

# Pennsylvania Libraries: *Research & Practice*

Practice

## Providing Quality Rights Metadata for Digital Collections through RightsStatements.org

Linda Ballinger, Brandy Karl, and Anastasia Chiu

*Linda Ballinger is Metadata Strategist, Cataloging and Metadata Services, Penn State University Libraries, [lmb49@psu.edu](mailto:lmb49@psu.edu)*

*Brandy Karl is Copyright Officer, Penn State University Libraries, [bak25@psu.edu](mailto:bak25@psu.edu)*

*Anastasia Chiu is Cataloging & Metadata Librarian, Stony Brook University, [anastasia.chiu@stonybrook.edu](mailto:anastasia.chiu@stonybrook.edu)*

In April 2016, the Digital Public Library of America (DPLA) and Europeana officially introduced RightsStatements.org, a new controlled vocabulary for international, standardized, machine-operable statements about the copyright status of digital resources. This article describes the rights statements, provides guidance and examples on how to apply the most common ones, and addresses common pitfalls. It outlines how assigning rights statements fits into the digital resources workflow at Penn State, including contributing to the DPLA through PA Digital, Pennsylvania's Service Hub. Finally, it offers suggested resources for libraries and other cultural heritage institutions to help make their own rights determinations when they lack the expertise of a copyright officer.

### Introduction

In April 2016, the Digital Public Library of America (DPLA) and Europeana officially introduced RightsStatements.org, a new controlled vocabulary for international, standardized, machine-operable statements about the copyright status of digital resources. RightsStatements.org was the culmination of two years' work by the International Rights Statements Working Group, which consisted of members from the DPLA, Europeana, and Creative Commons communities. Their white paper called for the establishment of RightsStatements.org, highlighting the problem of uncontrolled, free-text metadata about copyright and citing that, by April 2016, the DPLA had over 87,000 different rights statements (International Rights Statements Working Group, 2016). In a presentation given later that year, Gore and Hansen estimated there were 100,000 different rights statements in DPLA and noted there were "more words in the rights statement field than any other field," which would have skewed the results of keyword searches (Gore & Hansen, 2016, slide 5).

The International Rights Statements Working Group outlined quality goals for rights statements in their white paper, stating they should be simple, flexible, descriptive, accurate, and transparent. Citing the [visualization work of Farrell](http://deanfarr.com/viz/rights.php) (deanfarr.com/viz/rights.php), the working group concluded there were "reasonable ways to synthesize this information about copyright and re-use into several meaningfully distinct categories" (International Rights Statements

Working Group, 2016, p. 12). By simplifying the categories of copyright statements, researchers would be able to find and compare resources based on copyright status without needing to interpret each free-text statement. Rights statements need to be flexible enough to accommodate multiple jurisdictions and institution types, and researchers should find them descriptive enough to tell whose rights were being addressed—the rights of the digitizing institution or researcher. Finally, rights statements should be accurate and transparent, clarifying whether there are restrictions on reuse based on contracts or other non-copyright reasons.

The need for clearer rights metadata was described well before the inception of RightsStatements.org. In 2005, Coyle pointed out: “When users want to make use of an unprotected digital resource, they have to know the copyright status of the work in order to make a reasonable judgment,” and “yet, information about the copyright status of the work and contact information for the copyright holder is not generally included in the metadata for the digital material. In some cases, especially with archival materials, this information is not readily available.” She added, since most institutions place the responsibility for determining copyright on the researcher, “it is therefore a matter of good user service for the library to provide all available information relating to the copyright status of the work so that the determination can be made” (para. 3-4).

Schlosser (2009) agreed, stating, aside from mitigating libraries’ liability in case of copyright infringement by end users, “educating users about copyright is often considered a responsibility of libraries” (p. 373). However, she noted findings from her survey of copyright statements in digital collections showed significant gaps in providing key information, such as items’ copyright status and copyright owners. Prevalent trends from the survey included “The Vague Ownership Statement” (which does not state who, if anyone, owns copyright in digital objects, or claims ownership in digital reproductions) and “Protecting Ourselves and You” (which refers to statements that address the legality of the library’s use of the objects without substantially addressing reusability). Schlosser also pointed out the lack of “standard terminology or placement” for copyright status and terms of use in metadata as a confusing factor for users in navigating digital works, noting that “it is unlikely that users can distinguish between them” (p. 382). As McCormick noted in a 2014 literature review, their confusion was corroborated by multiple studies on the copyright law knowledge of students, faculty members, and academic librarians.

Mazzone (2011) lamented the lack of any public domain registry and suggested a public domain mark similar to the copyright mark. He called on libraries to lead the effort to protect the public domain, suggesting many ways for libraries to help promote and protect the public domain. His most practical suggestion was for the library catalog to “include an annotation as to whether a work (or part of it) falls in the public domain” (pp. 213-216). In essence, he advocated for clear and easy-to-interpret item-level rights metadata.

Bryant et al. (2015) identified additional issues with the way many cultural heritage organizations handled public domain works, such as claiming the right to require permission to use public domain materials. They stated that: “While consonant with common practices of assisting researchers and other users in obtaining licenses or permissions, such statements (and any purported ‘permission’ to use public domain materials) serve to muddy the waters of the public domain” (p. 390). Mazzone (2011) went as far as to call such practices “copyfraud,” which will be discussed further below. Sims (2017) discussed problematic practices in applying open licenses to public domain works, observing “some libraries and cultural organizations have attempted to apply Creative Commons licenses to scanned copies of public domain materials with the good intent of stimulating reuse. However, public domain works, by definition, have no rights holders, so nobody can apply a Creative Commons license to a public domain work” (p. 80). Sims also noted “RightsStatements.org is well on its way to addressing thorny problems caused by proliferating idiosyncratic ways of describing rights statuses in online collections” (p. 79).

The disorganized state of rights metadata described in literature thus far was also true of the Penn State Libraries’ digital collections in early 2016. That spring, we prepared some of our digital collections for harvesting into DPLA via Pennsylvania’s Service Hub, PA Digital. Reviewing our existing digital collections, we found that, of the 65

collections in our CONTENTdm repository, 27 had no rights information in the item-level metadata. Some of these collections had a lengthy, boilerplate rights statement on the collection homepage instead:

Digital collections are available for research, teaching, and learning. These collections are not all part of the public domain even though they have been made publicly accessible. Moreover, even if a collection is in the public domain you should still cite where it came from and/or credit the author. For those collections under copyright protection, texts and images from the digital collections may not be used for any commercial purpose or beyond the normal "fair use" guidelines without prior permission from Penn State University and/or those parties who own the copyright. Where copyright is of concern in any material in question, that right is owned either by Penn State University, or by the original creators of the object, or their descendants. Therefore, when using any of this material, it is your responsibility, as the user, to secure permissions and adhere to the stated access policy, copyright laws, and educational fair-use guidelines. (Penn State University Libraries, n.d.)

While such boilerplate statements save time for the libraries and provides some consistency, it makes the researcher's work harder. General warnings about fair use and citing sources give the researcher no guidance specific to the collection, much less each item. Such general information belongs in institutional copyright policy documentation, such as the portal mentioned in the statement above. A few of our collections used this or similar information in an especially unwieldy and noncommittal way (i.e. by having it in their item-level metadata). Several other collections offered terse statements, such as "no restrictions on access" or "no copyright notice found" without any guidance on what these statements meant for the researcher using the collections. Only seven collections were explicitly identified as being in the public domain, though once assessed, several more collections were public domain in the United States.

## The Problem of Unintentional Copyfraud

Copyfraud, or sometimes copywrong,<sup>1</sup> is a general term that refers to false claims about works with copyrightable subject matter and public domain works. Copyfraud includes claims of ownership of works to which a person or entity has no rights. Libraries and other cultural heritage organizations are amongst the largest perpetrators of copyfraud. Copyfraud is not always committed intentionally; however, well-meaning but incorrect copyright claims over digitized works ultimately limit the use of digitized collections.

The recent "Happy Birthday Song" case clearly illustrates the effects of false copyright claims. Warner/Chappel Music publishing company falsely represented for decades that they owned the copyright to the song, which was composed by two sisters in 1893 and eventually fell into Warner's hands after a series of acquisitions (and failures to register). After years of expensive litigation spearheaded by a documentary filmmaker, the court declared the song to be in the public domain and ordered Warner to repay \$14M for falsely representing that they owned the copyright in the work. In their memorandum in support of the settlement, the plaintiffs stated:

Untold thousands of people chose not to use the Song in their own performances and artistic works or to perform the Song in public. This has limited the number of times the Song was performed and used. After the Settlement is approved . . . the Song will be performed and used far more often than it has been in the past. While there is no way to make a reliable estimate of the increase that will result, there can be no dispute that the increase will be substantial. (Good Morning to You Productions, 2016, p. 12)

Most people can recall hearing an "alternative" birthday song performed in public or in a movie, because Warner falsely claimed to own the copyright in the more famous song. How often have users shied away from using

a digitized work because its copyright status was incorrectly labeled? Metadata that doesn't accurately reflect a work's copyright status can lead to confusion and low use of the very resources libraries would like to see used. By restricting reuse of works with well-intentioned but false claims of copyright, libraries' attempts to control downstream uses runs counter to the goal of making their collections more useful to the public.

Until recently, most libraries have not provided sufficient copyright information for digitized collections. Libraries have often provided copyright information that incorrectly indicates that the author is the rights owner, that the holding institution is the rights owner, or that institutional permission is required for use. Other institutions have incorrectly attached open licenses to digitized works when they have no authority to sublicense rights to others or when works are in the public domain.

One source of unintentional copyfraud is confusion over which object is being described for digitized materials in the repository. For some time, best practice was to describe the digital reproduction, rather than the original item. Following this logic, institutions recorded themselves as publisher and recorded the date digitized as the date created. Likewise, many institutions thought they were the rights holder when it came to the digital reproduction and claimed copyright for it. However, the 1999 *Bridgeman* case determined digital reproductions of two-dimensional works are insufficiently creative to justify their own copyright. While *Bridgeman* is just one district court, no other U.S. court has determined otherwise in the intervening years. Other, non-U.S. jurisdictions have come to different conclusions, however, so some confusion remains. The International Rights Statements Working Group chose to address this confusion for the aggregation platforms they represented by "strongly encouraging those contributing organizations that create scans to adopt the position that the scan should not create a new layer of copyright as a matter of policy and choice by participating institutions, and to the extent that rights are created, those rights are waived" (International Rights Statements Working Group, 2016, p. 10).

## What are the Rights Statements?

The RightsStatements.org initiative "establishes the vocabulary that every organization can use to talk to their audiences about copyright and related rights in a meaningful way" (International Rights Statements Working Group, 2016, p. 4). It does this by standardizing rights statements, making them applicable internationally, and making them machine-operable. This last characteristic is made possible by using Uniform Resource Identifiers (URIs) for the statements. URIs are defined by the World Wide Web Consortium (W3C) as "short strings that identify resources in the web: documents, images, downloadable files, services, electronic mailboxes, and other resources" (2006, para. 3). Uniform Resource Locators (URLs, or web addresses) are a type of URI. URIs can be used to uniquely identify and express relationships among resources on the web, in ways that search engines can understand. Free-text rights statements such as the example quoted above would be difficult, if not impossible, for computers to process in useful ways. By using URIs and a standardized set of statements, RightsStatements.org makes it easier to discover resources that meet certain reuse criteria. A digital resources repository could offer the rights statements as facets to narrow search results, for example. In aggregations of collections that use standardized rights statements, researchers could even discover the desired resources across institutions.

It should be noted that the RightsStatements.org website states:

These rights statements are high level summaries of the underlying rights status of the digital objects that they apply to. By providing a limited number of standardized statements we aim to facilitate the online discovery of cultural heritage objects based on their rights status. Our rights statements are intended to be used in addition to (more detailed) rights information that institutions already have and not to replace existing information. (n.d., para. 2)

Individual institutions can choose how they convey information beyond the rights statements, as we will discuss below.

The standardized rights statements fall into three categories: In Copyright, No Copyright, and Other. Because the statements are international in scope, a few apply either exclusively or primarily to European institutions. All 12 statements are outlined in Table 1. The unshaded cells indicate the statements that will likely be used most frequently in the United States. The next section provides some practical guidance on using these rights statements and gives examples from the Penn State collections while highlighting potential pitfalls.

**Table 1**  
*Outline of Rights Statements*

<b>In Copyright</b>	<b>No Copyright</b>	<b>Other</b>
<p>Use <b>In Copyright (InC)</b> for items in copyright, implying the item is in copyright, and the provider:</p> <ul style="list-style-type: none"> <li>• is the rights-holder,</li> <li>• received permission from the rights-holder(s) to provide the item,</li> <li>• or provides the work under an exception or limitation to copyright (including Fair Use).</li> </ul>	<p>If an item is clearly in the public domain worldwide, and has no other legal restrictions on use, the <b>Creative Commons Public Domain Mark</b> should be used instead of a rights statement.</p>	<p>Use <b>No Known Copyright (NKC)</b> for items whose copyright status has not been determined conclusively but when there is reasonable cause to believe the item is not under copyright or related rights. Do not use NKC for orphan works, which are assumed to be in copyright, or for items whose copyright status has not been investigated.</p>
<p>Use <b>In Copyright – Rights-holder(s) Unlocatable or Unidentifiable (InC-RUU)</b> for items in copyright but whose rights-holder(s) have not been identified or located after reasonable investigation. Only use this rights statement when reasonably sure the item is in copyright.</p>	<p>Use <b>No Copyright – United States (NoC-US)</b> for items copyright-free under U.S. law. Do not use NoC-US for orphan works, which are assumed to be in copyright, or for items whose copyright status has not been investigated.</p>	<p>Use <b>Copyright Undetermined (UND)</b> for items whose copyright status is unknown after investigation. Typically, use UND when missing essential facts to determining an accurate copyright status.</p>
<p>Use <b>In Copyright – Educational Use Permitted (InC-EDU)</b> for copyrighted items whose rights-holder(s) explicitly authorizes third parties to use the item for educational purposes without obtaining permission.</p>	<p>Use <b>No Copyright – Contractual Restrictions (NoC-CR)</b> for public domain items under contractual agreement restricting third-party use. To make NoC-CR conclusive, a page detailing the applicable contractual restrictions must be linked.</p>	<p>Use <b>Copyright Not Evaluated (CNE)</b> for items whose copyright status is unknown and whose copyright status has not been investigated.</p>
<p>Use <b>In Copyright – Non-Commercial Use Permitted (InC-NC)</b> for copyrighted items whose rights-holder(s) explicitly authorizes third parties to use the item for non-</p>	<p>Use <b>No Copyright – Non-Commercial Use Only (NoC-NC)</b> for public domain items that were digitized by public and private partners who have agreed to limit</p>	



commercial purposes without obtaining permission.	commercial use of the item’s digital representation by third parties.	
Use <b>In Copyright – EU Orphan Work (InC-OW-EU)</b> with orphan work items as defined in the terms of the EU Orphan Works Directive.	Use <b>No Copyright – Other Known Legal Restrictions (NoC-OKLR)</b> for public domain items that cannot be freely re-used due to known legal restrictions, such as cultural heritage or traditional cultural expression protections. For NoC-OKLR to be conclusive, a page detailing the legal restrictions that limit re-use of the item must be linked.	

Adapted from the [RightsStatements.org website](http://RightsStatements.org) (rightsstatements.org/en)

It is important to remember that rights statements from RightsStatements.org are not licenses, such as Creative Commons licenses. Unlike the Creative Commons Public Domain Mark, which RightsStatements.org does borrow, Creative Commons licenses can only be granted by the copyright holder. As explained in examples below, libraries and other institutions digitizing their collections are not necessarily the rights holder for their collections. In cases where the institution is the rights holder or hosts digital collections for a rights holder who wants to grant a Creative Commons license, the institution should provide Creative Commons license information instead of a RightsStatements.org rights statement. RightsStatements.org and Creative Commons licenses are not meant to be used together. Most importantly, an institution must not assign a Creative Commons Public Domain Mark while also assigning a Creative Commons license that limits use of the public domain item.

## Rights Statements: Recommendations and Examples

### In Copyright (InC)

**In Copyright (InC)** typically applies when the institution making the digital reproduction available is “[making] the Work available under an exception or limitation to copyright (including Fair Use) that entitles it to make the Work available” (Digital Public Library of America, 2016, p. 10).

- *Brent Wilson Papers*: These works were created between 1950 and 2013. While there are other papers in the collection, the digitized items consist only of Wilson’s unpublished art journals, mostly from 1990 to 2010. They are still in copyright due to the creation date.
- *Mira Dock Lantern Slides*: Although the dates of these teaching photographs range from 1897 to 1902, they were unpublished, and the creator’s death date at the time the rights statements were first assigned to this collection meant the slides were a few months short of being public domain. A year later, these statements could be updated to No Copyright – United States. This highlights a significant and growing problem with rights statements assignments: They are static and do not incorporate legal changes and copyright term expirations.

## In Copyright – Rights-holder(s) Unlocatable or Unidentifiable (InC-RUU)

**In Copyright – Rights-holder(s) Unlocatable or Unidentifiable (InC-RUU)** can be used when the institution making the digital reproduction available has made a reasonable effort to identify the rights holder but has not been successful. That essentially means an orphan work, but since the United States has no legislation covering orphan works, this rights statement is used instead.

- *Thomas W. Benson Political Protest Collection*: This collection consists of Vietnam War protest posters collected by Benson from the streets of San Francisco and New York. There are no author marks on most of the posters, and it would be impossible to identify an author for the few works that do bear a signature.

## In Copyright – Educational Use Permitted (InC-EDU) and In Copyright – Non-Commercial Use Permitted (InC-NC)

**In Copyright – Educational Use Permitted (InC-EDU)** and **In Copyright – Non-Commercial Use Permitted (InC-NC)** are only used when the digitizing institution either is the rights holder or when the deed of gift from the rights holder explicitly states that the work can be used by the public only under these circumstances. These two statements are directed at the end users' permissions, not the permissions for the digitizing institution. A common pitfall is that some institutions have interpreted these statements to mean a collection has been digitized and made available online based on a fair use determination.

## In Copyright – EU Orphan Work (InC-OW-EU)

In Copyright – EU Orphan Work (InC-OW-EU) is not used in the United States.

## Creative Commons Public Domain Mark

The **Creative Commons Public Domain Mark** should only be used when the work being reproduced is clearly in the public domain internationally. Since the longest copyright term worldwide is the life of the author plus 100 years, 1835 is the latest date that should be assumed in the public domain internationally, based on date alone.

- *Digital Map Drawer*: Several sub-collections in the Digital Map Drawer were assigned the Public Domain Mark based on their age. These include the Rare Maps collection with dates ranging from the 15th to 19th centuries.

## No Copyright – United States (NoC-US)

**No Copyright – United States (NoC-US)** indicates the work is in the public domain in the United States but makes no determination about its copyright status elsewhere in the world.

- *Digital Map Drawer*: One way materials could enter the public domain in the United States, but not necessarily elsewhere in the world, was failing to renew copyrights during a specific time period. Several of the Sanborn maps dated after 1923 were assigned the No Copyright - United States rights statement, because research indicated that their copyrights were not renewed. These Sanborn records also have a local field that briefly indicates the apparent failure to renew the copyright. U.S. government-authored maps are also in the public domain in the United States but not necessarily in the rest of the world, therefore, they are more accurately labeled No Copyright - United States.

- *Elisha J. Bracken Civil War Diary*: We have decided that No Copyright – United States more precisely conveys our copyright determinations for works not quite old enough for us to say with certainty that it is free of copyright restrictions worldwide. This United States Civil War diary is an example of such a work. Using the Creative Commons Public Domain Mark for such works is not incorrect, it is just less precise than No Copyright – United States. As one recent presenter pointed out, “It can be difficult to find and interpret the relevant law. But the statements serve an informational purpose, and it’s possible to draw lines that shape your basic decision-making process” (Zeller, 2017). Even for works published before 1835 and, thus, very likely to be public domain worldwide, there is some debate over whether they should be assigned a Creative Commons Public Domain Mark, which does not guarantee that the work is free of copyright around the world, but that it is “free of known copyright restrictions worldwide” (“What is the Public Domain Mark?,” 2012, para. 6).

## No Copyright – Contractual Restrictions (NoC-CR) and No Copyright – Non-Commercial Use Only (NoC-NC)

**No Copyright – Contractual Restrictions (NoC-CR)** and **No Copyright – Non-Commercial Use Only (NoC-NC)** both cover situations in which the work is in the public domain, but the digitizing institution has agreed to contracts with third parties that limit the type of uses they can allow. This situation arises when digitization is contracted out to a for-profit organization. Libraries may consider whether attempting to place restrictions on public domain works limits legal downstream uses, as the “Happy Birthday Song” case so aptly demonstrates.

## No Copyright – Other Known Legal Restrictions (NoC-OKLR)

**No Copyright – Other Known Legal Restrictions (NoC-OKLR)** is mostly intended for use in Europe. It covers works that are in the public domain but have legal restrictions beyond copyrights, such as those protecting cultural heritage. The Society of American Archivists Intellectual Property Working Group’s recommendation states “the U.S. does not have laws that could limit re-use of public domain material in this way, but a repository may wish to think about foreign laws if it is making a culturally sensitive collection available overseas” (Society of American Archivists Intellectual Property Working Group, 2016, p. 7).

## No Known Copyright (NKC)

**No Known Copyright (NKC)** indicates that there is substantial, but not complete, certainty that the work is out of copyright. NKC indicates the digitizing institution made an effort to identify the work’s copyright status.

- *World War I Glass Plate Stereographs*: These works were mostly made in Europe and then published in the United States with a bizarre chain of title-passing from one publisher to the next. There was an apparent failure to renew existing registrations or to register at all. So, after exhaustive research tracing the copyrights for these works, we concluded they are likely in the public domain in the United States, but we cannot provide further information of any certainty regarding the collection.

## Copyright Undetermined (UND)

**Copyright Undetermined (UND)** is used when key information for making the determination is missing, such as the date of publication or author’s death, missing pages or portions that may have held a copyright notice, or the inability to determine whether a work is published. UND indicates the digitizing institution



made an effort to identify the work's copyright status. If, when relying on the [Copyright Term and the Public Domain in the United States chart](#) ([copyright.cornell.edu/publicdomain](http://copyright.cornell.edu/publicdomain)) by Peter Hirtle to help your determination, you cannot narrow down the object to a single line, then you cannot determine the status unless the remaining available lines both indicate the same status.

## Copyright Not Evaluated (CNE)

**Copyright Not Evaluated (CNE)** should be avoided if possible. It indicates the institution making the digital reproduction available has done no research into the work's copyright status. However, institutions may choose to use it when remediating already-digitized collections, when it is not possible to retrospectively make an evaluation of the copyright status of the works. It should be the rights statement of last resort.

## General Workflow for Implementing Rights Statements

Penn State's workflow for digital collections has undergone multiple changes in the past year, and our implementation of RightsStatements.org has greatly influenced the processes we eventually developed for digitization and for updating existing digital collections. Not only did our widely-distributed digital collections team need a way to determine and record rights-related metadata for digital collections, we also needed a way to communicate that information and coordinate all production-related tasks involved in getting a digital collection online. At the same time, we needed a process for retrospectively assigning RightsStatements.org rights statements to existing collections.

In the spring of 2016, we held a digital collections retreat in the libraries, at which all the stakeholders gathered to determine our workflow, resulting in a high-level outline of every moving piece necessary for digitizing a collection. We made preliminary decisions on the order of some processes and which information people needed to communicate with each other. One person was designated as the digital collections coordinator, responsible for guiding collections throughout the evaluation and digitization process. This position is in the Preservation, Conservation, and Digitization department and is close to many of the activities that take place in the workflow. The digital collections coordinator organizes the master workflow in Trello, an online shared project management system, which provides far-flung digital collections team members with a common platform for sharing their contributions. It assures all the information is collated in one place.

Collections are selected for digitization in a variety of ways. The person most responsible for the collection (the collection "owner") may make the case for digitization, or the collection may be identified as a priority for digitization by the Special Collections Library or the Preservation, Conservation, and Digitization department. The collection owner works with the Digital Collections Coordinator to create a cover sheet for the collection, including as much information as possible about the collection to allow for the initial rights evaluation as well as to provide needed data for digitization.

Most existing digital collections are identified for rights and other metadata remediation, because they have been selected for contribution to DPLA and need their metadata updated to be compliant with current best practices. In other cases, the copyright officer and the Special Collections Library might identify an existing collection in need of remediation because of high patron usage and poor rights metadata, which typically leads to multiple copyright inquiries for the same few items. Whether driven by inclusion in DPLA, local demand, or collections strategy, when a collection is selected for digitization or remediation, it proceeds through the rest of the workflow essentially the same way. Keeping the processes uniform assists in speeding the workflow and reducing confusion.

## Copyright and Metadata Assessment

Once a collection has been identified for digitization or metadata remediation, it is passed to the copyright officer for a rights evaluation. We do not assign the RightsStatements.org rights statement at this step, but we consider basic questions about the collection's provenance and rights. We also assess the work involved in creating metadata for the collection and which remaining questions need to be researched. While basic initial research is performed in this step, the focus is on assessing the collection and providing a copyright-focused analysis of the collection.

Fair use determinations are considered in this step. If the collection is not public domain and fair use cannot support either all or part of its digitization for public access purposes (e.g. a collection may be still commercially valuable and providing access to it would not be fair use), permissions are sought, and the project may be put on hold until they can be obtained. The items may still be digitized for preservation and private access.

Frequently, we begin the copyright assessment phase assuming that a collection consists primarily of public domain works or that the libraries own the rights. However, an assessment may reveal this is not the case. Most often, a donor will transfer the rights in a collection via a Deed of Gift. The deed sometimes transfers the copyrights in the underlying works, though more often now, it transfers all the copyright *that the donor owns* in the works to the libraries to relieve the donor of making any determinations. The assessment phase has revealed a general misunderstanding, not specific to our own institution, of which rights the libraries gain when a donor transfers "all the rights that I have." Too often, a curator assumes that the libraries then own all the rights in the collection. A prime example of this misunderstanding is the Thomas W. Benson Political Protest Collection discussed above. A Deed of Gift covered the transfer of this collection as a subset of the War Posters Collection. Initial processing of this sub-collection indicated that Penn State Libraries owned all the rights in it, when in fact, the donor never owned any rights in these works. He merely preserved war protest posters he had collected off the streets. Situations like this have revealed the need to improve the preservation of permissions information with digitizing permissions-related documents and centralized storage.

For new digital collections, a format assessment is performed after the copyright assessment, which includes an evaluation of the collection's readiness for conversion, storage requirements, metadata needs, and other preservation and digitization considerations. Once the collection meets the minimum requirements of the major copyright, format, and metadata phases of the digitization process, the Preservation, Conservation, and Digitization department begins digitizing the collection. Metadata is provided by the Cataloging & Metadata Services department or by the Special Collections Library. This step happens either after digitization is complete or simultaneously, depending on the needs of the collection. For existing digital collections, Cataloging & Metadata Services assesses and remediates the descriptive metadata at this stage.

## Rights Determinations

After descriptive metadata is provided, the standardized rights statements are assigned—though it may also happen simultaneous to, before, or after digitization. The process of determining rights statements can begin as soon as the relevant metadata has been assigned, especially the author, date, publisher, description, and title fields. The digital collections coordinator or metadata strategist provides the copyright officer with a spreadsheet of all the items digitized (or to be digitized) and their metadata. The amount of necessary research that the copyright officer performs varies by collection. Some may require extensive research into its copyright owner, publication data, and copyright notice. A collection's process may also be very simple, such as a living donor or author signing over rights to the libraries for a collection of unpublished works for which they own all the rights.

Once the item-level rights statements have been determined, the copyright officer inserts the appropriate standardized rights statement URIs into the proper column in the metadata spreadsheet and returns it to the digital collections coordinator for workflow coordination and to the metadata strategist to load into CONTENTdm.

## Rights Determinations without a Copyright Officer

Not all institutions have the resources for a dedicated copyright librarian. While our institution assigns responsibility to the copyright officer to make the item-level standardized rights statement determinations, a specialized copyright librarian or attorney is not required to make determinations. Even at institutions with a copyright librarian or manager, that person is not necessarily responsible for performing the research for the rights statements assignments nor the one doing initial determinations.

Several resources are available to help determine rights. All copyright librarians and probably all copyright attorneys use [Copyright Term and the Public Domain in the United States chart](http://copyright.cornell.edu/publicdomain) (copyright.cornell.edu/publicdomain) by Peter Hirtle to make copyright duration determinations. The chart needs no copyright attorney to interpret: it is designed to be useful to anyone making a United States copyright determination. By locating the creation and/or publication date and the applicable conditions, the chart will indicate the copyright status of a United States work.

The Hirtle chart can be adapted for local needs as well, and the idea expanded to repositories that collect a significant number of non-U.S. works. Librarians at the University of Miami created a matrix “to standardize and simplify the decision making process to determine the rights status for various types of items. Crucial factors include: date of creation, creators’ date of death, publication status, and country of creation” (Capell & Williams, 2016). Their poster offers a copyright decision matrix that covers both the United States and Cuba. Similarly, the manual by Hirtle, Hudson, and Kenyon (2009) is a useful guide to non-specialists by “provid[ing] some basic information on copyright law and offer[ing] a structure for considering copyright issues in digitization projects” (p. x).

The Society of American Archivists Intellectual Property Working Group offers the “[Guide to Implementing Rights Statements from Rightsstatements.org](http://www2.archivists.org/standards/guide-to-implementing-rights-statements-from-rightsstatementsorg)” (www2.archivists.org/standards/guide-to-implementing-rights-statements-from-rightsstatementsorg). It is especially helpful for determining rights for unpublished materials, since a common pitfall is failing to recognize that a work has not been published, which can affect its copyright status. Although copyright law is rife with exceptions, the general rule is that publication has occurred when copies of a work have been distributed to the general public via sale, rental, lease, or lending. See Nimmer (1956) for a discussion of investive versus divestive publication and other general rules of publication under the 1909 Copyright Act.

Even though rights statements are assigned at the item-level, many items in a digitized collection have the same publication status, author, and general date range. Even larger collections typically contain distinct homogeneous sub-collections that make assigning rights statements faster, as less research and fewer determinations are required.

## Recording Rights Statements


Once the rights determination is completed, the metadata strategist coordinates recording that determination in the item-level metadata for the digital collection, based on the spreadsheet completed by the copyright officer. Currently, Penn State provides just the URI from RightsStatements.org in the rights field. This aligns with the technical needs of PA Digital, Pennsylvania’s DPLA Service Hub’s metadata guidelines, which require the URI be the only content in the metadata field that contains it (PA Digital Metadata/Standards Group, 2017). As Capell and Williams (2016) demonstrate, other institutions have chosen to supplement the URI by providing a separate local rights field containing human-readable text. For now, Penn State follows a stricter interpretation of the DPLA guidelines for using rights statements, which states that local rights fields can add additional information to the rights statement, but they


“must not contradict the standardized statement and . . . should not be merely a repetition of the information in the standardized statements” (Digital Public Library of America, 2016, p. 1). While we have not ruled out using local rights fields, we have not yet processed a digital collection that requires a local field that needs to be exposed for harvesting. We have used a rights field labeled “Copyright Status Notes” to record the research conducted to determine the copyright status of the post-1923 Sanborn maps that were not renewed. This field is primarily intended for internal use, so it is not mapped to a Dublin Core field within CONTENTdm, to ensure it is not harvested. Penn State uses Qualified Dublin Core, so we map the standardized rights statement to dc:rights, and if a future collection needs additional rights information, we would map it to dc:accessRights for the time when PA Digital becomes technically able to support outputting local rights information in addition to the RightsStatements.org URI.

When it comes to providing the URI for the standardized rights statement, we discovered through trial and error that, unlike URLs, URIs are case-sensitive. The PA Digital harvester was set up to normalize capitalization, so the harvester’s normalization process needed exceptions for the rights field to make the URI links work. Another mistake that later became obvious is that the web address appearing in the browser address bar of a RightsStatement.org page is a URL, not the URI. The URI must be copied from the text of the page. For example, the correct URI for the In Copyright statement is <http://rightsstatements.org/vocab/InC/1.0/>, not the URL address that appears in the browser, <http://rightsstatements.org/page/InC/1.0/?language=en>.

In CONTENTdm, only the URI displays, though it links to the appropriate page in RightsStatements.org. But in DPLA, the URI is resolved to display the official icon near the item’s thumbnail image, linking to the appropriate page in RightsStatements.org. It displays a brief, human-readable description of the rights statement in the body of the metadata record (Figure 1). DPLA also has plans to create a search facet in its interface, based on RightsStatements.org rights statements, so researchers can limit their search to items that match their plan for use.

Let there be peace and let it begin with me





Get full image from Pennsylvania State University

<b>Created Date</b>	Circa 1970
<b>Partner</b>	PA Digital
<b>Contributing Institution</b>	Pennsylvania State University
<b>Type</b>	image
<b>Subject</b>	Political posters--American Peace movements
<b>Language</b>	English
<b>Standardized Rights Statement</b>	<p>This Item is protected by copyright and/or related rights. However, for this Item, either (a) no rights-holder(s) have been identified or (b) one or more rights-holder(s) have been identified but none have been located. You are free to use this Item in any way that is permitted by the copyright and related rights legislation that applies to your use.</p> <p style="color: #C00000;"><a href="http://rightsstatements.org/vocab/InC-RUU/1.0/">http://rightsstatements.org/vocab/InC-RUU/1.0/</a></p>
<b>URL</b>	<a href="http://collection1.libraries.psu.edu/cdm/ref/collection/benson/id/218">http://collection1.libraries.psu.edu/cdm/ref/collection/benson/id/218</a>

Figure 1  
Record display in DPLA, showing the RightsStatements.org icon, a human-readable description of the rights statement, and the URI

## Providing Additional Information about the Works

In many instances, the libraries need or want to provide additional information about digitized works that is not rights-related. For example, the libraries may be contractually required to provide a suggested attribution or to credit the loaning institution. Some digital repositories generate recommended citations from a resource's metadata fields instead of storing a suggested citation in its own field. Credits may be generated the same way, and suggested attribution information could be mapped to `dcterms:bibliographicCitation`. If using Simple Dublin Core, this field is probably better left unmapped, as aggregators often need the broader field, `dc:Identifier`, for other purposes. Regardless of where the citation information is stored or how it is generated, it should not be stored in `dc:rights`.

Moreover, an institution-wide copyright policy statement can be used to provide fuller information on the rights of users and place them in context with the rights of the institution. A well-drafted copyright policy statement will put the library's digitization work into context, provide the institution's fair use justification for the digitization of its collections (e.g. for research, scholarship, and education), and point out that end users may need to seek permissions for reuse of the digitized materials if their uses are not fair use. This institution-wide copyright policy statement essentially represents what libraries were attempting to do with lengthy, free-text rights statements at the item or collection level: communicate to patrons that users' rights and responsibilities may be more complex than the standardized rights statement and to further develop the institution's relationship with the user. Ideally, RightsStatements.org would supplement its current right statement pages with practical examples and more specific guidance on determining which reuses are allowed for each rights status, insofar as its international scope allows. Meanwhile, cultural heritage organizations could provide such information in their copyright policy statement and share them with other organizations to use or adapt.

## Going Forward with Rights Statements

RightsStatements.org rights statements are not yet perfect. As the Mira Dock Lantern Slides example demonstrates, the statements are static. The assigning institution is responsible for updating the standardized rights statement when there is a change in an item's copyright status, which could be achieved by a reminder system. Potentially, by mapping worldwide copyright laws to computer-readable decisions trees,<sup>2</sup> it is possible to have at least partially automated rights statements determinations or updates in the future. While some organizations collect additional, hidden fields about works to help future rights determinations, similar to our Copyright Status Notes field, it is possible that advanced machine learning like Wolfram Alpha and other AI systems can help provide missing information for copyright determinations.

## Conclusion

A year after RightsStatements.org was introduced, DPLA reported that it "has been well received by cultural heritage professionals within the DPLA network and around the world. Digital libraries in Brazil, Australia, New Zealand, and India will be joining the project, with interest from additional libraries on every continent" (DPLAfest 2017: Post-Fest Wrap Up, 2017, para. 5). This interest in and rapid adoption of RightsStatements.org demonstrates cultural heritage institutions' recognition of the growing importance of presenting complete and accurate rights information about the resources in their collections, and that doing so promotes interest in and use of those resources. Implementing a rights statement workflow to achieve better rights metadata will be different at every institution and will depend on existing workflows and institutional dynamics. Moreover, every institution will make local choices when assigning rights statements to the items in their collections, which may vary depending on local preferences and

norms. Shared training and documentation will assist institutions in participating in this new movement to bring a broader set of services to researchers.

## Notes

<sup>1</sup> Copywrong has multiple meanings, including bad incentives present in the actual copyright law.

<sup>2</sup> See the [Durationator Project](http://www.limitedtimes.com) (www.limitedtimes.com).

## References

- Bridgeman Art Library v. Corel Corp., 36 F. Supp. 2d 191 (S.D.N.Y. 1999).
- Bryant, M. P., Chen, C. C., French, J., Levine, M., Liu, S. G., Roberts, M. G., Sims, N., Wythe, D., & Young, A. M. (2015). External uses. In Young, A. M. (Ed.), *Rights & reproductions: The handbook for cultural institutions* (Chapter 6). Retrieved from aam-us.org/ProductCatalog/Product?ID=5186
- Capell, L., & Williams, E. (2016, June). [Assigning rights statements to legacy digital collections](#). *Library Papers and Presentations*, Paper 2. Poster session presented at the Rare Books & Manuscript Section (RBMS) conference in Coral Gables, FL. Retrieved from scholarlyrepository.miami.edu/librarypapers/2
- Coyle, K. (2005). [Descriptive metadata for copyright status](#). *First Monday*, 10(10). Retrieved from journals.uic.edu/ojs/index.php/fm/article/view/1282
- Digital Public Library of America. (2016). [DPLA Standardized Rights Statements Implementation Guidelines](#). Retrieved from bit.ly/dpla-rights-guidelines
- [DPLAfest 2017: Post-fest wrap up](#). (2017, April 28). Retrieved from dp.la/info/2017/04/28/dplafest-2017-post-fest-wrap-up
- [Good Morning to You Productions Corp., et al. v. Warner/Chappell Music, Inc., et al.](#), No. CV-13-04460-GHK C.D. Cal. (Mem at 12, ECF at 301) (2016). Retrieved from docs.justia.com/cases/federal/district-courts/california/cacdce/2:2013cv04460/564772/301
- Gore, E. & Hansen, D. (2016). [RightsStatements.org \[PowerPoint slides\]](#). Retrieved from www.ala.org/advocacy/sites/ala.org.advocacy/files/content/copyright/RS%20Org%20ALA%20Copytalk\_final.pptx
- Hirtle, P., Hudson, E., & Kenyon, A.T. (2009). [Copyright and cultural institutions: Guidelines for digitization for U.S. libraries, archives and museums](#). Ithaca: Cornell University Libraries. Retrieved from ecommons.cornell.edu/bitstream/handle/1813/14142/Hirtle-Copyright\_final\_RGB\_lowres-cover1.pdf
- International Rights Statements Working Group. (2015, updated January 2016). [Rightsstatements.org white paper: Recommendations for standardized international rights statements](#). Retrieved from rightsstatements.org/files/160208recommendations\_for\_standardized\_international\_rights\_statements\_v1.1.pdf
- Mazzone, J. (2011). *Copyfraud and other abuses of intellectual property law*. Stanford, California: Stanford University Press.
- McCormick, A. (2014). [Copyright, fair use and the digital age in academic libraries: A review of the literature](#). *SLIS Student Research Journal*, 4(2). Retrieved from scholarworks.sjsu.edu/slissrj/vol4/iss2/5
- Nimmer, M. (1956). [Copyright Publication](#). *Columbia Law Review*, 56(2), 185-202. doi:10.2307/1119728
- Penn State University Libraries. (n.d.). [A. E. Bye drawings](#). Retrieved from libraries.psu.edu/about/collections/e-bye-drawings
- PA Digital Metadata/Standards Group. (2017). [PA Digital PA-DPLA metadata guidelines: requirements, recommendations, and best practices for preparing metadata for PA Digital's DPLA aggregator](#). Retrieved from bit.ly/padigitalmetadata2017



RightsStatements.org. (n.d.). [Use by cultural heritage institutions](#). Retrieved from [rightsstatements.org/en/documentation/#use-by-cultural-heritage-institutions](https://rightsstatements.org/en/documentation/#use-by-cultural-heritage-institutions)

Schlosser, M. (2009). [Unless otherwise indicated: A survey of copyright statements on digital library collections](#). *College & Research Libraries*, 70(4), 371-385. doi:10.5860/0700371

Sims, N. (2017). [Rights, ethics, accuracy, and open licenses in online collections: What's "ours" isn't really ours](#). *College & Research Libraries News*, 78(2), 79-82. Retrieved from [crln.acrl.org/index.php/crlnews/article/view/9620/11028](http://crln.acrl.org/index.php/crlnews/article/view/9620/11028)

Society of American Archivists Intellectual Property Working Group (2016). [Society of American Archivists guide to implementing rights statements from Rightsstatements.org](#). Retrieved from [www2.archivists.org/standards/guide-to-implementing-rights-statements-from-rightsstatementsorg](http://www2.archivists.org/standards/guide-to-implementing-rights-statements-from-rightsstatementsorg)

[What is the Public Domain Mark?](#) (2012). *PDM FAQ*. Retrieved from [wiki.creativecommons.org/wiki/PDM\\_FAQ#What\\_is\\_the\\_Public\\_Domain\\_Mark.3F](http://wiki.creativecommons.org/wiki/PDM_FAQ#What_is_the_Public_Domain_Mark.3F)

World Wide Web Consortium (W3C) (2006). *Learning About URIs*. Retrieved from <https://www.w3.org/Addressing/#background>

Zeller, M., (2017, June). [Turning Wrongs into Rights: Implementation of RightsStatements.org at Washington University](#). *University Libraries Presentations*, 29. Poster session presented at the Kraemer Copyright Conference in Colorado Springs, CO. Retrieved from [openscholarship.wustl.edu/lib\\_present/29](http://openscholarship.wustl.edu/lib_present/29)