Beyond Visualization

Dr. Lucy Nowell, Program Director
Data, Data Analysis & Visualization
Office of Cyberinfrastructure
US National Science Foundation

Presented by Dr. Jennifer Schopf
Office of Cyberinfrastructure
US National Science Foundation

Role of the Office of Cyberinfrastructure (OCI)

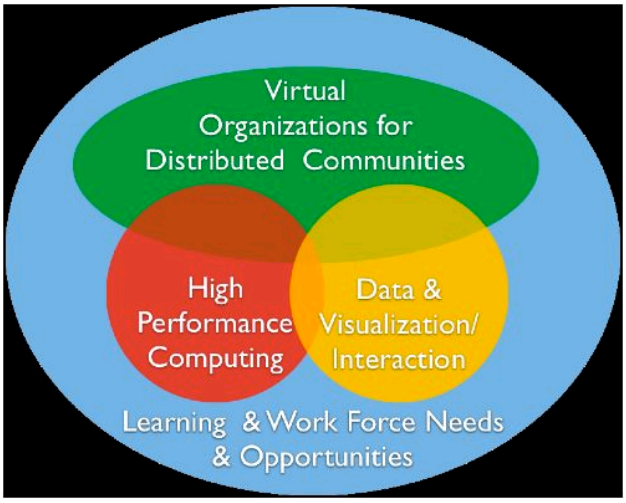
National Science Foundation
Director
Deputy Director

National Science Board

Office of Director
- CyberInfrastructure
- Integrative Activities
- Polar Programs
- International Science and Engineering

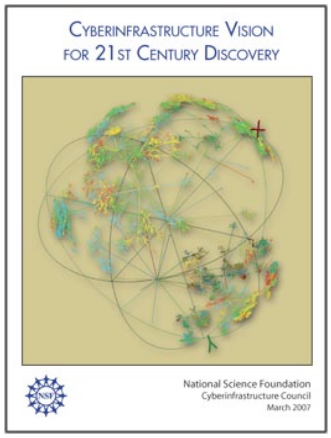
Research Directorates
- Biological Sciences
- Computer & Info. Science & Eng.
- Education & Human Resources
- Engineering
- Geosciences
- Mathematical & Physical Sciences
- Social, Behaviorial & Econ. Sciences

OCI Program Structure



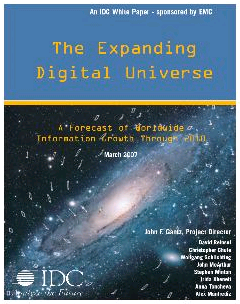
NSF Vision for Cyberinfrastructure

- "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."

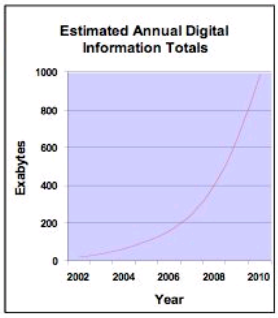


http://www.nsf.gov/pubs/2007/nsf0728/index.jsp

Context: IDC White Paper



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

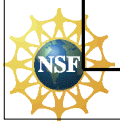


Appraising the future value of data and information presents one of the major challenges of our time.

http://www.emc.com/collateral/analyst-reports/expanding-digital-idc-white-paper.pdf

Enormous, Irreplaceable Data Sets

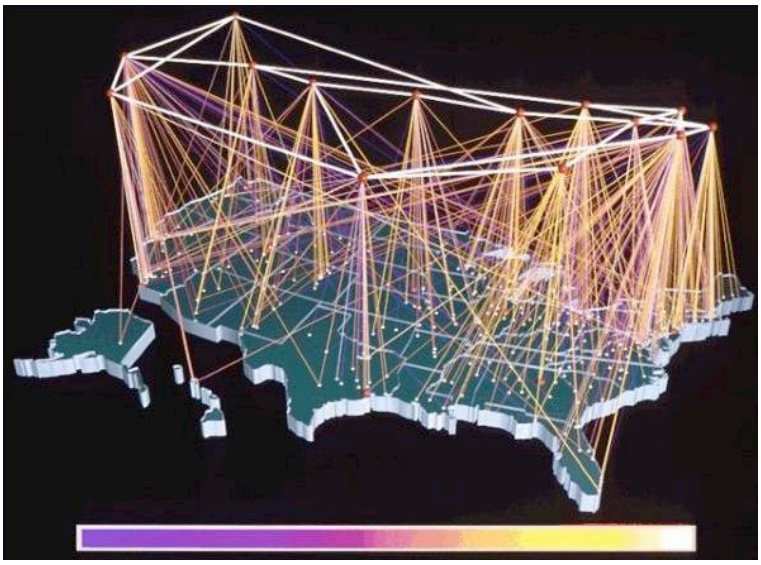
neon NATIONAL ECOLOGICAL OBSERVATORY NETWORK	~ 150 TB/Year
LSST	~ 30 TB/Night
Large Hadron Collider	~ 15 PB/Year
NASA E·O·S·D·I·S Earth Observing System Data & Information System	~ 64 TB/Year
earth scope	~ 40 TB/Year
Long tail of small science	?? TB/Year



OCI'S Current Data Focus

- Flagship Program: Sustainable Digital Data Preservation and Access Networks
  - DataNet
- Focus on data-level interoperability and data preservation
- Important context for any OCI programs in Visual Interaction.

Building a Network of Data Networks



DataNet : 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.
- Each project needed to develop a model for shared governance and the standards and protocols to enable interoperability.

Each DataNet Awardee Must…

- Integrate library and archival sciences, cyberinfrastructure, computer and information sciences, and domain science expertise.
- Work with multi-disciplinary science domains.
- Engage at the frontiers of computer and information science and cyberinfrastructure with research and development to drive the leading edge forward.

Why Preserve & Share Data?

- Broaden participation – core NSF value
- History of science – historical context
- Longitudinal studies to assess change, impact of policy/intervention, etc.
- Train and validate models and simulations
- Enable cross-disciplinary science with repurposing of data
- Good stewardship through re-use of costly data
- Accelerate the pace of scientific discovery and innovation
- Insure integrity of science by enabling replication of results

DataNet Status

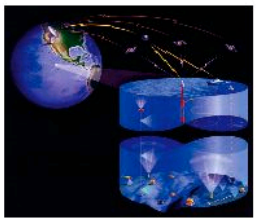
- Two award recommendations to be presented to National Science Board for approval Dec. 9-10, 2008.
- Each award ~$20 million ($4M per year for 5 years), with potential renewal for another $10M over 5 years.
- Strong international component, involvement of multiple science domains
- Second round pre-proposals were due Nov 13, and 2-3 additional awards are expected
- DataNet solicitation: http://www.nsf.gov/funding/pgm_summ.jsp?pims_id=503141&org=OCI&from=home
- Related information: http://www.nsf.gov/events/event_summ.jsp?cntn_id=110367&org=OCI and http://www.nsf.gov/pubs/2008/nsf08021/nsf08021.jsp

Why is DataNet Important to YOU?

- First – data preservation of 50 years+
- Second – DataNet is not just about data preservation - it is about re-use and repurposing of data, including new modes of data analysis (and visualization).
- Always – Grand challenge science in the 21st century - How can visualization/visual interaction help?

NSF Visualization Challenge

- You can *do* science without graphics. But it's very difficult to *communicate* it in the absence of pictures.
- Some insights can only be made widely comprehensible as images
  - Fractal geometry
  - Double helix
  - Solar flares
  - Cosmic microwave background
- To the general public, these and scores of other indispensable science 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

Making Sense of the Bits

- Most humans will read 1-5 gigabytes of text in lifetime of reading
  - And that's just reading
  - Only way to keep up with proliferation of information is through visualization
- Humans are not good at detecting patterns in numbers (think in tables)
  - Observatories are creating petabytes of readings a day

Visualizing Data?

- We collect data to help us understand some phenomenon
  - Data are simply a signal – evidence about the phenomenon
  - It is the phenomenon itself that is important to us, not the data per se.
- Visualizing data can help recognize outliers or detect errors in processing
  - Valuable, but not sufficient.
- Need to use the data to create visualizations of the phenomena that motivated us to capture the data

Vision and Visualization: Why Visualization Isn't Enough

- Cognitive psychologists tell us that humans learn best when multiple senses are engaged simultaneously
  - Personal experience, I know that I learn most effectively when I take detailed notes.
  - How about being a passenger in a vehicle?
  - Cognition study from studies with kittens
    - [Held, R. and Hein, A (1963) Movement produced stimulation in the development of visually guided behavior. Journal of Comparative and Physiological Psychology, 56, 872-876]
- We need to go beyond simple visualization
  - Support interaction to engage the human cognitive system.

What's the Point?!

- Beyond Visualization
  - Support interaction to engage the human cognitive system.
- Understand the Brains as well as the Bits
  - Recognize human cognitive limitations and perceptual characteristics
    - Change blindness
    - Impairments in color perception
    - Cognitive bias
      - Especially limitations in statistical reasoning

Where's the Viz?

- Why no visualization programs in OCI?
  - Where is the $20 million per year visualization program?
  - Where is the large program aimed at providing usable interfaces for science and engineering research?
- Where are they at NSF?
  - As many as 30 NSF Program Directors fund projects that develop visualizations

What Does It Take?

- How NSF creates new programs
  - Workshop reports
  - National Academies reports
- Steps thus far
  - Workshop at NSF in September 2007
  - Follow-on workshop at IEEE Vis, October 2007
  - Workshop at RENCI in April 2008

Enabling Science through Visual Exploration

- September 2007 workshop, held at NSF
  - http://www.visualizingscience.org/
- Organizers: Kelly Gaither, David Ebert, and Chris Gilpin
- Focus on science and what scientists need to facilitate “break through” science
- Goal: Discussing, organizing, and collecting the science and engineering challenges related to data analysis and visualization

Sept 2007 cont.

- Three themes
  - Grand Challenge science
  - Impediments to knowledge discovery
  - Designing sustainable models for integrating data visualization and analysis into the science pipeline
- "Visualizing science is not for the express purpose of generalizing imagery; rather it is a means to better understand and advance the underlying science"

Sept 2007 cont.

- Identified 10 areas to enable next generation science including:
  - Ubiquitous, portable, usable reliable viz tools
  - Interaction environments
  - Interoperability
  - Multi-scale interactions
  - Feature detection
  - Conversational interfaces

This workshop lead to a follow on at IEEE Viz 2008

Scientific Workflow with Immersive Interfaces for Visualization

- IEEE Visualization, October 2008
  - http://cave-wiki.dri.edu/vrvis/index.php/Main_Page
- Organizers: William Sherman, Patrick O'Leary, Oliver Kreylos, Rachael Brady
- Bring together domain scientists and visualization researchers collaborators
  - Looking to immersive interfaces as a means of better serving the domain scientist

Human-Computer Interaction for 21st Century Discovery

- April 2008 workshop, ReNCI
- Organizer: Marilyn Lombardi
- Focus was on building an understanding of the usability challenges scientific research env'ts
  - Discovery dependent on interactions
  - Collaboratory environments
  - Data Deluge

April 2008 Cont.

- Goal: Boundaries between the physical and the virtual shift and disappear
  - Computing becomes less obtrusive to discovery
  - HCI becomes far more effortless and intuitive

April 2008 cont.

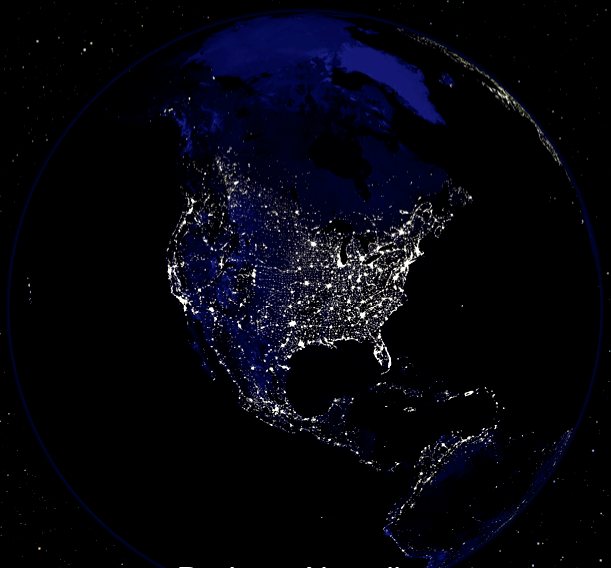
- Next steps push for development of “Third Paradigm” for collaboration at all levels
  - Inclusion of citizen scientists
- Networks of interactions
  - Platform independent
  - Based around communities of interest
  - Tools for seamless pattern seeking and new media creation

What Can You Do?

- Help to articulate a detailed research agenda for ultrascale visualization, especially the cross-cutting aspects that can be cyberinfrastructure.
- Propose more workshops aimed at bridging the gaps between the visualization community and the science/engineering user communities.
- Develop metrics that show the impact of visualization and visual interaction - and evaluate your systems!
- Propose to OCI's Strategic Technologies for Cyberinfrastructure (STCI) (Feb 12, 2009).

- The book Illuminating the Path: The Research and Development Agenda for Visual Analytics can be downloaded from the web site of the National Center for Visualization and Analytics at Pacific Northwest National Laboratory, at http://nvacl.pnl.gov/agenda.stm. The site also features a related movie.

Thank you!



Dr. Lucy Nowell
lnowell@nsf.gov