

# NATIONAL AGENDA FOR DIGITAL STEWARDSHIP

2014



**NDSA** 

# 2014 National Agenda for Digital Stewardship

## EXECUTIVE SUMMARY

Effective digital preservation is vital to maintaining the public records necessary for understanding and evaluating government actions, the scientific evidence base for replicating experiments, building on prior knowledge, and the preservation of the nation's cultural heritage. Substantial work is needed to ensure that today's valuable digital content remains accessible, useful, and comprehensible in the future — supporting a thriving economy, a robust democracy, and a rich cultural heritage. The 2014 National Agenda for Digital Stewardship integrates the perspective of dozens of experts and hundreds of institutions, convened through the Library of Congress, to provide funders and other executive decision-makers with insight into emerging technological trends, gaps in digital stewardship capacity, and key areas for development. It is meant to inform, rather than replace, individual organizational efforts, planning, goals, and opinions. Its aim is to offer inspiration and guidance and suggest potential directions and areas of inquiry for research and future work in digital stewardship.

The Agenda outlines the challenges and opportunities related to digital preservation activities in four broad areas: Organizational Roles, Policies, and Practices; Digital Content Areas; Infrastructure Development; and Research Priorities. The sections are arranged from the most comprehensive and encompassing topics to sequentially drill down to more specific challenges and recommendations. The Organizational Roles, Policies, and Practices section discusses the overarching challenges the digital preservation community faces. The Digital Content Areas section highlights specific kinds of content that need attention. The Infrastructure Development section identifies opportunities and makes specific recommendations for how the digital preservation community can respond. The Research Priorities section provides detailed recommendations to prioritize resource allocation towards areas of research that are critical to the advancement of both basic understanding and the effective practice of digital preservation.

### Organizational Roles, Policies, and Practices

Despite continued preservation mandates, it has become increasingly difficult to adequately preserve valuable digital content because of a complex set of interrelated societal, technological, financial, and organizational pressures, including:

## National Digital Stewardship Alliance

---

- Increased scope of responsibilities (data management, education of content creators, etc.)
- Growing financial pressures - increased costs and decreasing resources
- Lack of adequate staff, in numbers and expertise
- Increased complexity and volume of data
- Rapidly accelerating technological change
- Evolving data management, security and compliance policies
- Complex and evolving landscape of rights management
- Lack of prioritization of digital preservation by higher administration and those controlling budgets

As a community, we need to dramatically increase cross-organizational cooperation to increase the impact and leverage investments made by individual institutions. We must work together to raise the profile of digital preservation and campaign for more resources and higher priority given to digital preservation, and to highlight the importance of digital curation and the real costs of ensuring long term access. We must also coordinate to develop comprehensive coverage on critical standards bodies, and promote systematic community monitoring of technology changes relevant to digital preservation.

## Digital Content Areas

Both born-digital and digitized content present a multitude of challenges to stewards tasked with preservation: the size of data requiring preservation; the selection of content when the totality cannot be preserved; and the selection of modes of both content storage and migration to ensure long-term preservation. Areas of content of particular concern include:

- Electronic Records
- Research Data
- Web and Social Media
- Moving Image and Recorded Sound

Across these areas, content value and selection represent a core challenge that organizations need to address. Furthermore, digital stewardship planning must go beyond a focus on content we already have and technology already in use. Moreover, research is required to develop theoretically grounded and empirically tested models of information valuation.

## Technical Infrastructure Development

Successful digital preservation requires taking a broad view of technical infrastructure: the Agenda defines it generally as the set of interconnected technical elements that provides a framework for supporting an entire structure of design, development, deployment and documentation in service of applications, systems, and tools for digital preservation. This

definition includes hardware, software, and systems. Organizational policies, practice, and regulation inform many of the observations and recommendations for the development of digital stewardship infrastructure.

Specific priorities identified for infrastructure investment include:

- File Format Action Plan Development
- Interoperability and Portability in Storage Architectures
- Integration of Digital Forensics Tools
- Ensuring Content Integrity

There is a clear need for organizations to share their assessments of institutional risk and their plans for mitigating those risks; to develop use-case driven best practices for fixity in particular system designs and configurations; and to move the basic research in digital forensics tools from research to implementation in production workflows for organizations. Moreover, the need for integration, interoperability, portability, and related standards and protocols stands out as a theme across all of these areas of infrastructure development.

### Research Priorities

Research is critical to the advancement of both basic understanding and the effective practice of digital preservation. Research in digital preservation is under-resourced. In part this is because the payoff from long-term access occurs primarily in the medium-long term and tends to benefit broad and diverse communities.

We expect that research investment in five areas will yield unusually large impact:

- Applied Research for Cost Modeling and Audit Modeling
- Understanding Information Equivalence & Significance
- Policy Research on Trust Frameworks
- Preservation at Scale
- The Evidence Base for Digital Preservation

A common challenge running through this report is the limited amount of empirical evidence available. The digital preservation community is beginning to develop a shared evidence base that can be used to answer these and similar questions. However, these studies must be broadened and repeated over time to establish a robust evidence base from which generalizable guidance can be drawn. Furthermore, decision-makers should recognize that basic research in these areas often needs to be paired with the development, support, and evaluation of infrastructure.

## Conclusion

Effective digital stewardship is vital to maintaining the nation's cultural heritage, scientific evidence base, and the public records necessary for understanding and evaluating government actions. The 2014 National Agenda for Digital Stewardship identifies the key technological trends, gaps in digital stewardship capacity, and opportunities for future work for digital preservation professionals, decision-makers, and others interested in investing in the long-term management of digital content.

## ABOUT THE NDSA

Founded in 2010, the [National Digital Stewardship Alliance](#) (NDSA) is a consortium of institutions that are committed to the long-term preservation of digital information. NDSA's mission is to establish, maintain, and advance the capacity to preserve our nation's digital resources for the benefit of present and future generations. NDSA member institutions represent all sectors, and include universities, consortia, professional associations, commercial enterprises, and government agencies at the federal, state, and local level.

The NDSA comprises over 150 members. These members come from 45 states and include universities, consortia, professional societies, commercial businesses, professional associations, and government agencies at the federal, state, and local level. NDSA organizations have proven themselves committed to long-term preservation of digital information. Together, these institutions have contributed over 10,000 hours of expertise to NDSA projects.

Additional materials related to the *National Agenda* and the full document can be found at: <http://www.digitalpreservation.gov/ndsa/nationalagenda/index.html>

## BE A PART OF THE CONVERSATION

Comments welcome at [ndsa@loc.gov](mailto:ndsa@loc.gov) or #nationalagenda as well as @NDSA2

## ABOUT THE AUTHORS

The joint leadership group of the NDSA authored the report and engaged in discussions to identify significant trends and challenges. The membership of the NDSA contributed markedly to these discussions. This dialog was enriched by an extensive range of resources and current research. The joint leadership group is made up of the Coordinating Committee members, the Working Group co-chairs, and the NDSA facilitator:

<b>Micah Altman</b>	Director of Research, MIT Libraries, MIT; Non-Resident Senior Fellow, Brookings Institution.
<b>Jefferson Bailey</b>	Strategic Initiatives Manager, Metropolitan New York Library Council
<b>Karen Cariani</b>	Director of Media Library and Archives, WGBH
<b>Jim Corridan</b>	State Archivist and Director of the Indiana Commission on Public Records
<b>Jonathan Crabtree</b>	Director for Archives and Information Technology at the Odum Institute for Research in Social Science at UNC Chapel Hill
<b>Blaine Dessy</b>	Director, Federal Library and Information Center Committee (FLICC).

## National Digital Stewardship Alliance

---

<b>Michelle Gallinger</b>	Digital Programs Coordinator, National Digital Information Infrastructure and Preservation Program, Library of Congress
<b>Andrea Goethals</b>	Digital Preservation and Repository Services Manager, Harvard Library
<b>Abbie Grotke</b>	Lead Information Technology Specialist, National Digital Information Infrastructure and Preservation Program, Library of Congress
<b>Cathy Hartman</b>	Associate Dean of Libraries, University of North Texas
<b>Butch Lazorchak</b>	Digital Archivist, National Digital Information Infrastructure and Preservation Program, Library of Congress
<b>Jane Mandelbaum</b>	Manager of Special Projects, Office of the Director for Information Technology Services, Library of Congress
<b>Carol Minton Morris</b>	Director of Marketing and Communications, DuraSpace
<b>Trevor Owens</b>	Digital Archivist, National Digital Information Infrastructure and Preservation Program, Library of Congress
<b>Meg Philips</b>	Electronic Records Lifecycle Coordinator, National Archives and Records Administration
<b>John Spencer</b>	President and co-founder, BMS/Chace
<b>Helen Tibbo</b>	Alumni Distinguished Professor, School of Information and Library Science at the University of North Carolina at Chapel Hill
<b>Tyler Walters</b>	Dean of University Libraries, Virginia Tech
<b>Kate Wittenberg</b>	Managing Director, Portico
<b>Kate Zwaard</b>	Supervisory IT Specialist for Repository Development, Library of Congress