



Beyond Copyright: risk, benefit, and charting a course for action

Merrilee Proffitt & Ricky
Erway
OCLC Research

Annual RLG Partnership
Meeting
June 1, 2009

In this session

- Overview of (our) work to date
 - Registry of Copyright Evidence
 - SAA IP Working Group Best Practices
 - Google Book Settlement Analysis
- What is risk? What is reward?
 - Discussion about value proposition for research libraries
 - For print
 - For unpublished materials
- Help us plan for the future
 - Event?
 - What else?

A reminder of the salient facts

- Published materials
 - In the public domain if published in the US before 1923
 - 1923-1963: was it renewed?
 - To 1977: was it published without notice?
- Unpublished materials
 - Longer term: life of the author plus 70 years
 - In 2009, we have access to works from authors who died before 1939
 - If the date of death is unknown, still may be copyrighted after 120 years
 - TONS of “orphan works” in manuscript collections
 - Limited statutory damages for unpublished materials

A reminder of the salient facts

- Orphan work: a copyrighted work where it is difficult or impossible to find or contact the rightsholder
- Anonymous
- “Aunt Sally”
- Rights belong to heirs
- Rights belong to a new company
- The concept of “reasonable effort”
- The concept of a rights holders’ registry (in proposed legislation and in Google Settlement)
- Perpetually pending legislation: failed, considered bad (by all sides), could come back anew?

Copyright Investigation Summary Report

Participating RLG Program partner institutions and staff:

Cornell University
Peter Hittle

Emory University
Lisa Macklin

New York Public Library
Tom Lisanti

Rutgers University
Grace Agnew

Stanford University
Mimi Carter

University of California, Los Angeles
Angela Riggio

University of Michigan
Judy Aronheim

University of Texas at Austin
Dennis Dillon

Report prepared by:

Merrilee Proffitt, Arnold Arcolio, and Constance Malpas, RLG Programs

Background:

To support the RLG Programs work agenda, RLG Partners collaborate to identify new service areas and structures. This collaboration helps libraries, archives, and museums achieve a common understanding of the processes for which they should be responsible.

In this project, staff from eight partner institutions participated in copyright investigation interviews between August and September 2007 to share the ways in which their institutions currently obtain copyright permission to provide users with access to high-risk or special collection materials.

A few things to note about the participating partner institutions include:

- Some are involved in mass digitization projects, while others are not.
- Some are concerned with determining copyright status in order to provide access to digitized materials.
- All of these institutions are concerned with identifying rights-holders and securing specific permissions when scholars or businesses wish to use materials, often special collections materials, particularly photographs.
- In some cases, the same staff are responsible for both mass digitization and activities around archival or special collections materials.

www.oclc.org/programs/publications/2008-01.pdf
Proffitt, et al, for OCLC Programs and Research

March 2008
Page 1 of 6

January
2008,
6 p.

www.oclc.org/programs/publications/2008-01.pdf

Interviews from RLG Partnership (books)

- 8 institutions, not all working on “mass digitization”
- Identify “high risk materials” and eliminate them from pool in order to give quick access to the largest number of materials
- Books, published in the US, before 1923
- Not a lot of effort devoted to this work at this time
- Some well-established numbers from University of Michigan on costs for “low hanging fruit” and for identifying low risk materials to 1963.
- Left aside are riskier materials to 1963, materials published outside of US, materials after 1963
- Information stored in paper and electronic form, some storing information in MARC records

Interviews from RLG Partnership (unpublished)

- Mostly “permissions based”
- Work not transferable
- Information stored offline, not shared
- High effort, low return
- If efforts are non-conclusive, requests denied --
“We say no a lot”
- A profession paralyzed, collections shackled

OCLC Copyright Evidence Registry

- Open to anyone - you can contribute data!
- Draws primarily on information in WorldCat
 - Date and country of publication
 - Death dates of authors
- Additionally includes contributions from other sources
 - Stanford University's Copyright Renewal Database
 - University of Michigan (coming)
- "Copyright rules"
- Pilot continued through June
 - Contact Bill Carney (carneyb@oclc.org)

www.worldcat.org/copyrightevidence

Copyright Evidence Registry Advisory Group

- Initially interested in establishing best practices around orphan works
- Efforts stalled: participants in a wait-and-see mode
- Participants wanted to wait for outcomes of Google Book Settlement and assess impact of Book Rights Registry

Society of American Archivists IP Working Group (+)

- Group received sponsorship from OCLC Research
- Drafted in 2008, awaiting approval from SAA Council before publication
- Advises a step-wise approach for determining if unpublished materials may be considered as orphans
- Suggests sources, sets up a process
- Even though items may be in collections, focus is on item level assessment

Impact of the Google Book Settlement on Libraries

- <http://www.oclc.org/programs/publications/reports/2009-01.pdf>
- January 2009 (rev. March)
- Highlights
 - Most important for US in-copyright, not commercially-available books
 - Choices for contributing libraries
 - Research Corpus
 - For the rest of us
- But what does it all really mean?
 - Google and only Google has this arrangement
 - Appearance that Google has solved Orphan Works problem
 - Who's looking out for the people?

So what about print?

Efforts in monographs, published
in the US, to 1923, and 1923-
1963

Investment and reward?

The cost of “reasonable” effort

Transaction costs per book	Random, feasibility study		Fine & rare books study	
	Publisher	Title	Publisher	Title
Permission granted	\$228	\$197	\$251	\$65
Permission granted plus not located (OW)	\$129	\$109	\$147	\$54
Permission granted plus not located plus no response	\$81	\$67	\$138	\$50

Few titles
per publisher

Many titles
per publisher

Cost to digitize a book is \$25

Slide from Denise Covey Troll, CNI, April 2006

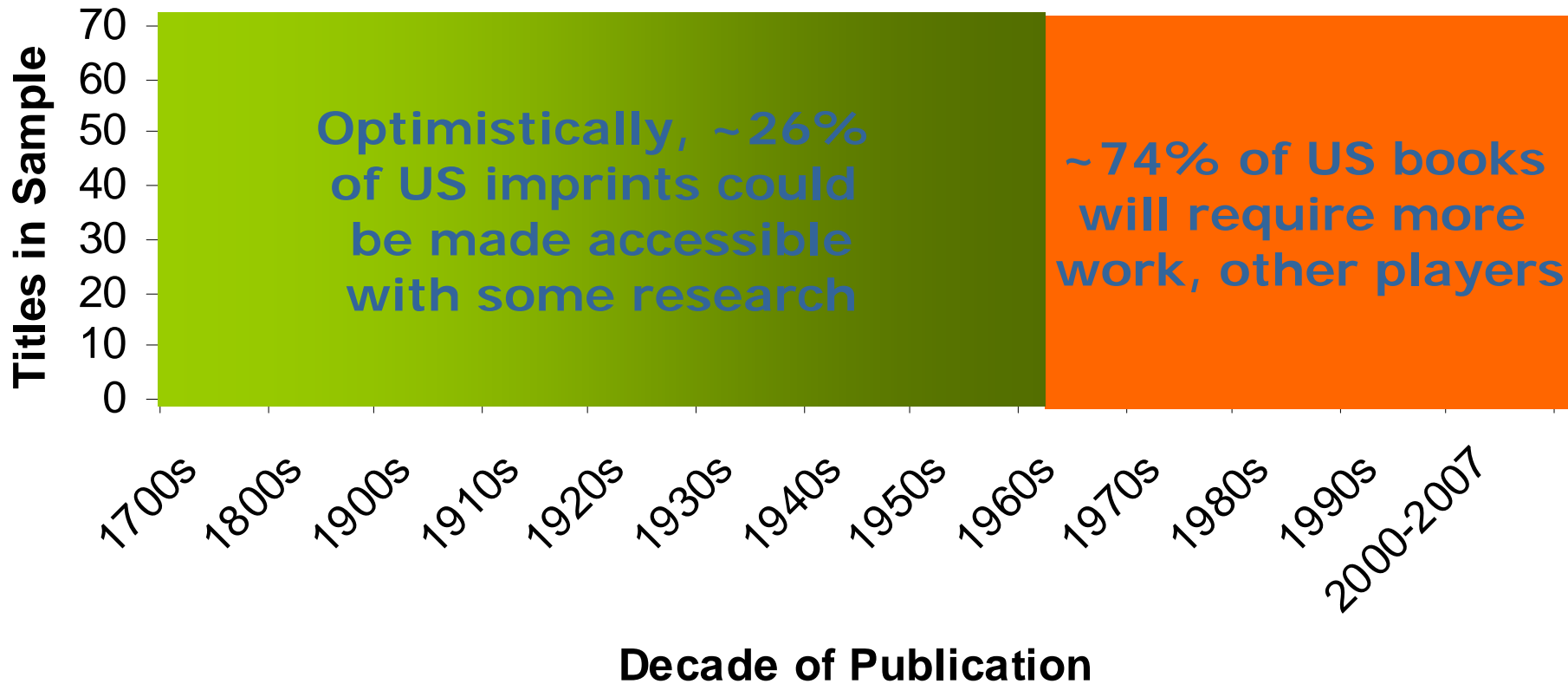
1923-1963: How much? What's the impact on research and teaching?

- Based on a January 2007 snapshot of WorldCat, we can estimate that ~15% of US imprints were published between 1923-1963; *~2M titles*
- Independent studies at Stanford and Michigan suggest that ~30% of US imprints are in copyright; *up to 70% may be in the public domain*
- An optimistic scenario: $\sim 2M * .70 = \sim 1.4M$ titles
- Add to this the pre-1923 books already in the public domain, est. ~15% of US imprints; *optimistically, a total of ~3.4M titles, or the volume equivalent of a mid-level ARL collection*

Suppose we go as far as we can with this?

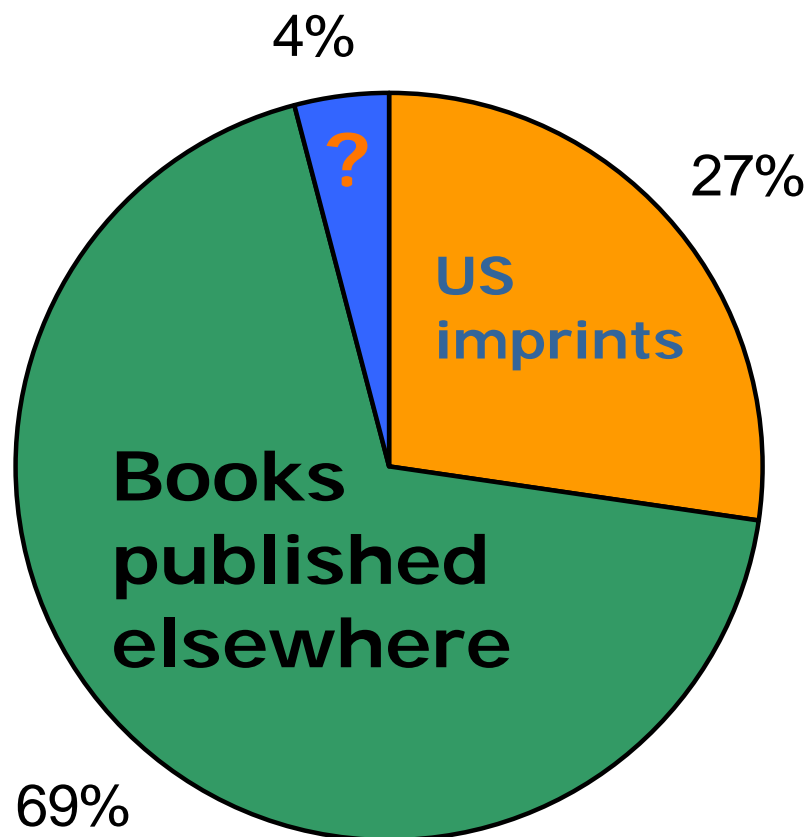
What's the likely impact?

US imprints in 1000 rec sample



What's missing from this picture?

Books published outside of the United States



Based on January 2007 snapshot of published print books in WorldCat
n = 48M titles

Demand: What access is needed to support scholarship?

Citations to US imprints (monographs only)	-1922	1923 - 1963	1964 - 1977	1978 - 1988	1989 -
<i>Lawrence and Aaronsohn</i> US imprints account for only 1/3 of works cited	8%	29%	12%	9%	41%
	8	28	12	9	40
<i>Shakespeare the Thinker</i> US imprints account for less than 1/4 of works cited	2%	38%	29%	19%	12%
	1	16	12	8	5
<i>The First Word</i> Almost all monographs cited published in the US. 2/3 of sources were from journal literature (not counted)		2%	6%	6%	85%
	0	2	5	5	70
	4%	21%	13%	9%	52%

Discussion

- Is clearing permissions or doing orphan works investigations a worthy investment for the research library?

So what about unpublished
materials?

Risk and reward?

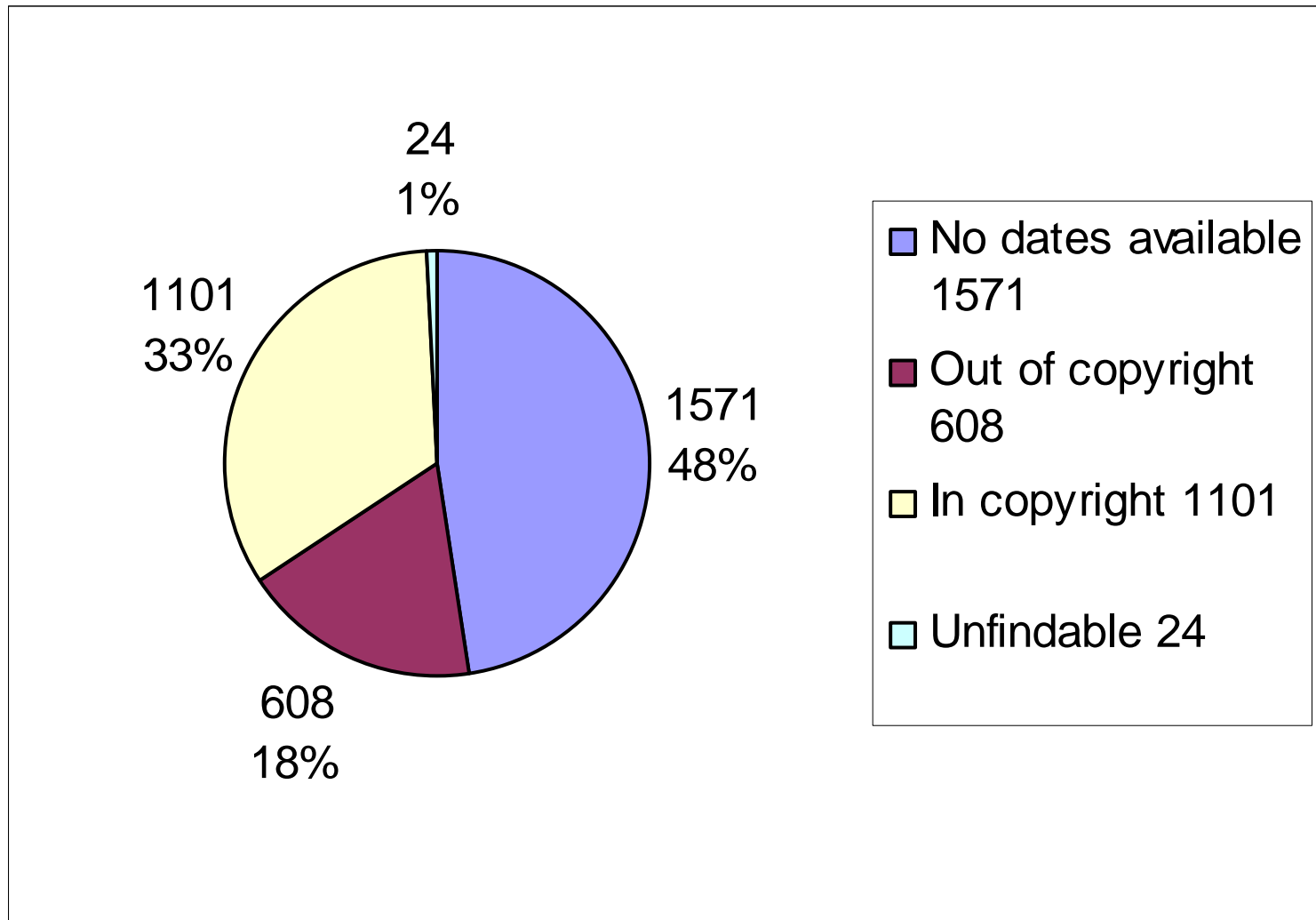
What diligence is “due”? What is reasonable effort?

University of North Carolina, Thomas E. Watson collection

- Privately funded two-year grant to digitize and publish
- 19.5 linear feet / 8500 letters, postcards, telegrams and notes
- Created by Thomas E. Watson, family, friends, and political and business colleagues: 3304 individuals!
- 1873-1986, bulk dates 1880s-1920s
- Sources included Wikipedia, Social Security Death Index, Ancestry.com, print reference works
- 14 weeks of FTE (or, 40 minutes per item)
- Results stunningly inconclusive

Thanks to Maggie Dickenson at UNC for providing this information

Watson results



What diligence is “due”? What is reasonable effort?

- Results questionable (sources, the “Aunt Sally” problem)
- No way to contact thousands of authors or their descendants
- My conclusion? A laudable experiment reveals a crazy amount of effort. No one should do this again.

Archival Code of Ethics

- IX “Archivists must **uphold all** federal, state, and local **laws.**”
- VII “Archivists **protect the privacy rights** of donors and individuals or groups who are the subject of records....”
- VI “Archivists strive to **promote open and equitable access** to their services and the records in their care without discrimination or preferential treatment...”
- III “Archivists should **exercise professional judgment** in acquiring, appraising, and processing historical materials....”

What is a balanced approach?

- Balance rights of copyright holders
- Minimize damage to reputation (and budget)
- Maximize access to collections
- Minimize damage to researchers of all types



CENTER FOR SOCIAL MEDIA

School of Communication American University



Newsletter

Subscribe to our free mailing list for event announcements, CSM news and our latest reports.

Your email address...

sign up

Feeds

- Subscribe to News RSS feed
- Subscribe to Blog RSS feed
- Get our video podcast
- Subscribe to iCalendar

Fair Use & Copyright Online Video Documentary Media Literacy

Fair use is the right, in some circumstances, to quote copyrighted material without asking permission or paying for it. Fair use enables the creation of new culture, and keeps current copyright holders from being private censors. With the Washington College of Law, the Center for Social Media creates tools for creators, teachers, and researchers to better use their fair use rights. Explore your fair use rights by clicking on one of the tabs above.

Publications

The Code of Best Practices in Fair Use for Media Literacy Education



This document is a code of best practices that helps educators using media literacy concepts and techniques to interpret the copyright doctrine of fair use. Fair use is the right to use copyrighted material without permission or payment under some circumstances—especially when the cultural or social benefits of the use are predominant. It is a general right that applies even in situations where the law provides no specific authorization for the use in question—as it does for certain narrowly defined classroom activities.

This guide identifies five principles that represent the media literacy education

Managing and accepting reasonable risk

- You are less likely get in hot water if you manage risk
- Assessment of collections along a variety of matrixes
 - Copyright
 - Violation of third party privacy and publicity rights
- Work with donors
 - Understanding of the creators of materials and copyright status of collection
 - Understanding of the content of the collection
- Some collections may require a “cooling period”
- Some collections may not be eligible for digitization
- Materials created for commercial purposes high(er) risk
- Record your decisions, share what you find

Managing and accepting reasonable risk (cont)

- Have good, clear take down policies
- Be responsive to complaints
- Web 2.0 may provide space for “the other side” to tell their story
- Establish a contingency fund for possible litigation
- Take one for the team! (We need case law.)

I am indebted to Maggie Dickenson, Laura Clark Brown, Peter Jaszi, Bill Landis, Aprille MacKay, Peter Hirtle, Denise Troll Covey, Dan Santamaria, and a cast of thousands for sharing their work and ideas with me.

So....

◀ Archives and Special Collections Program

Sharing Special Collections Project

Balance in Rights Management Project ▶

Streamlining Photography and Scanning

Managing Archival Collections Project

Barriers to Using EAD Project

Analyze Archival Descriptive Practice Project

Characterize "Hidden Collections" for Archives Project

Analyze Discovery Environments Project

RLG Programs : Our work : Managing the Collective Collection : Archives and Special Collections Program : Balance in Rights Management Project

Introduce Balance in Rights Management Project

Problem statement: Librarians and archivists often make extremely conservative judgments regarding the risk involved in copying unpublished collections. Many institutions have time-consuming, overly-cautious procedures to ensure vigorous compliance with copyright law—sometimes without a full understanding of the law or of the negative impact their procedures have on achieving their mission. If access is the goal, then any unnecessary restriction is counterproductive. The digital age has induced yet more caution, creating the ironic situation where, just when users ought to be getting improved services, they're not even getting as good a service as they could through interlibrary loan, in-person visits, and analog copying. The processes themselves are very costly, not just to the library or archive, but to society, in terms of what may be prevented from entering the scholarly record.

Impact: Streamlined rights procedures will maximize use of increasingly limited staff and financial resources and will increase service to researchers. Developing a community of practice will establish a baseline that can be followed with some degree of confidence, improving visibility of and access to special collections.

This activity will examine strategies for analyzing and developing acceptable risk behaviors and recommend practices for libraries and archives. We anticipate holding an invitational meeting, broadcast in real time, that collects imaginative thinking by experts from archives, special collections, time-based media and the law.

This project is lead by OCLC Research Senior Program Officers [Ricky Erway](#).

For more information

Ricky Erway
Program Officer
ricky_erway@oclc.org

Stray thoughts to spark discussion

- Working title: “Balancing act - responsible rights management with regard to unpublished collections”
- Focus on digitization of unpublished collections
- Invitational panel in San Mateo or LA - live-cast to all
- Encourage all speakers to have one or more provocative suggestions (and avoid rat-holes)
- Leave with a sense of what we can do now and what we should do next.

Possible agenda

- Mission-driven or risk-averse?
- Why things *have* to change
- Copyright with regard to unpublished materials (lawyer)
- The effect of following the rules
- Throwing down the gauntlet - what exactly is at stake?
- Examples of taking chances
- Balancing risk and benefit to society

More ideas?

- For the event?
- For follow-on activities?
- What positive thing are you doing now?
- What do you wish you could do?
- What troubles you?
- Any confessions to make?