

Rights Metadata Made Simple

Maureen Whalen

There are three common reactions when the issue of rights metadata arises:

1. “It’s too complicated and overwhelming.”
2. “We don’t have the staff or the money.”
3. “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.”

Here are some reasoned responses:

1. Yes, rights metadata can be complicated and overwhelming, but so is knitting a cardigan sweater until one simplifies the project by mastering a few basic techniques and following the instructions step-by-step.
2. Your institution is probably 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.
3. 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. Moreover, 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.

Usable, shareable, repurposable rights metadata can be obtained by capturing the following core information:¹

1. **The name of the creator** of the work or image, including the **nationality** and **date of birth**, and the **date of death**, if applicable. Ideally, this information should be copied automatically from an authority file. (*Generally, the “work” is the original work in the institution’s collection and not a digital surrogate. If the institution wants to create a rights metadata record for the digital surrogate, the approach described here would be valid, provided that the digital surrogate is described and differentiated from the original work.*)
2. **The year the work was created.** The year of creation may not be the year of publication. When two different dates exist, they should be identified separately. If the publication date is known, it should be recorded in the “publication status” field.
3. **Copyright status** (*one of the following five options can be selected from a controlled picklist by staff tasked with recording rights metadata*):
 - **Copyright owned by the institution.** The copyright is assumed valid and is owned by the institution that holds the work.
 - **Copyright owned by a third party.** The copyright is valid and is owned by someone or some entity other than the holding institution. If known, capture the name of the third party in a database field/metadata element designated for that purpose.²

¹ These suggestions for a simplified rights metadata approach are based on required rights metadata recommendations for copyrightMD, an XML schema for rights metadata developed by the California Digital Library (CDL). The copyrightMD schema is designed for incorporation with other XML schemas for descriptive and structural metadata (e.g., CDWA Lite, MARC XML, METS, MODS). See <http://www.cdlib.org/inside/projects/rights/schema/>. See also Karen Coyle, “Descriptive Metadata for Copyright Status,” *First Monday* 10, no. 10 (October 2005). http://www.firstmonday.org/issues/issue10_10/coyle/index.html.

N.B. The title of the work is not identified here as a rights metadata element per se; it is assumed that the title would be included in any metadata schema used to describe the work, and thus that element could be copied from the descriptive metadata record into the rights metadata schema in an automated manner.

² There may be certain conditions under which a license for certain specified uses of the work may have been granted to the institution. A license is not the same as ownership. If desired, when the copyright is known to be owned by a third party, the picklist could include an option for “license granted to the institution”; such a notation by itself, however, would not

- **Public domain.** If the work is determined to be in the public domain, it is helpful to identify the year in which the work entered (or will enter) the public domain, if known.
 - **Orphan work.** This is a work that may be protected by copyright law but for which the copyright owner or claimant cannot be identified or located.
 - **Not researched.**
4. **Publication status** (*one of the following four options can be selected from a controlled picklist by staff tasked with recording rights metadata*):
- **Published.** Include date, if known. Publication is defined in the Copyright Act as “the distribution of copies . . . of a work to the public by sale or other transfer of ownership, or by rental, lease, or lending.” Note that the offer to distribute copies, including the original work even if there is only one copy of it, constitutes publication.³
 - **Unpublished.** Some materials such as manuscripts and correspondence may be easily determined to be unpublished. Other works, however, such as speeches or paintings that are known to the public can still be considered “unpublished” under the Copyright Act definition.
 - **Unknown.** It is sometimes difficult to determine whether or not a work has been published, particularly for photographs of which there may be multiple prints or for manuscripts from which a work was later published.
 - **Not researched.**
5. **Date that rights research was conducted** (*if there are multiple dates on which rights research was conducted, best practice would be to include all of those dates, along with the initials of the researcher(s)*).

Gathering rights metadata and including it in an institutional information system or database⁴ will allow users with some basic copyright understanding to make thoughtful judgments about how the law may affect

be adequate to describe the various rights granted, or denied, or the specific term during which the license is valid, so a review of the specific licensed rights would be necessary.

³ 17 USC § 101.

⁴ There is increasing discussion about embedding rights metadata in the same file as the digital surrogate, thus avoiding the problem of two digital files that can and often do get separated during transmission. As of this writing, embedding rights metadata has been done only under limited circumstances and the software necessary to embed the data and provide users with access to it using a free, downloadable reader is not yet widely available.

use of the work in accordance with a legal exception.⁵ It may also help to guide determinations about how easy or how difficult it might be to obtain permission, if needed.

Table 1 gives specific examples of rights metadata for works in the public domain and works that are under copyright. Here are some examples of how the rights metadata elements articulated here can be applied in day-to-day decision making:⁶

- 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 for the work. To do the analysis and arithmetic, follow Peter Hirtle's excellent chart, *Copyright Term and the Public Domain in the United States*.⁷ Note: There are slightly different rules for works of foreign (non-U.S.) origin, including restoration of copyrights in works of foreign origin that may have been in the public domain for a period of time before restoration; that is why it is good practice to identify the nationality of the creator, if known.
- Unpublished works tend to have longer copyright terms than published works; therefore, if the work is assumed to be unpublished, the term of copyright protection should be calculated in accordance with the formula for unpublished works.
- While the Copyright Act specifically states that unpublished works qualify for fair use, courts tend to protect the creator's right to decide about first publication, so the standard for fair use of unpublished works is usually higher than for published materials.⁸

⁵ The U.S. Copyright Act includes a number of limitations on (rights holders') exclusive rights. The most well known of these limitations is fair use (Section 107 of the Act), whereby use of copyrighted works without permission of the rights holder is permitted if the use meets the statutory four-factor test. Another important exception applies to libraries and archives (Section 108 of the Act). Under this exception, libraries and archives are permitted to make copies of works in their collections under certain circumstances without permission of the rights holder, including replacement copies of published works, preservation and security copies for unpublished works, and copies for users provided that the copy becomes the property of the user and is for private study, scholarship, or research.

⁶ Examples include assumptions based on U.S. copyright law; examples and assumptions for non-U.S. jurisdictions are not provided here.

⁷ Available at http://www.copyright.cornell.edu/training/Hirtle_Public_Domain.htm. Also available as a PDF document at <http://www.copyright.cornell.edu/training/copyrightterm.pdf>.

⁸ "§ 107. Limitations on exclusive rights: Fair use. Notwithstanding the provisions of sections 106 and 106A, 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. In determining whether

If the rights metadata states that a work is unpublished, the user can assess the impact of that status on the fair use analysis.

- For works published in the United States between 1923 and 1963, renewal of the original copyright registration was required.⁹ Therefore, a work published in 1945 with the correct copyright notice and registration would require a renewal of the original copyright in 1973 (1945 + 28 = 1973) in order for that copyright to be valid today. One study indicates that 15 percent or less of the works in their original copyright terms between 1923 and 1963 were renewed.¹⁰ This means the majority of works initially protected by copyright during this period are now in the public domain. Of course, the more famous the work, the greater the likelihood that the original copyright registration was renewed. By contrast, renewals of registrations for more obscure works may be less likely.

the use made of a work in any particular case is a fair use the factors to be considered shall include—

- (1) the purpose and character of the use, including whether such use is of a commercial nature or is for nonprofit educational purposes;
- (2) the nature of the copyrighted work;
- (3) the amount and substantiality of the portion used in relation to the copyrighted work as a whole; and
- (4) the effect of the use upon the potential market for or value of the copyrighted work.

The fact that a work is unpublished shall not itself bar a finding of fair use if such finding is made upon consideration of all the above factors.” (Emphasis added; available at <http://www.copyright.gov/title17/92chap1.html#107>.)

Prior to passage of the Copyright Act of 1976, fair use was based on court decisions. In 1985 the U.S. Supreme Court in *Harper & Row Publishers, Inc. v. Nation Enterprises* (471 U.S. 539) ruled on the applicability of the fair use defense to unpublished works noting the “author’s right to control the first public appearance of his undissemated expression will outweigh a claim of fair use” (p. 555). In order to clarify how the unpublished nature of a work was to be evaluated under the fair use four-factor test set forth above, and to reverse a growing presumption that fair use was not available as a defense against an infringement claim for all unpublished works, Congress passed an amendment to the law in 1992, and the last sentence of this section was added—the one in boldface above. Notwithstanding this amendment, there is general legal consensus that courts will give greater weight to the unpublished nature of the work in fair use cases than would be given if the work had already been published.

⁹ All terms of original copyright run through the end of the 28th calendar year, making the period for renewal registration in the above example December 31, 1973, to December 31, 1974. When checking the Copyright Office renewal records, it is advisable to look at the years immediately preceding and following the calculated year for copyright term expiration. This will ensure that the work was not renewed properly in a different year.

¹⁰ William M. Landes and Richard A. Posner, “Indefinitely Renewable Copyright,” University of Chicago Law & Economics, Olin Working Paper No. 154 (August 1, 2002). <http://ssrn.com/abstract=319321>.

- Creation date may determine when the copyright term begins and ends; it is especially important when the author is unknown, the work is a work made for hire, or the work is one of corporate authorship, that is, a work created by a company such as a movie studio or record company.
- In 2006 the U.S. Copyright Office issued its report on orphan works.¹¹ Later that year, hearings on orphan works were held in both the U.S. House of Representatives and the U.S. Senate, and legislation amending the Copyright Act to reduce the legal liabilities relating to use of orphan works was introduced in the U.S. House of Representatives. While this legislation did not pass, many experts think that orphan works legislation will be enacted in the next few years. If so, the hope is that penalties and remedies for use of orphan works will be reduced or eliminated altogether. For that reason, it makes sense to identify orphan works as such. Moreover, regardless of whether or not orphan work legislation passes, it seems reasonable that if an institution attempts to identify and/or locate the copyright claimant and cannot do so despite diligent efforts, and this is explained to the court, there may be some recognition of this good faith activity by the judge if an infringement claim is brought by the emergent copyright claimant.
- Prior to 1978, the law required that a copyright notice be affixed to published works. Failure to include a legally sufficient notice put in the public domain American works that were published in the United States (without the notice). Therefore, an institution may decide to classify works as in the public domain if they were purchased before January 1, 1978, or were believed to have been offered for sale to the public before that date and there is no copyright notice affixed to the work.
- Obviously, if one knows a work is in the public domain or if the institution owns the copyright, permission to use the work is not required by law, although local policy may require internal authorization.

In order for catalogers and rights metadata analysts to be able to populate the recommended rights metadata elements, the institution will need some basic rules or assumptions to apply when copyright and publication status may not be clear and some suggestions for resources to help locate the sought-after information. There are numerous recommendations for where to look for the information requested. Currently, there

¹¹ *Report on Orphan Works: A Report of the Register of Copyrights*, January 2006, United States Copyright Office. <http://www.copyright.gov/orphan/>.

Table 1. Example of Core Elements for Rights Metadata

Metadata Element	Valid Data Values for This Element	Example—Public Domain Work	Example—Work Not in the Public Domain
Title	The data values for this element should be copied (preferably in an automated manner) from the title element from the descriptive metadata record for the work or item. Per <i>Cataloging Cultural Objects</i> , this element, which is repeatable, can contain translated titles, brief titles, display titles, etc., in addition to the title that is inscribed on the item or object, if one exists. Include a subelement for the parent object/work ("title larger entity") when applicable.	<i>Puzza in the Likeness of Isis, Seated on a Lotus Flower/Puzza sous une forme parallele à Isis, assise sur la fleur de lotos</i> from <i>Cérémonies et coutumes religieuses de tous les peuples du monde</i>	<i>San Diego Stadium (San Diego, California)</i> from <i>Julius Shulman photography archive</i>
Creator	The name of the creator of the original object or work, taken from a published controlled vocabulary (e.g., LCNAF, LCSH, ULAN) or local authority file whenever possible. The life dates in the case of individual creators, including the death date if applicable. Dates should be expressed according to a standard format, e.g., ISO 8601.	<i>Picart, Bernard</i> <i>b. 1673-11-06</i> <i>d. 1733-08-05</i>	<i>Shulman, Julius</i> <i>b. 1910</i>
Creation Dates	The date(s) of the creation of the work.* Dates should be expressed according to a standard format, e.g., ISO 8601.	<i>1723–1743</i>	<i>1967</i>
Creator Nationality	The nationality or culture of the creator of the work, if known	<i>French</i>	<i>American</i>
Copyright Status	Valid values for this element should be selected from a controlled list, e.g.: <ul style="list-style-type: none"> • Copyright owned by the institution that holds the original object/work or item • Copyright owned by a third party—Include a subelement for the name of the third party, taken from a published controlled vocabulary whenever possible. • Public domain • Orphan work • Not yet researched 	<i>public domain</i>	<i>copyright owned by institution</i> <i>© J. Paul Getty Trust</i>
Publication Status	Valid values for this element should be selected from a controlled list, e.g.: <ul style="list-style-type: none"> • Published—Include a subelement with the date of publication, if known, in a standard format, e.g., ISO 8601. Note that date of creation and date of publication are not necessarily identical. • Unpublished (in which case, the creator dates and/or date of creation are extremely important) • Unknown, after research and due diligence • Not yet researched 	<i>published</i> <i>1723–1743</i>	<i>not researched</i>
Date of Rights Metadata Research	This should be a repeating element, since metadata research is often necessarily an incremental process to which more than one individual contributes. The individual's name or initials should be provided by the information system, and associated with the relevant dates of research. Dates should be expressed according to a standard format, e.g., ISO 8601.	<i>2008-10-07 MTW</i>	<i>2007-09-13 MTW</i>

* Note that under current U.S. copyright law, a work is protected for the life of an individual author/creator plus 70 years regardless of the date of creation. The copyright term for corporate works and works made for hire is 125 years from the date of creation, or 95 years from the date of publication.

is no resource that sets forth commonly accepted practices regarding what is legally reasonable to assume about copyright or publication when only limited information is available, so institutions will need to draft their own guidelines.¹² Of course, local policy regarding use of material presumed to be protected by copyright and the institution's risk tolerance for infringement claims that arise in case the assumption is wrong will govern use decisions.¹³ With a little effort, however, the basic information needed to make informed decisions about rights for many works in an institution's collections can be easily available and accessible if the suggested rights information is captured.

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. Therefore, 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.

Now is the time to get started and not to be overwhelmed. Rights metadata can be made simple if everyone in the institution is aware of its long-term importance and there is a concerted, coordinated effort to research it, record it according to standards and best practices, and share it in fulfillment of the institution's mission in the digital age.

Author's Note

The rights metadata proposal and examples provided here are not legal advice. To answer specific questions of law or address policy matters with legal implications, professional advice from an attorney is always recommended.

¹² Drafting the assumptions to be applied locally should not be used as an excuse to delay capturing rights metadata. If necessary, institutions can start with the rights information that is known and agree on the assumptions over time.

¹³ Institutions may have zero risk tolerance or may have collections consisting primarily of works by living artists. In either case, local policy may be to seek permission. Others may feel that the good faith judgment based on reasonable assumptions applied to the law and the facts is sufficient to allow use and defend in cases of infringement claims.