

2014 DLF Forum

Big Collections in an Era of Big Copyright: Practical Strategies for Making the Most of Digitized Heritage

Tuesday, October 28, 9:00-10:30am

Salons 4,5,6, Georgia Tech Hotel and Conference Center

Session Leaders

John Mark Ockerbloom, University of Pennsylvania

Melissa Levine, University of Michigan

Jeremy York, HathiTrust

Mark Matienzo, Digital Public Library of America

Slides

<http://www.diglib.org/wp-content/uploads/2014/11/Big-Collections-in-an-Era-of-Big-Copyright.pdf>

Notes

Introduction by John Mark Ockerbloom -

1922 questions of copyright become tricky. Before, simple, after complicated.

In Europe copyright went into effect earlier.

Those of use building digital libraries are aware of the copyright issues inherent in digital collections. Education is allowed to use copyright controlled materials under the 4 rules.

up until 1989 works needed a copyright notice, or else they were open and out of copyright.

1998 Congress froze the public domain. It was challenged in court and went to the supreme ct. LibCongress did a complete scan of copyrights and renewal info. This renewal info was put online and used by many. Project Gutenberg produced a transcription for their work.

Stanford Univ then took this info and made a database called the Determinator (now used by HathiTrust).

Once this info is searchable, one can then do analysis and other exploratory work. The next step was to explore which journals were available and which contributions to said periodicals were under copyright. Penn has database of periodicals "First Copyright Renewals for Periodicals"

This is work that can be constantly built upon. On WikiSource, pulp magazine runs (Amazing Stories) are made available thanks, in part, to copyright info. Science fiction and utopian literature represent gender ideals or fears -- there was strong representation by women in both fields. Woman, sometimes, represented science fiction domestic lives. Lisa Yaszek's text, Galactic Suburbia, recovers women's science fiction. Many of the stories she mentions

are in public domain and 5 yrs after publication. Yaszek's text will be full-text available. These are examples on how to build deep collections in light of copyright law.

Jeremy York, Hathitrust.

The HT digital repository came about in 2008. 12.5 million volumes online. 4.6 million volumes online and publicly available. UC and Michigan are the largest contributors. a little less than half of the collection is in English. More than 2/3 of the HT collections are published outside of the US.

Missions: preserve materials, Enable access

Strategy to ensure use

take full advantage of the law allowing access

Framework:

- law
- Identification
- access policies
- technical infrastructure

There is an automatic rights determination done based on metadata

There is also a manual rights determination

Lawful Uses

Full text search

access for users w/disabilities

Take downs and deletions

- Take down
 - remove immediately
- investigate
 - reopen or keep closed
 - Delete

Got kicked off of wireless, lost the later speakers.

Levine, UMich Copyright Officer

CRMS - Copyright Review

HT provides oppty to resolve barriers of copyright indeterminacy (uncertainty of copyright status) to set works free. Going from Public Domain? To Public Domain. Their carefully and intentional system of determining PD status of a work relies on skilled and knowledgeable copyright researchers plus clever and well designed tools for automating ingest of copyright status information. This combination of approaches has allowed HT to clearly set free works

otherwise left behind access walls due to legal doubt. Next phase of project will export protocols and tools for use by other institutions needing to analyze copyright status of their own collections.