
- A non-commercial, non-governmental entity
- Led and driven by the research community and data users
- Aimed at near-term results (18 months)

Initial PIs funded to support RDA

US:

Fran Berman – RPI

Bill Michener – *DataOne*

Beth Plale – Indiana University

Sayed Choudhury – Johns Hopkins University

Australia:

Ross Wilkinson – Australia National Data Service

Andrew Treolar – Australia National Data Service

Europe:

Leif Laaksonen (iCORDI/CSC)

Peter Wittenburg (iCORDI/Max Planck Institute)

Juan Bicarregui (iCORDI/STFC)

Research Data Alliance (RDA)

- Kickoff plenary meeting: March 18-20, Gottenberg, Sweden

Next: RDA Update by Dr. Fran Berman

Thank you!

egrumbli@nsf.gov