

***Update:* Blue Ribbon Panel on  
Sustainable Digital Preservation and  
Access**

***Dr. Fran Berman***

*BRTF-SDPA Co-Chair*

*with thanks to Sayeed Choudhury, Brian Lavoie, Amy Friedlander, David  
Minor, Ardys Kozbial, and Robert McDonald*

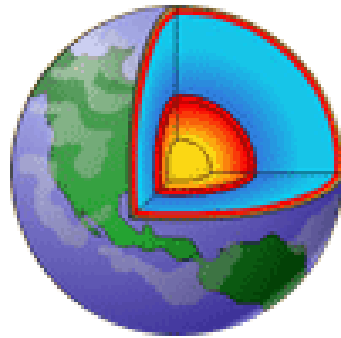
## Presentation Outline

- Digital preservation and economic sustainability
- Blue Ribbon Task Force
  - History, membership, charge and focus
  - Deliverables
  - Activities
- A Call to Action

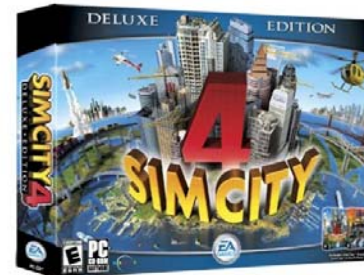
# The Digital World



**Research and Education**



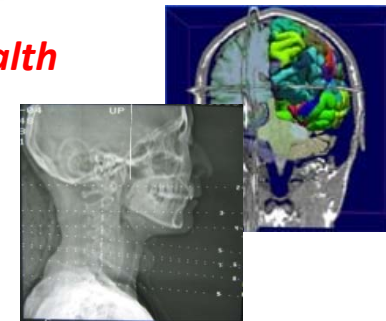
**Shopping**



**Entertainment**



**Health**



**Information**



**Business**

Blue Ribbon Task Force on Sustainable Digital Preservation and Access

# Digital Data is the natural resource of the Information Age

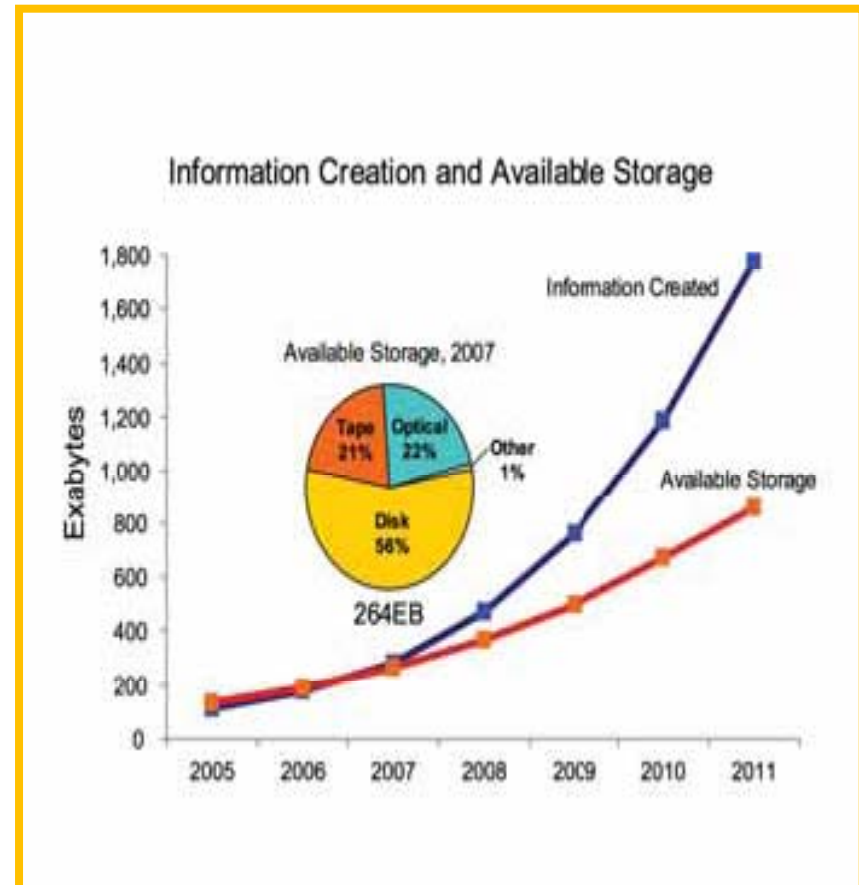
- **Digital data is a natural resource**
  - Fragile— dependent on rapidly changing technologies
  - Valuable data lost or damaged often cannot be recovered
- **100 years of digital data represents**
  - Dozens of new generations of technologies
  - 100's of new data standards and formats in many communities
  - Thousands+ of new valued collections
  - Millions of potential users with as yet unknown information needs and workflows



**Blue Ribbon Task Force on Sustainable Digital Preservation and Access**

## IDC Projects Indicate that the Time for Strategic Planning for Data Preservation is *NOW*

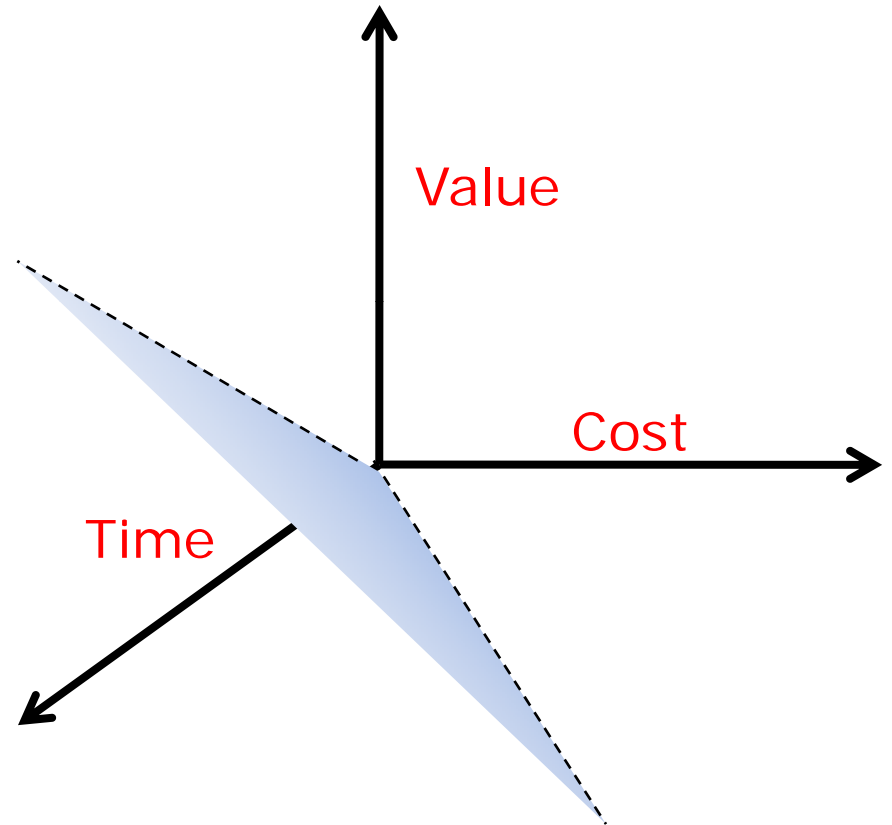
- **2007 was the “crossover year”** where the amount of digital information is greater than the amount of available storage
- **Increasing requirements for data retention** in private, public, academic sectors render data infrastructure and preservation policies more critical
  - Agency requirements for data management and preservation, Sarbanes-Oxley, etc.
  - Trust management, etc.



Source: “The Diverse and Exploding Digital Universe” IDC Whitepaper, March 2008

# Data Preservation

- Key Questions:
  - 1) What should we save it?
  - 2) How and who should save it?
  - 3) Who should pay for it?
- The answer to 3) depends in part on the answers to 1) and 2).



# What should we save?



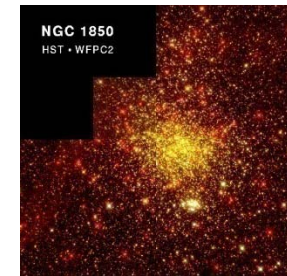
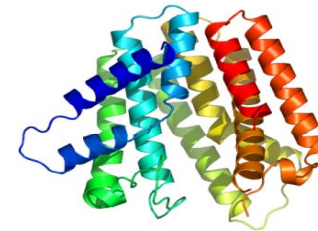
## Data we\* want to keep over the long-term:

- **We = “Society”**
  - Official and historically valuable data (Census information, presidential emails, Shoah Collection, etc.)
- **We = Research Community**
  - Protein Data Bank, National Virtual Observatory, PSID, etc.
- **We = Me**
  - My medical record, my Quicken data, digital photos of my Mom’s 80<sup>th</sup> birthday, etc.



Area	Population	Change	Rate	Rate	Rate
Alameda	1,100,000	100,000	10%	10%	10%
Alameda	1,100,000	100,000	10%	10%	10%
Alameda	1,100,000	100,000	10%	10%	10%
Alameda	1,100,000	100,000	10%	10%	10%
Alameda	1,100,000	100,000	10%	10%	10%
Alameda	1,100,000	100,000	10%	10%	10%
Alameda	1,100,000	100,000	10%	10%	10%
Alameda	1,100,000	100,000	10%	10%	10%
Alameda	1,100,000	100,000	10%	10%	10%
Alameda	1,100,000	100,000	10%	10%	10%

**Sarbanes-Oxley**  
Financial and Accounting Disclosure Information



**Blue Ribbon Task Force on Sustainable Digital Preservation and Access**

# Who should pay?

- The “Free Rider” *non*-solution: “Let X do it” where X is:

- The Government
- The Libraries
- The Archivists
- Google, Yahoo, Microsoft, etc.
- Data users
- Data owners
- Data creators, etc.

- For valued digital data that is a “public good”, creative partnerships needed to provide preservation solutions with
  - Trusted stewards
  - Feasible costs for users
  - Sustainable costs for infrastructure
  - Very low risk for data loss, etc.

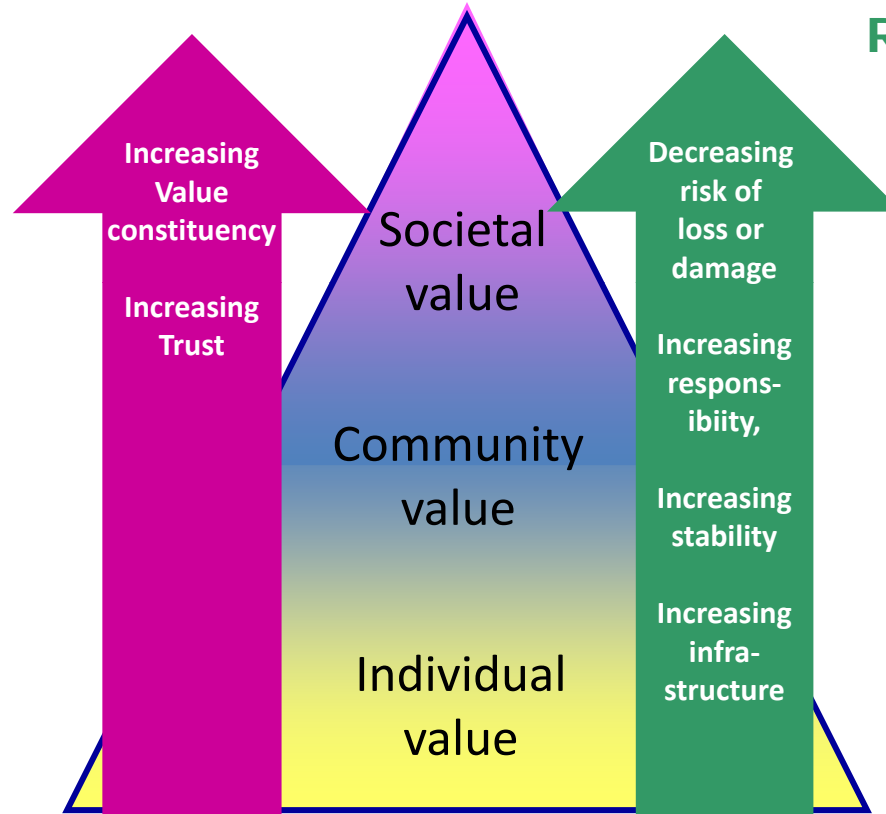
# A Framework for Thinking about Digital Stewardship and Preservation

## Digital Data Collections

Reference, nationally/ internationally important, irreplaceable data collections

Key research and community data collections

Personal data collections



## Repositories / Facilities

National / international scale repositories, libraries, archives

“Regional” scale libraries and targeted data archives / centers

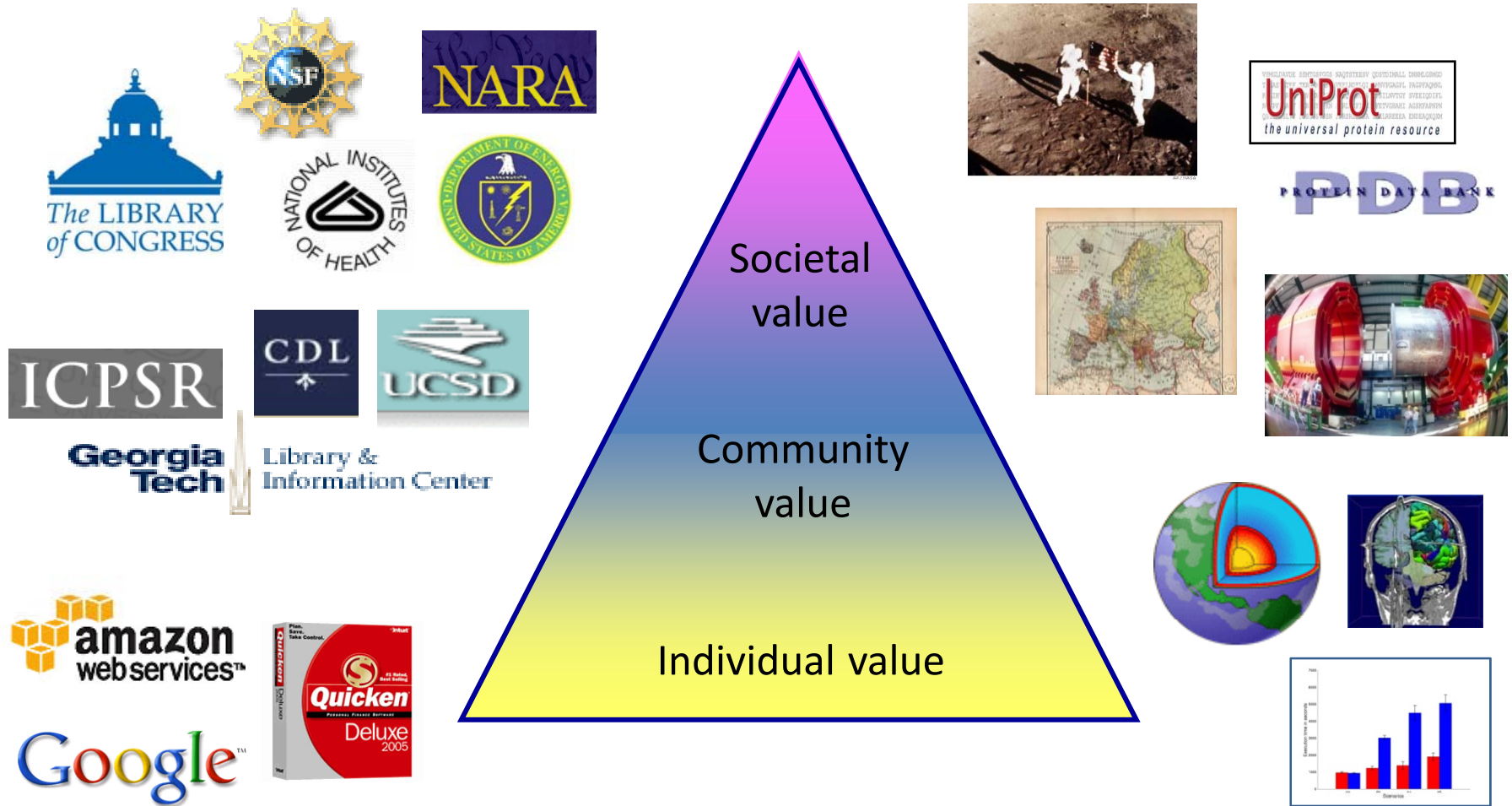
Private repositories

The Data Pyramid

Blue Ribbon Task Force on Sustainable Digital Preservation and Access

# Why not just let **GOOHOOsoft** do it?

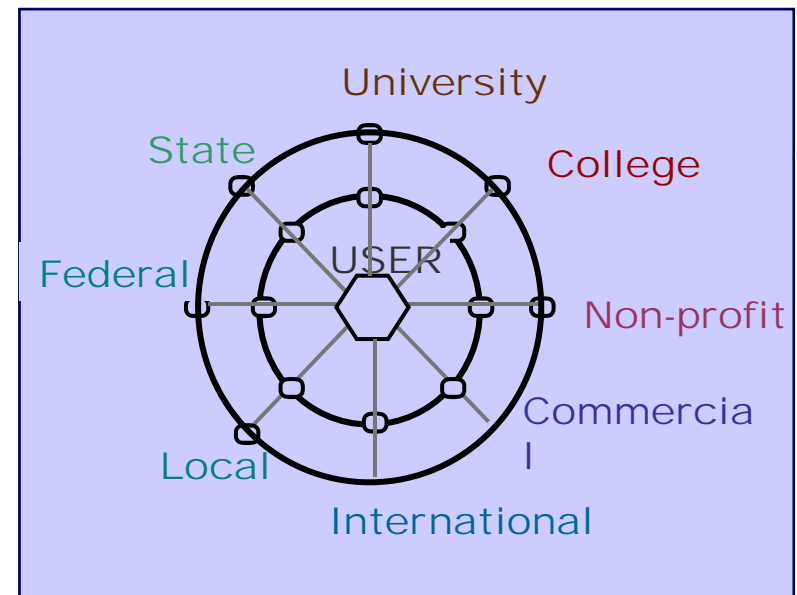
## Multi-sector solutions required for data in the public interest



Blue Ribbon Task Force on Sustainable Digital Preservation and Access

# Blue Ribbon Task Force on Sustainable Digital Preservation and Access

- **History:**
  - International **Blue Ribbon Task Force (BRTF-SDPA)** initiated to study issues of economic sustainability of digital preservation and access
  - Proposals developed during Spring-Summer 2007
  - BRTF-SDPA pre-planning took place Fall 2008
  - Kick-off meeting in DC held in January 2008
- **Task Force meeting 2008-2009**
- **BRTF is an independent group who will report to the community**



*Image courtesy of Chris Greer,  
October 2006*

Blue Ribbon Task Force on Sustainable Digital Preservation and Access

## BRTF-SDPA Charge

1. To conduct a **comprehensive analysis of previous and current efforts** to develop and/or implement models for sustainable digital information preservation; **(First year report)**
2. To identify and **evaluate best practice** regarding sustainable digital preservation among existing collections, repositories, and analogous enterprises;
3. To **make specific recommendations for actions** that will catalyze the development of sustainable resource strategies for the reliable preservation of digital information; **(Second Year report)**
4. Provide a **research agenda** to organize and motivate future work.

# BRTF-SDPA Participants

## Blue Ribbon Task Force:

- Paul Ayris, University College London
- Fran Berman, SDSC/UCSD
- Bob Chadduck, NARA Liaison
- Sayeed Choudhury, Johns Hopkins University
- Elizabeth Cohen, AMPAS/Stanford
- Paul Courant, University of Michigan
- Lee Dirks, Microsoft
- Amy Friedlander, CLIR
- Chris Greer, NITRD Liaison
- Vijay Gurbaxani, UC Irvine
- Anita Jones, University of Virginia
- Ann Kerr, Consultant
- Brian Lavoie, OCLC
- Cliff Lynch, CNI
- Dan Rubinfeld, UC Berkeley
- Chris Rusbridge, DCC
- Roger Schonfeld, Ithaka
- Abby Smith, Consultant
- Anne Van Camp, Smithsonian

## Sponsoring Agencies/Institutions:

- National Science Foundation
- Mellon Foundation
- Library of Congress
- National Archives and Records Administration
- CLIR
- NITRD
- JISC
- Member institutions

## Specific Responsibilities

- Fran Berman / co-Chair
- Amy Friedlander / First Report Editor
- Ann Kerr / January Panel Rapporteur
- Brian Lavoie / co-Chair
- Susan Rathbun / Task Force Support
- Abby Smith / Second Report Editor
- Jan Zverina / Communications Lead
- Lucy Nowell / NSF Program Officer
- Don Waters / Mellon Program Officer
- Laura Campbell, Martha Anderson / LC representative

## BRTF Deliverables

- **First Year (2008) Report (positive, “what is”):**
  - Describe past and current models (case studies, etc.);
  - Identify points of convergence/divergence; “lessons learned”;
  - What we know so far, and what our key knowledge gaps are.
- **Second Year (2009) Report (normative, “what should be”):**
  - General cost framework: key cost categories of digital preservation
  - Set of economic models/“scenarios”: alternate ways of organizing digital preservation activities, within the context of the cost framework
  - Describe each model: features, pros, cons, trade-offs, etc.
  - List real world conditions for which each model is best suited.
    - “If your digital preservation context is X, we recommend you consider using model Y to organize your activities in a sustainable way.”

# BRTF-SDPA Working Definition of Economic Sustainability

- **Economic sustainability in a digital preservation context:** *The set of business, social, technological, and policy mechanisms that*
  1. *Encourage the gathering of important information assets into digital preservation systems, and*
  2. *Support the indefinite persistence of the digital preservation systems, thus securing access to and use of the information assets into the long-term future*
  
- **Economically sustainable digital preservation requires**
  - Recognition of the benefits of preservation on the part of key decision-makers, as part of a process of selecting digital materials for long-term retention;
  - Appropriate incentives to induce decision-makers to act in the public interest;
  - Mechanisms to secure an ongoing allocation of resources, both within and across organizations, to digital preservation activities;
  - Efficient use of limited preservation resources;
  - Appropriate organization and governance of digital preservation activities.

## Scope of BRTF-SDPA

- **BRTF focusing on economic sustainability of digital data for which there is a clear public interest**
  - Materials may come from a variety of sectors and in a variety of forms (text, audio, video, website, scientific data set, e-journals, etc.)
- **Some of this data can/will be identified at its inception as data of long-term value, but for some, retention value only becomes apparent over time.**
  - Economic sustainability models must be developed / identified in the context of changing arrangements between multiple missions and models for various sectors
  - Target audiences and relevant domains may also vary over time
  - *Key topic for consideration:* transfer of preservation responsibilities and associated digital content from one party to another party (or sets of parties) based on trigger or milestone events and changes in perceptions of value.

## Challenges in Developing an Economic Model

- Difficult to compare and contrast different studies given diverse definitions of costs, units of measurement, scope or accounting methods
- Need to include both costs associated with human dimensions (e.g., management) and infrastructure elements of storage (e.g., power and cooling) relevant
- “Format matters; scale matters”
- *Key related work:* relevant and related recent studies focusing on various “parts of the elephant” (e.g. LIFE, BCL cost models which focus on entire life-cycle of costs as well as inter-temporal considerations (e.g., inflation/deflation, interest rates))

## Key Deliverable for Year 1: Report which describes Current and Best Practices for Economic Sustainability

- **2008 Report will focus on**
  - Key issues for economic sustainability of digital preservation and access
  - Key concepts and terminology
  - Relevant literature
  - Compilation and analysis of evidence on existing efforts to implement economically sustainable digital preservation activities
- **The purpose is to provide a thorough survey of the problem space as it exists today,** which in turn will provide context for and inform the work of BRTF in the second year.
- BRTF Year 1 Report is currently being drafted by Task Force members and Editor Amy Friedlander. Due out by January 2009.

## What We'd Like to Know from the Community: Questions on Economic Sustainability for Digital Preservation

- **What is the nature of the digital materials being preserved** (source(s), content, volume, format, copyright restrictions, frequency of use ...)?
- **Who are the stakeholders for these materials?** (Who benefits from the preservation of the target materials?)
- **What is the “value proposition” for this preservation effort?**
  - Why are stakeholders interested in the long-term preservation of the materials?
  - What are the anticipated future uses of the materials?
  - Is the “value proposition” perpetual, or does it expire within a finite time frame?
- **What are the key features of long-term preservation for these materials?**
  - What are the elements of “successful preservation” for these materials? E.g., bit preservation only? Format migration? Rich description/documentation with metadata?
  - What are the key features of a preservation service that would meet the preservation requirements of the stakeholder community.
- **What are the “economic aspects” of digital preservation?**
  - Given the preservation requirements for the materials, what are the key cost categories associated with a preservation strategy aimed at meeting these requirements?
  - Does the organization responsible for preservation also have a stake in the ongoing availability of the materials, or is it performing a “service for hire”?
  - What are the mechanisms by which funding is generated for digital preservation activities? How are “paying stakeholders” identified?

## 2008 BRPanels

### ■ *January Panel*

#### *Speakers:*

- John Gantz, IDC
- Stijn Hoorens, RAND
- Andy Maltz, AMPAS
- Edmund Mesrobian, RealNetworks

### ■ *July Panel Speakers:*

- Helen Berman, Protein Data Bank
- Kris Carpenter, Internet Archive
- Myron Gutmann, Interuniversity Consortium for Political and Social Research
- Rick Luce, Emory Library
- Paul Ratnaraj, Wharton Research Data Services
- Nan Rubin, PBS
- Rick Zuray, Boeing
- TBD, Microsoft

## Other Venues for Information Gathering

- **Testimony** – template will be available on the website for community input on economic sustainability
- **In-depth interviews** – with community leaders
- **Literature search** – Existing and new raft of articles on economic sustainability for digital preservation and other materials provide a rich resource
- **Task Force members** – diverse, international group representing broad spectrum of sectors, perspectives, institutions

# Other Deliverables: BRTF-SDPA Website ([brtf.sdsc.edu](http://brtf.sdsc.edu))

Blue Ribbon Task Force  
on Sustainable Digital Preservation and Access

About Us | Members | Bibliography | News Center | Intranet | Contact Us

*This is the only group I know of that is chartered to help the community understand the economic issues surrounding sustainable repositories and identify candidate solutions*

— Lucy Nowell,  
Program Director  
Office of Cyberinfrastructure, NSF

### Goals

- Conduct an analysis of previous and current models for sustainable digital preservation, and identify current best practices among existing collections, repositories and analogous enterprises.
- Develop a set of economically viable recommendations to catalyze the development of reliable strategies for the preservation of digital information.
- Provide a research agenda to organize and motivate future work in the specific area of economic sustainability of digital information.

### Sponsors

- NSF
- SDSC  
SAN DIEGO SUPERCOMPUTER CENTER
- The Andrew W. Mellon Foundation
- OCLC
- Digital Preservation
- ERA
- JISC
- Council on Library and Information Resources

About Us | Members | Bibliography | News Center | Intranet | Contact Us

unded by the National Science Foundation and the Andrew W. Mellon Foundation, in partnership with the Library of Congress, the Joint Information Systems Committee of the United Kingdom, the Council on Library and Information Resources, and the National Archives and Records Administration.  
© 2008 Blue Ribbon Task Force on Sustainable Digital Preservation and Access

Internet

- BRTF-SDPA website being developed as a community resource for
  - Information about the Task Force
  - Input from the community
  - literature related to Digital Preservation

- *Templates for “testimony” to be available on the website*
- *Suggestions for inclusions for the bibliography encouraged*

Blue Ribbon Task Force  
on Sustainable Digital Preservation and Access

About Us | Members | Bibliography | News Center | Intranet | Contact Us

### Bibliography

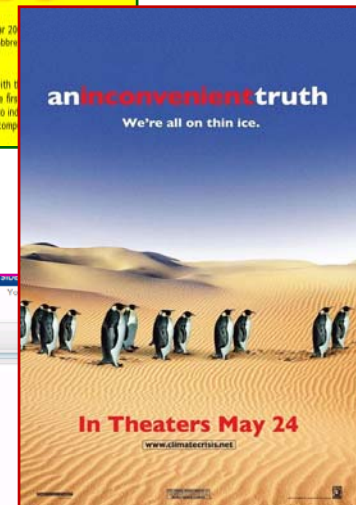
- 100 Year Archive Task Force (2004)  
Defining Storage Practices and Standards for Long-term Digital Information Retention  
SNIA Data Management Forum.
- 100 Year Archive Task Force - Overview  
SNIA, September, 2006.
- APSR Sustainability Issues Discussion Paper  
BRADLEY, K.; National Library of Australia (2005)
- Addressing the uncertain future of preserving the past  
Towards a robust strategy for digital archiving and preservation  
Hoarens, Rothenberg, van Oranje, van der Mandele, Levin, 2007.
- Agenda for Developing E-Science in Research Libraries  
ARL e-science report, November 2007.

Trusted sites

Blue Ribbon Task Force on Sustainable Digital Preservation and Access

# Call to Action for both BRTF and the NDIIPP Community: Get the message out beyond the “usual suspects”

- Does your dry cleaner or your neighbor know what digital preservation is?
- We can help focus attention and encourage new resources from decision makers and the general public if we get the message out



Blue Ribbon Task Force on Sustainable Digital Preservation and Access