



Special Collections in Our Digital World

The Future Is Now

Jackie Dooley

Program Officer

OCLC Research

LIBER workshop: Heritage Collections
in the Digital Future

Barcelona, 28 June 2011

Overview

1. OCLC Research
 - Who are we?
2. OCLC Research survey results (2009)
 - Digitized special collections
 - Born-digital special collections
3. Digitization
 - Our published work to date
4. Born-digital special collections
 - Our latest project

OCLC Research: *Who we are*

- ~45 FTE with offices in Ohio, California and the UK
- Sponsored by OCLC and a partnership of research libraries around the world that share:
 - A strong motivation to effect *system-wide change*
 - A commitment to *collaboration* as a means of achieving collective gains
 - A desire to *engage internationally*
 - Senior management ready to provide *leadership* within the transnational research library community
 - Deep and rich *collections* and a mandate to make them accessible
 - The *capacity* and the will to contribute

1. OCLC Research: Current portfolios

- Mobilizing unique materials
- Research information management
- Metadata support and management
- Infrastructure and standards support
- System-wide organization
- User behavior studies and synthesis

- ... & the OCLC Research Library Partnership

2. OCLC Research survey results (2009)

- Digitized special collections
- Born-digital special collections

Taking Our Pulse: The OCLC Research Survey of Special Collections and Archives



Taking Our Pulse: The OCLC Research Survey of Special Collections and Archives

Jackie M. Dooley
Program Officer

Katherine Luce
Research Intern

OCLC Research



A publication of OCLC Research



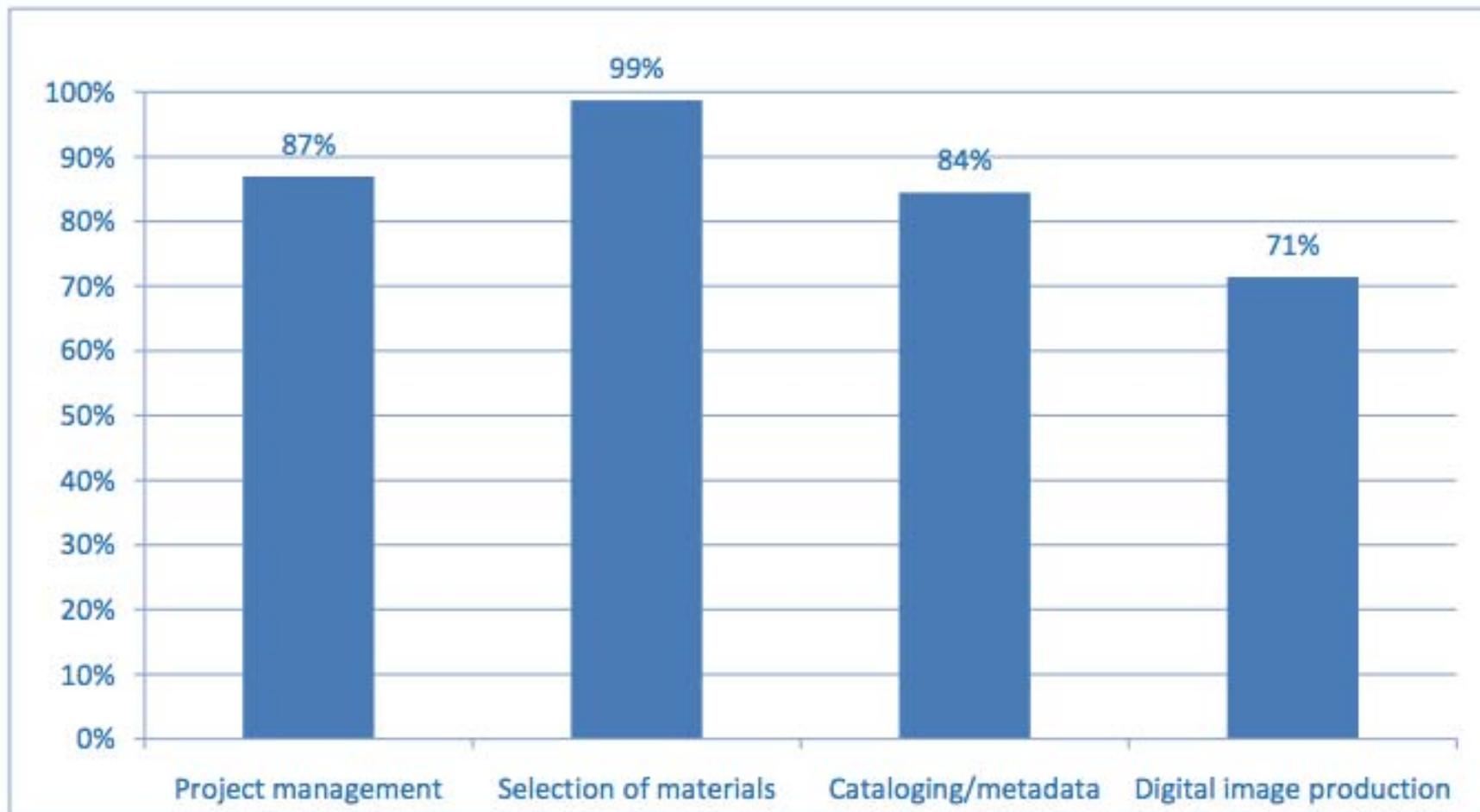
Key Findings

- The top three “most challenging issues”
 - Space
 - **Born-digital**
 - **Digitization**
- Tough economy renders “business as usual” impossible; 75% of library budgets diminished
- User demand for **digitized** collections is insatiable
- Management of **born-digital** materials is in its infancy

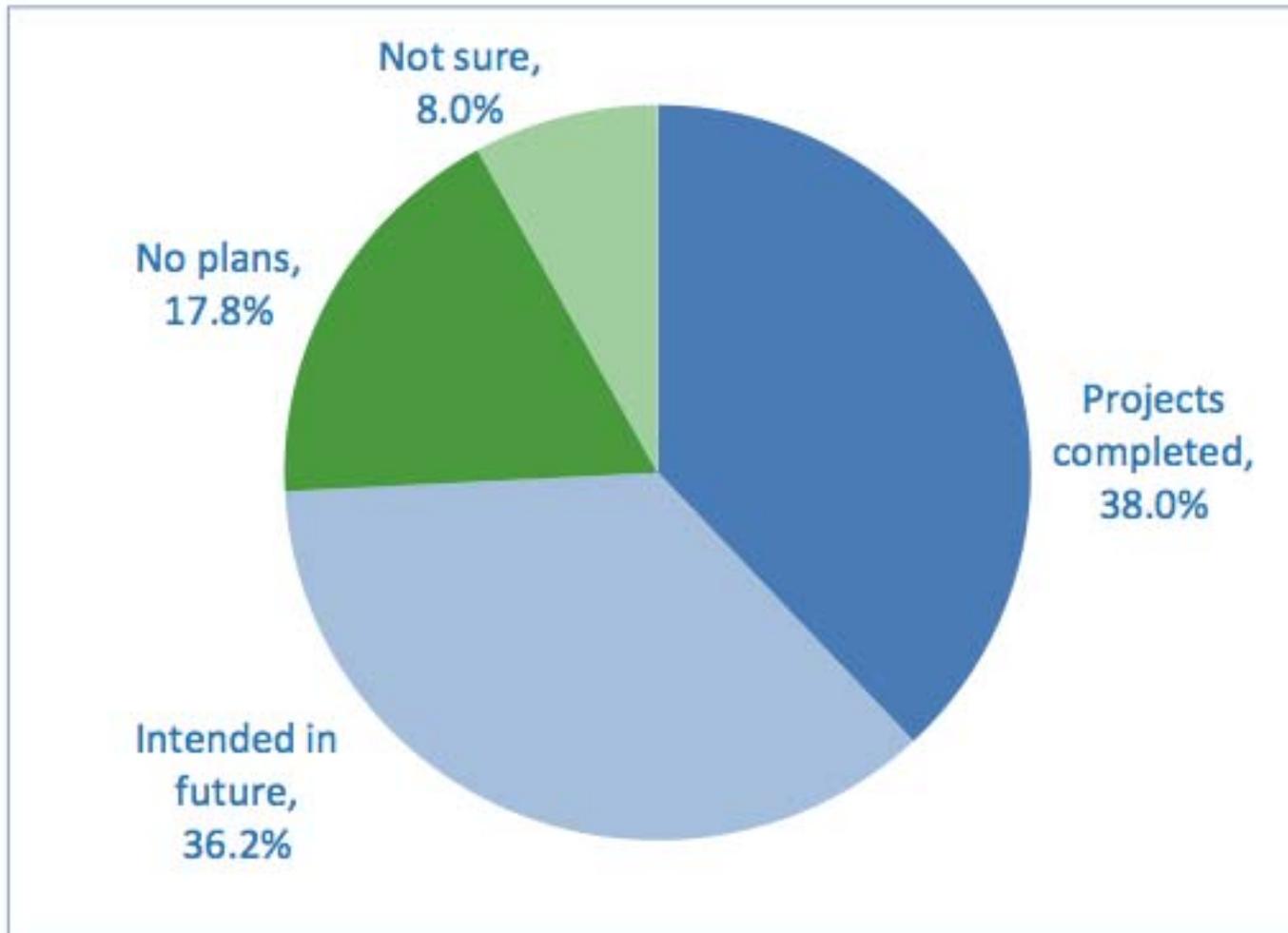
Top education and training needs

1. Born-digital materials: 83%
1. Information technology: 65%
2. Intellectual property: 56%
3. Cataloging and metadata: 51%

Special collections involvement in digitization projects



Large-scale digitization



Definition: Systematic reproduction of entire collections using streamlined production methods that account for special handling.

Born-digital archival materials

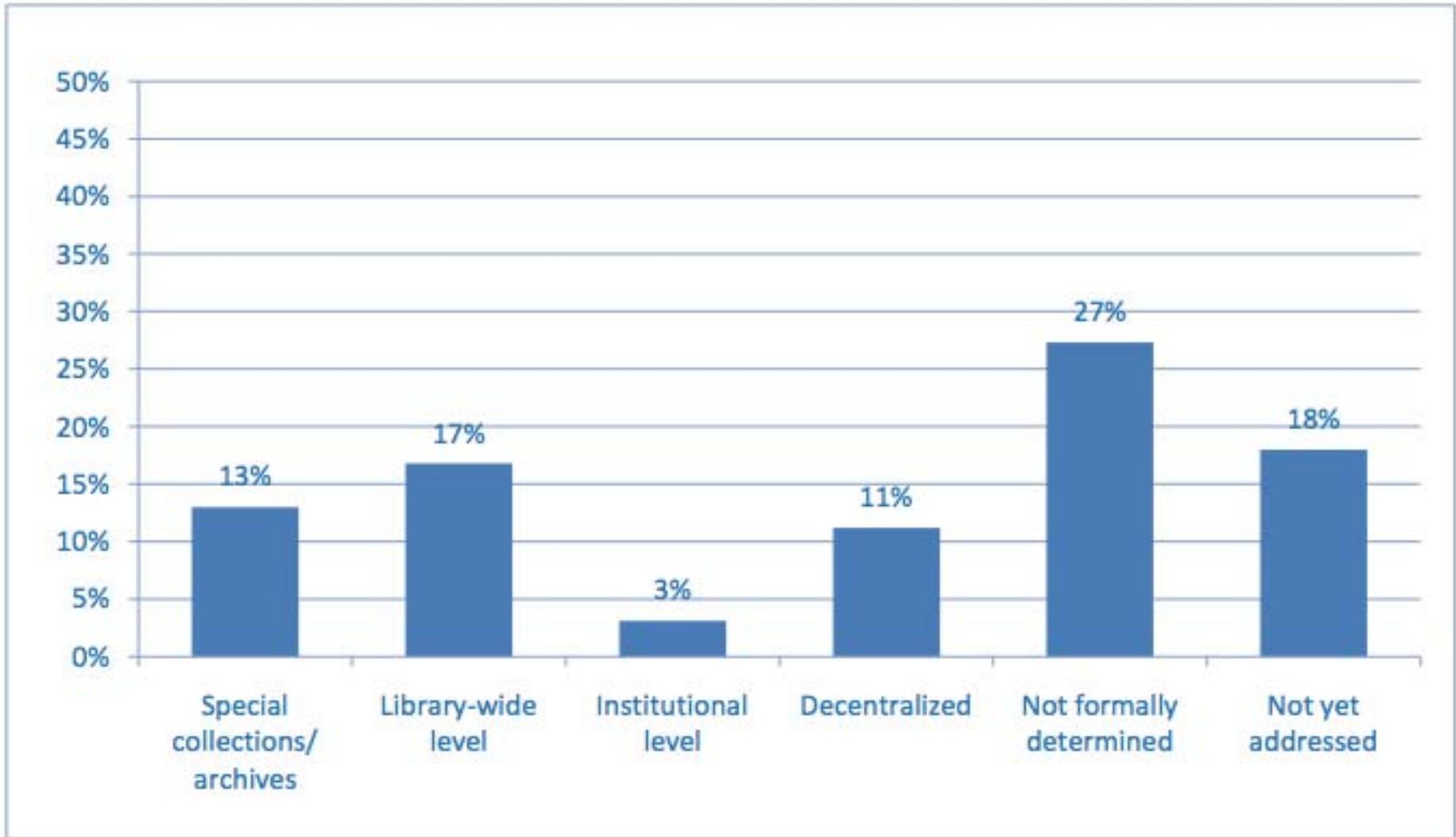
In a nutshell ...

- Undercollected
- Undercounted
- Undermanaged
- Unpreserved
- Inaccessible

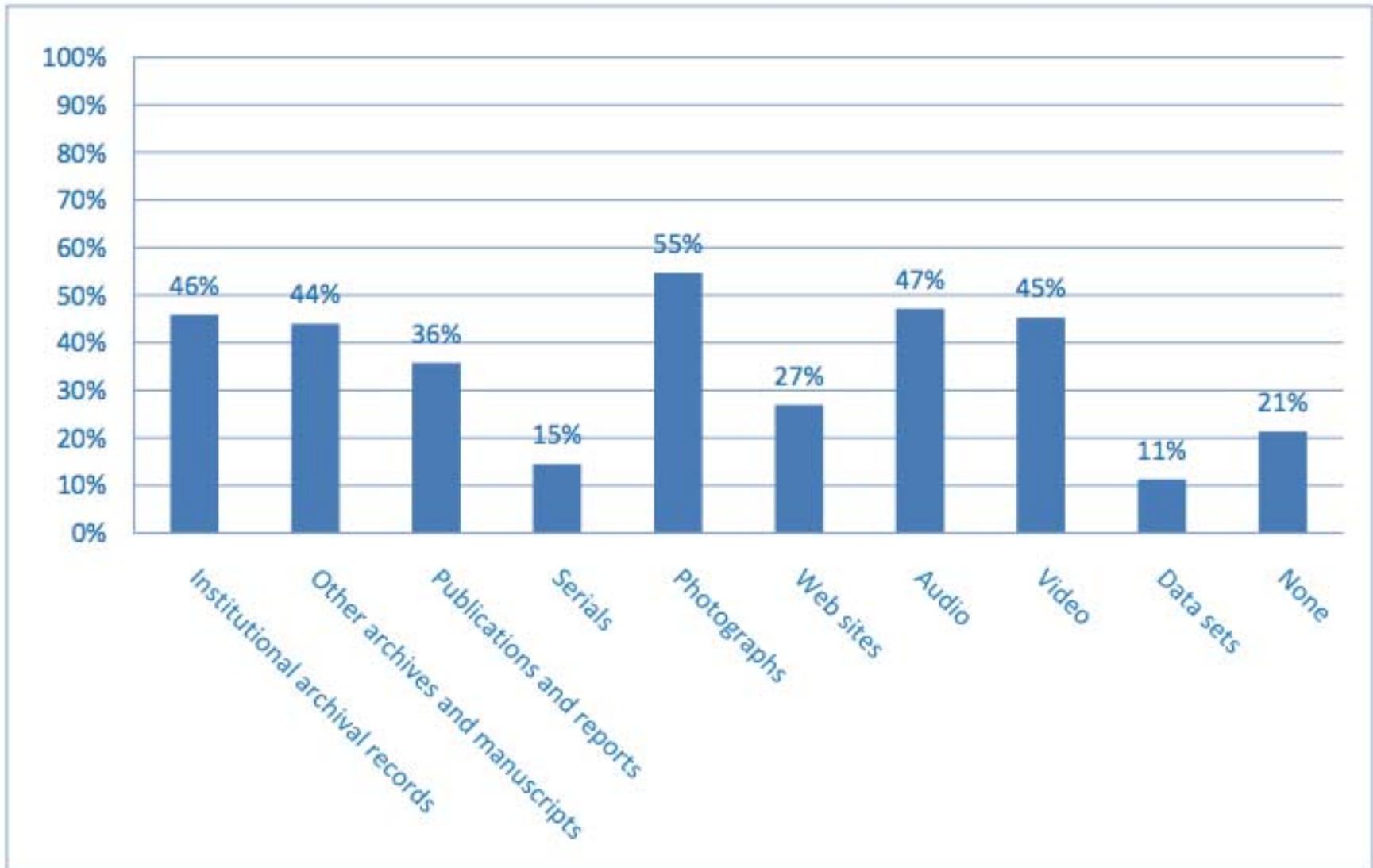
Born-digital archival materials

- Holdings reported by: **35%**
 - Percent held by top two libraries: 51%
 - Percent held by top 13 libraries: 93%
- Digital materials currently held by: **79%**
- Assignment of responsibility for born-digital management made by: **55%**
- Education/training needed by: **83%**
- We surmise that collecting is generally reactive, sporadic, limited.

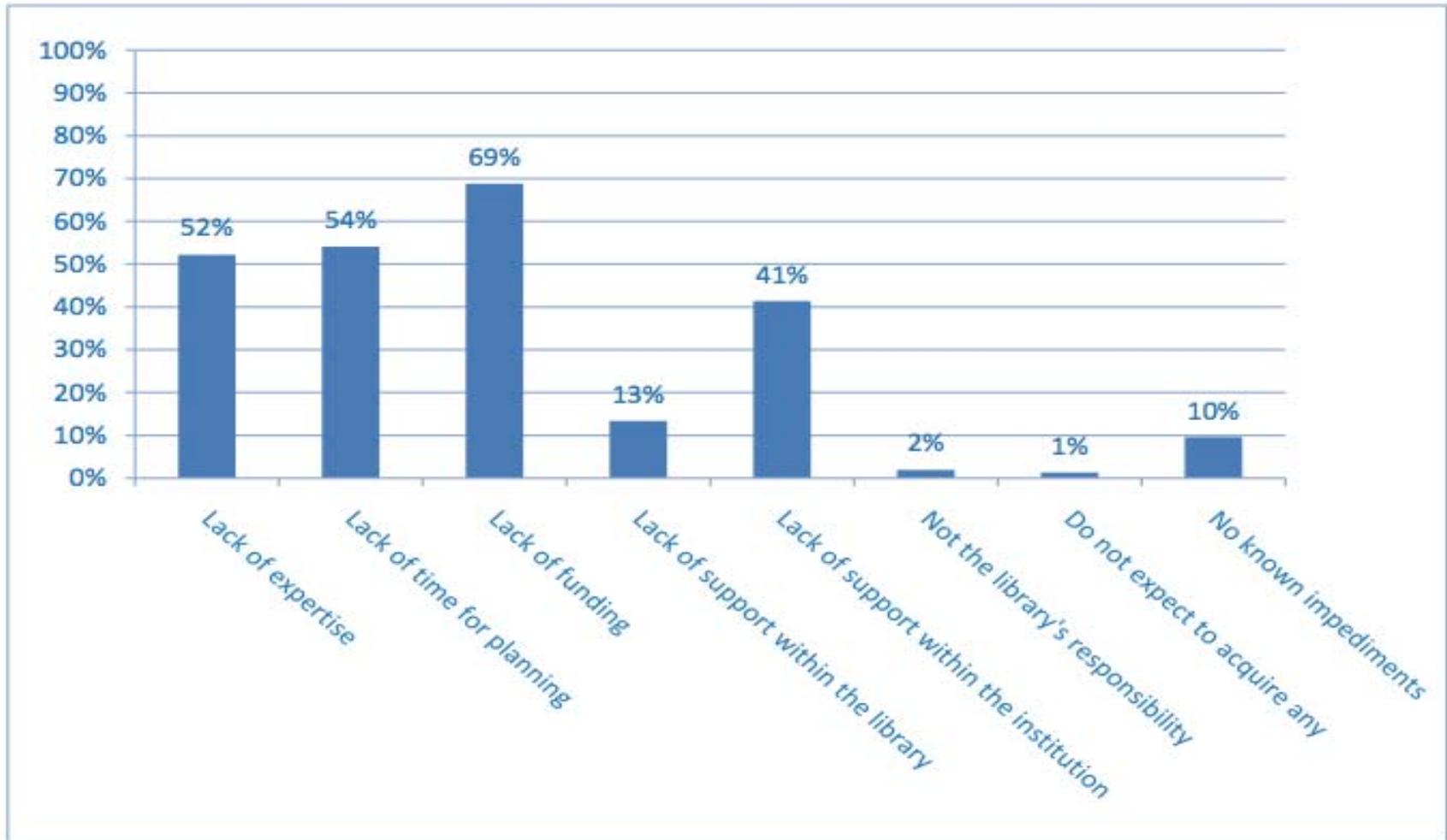
Born-digital: Assignment of responsibility



Born-digital: Materials already held



Born-digital: Impediments



Note: Respondents could check all that apply.



Recommended actions: Digital special collections

- **Digitization**
 - Develop models for large-scale digitization
 - Scope the existing corpus of digitized rare books to aid identification of new projects
- **Born-Digital Archival Materials**
 - Identify relationship between “special collections” and “born-digital”
 - Outline “baby steps” approach to creating a program for management
 - Develop use cases and cost models for selection, management, preservation

3. Digitization: Our Work to date

- Digitization is more for access than for preservation
- Balancing risk analysis and access, with regard to copyright and privacy
- Guidance for commercial partnerships
- **Scaling up** capture of archival formats
- Building the digital library with researcher-initiated digitized content

Shifting Gears: Digitization for Access



Shifting Gears: Gearing Up to Get Into the Flow

Ricky Erway and Jennifer Schaffner

Program officers
OCLC Programs and Research

A publication of OCLC Programs and Research





Shifting Gears: Recommendations

This project catalyzed libraries to switch from highly selective “boutique” digitization projects to thinking about ways to scale up.

- Access vs. preservation
- Quantity vs. quality
- Selection has already been done
- Do it ONCE (then iterate)
- Programs, not projects
- Archival description practices as model
- Discovery happens elsewhere
- Partnerships and funding

“Good Terms”: Improving Commercial-Noncommercial Partnerships for Mass Digitization

[Search](#) | [Back Issues](#) | [Author Index](#) | [Title Index](#) | [Contents](#)

ARTICLES

D-Lib Magazine
November/December 2007

Volume 13 Number 11/12
ISSN 1082-9873

Good Terms - Improving Commercial-Noncommercial Partnerships for Mass Digitization

A Report Prepared by Intelligent Television
for RLG Programs, OCLC Programs and Research

[Peter B. Kaufman](#) and [Jeff Ubois](#)
Intelligent Television

Executive Summary

In 2007, OCLC Programs and Research engaged Intelligent Television to study the partnership agreements between cultural institutions and for-profit companies for the mass digitization of books and other media. This report presents the findings of that study.

Libraries have been digitizing portions of their collections for more than twenty years, but recent opportunities to work with private partners, such as Google, Microsoft, and others, on mass digitization has opened up possibilities that were unimaginable just a few years ago. Private funding, commercially developed technology, and market-oriented sensibilities together may generate larger aggregations of digitized books far sooner than the library community had dreamed possible. There are many efforts underway to assess various aspects of these partnerships; this paper focuses on the terms in mass digitization agreements that affect research-community-centered outcomes.

The libraries and other cultural institutions that private companies first approached saw significant potential in these overtures; they were diligent in seeing that near-term local needs were met. Only when it became clear that a significant number of these partnerships were underway did the library community as a whole begin to think about the overall impact of these business relationships on the future of scholarship.

When we fantasize about that future, we imagine a single way to search all digitized books, journals, and other media; a combined index of all the full texts that will enable research that is otherwise



Detail of illustration by Dan Beard in Cosmopolitan Magazine, December 16, 1893. Public domain



Good Terms: Recommendations

Negotiate for:

- Limited confidentiality
- More complete deliverables
- More open access allowing the development of new applications
- Less restricted distribution
- Responsible treatment of usage data
- Limited duration and survivability

Well-intentioned practice for putting digitized collections of unpublished materials online



Well-intentioned practice for putting digitized collections of unpublished materials online

The primary responsibilities of cultural materials repositories - stewardship and support for research and learning - require us to provide access to materials entrusted to our care. This document establishes a reasonable community of practice that increases and significantly improves access to collections of unpublished materials by placing them online for the purpose of furthering research and learning. Although it promotes a well-intentioned, practical approach to identifying and resolving rights issues that is in line with professional and ethical standards, note that this document does not concern itself with what individuals who access particular items may do with them. While the document was developed with US law in mind, it is hoped that the spirit of the document will resonate in non-US contexts.

If your institution has legal counsel, involve them in adopting this approach; after the approach has been adopted, only seek their advice on specific questions.

Select collections wisely

- Keep your mission in mind and start with a collection of high research value or high user interest.
- Assess the advantages and risks of relying on fair use (in the US) to support public access.
- Some types of materials may warrant extra caution when considering rights issues, such as
 - Contemporary literary papers
 - Collections with sensitive information, such as social security numbers or medical data
 - Materials that are likely to have been created with commercial intent (because they are more likely to have economic value)
 - Very recent materials that were not intended to be made public
- If research value is high and risk is high, consider compromises, such as making a sensitive series accessible on-site only, until a suitable time has passed.

Use archival approaches to make decisions

- Check donor files and accession records for permissions, rights, or restrictions.
- Assess rights and privacy issues at the appropriate level, most often at the collection- or series-level.
- Attempt to contact and get permission from the rights-holder, if there's an identifiable rights-holder at that level.
- Include what you know about the rights status in the description of the collection, including if the collection is in the public domain, if the institution holds the rights, or if the rights-holder has given the institution permission to place the digitized collection online.
- Document your processes, findings, and decisions and share them with your professional community.

Provide take-down policy statements and disclaimers to users of online collections

➤ Adopt a liberal take-down policy, such as: "These digitized collections are accessible for purposes of education and research. We've indicated what we know about copyright and rights of privacy, publicity, or trademark. Due to the nature of archival collections, we are not always able to identify this information. We are eager to hear from any rights owners, so that we may obtain accurate information. Upon request, we'll remove material from public view while we address a rights issue."

- Use an appropriate disclaimer at the institutional level, such as "[institution] makes digital versions of collections accessible in the following situations:
 - They are in the public domain
 - The rights are owned by [institution]
 - [institution] has permission to make them accessible
 - We make them accessible for education and research purposes as a legal fair use, or
 - There are no known restrictions on use

To learn what your responsibilities are if you'd like to use the materials, go to [link]"

Prospectively, work with donors

- Identify possible intellectual property issues and get relevant contact information.
- Ask donors to state any privacy concerns and identify sensitive materials that may be in the collection.
- Suggest that donors transfer copyright to the institution or license their works under a Creative Commons CC0 license.
- Include statements in your collecting policies and in your deeds of gift or transfer documents that:
 - ensure that no restrictions are placed on content that is already in the public domain,
 - grant license to digitize the materials for unrestricted access even when donors retain the rights,
 - and guard against limitations or restrictions on fair use rights.

For more about this document, see <http://www.ocic.org/research/activities/rights>
Revised 03/28/2010

United States Air Force. "DAYTON, Ohio (10/27/2009) -- Members of Boy Scout Troop...at the National Museum of the U.S. Air Force" Public domain

Well-intentioned practice: Balancing rights & access

This project identified strategies for analyzing and developing acceptable risk behaviors and recommending rights clearance and digitization practices for libraries and archives.

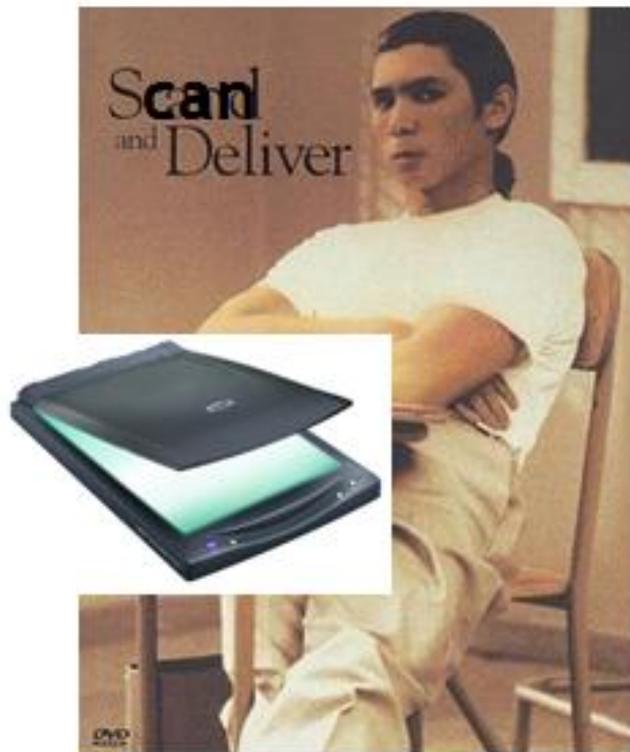
- Select collections wisely
- Use archival approaches to make decisions
- Provide take-down policy statements and disclaimers to users of online collections
- Prospectively, work with donors

Rapid capture

This project investigated digital capture of special formats *at scale*. Case studies describe various methodologies, including in-house capture, outsourcing solutions, hardware (scanners, cameras, materials handling tools, and lighting equipment), staffing, and throughout rates.

- Audio
- Correspondence
- Medieval manuscripts
- Microfilm
- Mixed archival materials
- Photographs
- Scrapbooks and albums
- University records

Managing User-Initiated Digitization in Special Collections and Archives



Scan and Deliver:

**Managing User-initiated Digitization in
Special Collections and Archives**

Jennifer Schaffner
OCLC Research

Francine Snyder
Solomon R. Guggenheim Museum

Shannon Supple
University of California at Berkeley



Managing user-initiated digitization

Sample workflows for creating digital versions of original materials in response to users' requests, ranging from a simple method that requires minimal staff resources to an elaborated workflow that facilitates addition of the files to the institution's digital collections.

Simplify each stage in the process

- Review
- Decide
- Scan
- Deliver

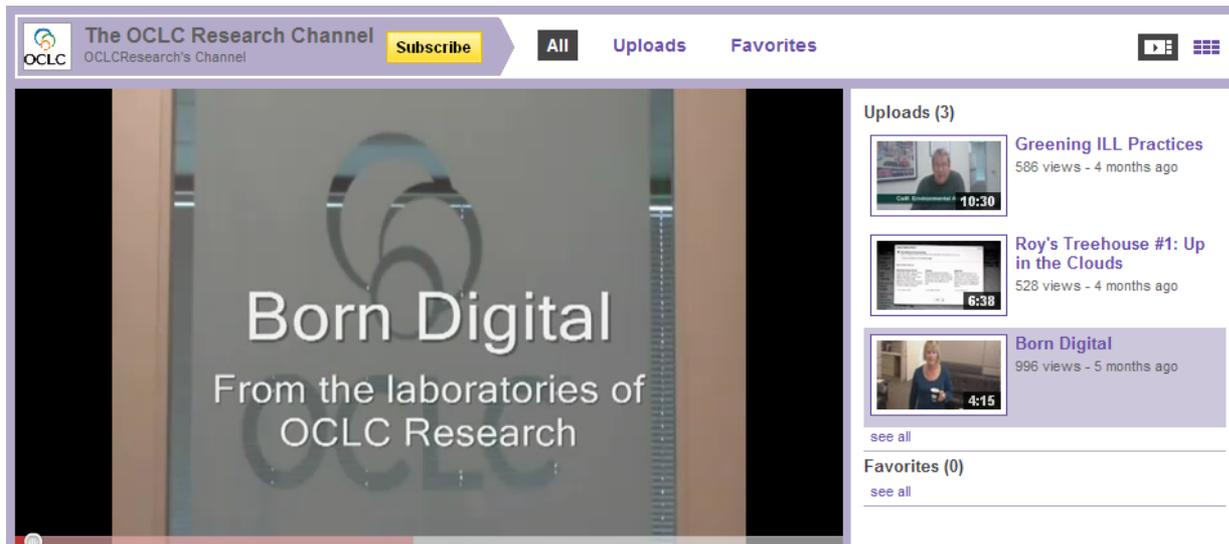
Born-digital special collections: Our latest project

- Types of material
- Skills and expertise
- Initial steps



Born-Digital Materials are ...

- *Digital art*
- *Digital documents*
- *Digital manuscripts*
- *Digital media publications*
- *Digital photographs*
- *Dynamic data*
- *Electronic records*
- *Static data sets*
- *Websites*
- *... and more*



Skills and expertise

This new project focuses on effective management of born-digital materials as they intersect with special collections and archives practices in research libraries. We will:

- Explore the array of **skills and expertise** held by special collections librarians and archivists that are crucial to effective management of many types of born-digital materials.
- Describe how those areas of expertise do or don't pertain to various types of born-digital materials.

Initial steps

A third element of the new project will outline some of the most **basic initial steps** that an archives can take to begin managing born-digital materials.

Planning

- Connect with key stakeholders
- Consider key elements of donor negotiations

Managing

- Inventory what you have
- Transfer from physical media to secure storage
- Learn “do no harm” file management steps



Thank you

Jackie Dooley

dooleyj@oclc.org