

# NDSA



Content



Standards



Infrastructure



Innovation



Outreach

NDSA: Advancing the Capacity to Preserve  
our Nation's Digital Resources



Prepared for

*Digital Preservation, 2014*

July 2014

# An Overview of *NDSA* Advancing the Capacity to Preserve our Nation's Digital Resources



Content



Standards



Infrastructure



Preservation



Outreach

Presented by:

Micah Altman, <[escience@mit.edu](mailto:escience@mit.edu)>

Director of Research, MIT Libraries

Chair, Coordination Committee, National Digital Stewardship Alliance

# DISCLAIMER

These opinions are my own, they are not the opinions of MIT, Brookings, any of the project funders, nor (with the exception of co-authored previously published work) my collaborators

*Secondary disclaimer:*

“It’s tough to make predictions, especially about the future!”

-- Attributed to Woody Allen, Yogi Berra, Niels Bohr, Vint Cerf, Winston Churchill, Confucius, Disreali [sic], Freeman Dyson, Cecil B. Demille, Albert Einstein, Enrico Fermi, Edgar R. Fiedler, Bob Fourer, Sam Goldwyn, Allan Lamport, Groucho Marx, Dan Quayle, George Bernard Shaw, Casey Stengel, Will Rogers, M. Taub, Mark Twain, Kerr L. White, etc.

# Collaborators & Co-Conspirators

- The over 160 institutional members of NDSA, and the tens of thousands of hours contributed by their representatives to NDSA working groups, meetings and reports
- National Agenda Authors:

Jefferson Bailey, Karen Cariani, Jim Corridan, Jonathan Crabtree, Blaine Dessy, Michelle Gallinger, Andrea Goethals, Abigail Grotke, Cathy Hartman, Butch Lazorchak, Jane Mandelbaum, Carol Minton Morris, Trevor Owens, Meg Phillips, Abbie Potter, John Spencer, Helen Tibbo, Tyler Walters, Kate Wittenberg, Kate Zwaard

# Who needs digital preservation?

# No one needs digital preservation... for its own sake

- And no one needs drills ...
- ... microscopes ...
- ... fiber optic cable ...
- ... TCP/IP protocols ...
- ... portland cement ...
- ... the world-wide-web...

# Preservation = Durable, Long-Term Access

- The rule of law require maintaining authentic public records
- Scientific advances rely on a cumulative, traceable evidence base
- Art, history, culture require durable access to national heritage information
- Our nation needs durable access to a strategic information reserve
- Humanity needs durable long-term access information in order to *communicate to future generations*

Why do we need  
collaboration and  
coordination in  
stewardship?



# Some General Trends in Information (aka “more, more, more”...)

## Lots More Data

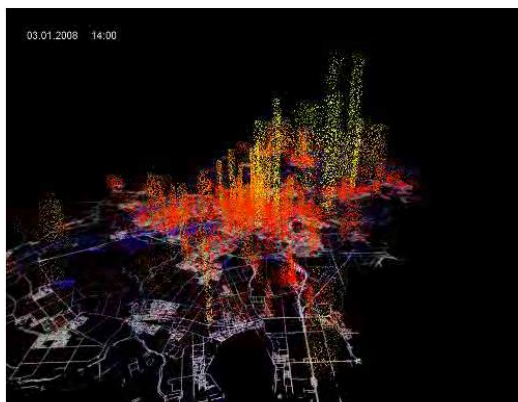


## More Open

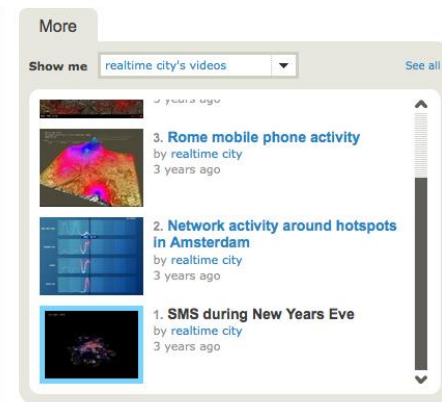


Open Knowledge  
Foundation

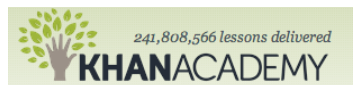
## Shifting Evidence Base



Visualization of SMS activity in Amsterdam during New Years Eve



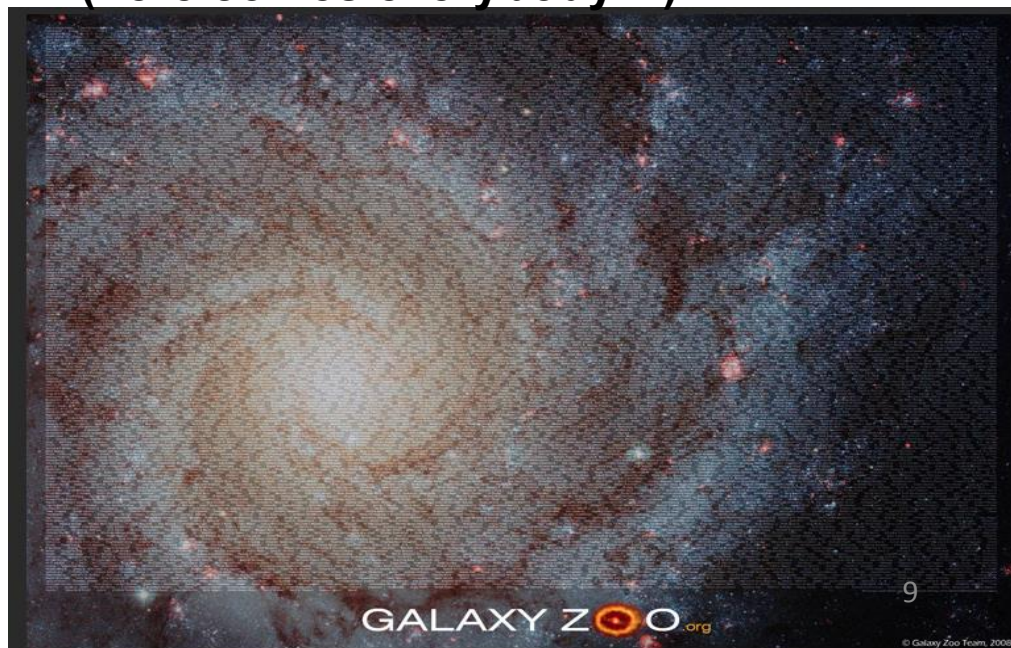
## More Learners



## Publish, then Filter



## High Performance Collaboration (here comes everybody...)



# Digital Stuff is Different

- Accessible
- Replicable
- Computable
- Changeable
- ...

# No Single Organization can Preserve all the Information upon which it Relies

- More production of digital content
- More publishing, filtering and access
- More learners and collaborators
- More attention to public information
- More embedding of information you want, in a larger context required to understand it

# Why can't someone else do it?



- Why not put everything in Amazon?
- Amazon claims reliability of 99.9999999999% (Better odds than winning Powerball<sup>®</sup>, being struck by lightning, and finding alien life... combined)

# What's left out of the Eleven Nines?

- What are the units? - Collection? Object? Bit?
- How was the failure rate calculated? (It's theoretical)
- What is the empirical evidence for MBTF?
- What threats are assumed away?
- Do SLA's or audits back up "design" reliability claims?
- How much does it really cost? Storing data in Glacier may be relatively cheap --- Getting it back is often more expensive

# Diverse Risks



Physical & Hardware



Insider & External Attacks

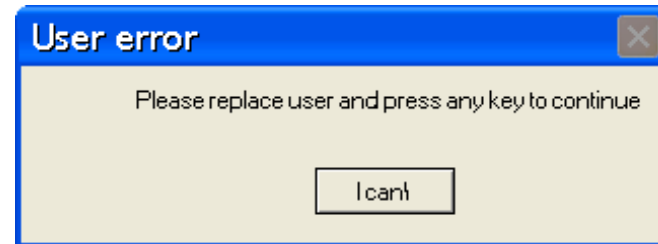


Software



**CLOSED**

Organizational Failure



Curatorial Error

# Single Institutions Cannot Counter all Risks – Digital Offers Opportunity to Diversify

- Distribute to mitigate external risks
  - Third party attacks
  - Institutional funding
  - Change in legal regimes
- Distribute to mitigate internal risks (“Quis custodiet ipsos custodes?”)
  - Unintentional curatorial modification
  - Loss of institutional knowledge & skills
  - Intentional curatorial de-accessioning
  - Change in institutional mission

# Essential Knowledge is Distributed

*Digital information and communication technologies change frequently and cross many traditional boundaries*

- Effective digital preservation requires knowing:
  - Research
  - Practice
  - Organizational collaborations
- State of the art, practice, collaboration is distributed across:
  - disciplines
  - sectors
  - communities of practice



# A roadmap towards coordinated stewardship of digital information

# NATIONAL AGENDA FOR DIGITAL STEWARDSHIP

2015



**NDSA** 

# National Agenda for Digital Stewardship

*Identifies high-impact opportunities for the state of the art, practice & collaboration*

*Aimed at senior institutional decision makers.*

- Digital Content Areas
- Organizational Roles, Policies, and Practice
- Technical Infrastructure
- Research

# What's New: Extending the Roadmap to Actionable Opportunities

- First edition of Agenda identified
  - Key trends
  - Key challenges
- 2015 edition calls out
  - Action steps

# Developing the Evidence Base for Coordinated Action

- Digital Content Areas
  - What content is at risk?
  - What content is safeguarded?
- Organizational Roles, Policies, and Practice
  - What practices are used in the community?
  - How are we managing risk?
  - Where is there community capacity?
  - Who's engaging with critical tools and standards?
- Technical Infrastructure & research
  - What infrastructure is in use by the community?
  - What is being developed?
  - What critical tools and standards do we need to monitor and engage with?
- Research
  - What works – and how do we know?
  - What information is valuable – and how do we know?

How NDSA advances our  
capacity for durable  
information access.

# *About the NDSA*



Content



Standards



Infrastructure



Innovation



Outreach



NDSA: Advancing the Capacity to Preserve  
our Nation's Digital Resources

# NDSA Mission

**To establish, maintain, and advance the capacity to preserve our nation's digital resources for the benefit of present and future generations.**

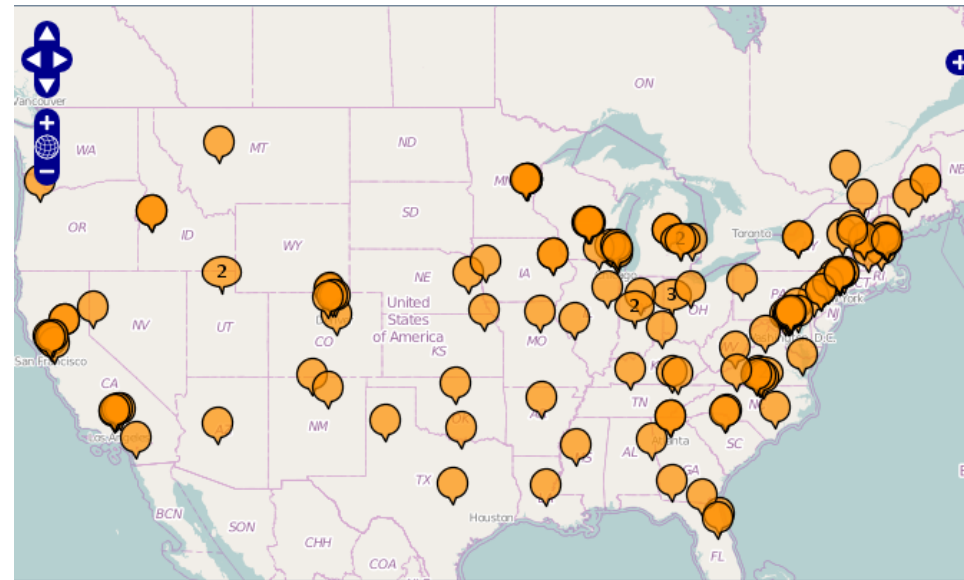


# (Brief) History the NDSA –

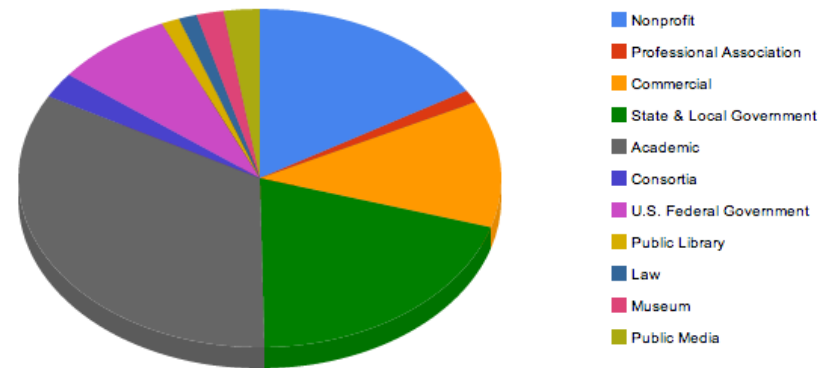
- The *National Digital Stewardship Alliance* (NDSA) **was launched in 2010** as an initiative of the *National Digital Information Infrastructure and Preservation Program* (NDIIPP), with a core of NDIIPP partner organizations
- Over the last four years, the NDSA has **grown rapidly**, to include *165* organizations that represent all sectors, and include universities, consortia, professional associations, commercial enterprises, and government agencies at the federal, state, and local levels..

# How NDSA Works

- **Based** on collaborative community effort and partnership -- there are no fees for NDSA membership.
- The **Library of Congress** has provided essential organizational support, logistical support, and substantive collaboration and expertise.
- Each NDSA member organization joins by being vetted, committing to the NDSA principles
- Each member *contributes* efforts to working groups, reports, surveys, meetings and other NDSA initiatives.



Items grouped by sector



[digitalpreservation.gov/ndsa/memberslist.html](https://digitalpreservation.gov/ndsa/memberslist.html)

NDSA: Advancing the Capacity to Preserve  
our Nation's Digital Resources

# NDSA Institutional Values

- Stewardship
  - Members of the NDSA are committed to managing digital content for current and long-term use. The members of the NDSA are actively ensuring sustained access to the digital content that constitutes our national legacy and empowers us as leaders in the global knowledge economy...
- Collaboration
  - Collaborative work is the centering value of the Alliance; it is a value shared by all members and a priority in work with all organizations and associations...
- Inclusiveness
  - The NDSA is a collaborative effort to preserve a distributed national digital collection for the benefit of current and future generations. We value the range of experience, the potential for innovation, and the fault-tolerance that heterogeneity brings...
- Exchange
  - Members of the Alliance encourage the open exchange of ideas, services, and software....  
The Alliance is committed to transparency and all products generated or produced by the Alliance will be circulated under open licenses.

# NDSA Initiatives

## Working Groups



### Content

selection and acquisition of digital collections



### Standards

digital formats and best practices



### Infrastructure

development and maintenance of tools



### Innovation

cutting edge research and development



### Outreach

relationships and messaging

## Extending Knowledge

- Preservation Storage Survey
- Web Harvesting Survey
- Preservation Staffing Survey
- Geospatial Selection & Appraisal report
- Content case studies
- NDSA Interview Series

## Tools for Practice

- Levels of Preservation
- Digital Preservation in a Box
- Digital Preservation on Wikipedia

## Dissemination

- National agenda for digital stewardship
- NDSA Innovation Awards
- NDSA Social Media

## Recent Outputs

# Extending Community Knowledge

- *NDSA Storage Survey* – biannual surveys of preservation storage capacity and practices
- *Staffing Survey* – providing a baseline on organization staffing and capacity for digital curation
- *Web Archiving Survey* – revealing the landscape of web archiving activities in the United States—including organizations, tools, content.
- *Content Case Studies*: Specific case studies developed around a variety of types of content to share compelling stories that demonstrate the value of digital preservation in our communities.
- *NDSA Interview Series*: bring individuals outside of the organizations and specialties currently participating in the NDSA into conversation with the field of digital preservation to inspire new approaches to current challenges. Interviews are published to The Signal, the Library of Congress digital preservation blog.

# Advancing Capacity through Community Practices

- Reports
  - *The National Agenda for Digital Stewardship*
  - *"The Benefits and Risks of the PDF/A-3 File Format for Archival Institutions"*
  - *"Issues in the Appraisal and Selection of Geospatial Data."*
- Tools
  - *The Levels of Preservation*: A tiered set of recommendations on how organizations should begin to build or enhance their digital preservation systems.
  - *Digital Preservation in a Box*: A list of resources that can be used to provide introductory-level education to those who may have little to no knowledge of digital preservation and digital curation issues, and to help them steward their own digital information, which might include photographs, audio, video, electronic mail, personal documents, and web archives.
  - *Digital Preservation information on Wikipedia*:. Started a community of wikipedians to maintain and improve the information about digital preservation through WikiProject Digital Preservation.

# Raising Awareness & Sharing Knowledge

- *Regional meetings.* NDSA regional meetings bring new institutions to recognize and commit to principles of digital stewardship, and been a resource for practitioners who are seeking a wider community to inform their institution's stewardship plans and practices.
- *NDSA Innovation Awards:* the Annual Innovation Awards were established by the National Digital Stewardship Alliance to recognize and encourage innovation in the field of digital preservation stewardship and are promoted .
- *Social media.* The NDSA has actively engaged in raising awareness of digital preservation needs, practices, and activities through social media. The NDSA contributes extensively to the Signal an active and well-read blog—hundreds of posts have generated thousands of comments, and averages over 35,000 page-views per month
- *Outreach.* NDSA member experts make their time available to stewardship organization, funders and the media through NDSA

# Safeguarding the Information Reserve

- There are over **100 Petabytes** of content in the NDSA member network, based on survey estimates
- NDSA members report that they expect content they manage to more than double over the next three years
- NDSA members content is *expertly selected and intensively curated*
- NDSA members commit to durable long-term access to content
- *NDSA members are stewards of content for communities beyond their institutional boundaries*
  - *NDSA member content is part of the nation's heritage, public record, and scientific evidence base*



# Want more information?

Contact NDSA for...

- Briefings, webinars, and consultations on the Agenda or other NDSA work
- Assistance in gathering comments on National policies and programs
- Assistance in recruiting experts for review and discussion panels
- Referrals to content stewards in specific areas

# *More Information*

[digitalpreservation.gov/ndsa/](https://digitalpreservation.gov/ndsa/)  
[ndsa@loc.gov](mailto:ndsa@loc.gov)