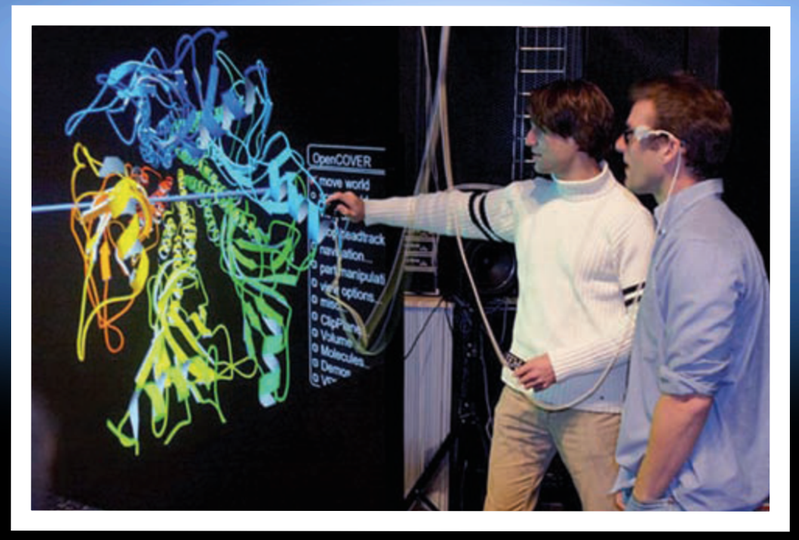
Data & Interaction: Foundations for Science

Lucy Nowell, Program Director

Data, Data Analysis, and Visualization/Interaction

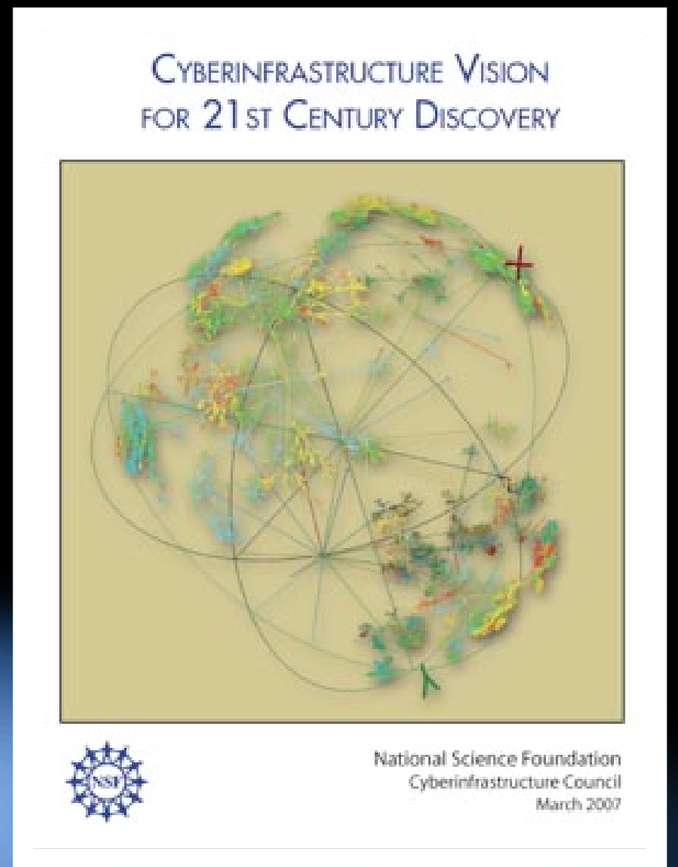


National Science Foundation
Where Discoveries Begin

Office of Cyberinfrastructure

- "Science and engineering digital data are routinely deposited in a well-documented form, are regularly and easily consulted and analyzed by specialists and non-specialists alike, are openly accessible while suitably protected and are reliably preserved."

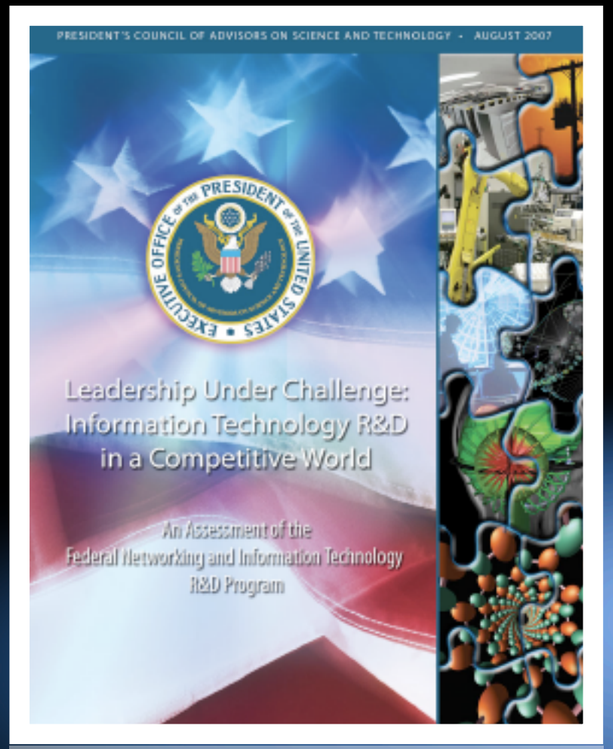
- "Scientific visualization, including not just static images but also animation and interaction, leads to better analysis and enhanced understanding."



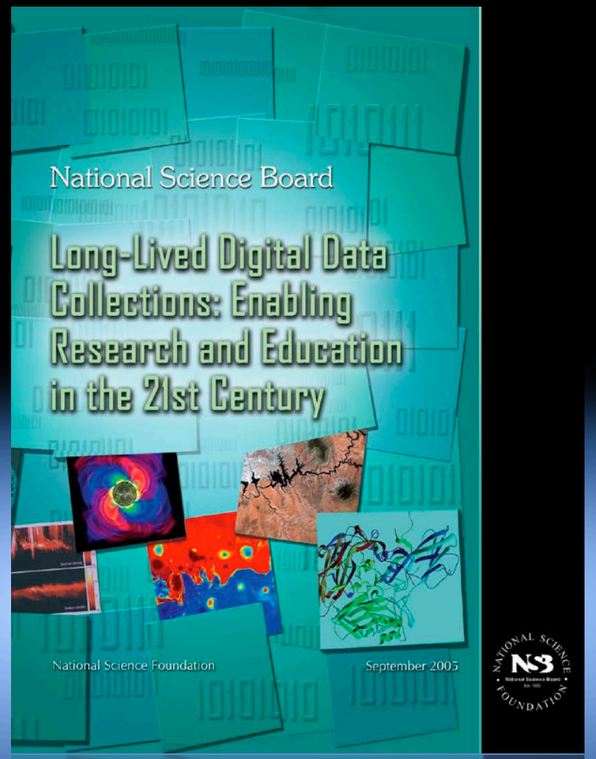
NSF Cyberinfrastructure Vision for 21st Century Discovery

Lucy Nowell
lnowell@nsf.gov

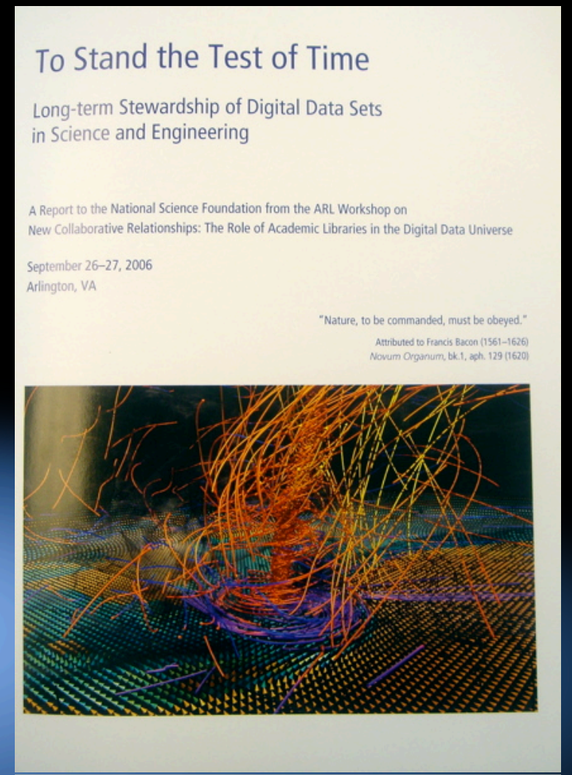
PCAST Recommendations on Digital Data



NSB Report: Long-Lived Digital Data Collections Enabling Research and Education in the 21st Century



NSF Supported Experts Study

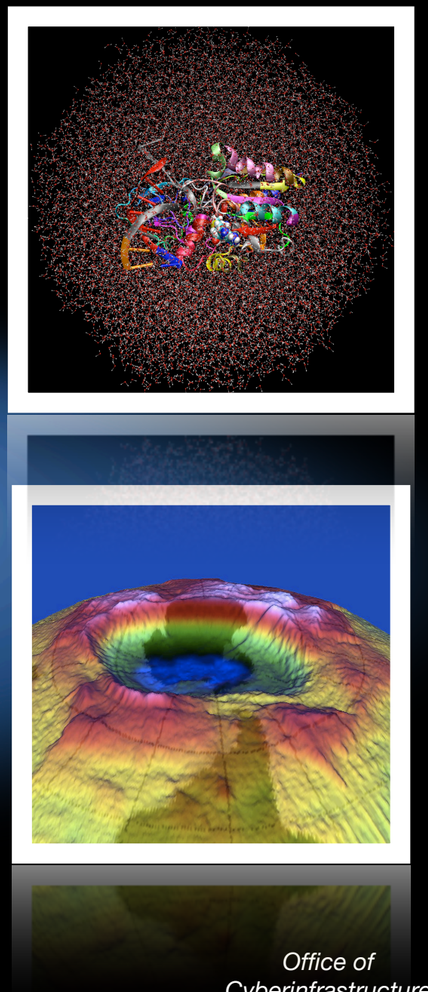


Examples of Input Informing Data Activities

Lucy Nowell
lnowell@nsf.gov

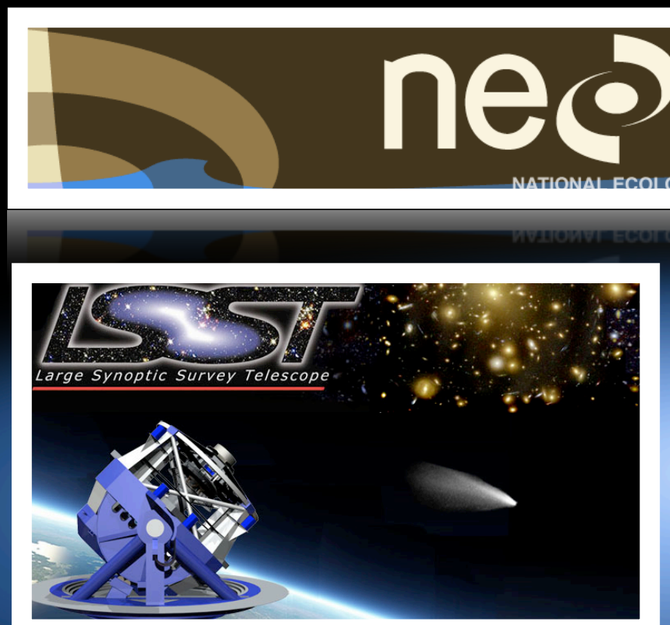
Drivers for OCI Data & Data Interoperability Strategy

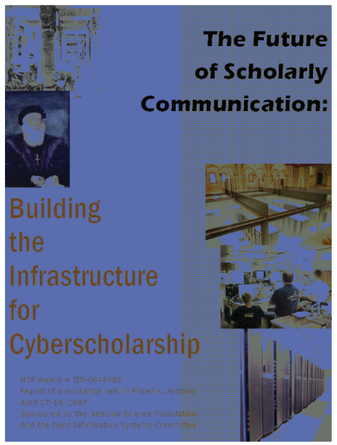
- Increased scale & heterogeneity of Data
- Demand for federation & semantic interoperability
- Increased expectations for sharing & openness
- More systematic quality control & long-term access.

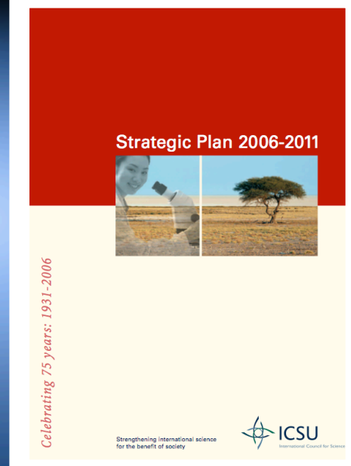


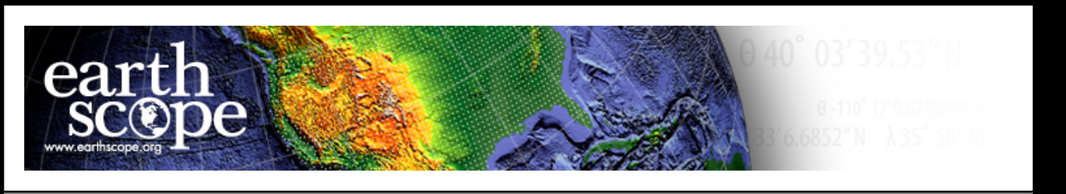
Lucy Nowell
lnowell@nsf.gov

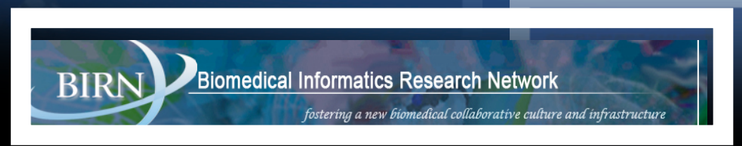
Examples of Science Drivers for OCI Data Activities

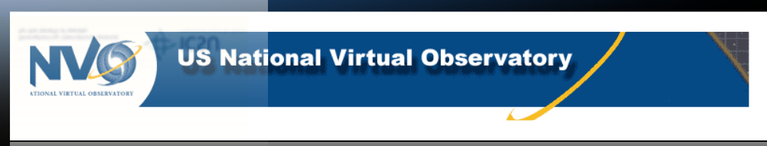












National Science Foundation
Where Discoveries Begin

Lucy Nowell
lnowell@nsf.gov

Office of Cyberinfrastructure

OCI Data Strategy

Mechanisms

Activities

	Funded Programs	Policy Consideration	Partnerships
National Framework	DataNet	Digital Data Coordinating Group	Interagency Data Working Group Blue Ribbon Task Force on Sustainable Data Repositories NSF-wide Data Working Group NAS, CODATA, UNESCO
Culture Change	INTEROP	Exec-level Review of NSF Data Policies	
Tools/Resources	SDCI, STCI	NSF Data Policies & Registry Services	
Opportunity Exploration	Federal Agencies, Academia, Library & Preservation Sector, Foundations & Non-profits, Commercial Sector, National Laboratories, International Agencies		

Lucy Nowell
lnowell@nsf.gov

NSF Data Management Policy

- **Current NSF Policy:**
  - ***"Investigators are expected to share with other researchers, at no more than incremental cost and within a reasonable time, the primary data … created or gathered in the course of work under NSF grants."***

NSF Proposal and Award Policies and Procedures Guide, 2007

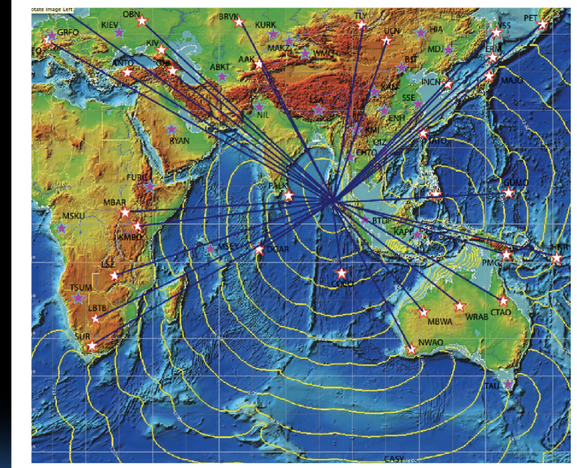
- NSF Data Working Group, with senior representatives from all Directorates, is working to strengthen the policy and implementation thereof.

- **Goal: Provide for clear, effective, and transparent implementation of long-standing NSF policy**

Lucy Nowell
lnowell@nsf.gov

Office of Cyberinfrastructure

- History of Science
- Ensuring the Integrity of Science via Replication of Results
- Longitudindal Science: Change and the Impact of Human Activity
- Training and Validating Predictive Models and Simulations
- Facilitating Cross-Disciplinary Science through Re-use and Repurposing of Data
- Accelerating the Pace of Scientific Discovery and Innovation
- Economies of Scale and Good Stewardship of Resources
- Broadening Participation



Why Preserve & Share Data?

National Science Foundation
Where Discoveries Begin

Lucy Nowell
lnowell@nsf.gov

Office of Cyberinfrastructure

Storage Networking Industry Association (SNIA) *100 Year Archive Requirements Survey Report*

"The results confirmed our view that there is a pending crisis in archiving…If we're going to be able to avoid this crisis we have to create long-term methods for not only preserving information, but also for making it available for analysis in the future."

Paperwork Reduction Act*:

The purposes of this subchapter are to ….

(2) Ensure the greatest possible public benefit from and maximize the utility of information created, collected, maintained, used, shared, and disseminated by or for the federal government; ….

(7) Provide for the dissemination of public information on a timely basis, on equitable terms, and in a manner that promotes the utility of the information to the public and makes effective use of information technology

Office of Management and Budget (OMB) Circular A-130: Management of Federal Information Resources

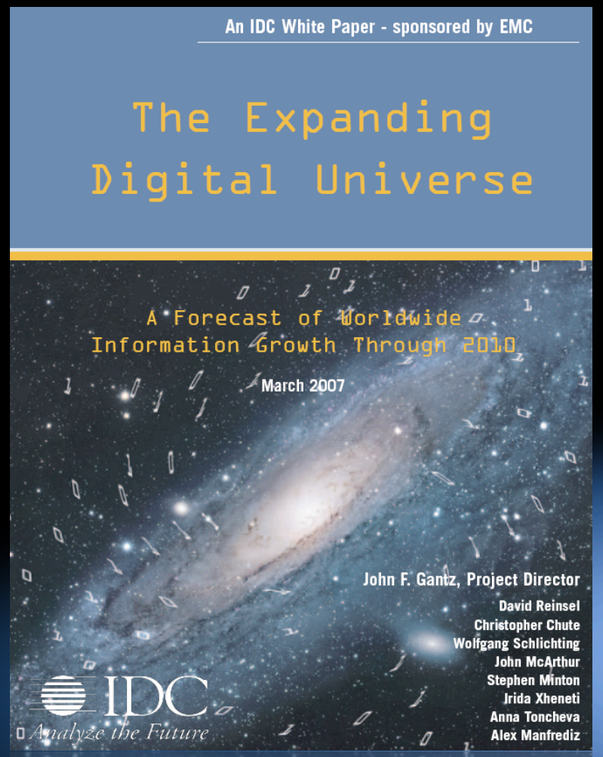
Part 7. Basic Considerations and Assumptions:….

k. The open and efficient exchange of scientific and technical government information, subject to applicable national security controls and the proprietary rights of others, fosters excellence in scientific research and effective use of Federal research and development funds.

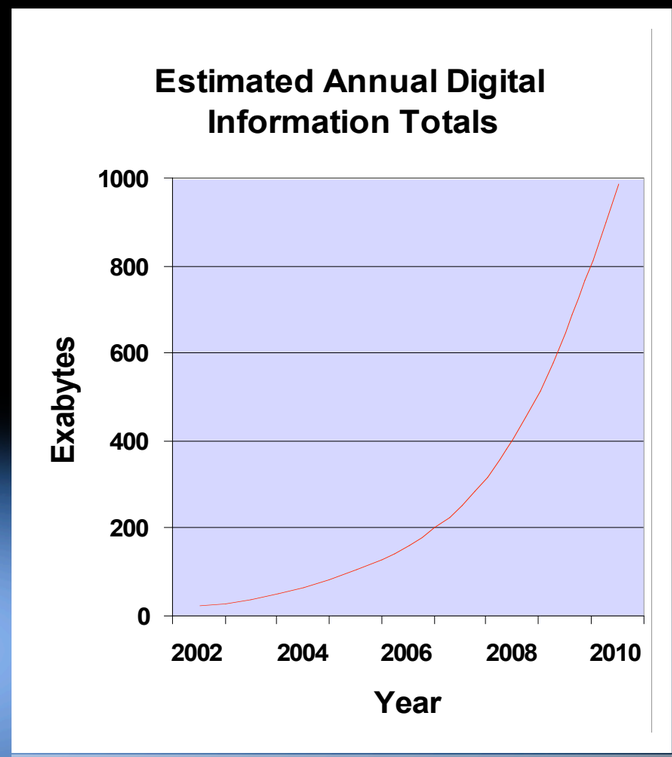
More Examples of Input Informing Data Activities

Lucy Nowell
lnowell@nsf.gov

IDC - Global Marketing Intelligence Company



"In 2007, the amount of information created will surpass, for the first time, the storage capacity available."



We Must Choose What to Keep

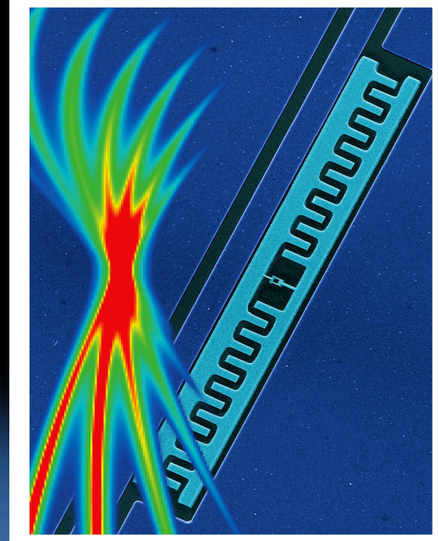
- **Long-Lived Repositories** project by Bernard Reilly of the Center for Research Libraries

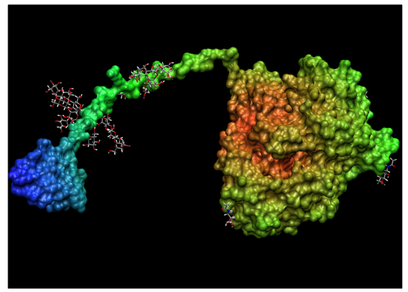
- **Blue Ribbon Task Force on Sustainable Digital Preservation and Access**, chaired by Fran Berman and Brian Lavoie, in cooperation with the Andrew W. Mellon Foundation, Library of Congress, National Archives and Records Administration, the Council on Library and Information Resources, and the Joint Information Systems Committee (JISC) of the United Kingdom

Related Activities

DataNet Program Primary Goals

- **Provide** reliable digital preservation, access, integration, and analysis capabilities for science/engineering data over decades-long timeline.
- Achieve **long-term preservation and access** capability in an environment of rapid technology advances.
- Create systems and services that are **economically and technologically sustainable**.
- **Empower science-driven information integration** capability on the foundation of a reliable data preservation network.



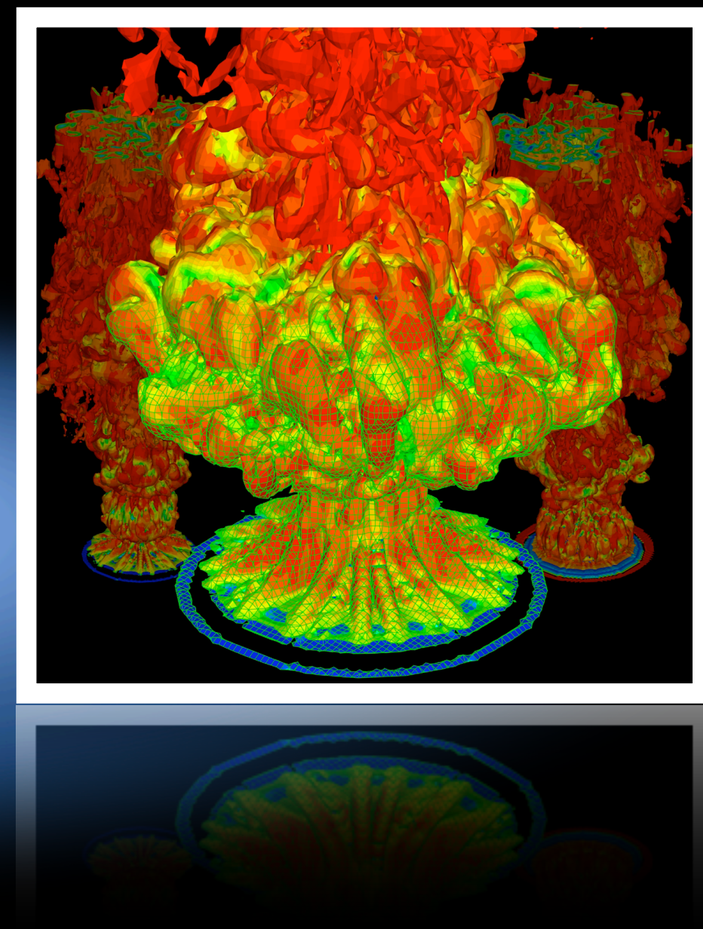


Lucy Nowell
lnowell@nsf.gov

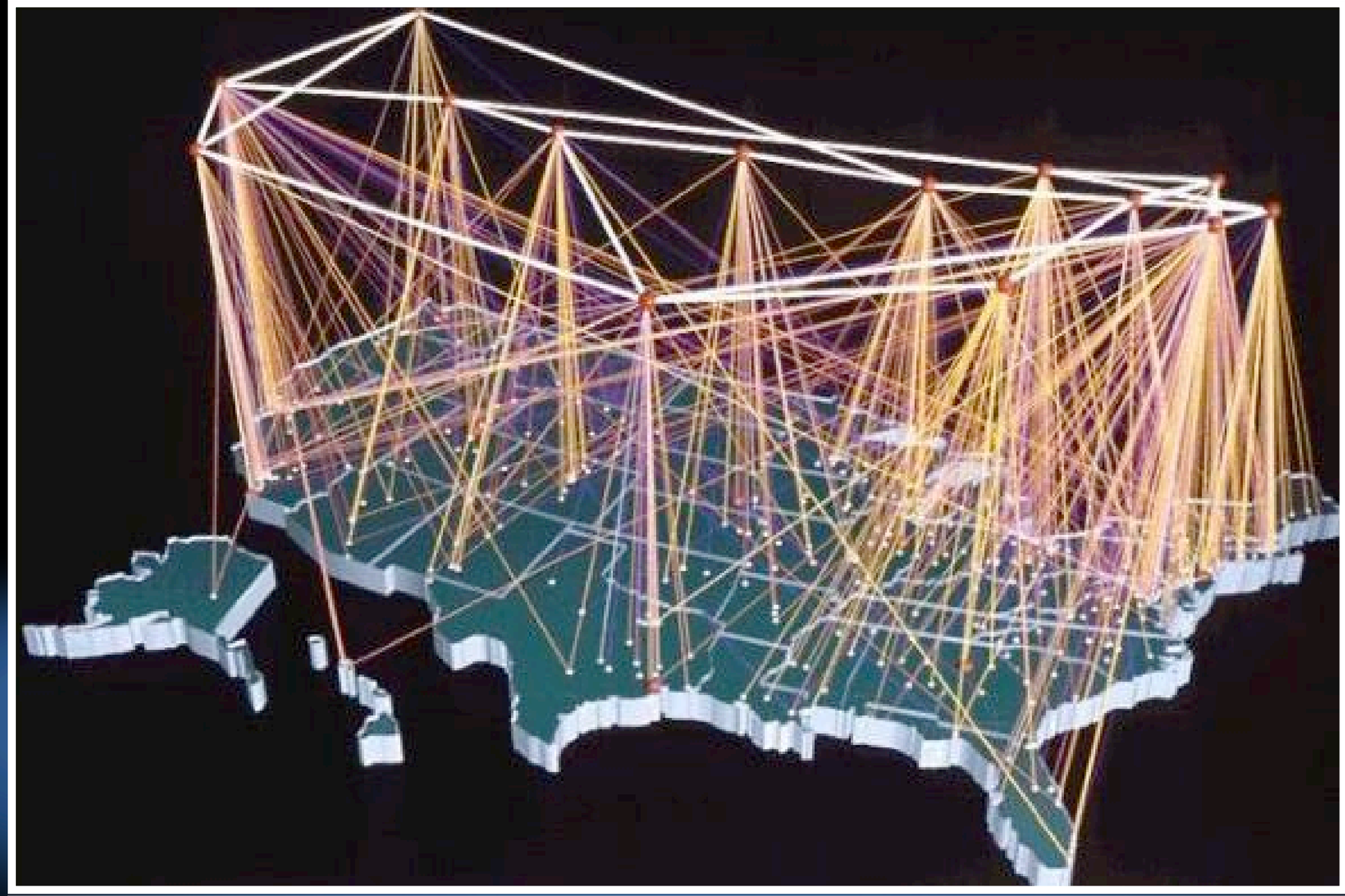
Office of Cyberinfrastructure

DataNet Approach

- 5 awards, 2 this year and 3 next year
- Explore, demonstrate and understand diverse approaches to developing in sustainable ways with diverse content.
- Initial focus on several disciplinary areas, with active outreach to more communities and more disciplines



Lucy Nowell
lnowell@nsf.gov



Goal: A DataNet Partners Network of Networks

Lucy Nowell
lnowell@nsf.gov

Building the Global Science Data Network Infrastructure

- European Heads of Research Councils (EUROHORCs)
- European Science Foundation
- UK's Joint Information Systems Council (JISC)
- Alliance for Permanent Access to the Records of Science
- Australian Partners for Sustainable Repositories (APSR)
- Data Management Canada, National Research Council and others
- Chinese Academy of Science (preliminary)

International Outreach

Lucy Nowell
lnowell@nsf.gov

- Goal: Identify research grand challenges in data preservation and sharing, providing a basis for coordinated international research efforts.

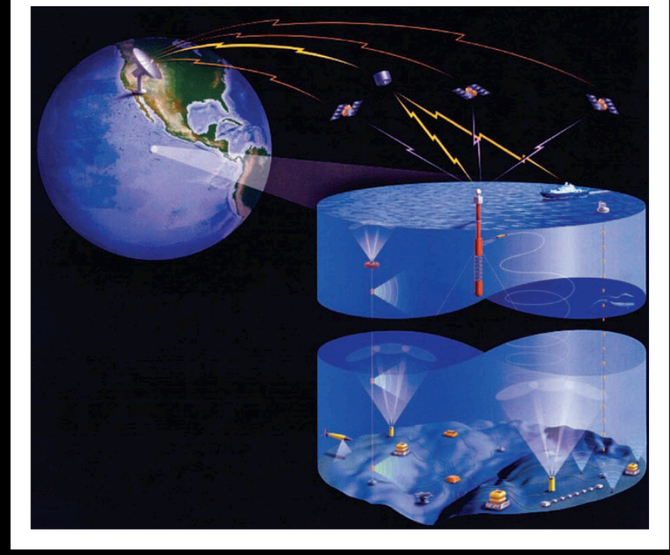
- Potential first meeting in May 2009, in conjunction with Open Repositories Conference in Atlanta, GA; representatives of funding agencies only, to set direction.

- Two additional meetings contemplated, six months apart, one in Europe and one in Australia. Include invited non-funding participants with technical expertise.

- Video-conference with interested funding parties? TBD.

New International Task Force on Data Preservation & Access Research Agenda?

- You can *do* science without graphics. But it's very difficult to *communicate* it in the absence of pictures. Indeed, some insights can only be made widely comprehensible as images. How many people would have heard of fractal geometry or the double helix or solar flares or synaptic morphology or the cosmic microwave background, if they had been described solely in words?



- To the general public, whose support sustains the global research enterprise, these and scores of other indispensable concepts exist chiefly as images.

NSF Science and Visualization Challenge 2007, Special Report

http://www.nsf.gov/news/special_report/scivis/index.jsp?id=challenge

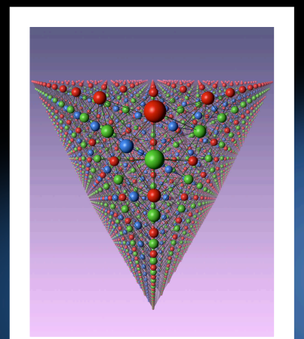
Visualization & Interaction

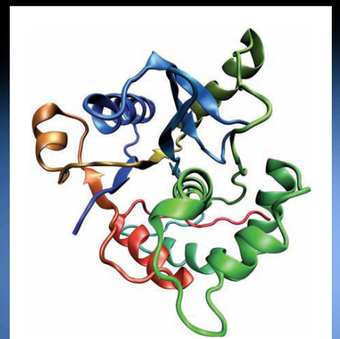
Cyberinfrastructure

National Science Foundation
Where Discoveries Begin

Lucy Nowell
lnowell@nsf.gov

“Without investments made by previous generations, we would not enjoy the seemingly invisible infrastructure that makes our modern lives possible. It goes without saying that if we don’t make similar investments now we will rob future generations of the quality of life they should enjoy.”





Princeton University Engineering School, Feb. 19, 2008, Greatest Technological Research Challenges of the 21st Century Identified by Expert Panel. *Science Daily*

http://www.sciencedaily.com/release/2008/02/080215151157.htm

Greatest Technological Challenges of the 21st Century

Lucy Nowell
lnowell@nsf.gov


Thank
you!