



D | C | C

Because good research needs good data

Sustainable Digital Preservation and Access

Blue Ribbon Task Force

Presentation to Alliance for Permanent
Access

Chris Rusbridge (based on slides by Brian
Lavoie, OCLC and Fran Berman, RPI)

Funded by:

JISC



D | C | C

Because good research needs good data

Roadmap

- Blue Ribbon Task Force on Sustainable Digital Preservation and Access
- Sustainability (economically speaking)
- Interim Report highlights
- Economic Framework
- Scenarios
- Research data
- Recommendations



D | C | C

Because good research needs good data

Blue Ribbon Task Force

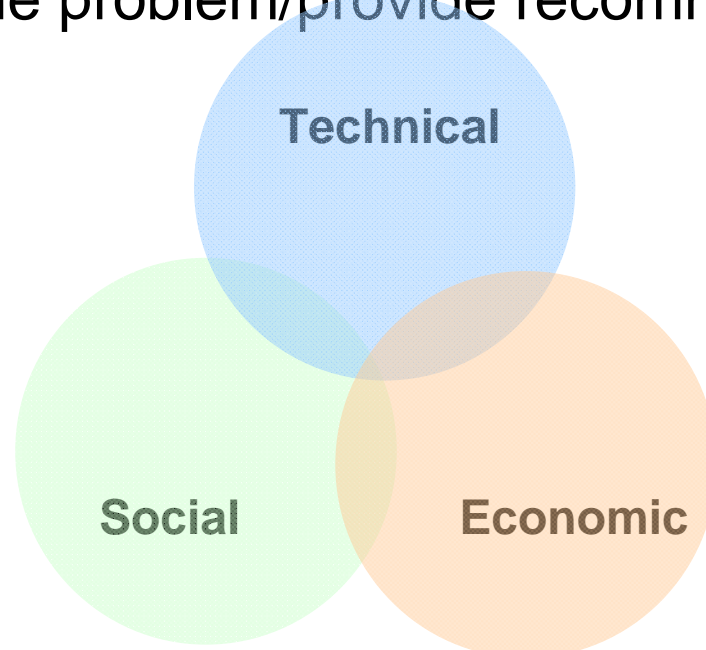
- Task Force:
 - Supported by NSF, Mellon, LoC, JISC, CLIR, NARA
 - Co-chairs: Brian Lavoie (OCLC), Fran Berman (SDSC/RPI)
 - Cross-domain, cross discipline. <http://brtf.sdsc.edu/>
- Comprehensive analysis of sustainable digital preservation
- Identify and evaluate best practice
- Make specific recommendations for action
- Provide a research agenda for further work



D | C | C

Because good research needs good data

- Frame digital preservation as *sustainable economic activity*
 - Economic activity: deliberate allocation of resources
 - Sustainable: ongoing resource allocation over long periods of time
 - Articulate the problem/provide recommendations & guidelines



•Lavoie

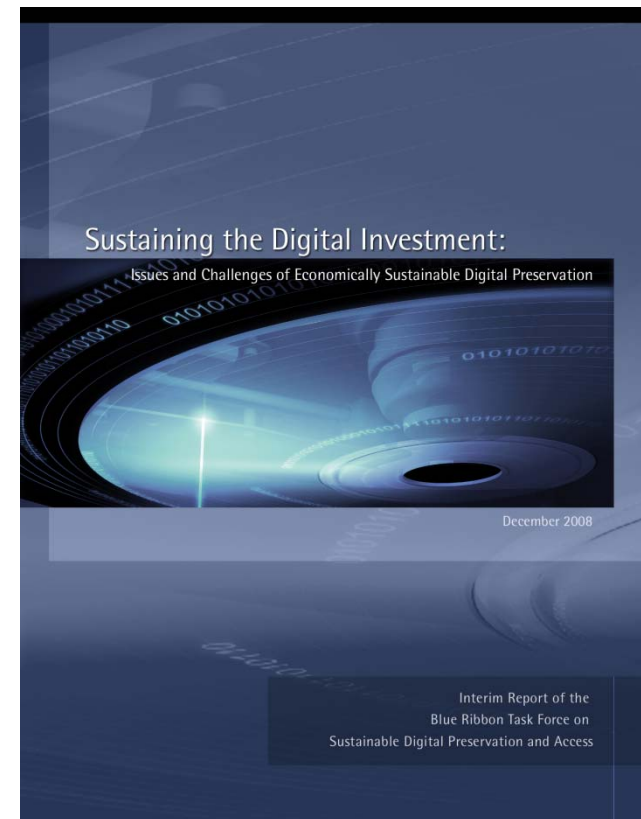


D | C | C

Because good research needs good data

Task Force Interim Report (December 2008)

Sustaining the Investment:
Issues and Challenges of
Economically Sustainable Digital
Preservation



http://brtf.sdsc.edu/biblio/BRTF_Interim_Report.pdf •Lavoie

Alliance for Permanent Access November 2009



D | C | C

Because good research needs good data

Definition: economic sustainability

Economically sustainable digital preservation requires:

- *Recognition of benefits*
- *Incentives for decision-makers to act*
- *Selection*
- *Mechanisms to support ongoing, efficient allocation of resources*
- *Appropriate organization and governance*

•Lavoie



Because good research needs good data

Benefits & Incentives

- Clearly articulate benefits of digital preservation activity
 - “Value proposition” for digital preservation
 - Benefits should emphasize outcomes
 - Articulate benefits → cultivate sense of value, “willingness to pay”
- Clearly articulate incentives for decision-makers to act
 - Accept responsibility to undertake preservation
 - Identify and leverage institutional “self-interest”: e.g., business opportunity; mission-driven; policy compliance
 - Orchestrate incentives over complete digital lifecycle

•Lavoie



D | C | C

Because good research needs good data

Selection & Allocation of Resources

- Selection: can't "preserve everything for all time"
 - Prioritization: allocate resources where they generate most value
 - Circumscribed set of materials; realistic preservation goals
 - Manage expectations; align expectations and capacity
- Support ongoing, efficient allocation of resources
 - Coordinate resource transfer from those who are willing to pay to those who are willing to preserve (pricing, donations, fees/taxes)
 - Efficiency: productive use of resources; leverage economies of scale, economies of scope

•Lavoie



Because good research needs good data

Organization & Governance

- Preservation activities can be managed through a variety of **organizational forms**, e.g.:
 - Organization with no private interest in preservation (e.g., third party service)
 - Organization with private interest in preservation; preserves on behalf of itself and other organizations (e.g., research library)
 - Organizations with mandate to preserve, conferred by public policy, to fulfill stated public interest (e.g., national archive)
- **Governance**: strategy, responsibility, accountability
- Organization/governance → **trust**

•Lavoie



D | C | C

Because good research needs good data

Economic sustainability: problem space

Demand-side
VALUE

Beneficiaries

Providers

Supply-side
INCENTIVES

Process
SELECTION

ONGOING/EFFICIENT
RESOURCE ALLOCATION

ORGANIZATION/
GOVERNANCE



Digital Preservation
Activity

•Lavoie



D | C | C

Because good research needs good data

Challenges

- Long-term preservation activities funded by **short-term resource allocations**
- Lack of clear responsibility for digital preservation, and a prevailing assumption it is **someone else's problem**
- **Misaligned incentives** between those who are in a position to preserve, and those who benefit
- **Little coordination** of preservation activity across diffused stakeholder communities
- **Challenges in valuing/monetizing** benefits of digital preservation, to attract funding and investment

•Lavoie



D | C | C

Because good research needs good data

Lessons learned

- Separating preservation costs from other costs is difficult
 - No clear distinction between process of “making things available now” v. “making things available in the future”
 - Presents challenges for segregating digital preservation as separate activity and answering questions like “what does it cost?”
- Monetizing and charging for a “social good”:
 - Public-spirited, mission-driven institutions sometimes resistant to charging for content & services
 - Compelling value expressed in monetary terms, coupled with mechanism for charging reasonable fee to those who share in value

•Lavoie



D | C | C

Because good research needs good data

Lessons learned

- Digital preservation is not just “for the future”
 - Incur costs now for **current and** future benefits
 - Perception: Digital preservation separable from interests of today’s stakeholders; focused on future stakeholders
 - Reality: Digital preservation more about ensuring digital assets are **handed off in good condition** to next stewards 5/10 years from now, than investing to benefit users 100 years hence
- **Non-monetary incentives can be important**
 - Preservation bestows societal benefits to research, learning, culture
 - Engage private enterprise in supporting provision of these benefits
 - Leverage corporate recognition and reputation enhancement

•Lavoie



Because good research needs good data

Real world to theory to recommendations

- Final report: practical recommendations for decision-makers associated with digital preservation activities
- Approach:
 - Collect, organize, synthesize key information from real world digital preservation case studies ...
 - ... to which we can apply economic theory/insight to isolate important economic implications ...
 - ... from which we can derive practical recommendations/guidelines to support economic sustainability

•Lavoie



D | C | C

Because good research needs good data

Key Deliverable for Year 2: BRTF Final report

Represent digital preservation as set of economic factors and conditions that impact sustainability

Economic Framework

Real World Scenarios

Based on a particular set of economic factors and conditions, identify and understand the economic implications in terms of prospects for long-term sustainability

Economic Analysis

Sustainability Definition

Based on economic analysis, provide Practical recommendations for eliminating or mitigating weaknesses or risks for achieving sustainability

Recommendations

•Berman



D | C | C

Because good research needs good data

Economic framework

Describe a digital preservation activity in terms of **essential elements** relevant for thinking about long-term economic sustainability

Digital
Preservation
Activity

CORE ATTRIBUTES

- Intrinsic to all d.p. activities
- Identified by our econ experts

DOMAIN-SPECIFIC ATTRIBUTES

- Apply in some (not all) d.p. activities
- Identified by our domain experts

CHOICES

- For decision-makers/stakeholders
- Identified by our domain experts

•Berman



Because good research needs good data

Generalized digital preservation scenarios

- Base economic analysis on information gathered from real-world digital preservation case studies. But:
 - Every activity “differs in the details”
 - Analysis/recommendations should have wide applicability
- Roll individual case studies up into **generalized scenarios**:
 - Categories of digital preservation activities that at a reasonable level of abstraction, share roughly the same characteristics
- Apply economic analysis to/derive recommendations for the generalized digital preservation scenarios

•Lavoie



D | C | C

Because good research needs good data

Four generalized scenarios

Research Data
(e.g., data sets)

By researchers
For researchers

Scholarly Discourse
(e.g., e-journals)

By researchers
For researchers

Collectively-produced
Web Content
(e.g., blogosphere)

By ... everyone
For ... everyone?

Commercially-owned
Cultural Content
(e.g., movie studio output)

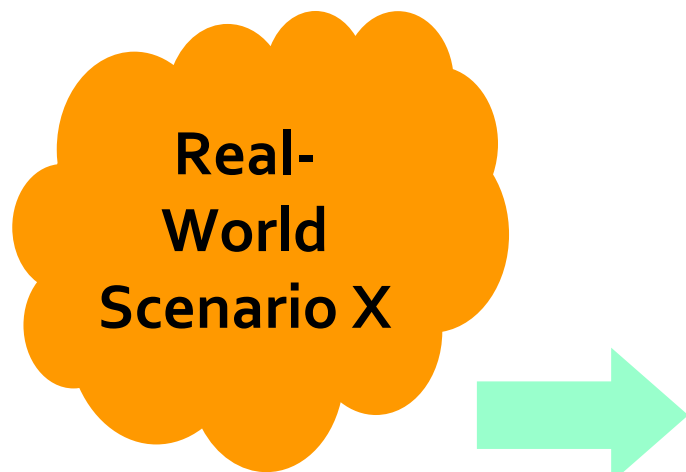
By private entities
For privately controlled
purposes



D | C | C

Because good research needs good data

Abstract the essentials ...



Core Attribute 1 Derived demand
Core Attribute 2 Depreciable durable good
Core Attribute 3 Nonrivalrous consumption
Core Attribute 4 Temporally dynamic
path-dependent

Domain-specific attribute 1
Domain-specific attribute 2

...

Choice 1

Choice 2

...

•Berman



D | C | C

Because good research needs good data

Research data

- Most research generates data
- Value varies widely
 - So preservation need varies widely
- Types of data...
 - Observational
 - Experimental
 - Computational
 - Reference
 - Documentary/metadata



D | C | C

Because good research needs good data

Demand-side

- Who deposits?
- Who re-uses?
- Mostly the same community
 - I.e. similar subset of researches
 - Sometimes researchers from other fields
 - Occasionally the public
- IPR mostly not an issue?
 - Compared with other scenarios, anyway!
 - Privacy/ethical issues sometimes major issue



D | C | C

Because good research needs good data

Supply side

- Limited set of discipline archives
- Small set of national-scale services
- Emerging institutional role
 - Needs developing
 - Faculty? Department?
 - Not research group?
- Journal/publisher role???
- Issues
 - Scale
 - Complexity
 - Disciplinarity



D | C | C

Because good research needs good data

Economic analysis

- Under way now
- Market mechanisms unlikely
 - Desire for openness
- Quantity versus quality issues?
- Economies of scale?
 - And scope?
- Research as zero sum game?
 - Preservation funds subtract from research funds...
 - Preservation for re-use
 - More curation than preservation
 - Mandate role critical... but insufficient?



D | C | C

Because good research needs good data

Economic analysis

- Lifecycle approach needed
 - Future stakeholders not well represented
 - Handoffs a serious issue
- Need diversity of funding streams
 - Money supply a single point of failure
 - Eg Aradopsis resource
 - “US National Science Foundation (NSF) is phasing out funding after 10 years as the [TAIR] data resource's sole supporter.”
 - “A global, ELIXIR-like initiative is urgently needed, run perhaps by an international, relatively unbureaucratic organization such as the Human Frontier Science Program.” Nature 462



D | C | C

Because good research needs good data

Recommendations

- Some will be generic
- Some will be specific (eg per scenario)
- Likely to be rooted in context
 - Mostly US
 - Some UK
- Will not fully address European requirements
 - More work needed!



D | C | C

Because good research needs good data

Conclusion

- Final Report: January 2010
- Report will be successful if:
 - Recommendations are of practical use to decision-makers
 - It catalyzes further work on economically sustainable digital preservation/curation
- Economic sustainability = risk management
 - Many threats to long-term future of digital materials
 - Economic threat is pervasive and immediate
 - Sustainability strategy is a means to mitigate economic risk

•Lavoie



Because good research needs good data

For More Information

- BRTF Website: brtf.sdsc.edu
 - Interim Report
 - Preservation Bibliography
 - Task Force Information
 - Upcoming Final Report, etc.

The screenshot shows the website for the Blue Ribbon Task Force on Sustainable Digital Preservation and Access. The header includes navigation links: About Us, Members, Bibliography, News Center, Intranet, and Contact Us. The main banner features a blue and black background with binary code and a glowing sphere. Text on the banner reads: "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." Below this, it identifies Lucy Nowell as Program Director at the Office of Cyberinfrastructure, NSF. The page is divided into two columns: "Goals" and "Sponsors".

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

Logos for sponsors include: NSF, SDSC (San Diego Supercomputer Center), The Andrew W. Mellon Foundation, Digital Preservation, JISC, ERA, and Council on Library and Information Resources.

Footer text: "About Us | Members | Bibliography | News Center | Intranet | Contact Us" and "Funded 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".

Alliance for Permanent Access November 2009