Digital Rights Management

Copyright Status Information in Metadata

DRM
• DRM Starts with copyright
  ---But it doesn’t stop there
• Libraries have always mediated between copyright holders and end users for broad yet appropriate use
• DRM is about seamless resource access, regardless of rights status

DRM
• Recording of copyright-related information has often not been part of cataloging/metadata creation tradition in libraries
  – MARC field 542 for information related to copyright status
• But essential for digital projects/collections

DRM
• Previously, Digital Rights Management (DRM) focused on security and encryption as a means of solving the issue of unauthorized copying, that is, lock the content and limit its distribution to only those who pay.
• This was the first-generation of DRM, and it represented a substantial narrowing of the real and broader capabilities of DRM.
• The second-generation of DRM covers the description, identification, trading, protection, monitoring and tracking of all forms of rights usages over both tangible and intangible assets including management of rights holders relationships.
• Additionally, it is important to note that DRM is the "digital management of rights" and not the "management of digital rights".
• That is, DRM manages all rights, not only the rights applicable to permissions over digital content.

Rights metadata
• Rights schemas with limited scope
• Full blown Rights Expression Languages (REL) for managing intellectual property rights, particularly by rights owners
• Rights information is not well understood
  – Different laws in different jurisdictions
  – Machine actionable vs. human understandable
• Rights take different forms
  – legal statutes, e.g. copyright
  – contractual rights, e.g. licenses

Rights schemas with limited scope
• METS Rights
  – Access rights for use with METS objects
  – Rights declarations
  – Rights holder
  – Context
• CDL copyright schema
  – Specifically copyrights, not other intellectual property rights
  – Information you need to know to assess copyright status (e.g. creators, rights holders, dates, jurisdiction)
Rights schemas with limited scope

- **PREMIS Rights**
  - Focused on rights for preservation rather than access
  - Revision of PREMIS data dictionary expanded this area
  - Allows for extensibility, i.e. inserting another rights schema
- **Creative commons**
  - Allows creators to choose a license for their work
  - Simple rights statements that fit a lot of situations

Rights metadata for specific object types

- **PLUS for images**
- **MPEG-21 REL for moving images, etc.**
- **ONIX for licensing terms**
- **Full Rights Expression Languages**
  - XRML/MPEG 21
  - ODRL

Building a Rights Framework

- Libraries and archives exist in increasingly complex digital environment
- Need for enterprise-wide solution
- Conceptual framework—a work-in-progress (overview, context, concepts, stages)
- Descriptive metadata for copyright

Private to Public

- Libraries and archives on both sides of issue
- What do you own?
- Once digitize what are rights?
- Stewardship and service
- Code as Control (digital work)
- Orphan Works

Building a Rights Framework

- Provide the broadest set of services to the greatest number of users
- Layered service model
- Interoperability
- Generalizable, scaleable solutions
- Preserving cultural heritage resources

Digital is Different

- Legal issues
- Analog to digital requires “copy”
- Reading and rendering require “copy”
- Posting on web require “copy” and “display”, “distribution”
- Ease and ubiquity of copies
- Rights change over time
- Recordkeeping (documenting) rights
Digital is Different

“Digital copies are perfect copies of the original. For digital content, production is reproduction.”


Digital is Different

“It is possible only to preserve the ability to reproduce the electronic record, rather than to preserve the electronic record ‘itself’.”

- Ken Thibodeau, InterPARES Preservation Task Force

Need for a Solution

- Digital libraries appraise, select, acquire, describe, manage, preserve and make available materials to users over time
- Rights issues arise throughout processes, workflows, services from appraisal, selection, ingest, preservation and access
- Acquired and built content

Rights Management Implementation Criteria

- Indemnifiable
- Manageable
- Maximally Accessible
- Transparent
- Well Documented
- Affordable
- Effective

Why Capture and Record Rights Metadata?

- To aggregate rights information obtained from multiple sources such as collection files, contracts, acquisition agreements, deeds of gift, published materials and databases, as well as curators, librarians, scholars, and users.
- To reduce the time people must spend independently researching the copyright status of a work.
- To identify and provide contact information for rights holder(s).

“Rights” and controls over access and uses of digital information

- Copyright
- Contracts or Licenses
- Technological controls
- Other “rights” and or controls including privacy, right of publicity
Copyright

- Congress has the power “to promote the Progress of Science and the useful Arts, by securing for Limited Times to Authors and Inventors the exclusive Right to their respective Writings and Discoveries”.
- The first concern copyright, the second patents.

The Meaning of Copyright

Ownership
- Owners hold specific rights. Copyright owners do not have all rights.
- Author is the copyright owner.
- Employer may be the copyright owner.
- Copyrights can be transferred.

Things that can not be copyrighted
- Ideas can not be copyrighted. Expressions of ideas can be copyrighted.
- Facts can not be copyrighted
- Databases are a grey area.
  - Feist vs Rural telephone said that the phonebook is not copyrightable. It does not have the creative element
  - database industry is pressuring for copyright
- in many instances the copyright holder is not that important, but the licensing is.

Fundamental Principles of Copyright

- Copyright law applies to nearly all creative and intellectual works.
- Works are protected automatically, without copyright notice or registration.
- Copyright protection lasts for many decades.

Fair Use and Other Rights of Use

- Activities within fair use are not infringements.
- Fair use is one of many statutory rights to use copyrighted works.
- Uses are also allowed with permission.

What is copyrightable

- literary works
  - books, articles, software, web pages and sites
- musical works and their words
- drama, pantomime, choreography
- pictures, graphics, sculpture
- motion pictures and other audiovisual works
- sound recordings
- architectural works
Rights of copyright holders

• main rights are
  – right to reproduce
  – right to distribute copies
  – right to prepare derivative works
  – right to perform or display the work publicly

• the only important limit are
  – the concept of fair use
  – and section 108 that deals with libraries.
    • permits libraries to make copies of materials for preservation and security, to give copies to users for their private study or research, and to send copies through interlibrary loan.

Doctrine of fair use

• in the US, it is in section 107: “The fair use of a copyrighted work, including such use by reproduction in copies or phonorecords or by any other means specified by that section, for purposes such as criticism, comment, news reporting, teaching (including multiple copies for classroom use), scholarship, or research, is not an infringement of copyright.”

How to determine fair use

• In determining whether the use made of a work in any particular case is a fair use the factors to be considered shall include
  – the purpose and character of the use, including whether such use is of a commercial nature or is for nonprofit educational purposes;
  – the nature of the copyrighted work;
  – the amount and substantiality of the portion used in relation to the copyrighted work as a whole; and
  – the effect of the use upon the potential market for or value of the copyrighted work.

Change in Copyright Law

• Before 1986: duration based on date of copyright or creation
• Prior 1989: notice required
• After: duration based on life of the author
• Now: life of the author plus 70 years

"a big mistake, but one that Congress can make”

Marybeth Peters, Register of Copyrights, on the latest extension to the copyright term
Public Domain

- Published in the U.S. before 1978 with No Copyright Notice = PUBLIC DOMAIN
- Published in the U.S. before 1964 without Renewal = PUBLIC DOMAIN
- Rule of Thumb: Sail the Ocean Blue through 1922
- Rocks at the Edges
  - Don't Use Newer Versions of Expired Works
  - Don't Use Trademarks: Like Diamonds, Trademarks Last Forever
  - Special Concerns about Digitizing Foreign Works

Section 108

- SECTION 108 allows libraries and archives to digitize works in the last twenty years of their copyright term.
- Section 108 has additional exceptions for preservation, security and replacement.
- These latter provisions allow for digitization, but you can't put the copies on the web. They must stay in the library.

Fair Use

- Courts determine whether or not a use is a Fair Use on a case by case basis.
- We don't know whether a library digitization project (or a particular item) is Fair until one of us gets sued and wins (or loses, but then at least we'll have some guidance).

Get Permission

- The good news is that getting PERMISSION gives you a green light to go ahead without reservation.
- The bad news, is it's really, really hard to locate the rights' holders because they're likely dead and gone, and you are tracking down probate records and heirs
- Might take a fresh look at FAIR USE

Why would libraries and users benefit from copyright metadata schema now?

- Copyright controls “copying” and “copies” of creative expression
- Every use in digital environment creates a “copy”
- Every use potentially implicated or controlled by copyright
- Libraries throughout time mission provide long-term access and use of collections
- Copyright metadata schema designed to support use and uses of digital information

Goal:

“Save the time of the reader”

S R Ranganathan
5 Laws of Library Science
1931
Copyright status is not static

- Circumstantial
  - Depends on user and uses
- Changes over time
  - Assessment needs to be made for each use

Copyright Clearance

- Copyright is an automatically assigned right. It is therefore likely that the majority of works in a digital collection will be covered by copyright, unless explicitly stated. The copyright clearance process requires the digitizer to check the copyright status of:
  - Published, unpublished and Web site articles
  - Photographs and illustrations
  - Dynamic media (sound, video)
  - Software components
  - Database usage

Copyright Clearance

- Am I using unaccredited copyrighted material produced by others? - Published work that contains unaccredited material infringe upon the intellectual property of others. The results of such discovery will vary: the unaccredited author may request they are credited or a correction is published; the author may request their work is removed; or they make take legal action against the author. To avoid such issues, document all research made during investigation.

Copyright Clearance

- Copyright clearance should be established at the beginning of a project. If clearance is denied after the work has been included in the collection, it will require additional effort to remove it and may result in legal action from the author.
  - Maintain a negotiation log - A log will document all meetings, outlining subjects of discussion, objections and agreements by either party. This will enable the organization to refer to the relevant section to establish they have gained copyright clearance and refer to a detailed description of the meetings that took place.
Copyright Clearance

- **Identify who the author is and when it was produced** - Current copyright law indicates the author's lifespan plus 70 years as the limit for copyright. Therefore it is possible that a collection may consist of works that are outside current copyright laws (such as the entire works of Shakespeare, Conan Doyle, etc.). If the author is still alive, they must be contacted to gain permission to use their work.

Copyright Clearance

- **Establish long-term access rights** - Internet content may appear in a site archive for several years after it was published. When meeting the author, establish any time factors in use of their work, indicating the length of time that work can be used. If the goal of the project is to enable long-term preservation of work, persuade the individual/s to allow the repository to host work indefinitely and translate it to modern formats when required.
  - In the event that an author, or authors, is unobtainable, the project is required to demonstrate they have taken steps to contact them.

Three common reactions when the issue of rights metadata arises:

- "It's too complicated and overwhelming."
- "We don't have the staff or the money."
- "It's not the library's [or archive's, or museum's] job; it's up to users to figure out rights information if they want to publish something from our collections"

The Reality

- Institutions are already spending staff time and money on rights research.
- Capturing rights metadata in a shared information system as a routine, programmatic activity with structured data rules and values and an established work flow should not cost any more than ad hoc rights research, and it will provide longer-lasting benefits.
- In a world where "if it's not digital, it doesn't exist," libraries, archives, and museums have new roles with respect to their users, as well as the creators and authors of the works in their collections.

The Reality

- Cultural heritage institutions need rights information for their own uses of the works in their collections.
- Rights metadata is not just about compliance with intellectual property laws; it is also about being responsible stewards of the works in our collections and the digital surrogates of those works that we create—and in a digital world.
- It is crucial to a memory institution's broader mission of collection, preservation, and access.

Core Information

- Name of creator of work or image, including nationality, date of birth, date of death. Ideally obtained from authority file.
- Year work was created. May not be the publication date. If two dates exist, identify separately.
Core Information

• Copyright status
  – Copyright owned by the institution
  – Copyright owned by third party
  – Public domain
  – Orphan work
  – Not researched

• Publication status
  – Published
  – Unpublished
  – Unknown
  – Not researched

• Date that rights research was conducted

Copyright

• Knowing the birth and death dates of the creator, or the year(s) in which the work was created and published, will allow for quick calculations about the copyright term of the work
• Unpublished works tend to have longer copyright terms than published works
• Standard for fair use of unpublished works is usually higher than for published materials
• For works published in the U.S. between 1923 and 1963, renewal of the copyright registration was required

• Creation date may determine when the copyright terms begins and ends
• Identify orphan works as such
• Prior to 1978, the law required that a copyright notice be affixed to published works

Rights Metadata

• Any rights metadata effort should be viewed as dynamic and ongoing.
• New information may come from various sources: a user, a curator, a librarian, or even the creator of the work.
• Rights information needs to be updated and augmented, and additional information will need to be captured for works with more complicated rights situations, such as audiovisual materials.
• It is important that staff tasked with inputting rights metadata be identified to all those involved in cataloging and digitization efforts so that when new rights information is discovered, it can be input into the institutional database.

What are the metadata elements that can support the copyright status of the item?
Available Metadata Formats

- MODS
- Dublin Core
- METSrights
- MARC21

“The usual suspects”

Data Elements

- Creator(s)
- Rights holder(s)
- Publishing information
- Creation information

Services

Publishing Information

- Published/unpublished?
- Publisher name
  - Publisher contact information
- Date of publication
- Country of creation or publication
Creator(s)

- Creator name
  - Personal or corporate?
- Personal creator death date

Rights Holder(s)

- Rights holder name
- Rights holder contact (holder or agent)
- Rights statements from work

Services Contact

- Contact information
- Notes
Dealing with Unknowns

- Information is unknown
  - E.g. photo with no information about photographer, location, date
- Information is not provided in metadata
  - E.g. archive was unable to examine each piece
- Information exists but is not for public display
  - “Contact archive for more information”

Dealing with Uncertainty

- Date is...
  - Exact
  - Approximate
  - Unknown

What’s required?

```xml
<copyright
copyright.status="copyrighted" publication.status="unknown"/>
```

copyrightMD Schema

http://www.cdlib.org/inside/projects/rights/schema/
Creative Commons

• Creative Commons is a particularly popular licensing model available to all creative works.
• It is therefore usual to find it applied to Web sites, scholarship, music, film, photography and literature that are not traditionally covered by similar distribution schemes.
• Creative Commons (CC) refers to a movement started in 2001 by US lawyer Lawrence Lessig that aims to expand the collection of creative work available for others to build upon and share.

Creative Commons

• The Creative Commons model makes a distinction between the big C (Copyright) meaning All Rights Reserved and CC meaning Some Rights Reserved.
• It does so by offering copyright holders licenses to assign to their work, which will clarify the conditions of use and avoid many of the problems current copyright laws pose when attempting to share information.

Creative Commons

• Creators choose a set of conditions they wish to apply to their work.
  – Attribution. You let others copy, distribute, display, and perform your copyrighted work — and derivative works based upon it — but only if they give credit the way you request.
  – Noncommercial. You let others copy, distribute, display, and perform your work — and derivative works based upon it — but for noncommercial purposes only.
  – No Derivative Works. You let others copy, distribute, display, and perform only verbatim copies of your work, not derivative works based upon it.
  – Share Alike. You allow others to distribute derivative works only under a license identical to the license that governs your work.