Increasing Access to Special Collections

Ricky Erway, OCLC Research

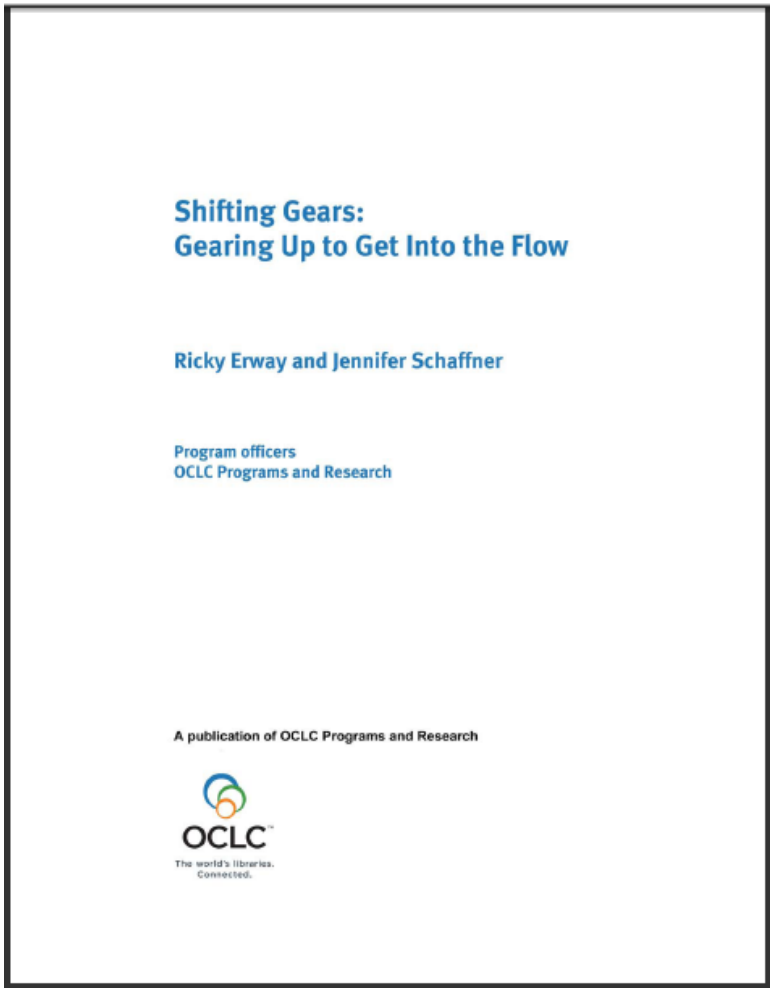
LIBER Annual Conference, Barcelona

29 June 2011

OCLC Research: Mobilizing Unique Materials

Zeets Jones. Light at the End of the Tunnel. 2007. (CC-BY-NC-ND)

Shifting Gears:
Gearing Up to Get into the Flow



Shifting Gears – Recommendations

- It's about access
- It's about quantity
- Selection has already been done
- Iterate
- Programs not projects
- Partnerships
- Adopt archival description practices
- Discovery happens elsewhere

Selecting for digitization

If we could we'd digitize everything en masse

- Look for high-demand collections or series
- Compromise by digitizing samples to gauge the interest

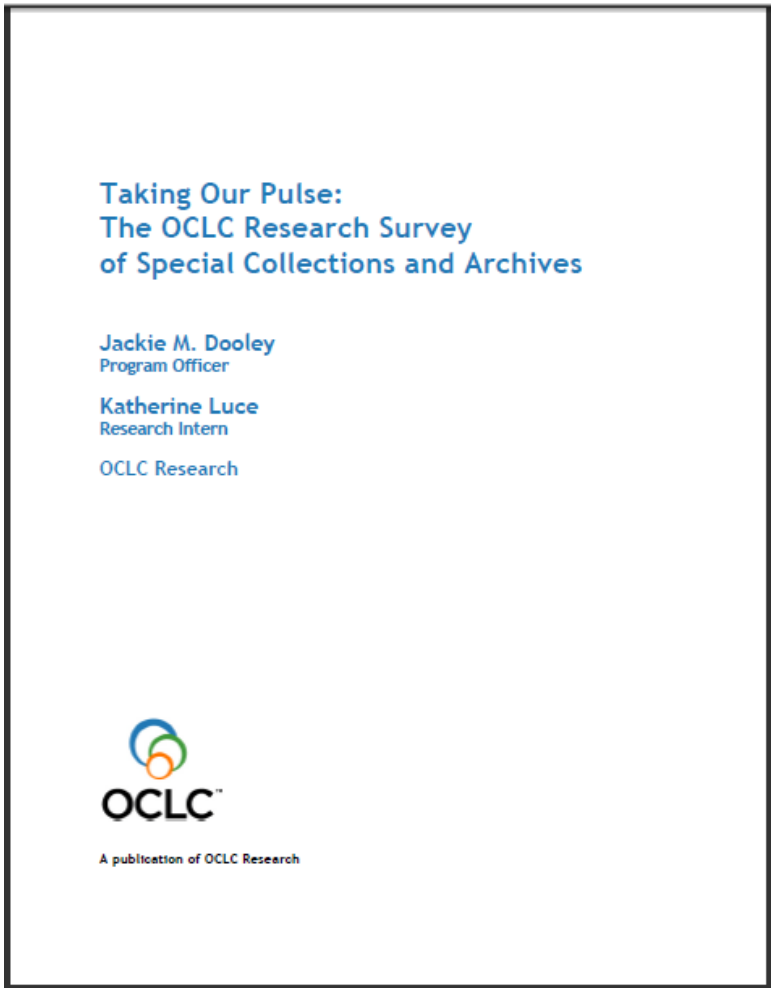
Sometimes we have to do selections

- (Is it worth it?)

Put your efforts into the larger context

Can you get the best of all worlds?

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



Taking Our Pulse – Key Findings

- Many rare and unique material are undiscoverable
- Tough economy renders “business as usual” impossible
- Staff are having trouble keeping up

- The top three “most challenging issues”
  - Space
  - Born Digital
  - Digitization

Taking Our Pulse – Recommendations

- User Services
  - Promote policies that facilitate (rather than inhibit) access
- Born-Digital Archival Materials
  - Identify practices of archivists useful for born digital
  - Basic steps for managing born digital
- Digitization
  - Scope the digitized rare book corpus
  - Develop models for large scale digitization

Well-intentioned practice for putting digitized collections of unpublished materials online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for putting digitized collections of unpublished materials online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.oclc.org/research/activities/rights
Revised 05/28/2010

Well-Intentioned Practice – Main Points

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

Good Terms: Improving Commercial-Noncommercial Partnerships for Mass Digitization

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

Good Terms - Recommendations

Negotiate for

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

  and
- Limited duration and survivability

Rapid Capture: Faster Throughput in Digitization of Special Collections

Rapid Capture:

Faster Throughput in Digitization of Special Collections

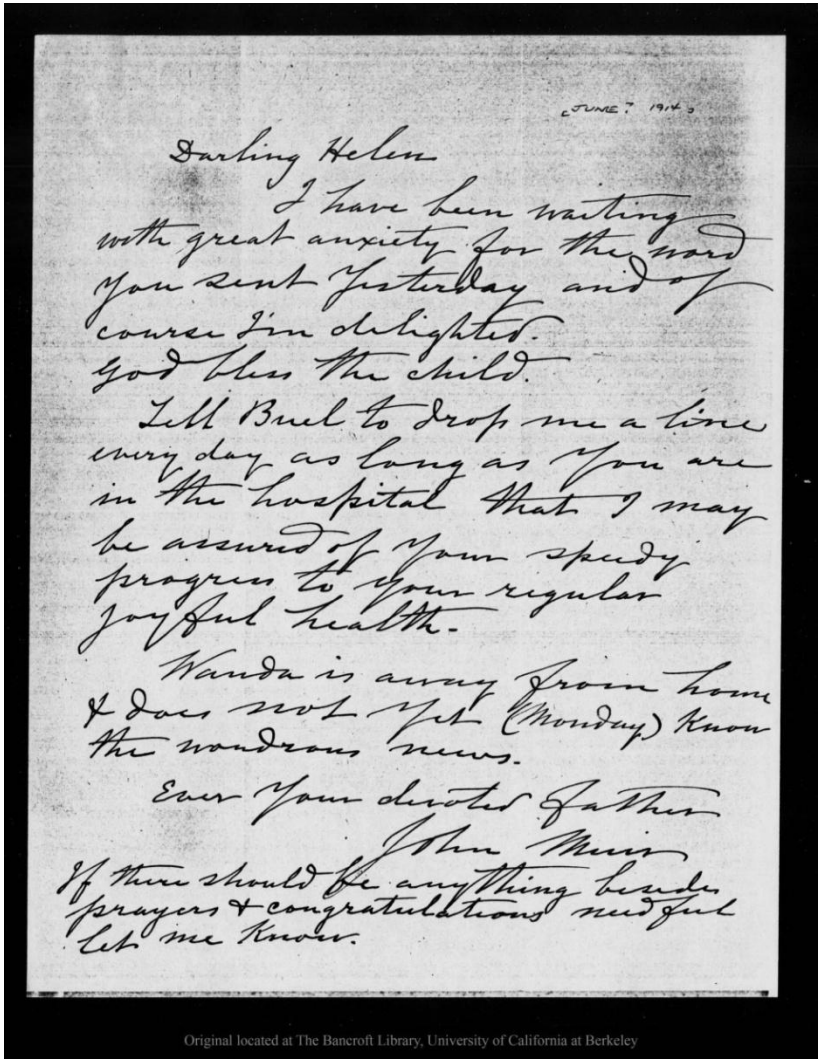
Ricky Erway

Senior Program Officer
OCLC Research

OCLC

A publication of OCLC Research

University of California, Berkeley and University of the Pacific



JUNE 7 1914

Darling Helen

I have been waiting with great anxiety for the word you sent yesterday and of course I'm delighted. God bless the child.

Tell Buel to drop me a line every day as long as you are in the hospital that I may be assured of your speedy progress to your regular joyful health.

Wanda is away from home & does not yet (Monday) know the wondrous news.

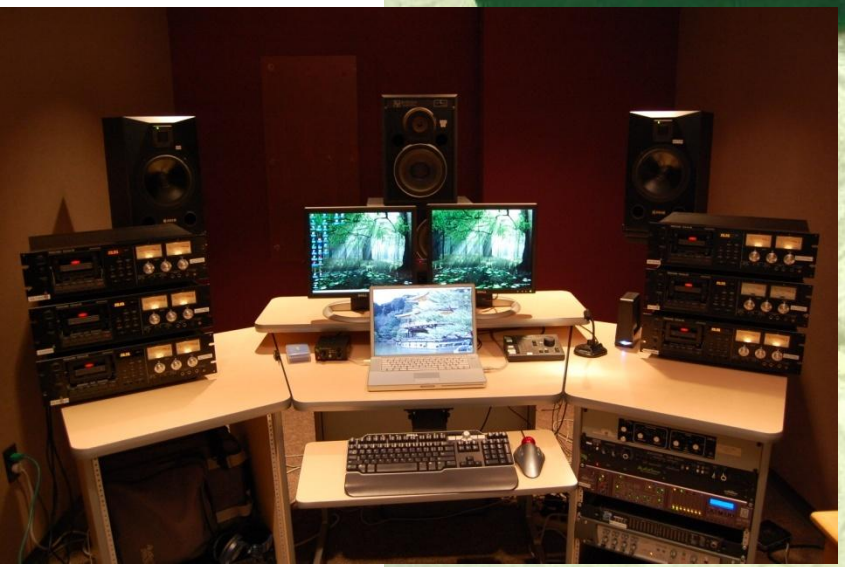
Ever your devoted Father

John Muir

If there should be anything besides prayers & congratulations needful let me know.

Original located at The Bancroft Library, University of California at Berkeley

Archives of Traditional Music, Indiana University



University Archives
University of Minnesota

University of Minnesota News Service

Sports Release

1944 Copy 1

Minneapolis, Jan. 7/44 One of the lightest hockey schedules ever drawn by a University of Minnesota team will get under way January 7 - 8 when the Gophers face the fast traveling, experienced St. James Club of Winnepeg at the Minneapolis Arena.

At present the St. James series is the only set of games definitely scheduled for Coach Larey Armstrong and his puckxters. Dates and arrangements have not been completed as yet for a meeting with the Canadian Navy team of Port Arthur, however Coach Armstrong feels confident that the dates of January 14 - 15 will be approved by Canadian immigration authorities and military officials which govern the passage of Canadian teams across the border, thus enabling the Port Arthur team to keep it's appointment in Minneapolis for their encounters.

Minnesota's old rival in the ice sport, the University of Michigan, may cancel it's games set for the Arena the later part of January due to a manpower problem facing the Wolverines this year. Most of the Michigan puck candidates are naval trainees and as such are under the forty-eight hour rule which forbids navy men to be away from their base in excess of the allotted time. On word received from Michigan athletic officials they claim it may prove impossible for them to send a team on the road for a two day stand and still abide by navy orders.

All this, however, has not discouraged Coach Armstrong for he has his squad going through daily drills for the past seven weeks in hope that additional games will be added to the schedule in the near future.

This year's schedule is quite a contrast to the fifteen game campaign of last year in which Minnesota won ten, lost four and tied one, winning

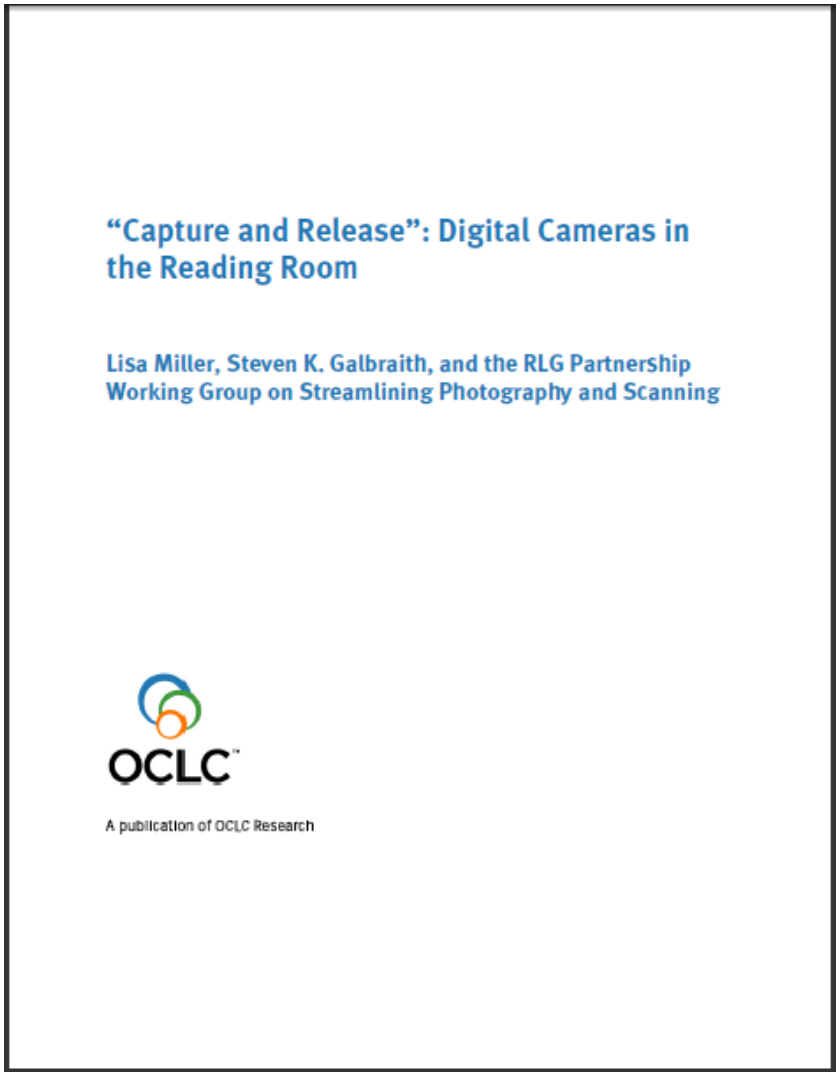
Sharing Special Collections

Advisory group made up of special collections and ILL staff

- Streamlining workflows for loan requests
- Model policy for lending special collections

Pascal Calarco. Interlibrary loan dept. KU Leuven library, Leuven, Belgium. 2008. (CC-BY-NC-SA)

Capture and Release



Helti. mladica / hucho hucho - catch and release. 2008. (CC-BY-NC-ND)

Capture and Release – Benefits

- Digital camera
- Digital camera
- Digital camera
- Digital camera room checkou
- Digital camera
- Digital camera infringement
- Repositories s
- Digital camera

Facet	Shutter-bug	Exposed	Camera-shy
Traditional photocopying (possible baseline for digital camera policy)	• Self-service	• Self-service after staff review	• All copying done by staff
Equipment	• No flash, no lights • Allow flatbed scanners • Allow and/or provide copy stand, tripod, extension cords, stepstool, etc. • Repository supplies camera or self-service overhead book scanner in addition to allowing patron's camera	• Patron's camera • Limits on supporting equipment (copy stands, tripods, cords, etc.)—some pieces allowed, others not • No flatbed scanners	• Repository's camera only (and possibly other equipment supplied by repository)
Photography space	• In reading room at any station	• In reading room at designated stations, usually close to reference desk	• Separate room
Photography rules	• No standing on tables or chairs • No rearrangement of furniture • No materials on floor • Remain behind table, facing forward at all times • Set camera to "mute" • Do not disturb others • No photographs of reading room, staff, or patrons		
Appointments	• Appointment not required	• Appointment made during visit	• Appointment made in writing in advance
Staff review of collection materials	• Part of standard staff surveillance of patrons in reading room	• Patron must verbally notify reference attendant each time camera is used and show attendant the materials being photographed	• Patron must formally indicate and curatorial staff formally review all materials • Camera stays in locker until approval is given • Same-day approval may

Scan and Deliver: 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

OCLC

Scan and Deliver – tiered workflow

A. Review

- *How do users request digital reproductions?*
- *How are requests for copies approved*
- *How many people need to approve?*
- *Have the items already been digitized?*
- *How do you manage copyright & legal issues?*

B. Decide

- *Will you keep the images?*
- *Is it useful for you to keep the images?*
- *Do you have sufficient resources to keep them?*
- *What resolution is needed for scanning?*
- *Is the request for high-resolution?*
- *What metadata will you create?*
- *Will you scan the whole volume?*

C. Scan

- *Who digitizes and describes materials?*
- *Will you perform quality control?*

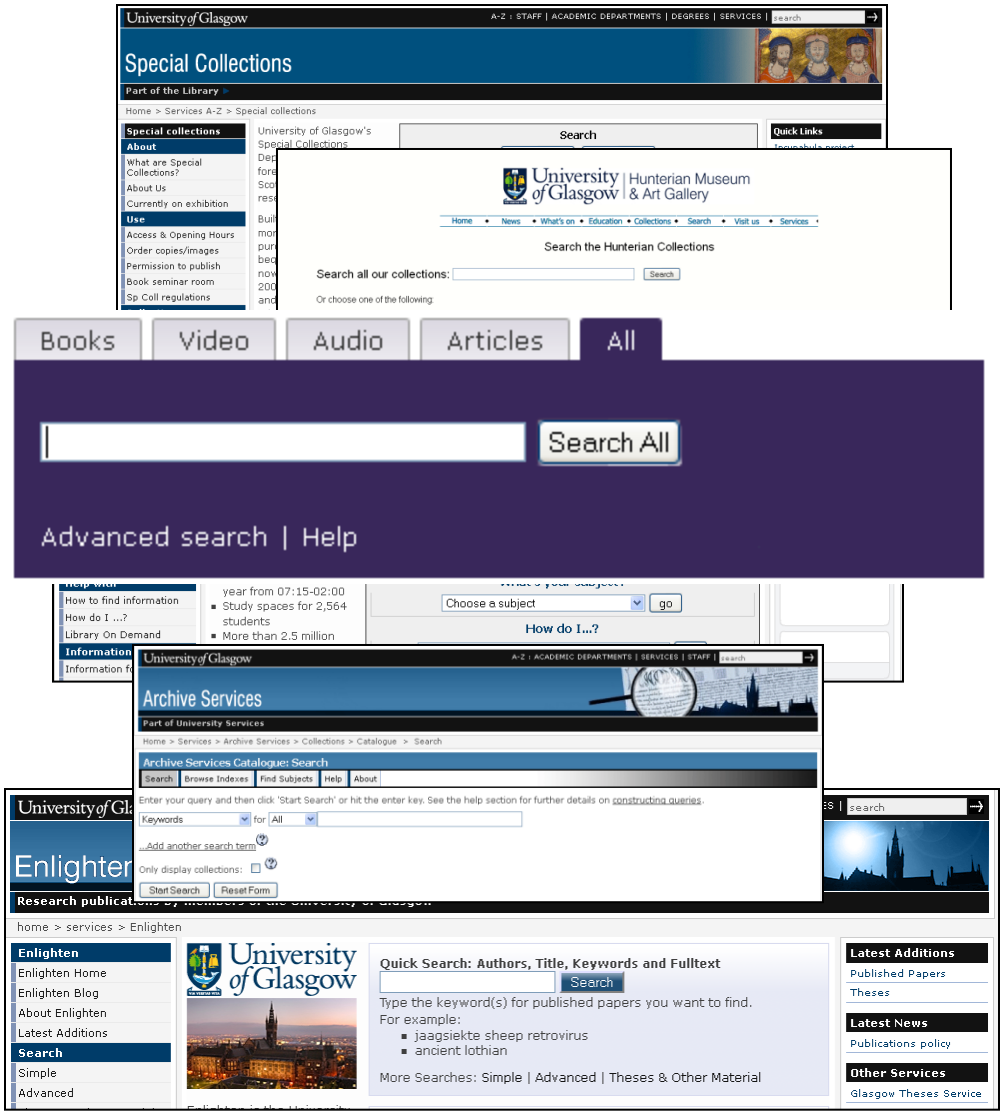
D. Deliver

- *How do you deliver digital copies to the user?*

	INSIDE TRACK	MIDDLE TRACK	OUTSIDE TRACK
REVIEW			
REQUEST	Verbal request	Simple form	Detailed forms
APPROVE	One staff member	Two staff	Three or more staff
SEARCH FOR EXISTING SURROGATE	Don't take the time	Quick check (local files or online)	Search: • Google Books • EEBO & ECCO • HathiTrust • Etc…
RIGHTS	User's responsibility	User's responsibility	Search, control, & monitor thoroughly
DECIDE			
STAFF	One staff member	Two staff	Multiple specialists: • Curator • Conservator • Metadata creator • Image professional • Etc…
RESOLUTION	Easiest option	Middling to high quality	High quality
METADATA	None	Filename and keyword or two	Metadata & citation (choose standards)
WHOLE / PART / ITEM	Will scan whole unit if it takes up to "x" minutes	Entire chapter or series	Entire volume or collection (large-scale digitization)
SCAN			
STAFF	One staff member	Two staff	Multiple staff
QUALITY CONTROL	Trust your staff!	Quick review	Thorough review by multiple specialists (color-correction, etc.)
DELIVER			
DELIVERY METHOD	Deliver forthwith to user	Keep in local files and Web site	• Digital library • Digital repository • DAMS • Consortia • Aggregations

Figure 1. Tiered Workflow for User-initiated Digitization

Single Search



The Quest for the Holy Grail:

Single Search across an Institution's Collections

Leah Prescott
Digital Projects Coordinator
Getty Research Institute

Ricky Erway
Senior Program Officer
OCLC Research

OCLC

A publication of OCLC Research

Single Search - Implementers Share Experiences

Victoria & Albert Museum

Yale Center for British Art

Yale University

University of Calgary

Minnesota Historical Society

The Magnes Collection of Jewish Art and Life

Getty Research Institute

Wellcome Trust

Smithsonian Institution

Motivation

Organizational structure

Technical approaches

Crosswalks

Gathering data

Vocabularies

Managing digital assets

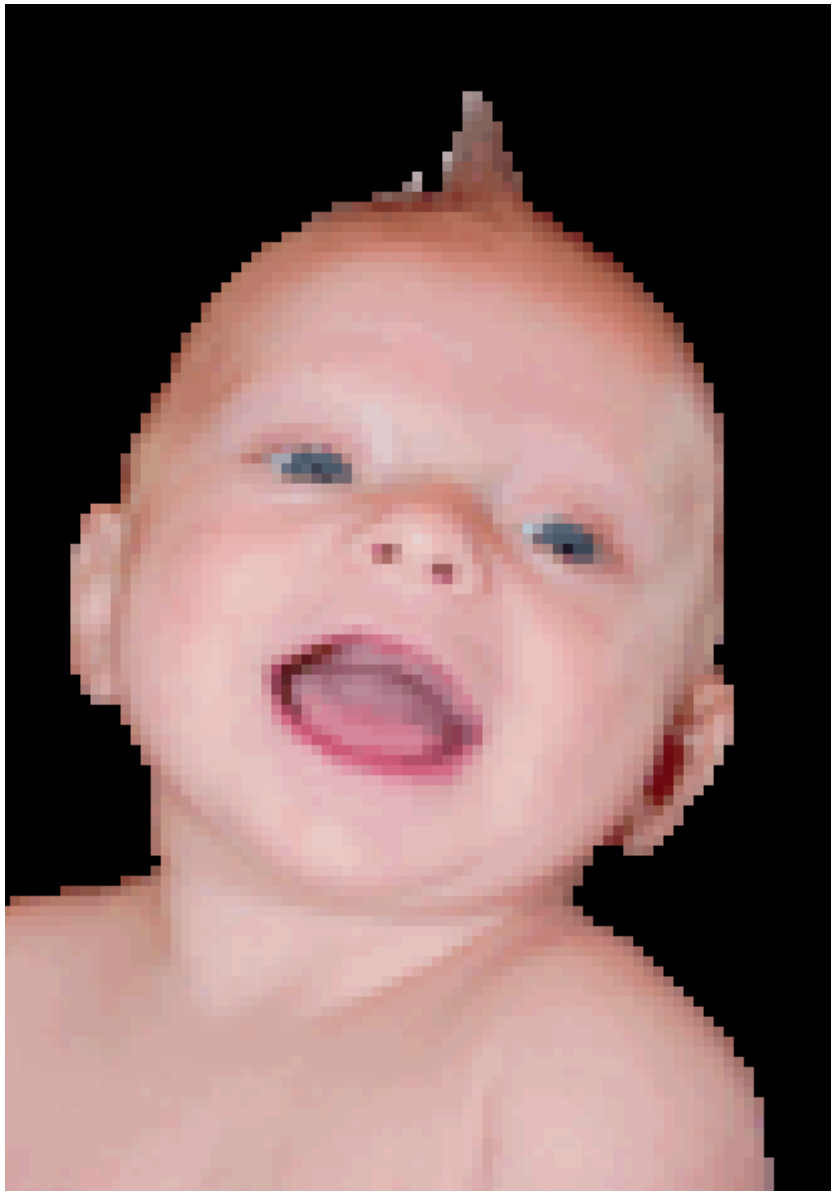
Access

Born Digital

Definition

Archival practices

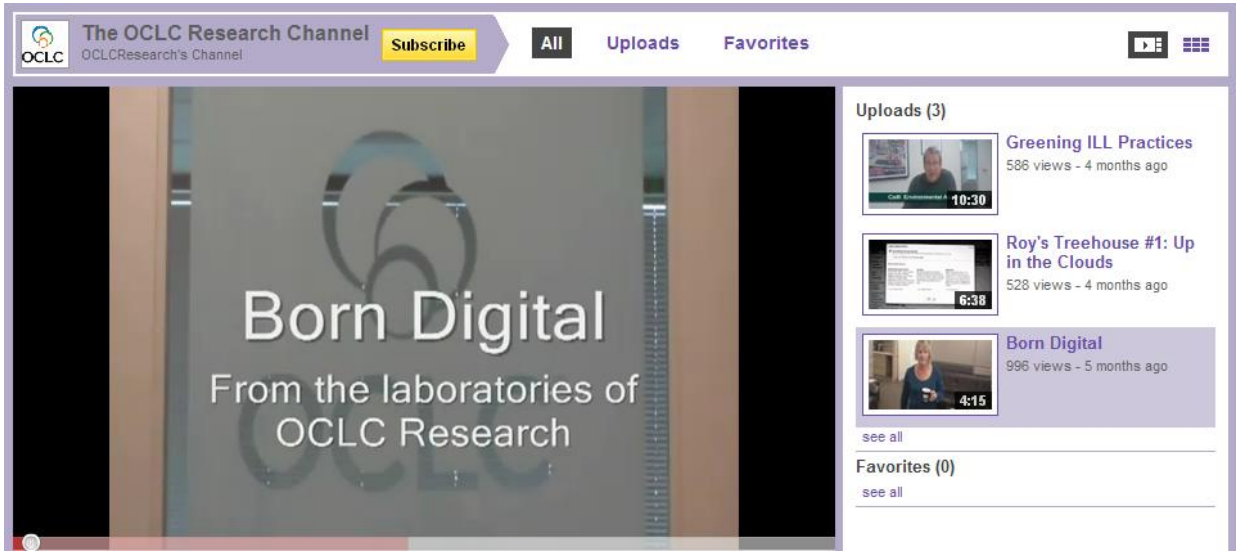
First steps



D Sharon Pruitt. Innocent Baby Laughing. 2009 (CC-BY)

Types of Born-Digital Materials

- *Digital photographs*
- *Digital documents*
- *Harvested Web content*
- *Digital manuscripts*
- *Electronic records*
- *Static data sets*
- *Dynamic data*
- *Digital art*
- *Digital media publications*



The Metadata *is* the Interface: Better Description for Better Discovery

The Metadata *is* the Interface
Better Description for Better Discovery of Archives and Special Collections, Synthesized from User Studies

Jennifer Schaffner

Program Officer
OCLC Research

OCLC

A publication of OCLC Research

Getting the most out of metadata

- Reduce the amount of effort we put into metadata, while improving discovery
- Do high-level description
- Iterate
- Describe the “aboutness”
- Encourage tagging
- Involve researchers
- Be where the users are

There is light at the end of the tunnel…

Eric Jones. Lon Peris cycle track seen from the other end of the tunnel. 2007. (CC-BY-SA)

OCLC Research: Mobilizing Unique Materials

Zeets Jones. Light at the End of the Tunnel. 2007. (CC-BY-NC-ND)

Links and contact info

- Shifting Gears http://www.oclc.org/research/publications/library/2007/2007-02.pdf
- Taking our Pulse www.oclc.org/research/publications/library/2010/2010-11.pdf
- Well-Intentioned Practice http://www.oclc.org/research/activities/rights/practice.pdf
- Good Terms http://dlib.org/dlib/november07/kaufman/11kaufman.html
- Rapid Capture http://www.oclc.org/research/publications/library/2011/2011-04.pdf
- Sharing Special Collections http://www.oclc.org/research/activities/sharing/default.htm
- Capture and Release www.oclc.org/research/publications/library/2010/2010-05.pdf
- Scan and Deliver http://www.oclc.org/research/publications/library/2011/2011-05.pdf
- Single Search – http://www.oclc.org/research/publications/library/2011/2011-17.pdf
- Born Digital Special Collections http://www.oclc.org/research/activities/borndigital
- Defining "Born Digital" http://www.youtube.com/user/OCLCResearch
  http://www.oclc.org/research/activities/hiddencollections/borndigital.pdf
- The Metadata *IS* the Interface
  http://www.oclc.org/research/publications/library/2009/2009-06.pdf

oclc.org/research/publications

erwayr@oclc.org