

Berlin 5

**Padova, Italy
September, 2007**

Chris Greer
U.S. National Science Foundation



U.S. President

Office of Management and Budget

Science Advisor
Office of Science and Technology Policy

Other boards, councils, etc.

Major Departments

Agriculture

Health and Human Services

Interior

Homeland Security

Defense

Energy

Commerce

Independent Agencies



National Aeronautic and Space Administration

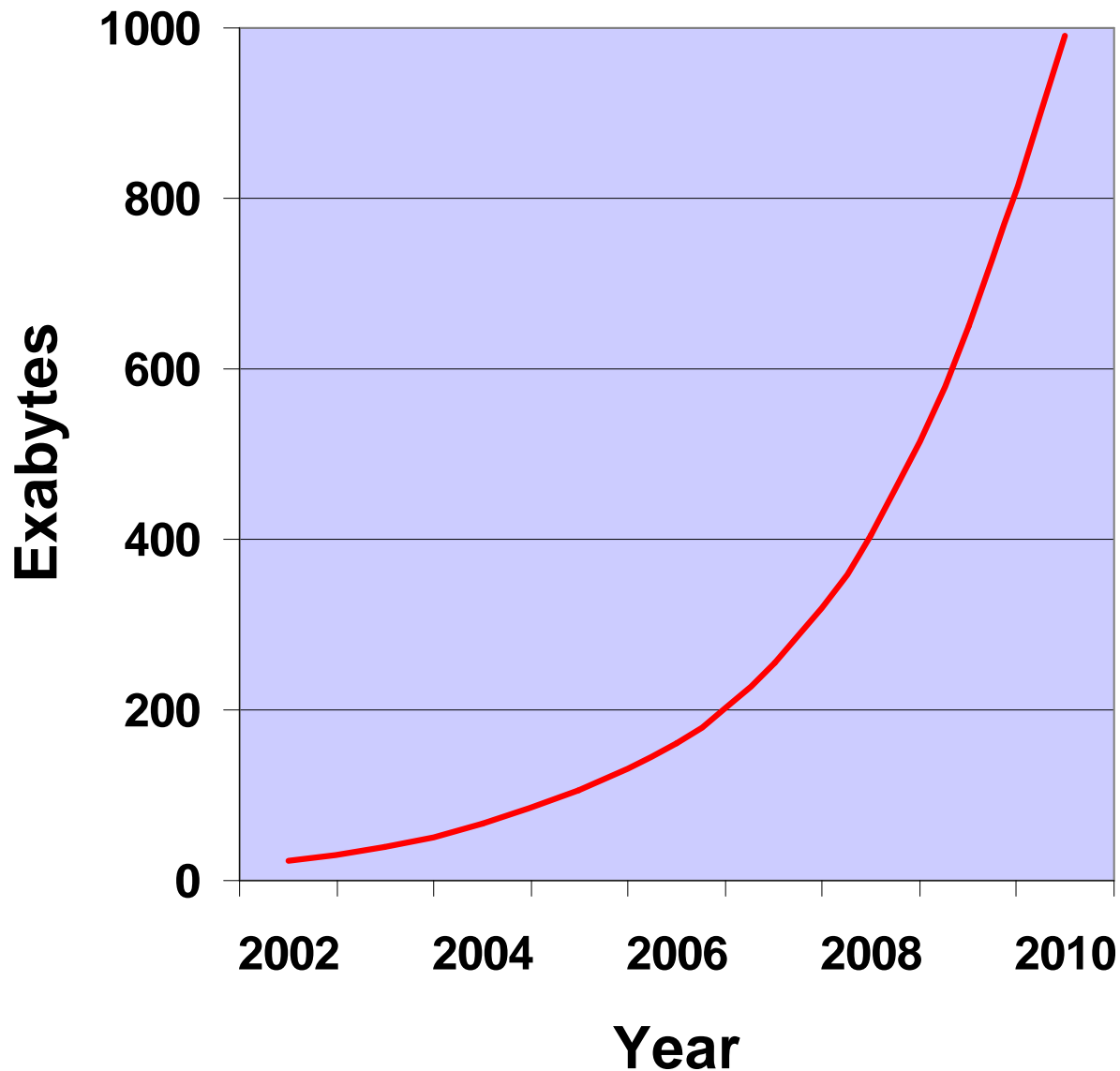
Environmental Protection Agency

Smithsonian Institution

Nuclear Regulatory Commission

Other agencies

Estimated Annual Digital Information Totals

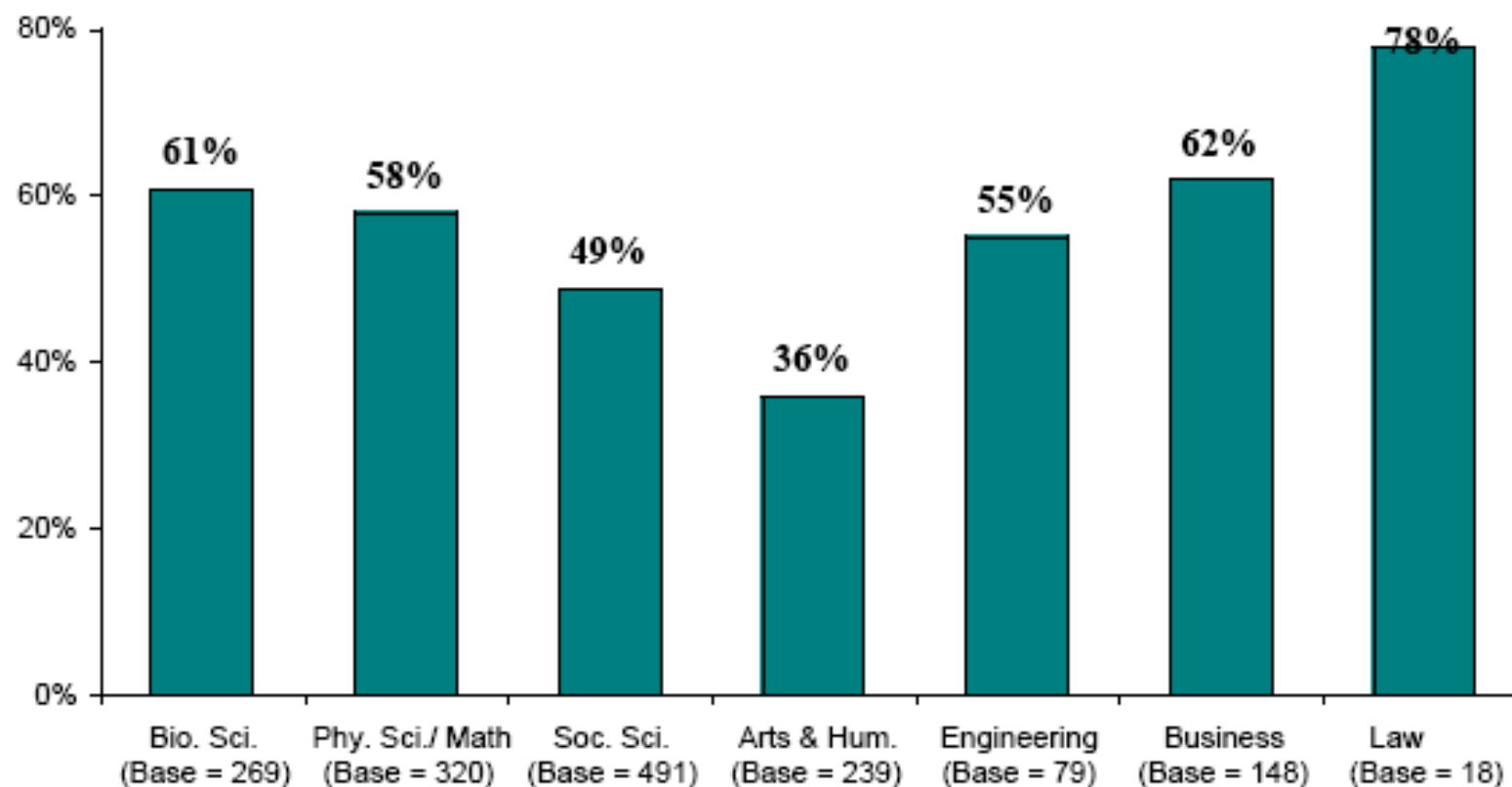


Sources:

- The Expanding Digital Universe, March 2007, IDC White Paper sponsored by EMC Corporation, www.emc.com/about/destination/digital_universe/

How Much Information? 2003, Peter Lyman and Hal Varian, Berkeley School of Information, <http://www2.sims.berkeley.edu/research/projects/how-much-info-2003/>

Percent of Research Information Retrieved Online



Source: DLF/Outsell Survey, 2002, Cited in Access in the Future Tense, CLIR, 2004

What are the challenges for preservation and access to primary digital scientific data (including raw data, pre-publication data, materials, algorithms, and computer codes)?

What should be saved, for how long, and who decides?

What means can be used to assess the value of data whose most important use may lie years in the future for unforeseen purposes?

For preservation of software (including models and simulations):
When is emulation adequate and when is preservation of the original hardware and operating systems required?

What information (metadata) about data processing, filtering, transformations, workflows, and other manipulations should be saved and how can it be linked reliably to the original data?

What physical media are the best choice for digital archiving?

Who owns (NSF-funded) research data?

Bayh-Dole Act* or University and Small Business Patent Procedures Act

“Since 1980, the Bayh-Dole Act and subsequent executive actions generally have given funding recipients the option to retain ownership rights to, and profit from, commercializing the inventions they create as part of federally sponsored research projects. In return for these rights, they are required to file for patent protection, pursue commercialization of the inventions, give preferences to small businesses in licensing, ensure that any products resulting from the inventions are substantially manufactured in the United States, and comply with certain reporting requirements.”#

*35 U.S.C. § 200-212; 37 C.F.R. 401 December, 1980

#GAO-03-536 Technology Transfer

Grant Proposal Guide (NSF 04-23)

L. LEGAL RIGHTS TO INTELLECTUAL PROPERTY

NSF normally allows grantees to retain principal legal rights to intellectual property developed under its grants. This policy provides incentive for development and dissemination of inventions, software and publications that can enhance their usefulness, accessibility and upkeep. **It does not, however, reduce the responsibility of researchers and organizations to make results, data and collections available to the research community.**

To what extent is a (NSF-funded, non-government) scientist responsible for supplying research data to other scientists?

Grant Proposal Guide (NSF 04-23)

NSF expects significant findings from supported research and educational activities to be promptly submitted for publication with authorship that accurately reflects the contributions of those involved. **It expects PIs to share with other researchers, at no more than incremental cost and within a reasonable time, the data, samples, physical collections and other supporting materials created or gathered in the course of the work.** It also encourages grantees to share software and inventions, once appropriate protection for them has been secured, and otherwise act to make the innovations they embody widely useful and usable.

Grant Proposal Guide (NSF 04-23)

NSF program management will implement these policies, in ways appropriate to field and circumstances, through the proposal review process; through award negotiations and conditions; and through appropriate support and incentives for data cleanup, documentation, dissemination, storage and the like. Adjustments and, where essential, exceptions may be allowed to safeguard the rights of individuals and subjects, the validity of results and the integrity of collections, or to accommodate legitimate interests of investigators.

What are the policies for making technical and scientific government information available?

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.

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;

*44 U.S.C. 3501 *et seq.*

What are the current standards for accessing and maintaining research data and how should these evolve in the future?

Enabling Communities of Practice

Policies and strategies to facilitate the management, preservation, and sharing of digital data will have to fully embrace the essential diversity in digital data collections. This diversity arises from many sources including ... differences in user needs, expectations, and access procedures.

Examples of Data Policies

- Science of Science Policy
<http://www.nsf.gov/pubs/2007/nsf07547/nsf07547.htm>
- Earth Sciences
http://www.nsf.gov/geo/ear/EAR_data_policy_204.pdf
- Social and Economic Sciences
<http://www.nsf.gov/sbe/ses/common/archive.jsp>
- Polar Programs
http://www.nsf.gov/publications/pub_summ.jsp?ods_key=opp991
- Long Term Ecological Research (LTER) Community Data Policy
<http://www.lternet.edu/data/netpolicy.html>
- Ocean Sciences
http://www.nsf.gov/pubs/2004/nsf04004/nsf04004_1b.htm

Examples of Data Policies

Science of Science Policy: To ensure efficient accessibility of new data, metrics and indicators that are developed via this SciSIP competition, all research proposals that develop new datasets must include a data management plan.

Requirements for the data management plan:

- Statement regarding where data will be archived.
- Identification of the data management point of contact.
- Clear indication of which data are community data. Community data must be made available through an openly accessible data management system as soon as data are collected and verified.

Examples of Data Policies

LTER Network Community Data Release Policy

Data and information derived from publicly funded research in the U.S. LTER Network, totally or partially from LTER funds from NSF, Institutional Cost-Share, or Partner Agency or Institution where a formal memorandum of understanding with LTER has been established, are made available online with as few restrictions as possible, on a nondiscriminatory basis.

Overall philosophy for NSF Data Policies

NSF Grant Proposal Guide (NSF 04-23):

[NSF] expects PIs to share with other researchers, at no more than incremental cost and within a reasonable time, the data, samples, physical collections and other supporting materials created or gathered in the course of the work.

U.S. Interagency Working Group on Digital Data

IWG Co-Chairs:

Cita Furlani, NIST

Chuck Romine, OSTP

Chris Greer, NSF

Membership:

- Department of Agriculture
- Department of Commerce
- Department of Defense
- Department of Education
- Department of Energy
- Department of Health and Human Services
- Department of Homeland Security
- Department of the Interior
- Department of Labor
- Department of Justice
- Department of State
- Department of Transportation
- Department of the Treasury
- Department of Veterans Affairs
- Central Intelligence Agency
- Environmental Protection Agency
- Library of Congress
- National Aeronautics and Space Administration
- National Archives and Records Administration
- National Science Foundation
- The Smithsonian Institution
- US Army Corps of Engineers
- Council on Environmental Quality
- Domestic Policy Council
- Homeland Security Council
- National Economic Council
- National Security Council
- Office of Management and Budget
- Office of Science and Technology Policy

Purpose:

The purpose of the IWG is to develop and promote the implementation of a strategic plan for the Federal government to cultivate an open interoperable framework to ensure reliable preservation and effective access to digital data for research, development, and education in science, technology, and engineering.