Unraveling Challenges: Rights Statements in Digital Cultural Heritage Collections

Hannah Stitzlein\textsuperscript{a}, Myung-Ja K. Han\textsuperscript{b}, and Sara R. Benson\textsuperscript{c}

\textsuperscript{a}Metadata Librarian, University of Cincinnati, Cincinnati, Ohio, USA; \textsuperscript{b}Metadata Librarian, University of Illinois at Urbana-Champaign, Champaign, Illinois, USA; \textsuperscript{c}Copyright Librarian, University of Illinois at Urbana-Champaign, Champaign, Illinois, USA

ABSTRACT
Describing copyright for cultural heritage objects in a rights statement can be a complicated undertaking when there is not enough information about the object. Similarly, the complexity and misunderstanding of copyright law can lead to erroneous claims of copyright ownership when objects are digitized. These challenges are further compounded by confusingly written and poorly implemented rights information, leaving the user in doubt as to how they may interact with, or reuse a digital object. An audit of rights statements in the Illinois Digital Heritage Hub (IDHH) was performed to determine how standardized rights statements from RightsStatements.org can mitigate some of the issues surrounding information used in the rights field of metadata for digital objects. The audit highlighted common issues found throughout the digital object metadata harvested by the IDHH, which were then analyzed to understand how contributors use and apply the rights field. Possible solutions to these issues are included, such as copyright education and outreach programs that help contributing institutions understand copyright laws and the use of proper rights statements.

KEYWORDS
Copyright for digitized objects; metadata; digital cultural heritage materials; standardized rights statements; digital libraries

Introduction
Copyright law is complex, and identifying the copyright owner can often be an equally complicated undertaking. Because of such challenges, creating or selecting the appropriate rights statement for digitized and/or born-digital objects can be a daunting task. As many aggregated digital service providers, including the Digital Public Library of America (DPLA), require that the metadata for digital objects should include a rights statement, most digital objects have some form of a rights statement that informs users how they may engage with the object or who owns the copyrights. Until
recently, however, there has been little conversation around how copyright information should be described for digital objects, and whether the description is helpful and understandable to a user that desires to interact with those objects in a digital environment.

To promote the use of standardized rights statements, the DPLA, Europeana, Creative Commons, and other community leaders came together to create the Recommended for Standardized International Rights Statements,¹ to address how standardized rights statements can be used in digital collections, and why they are needed. The result of that work is a set of 12 standardized rights statements that offer clear and concise rights language for use by cultural heritage institutions so that they may convey the copyright status and the permissions of use of the objects in their digital collections.² However, the actual implementation of standardized rights statements is a challenge for many cultural heritage institutions because of the resources needed to assess collections and the current rights information attached to digital objects within those collections. To use the standardized rights statements for already existing collections, collection owners must first understand the copyright statuses in their digital collections, and determine if their current rights statement accurately reflects whether the individual items therein are or are no longer in copyright. Secondly, collection owners must identify a corresponding standardized rights statement if there is one, and replace the current rights statement with the new standardized rights statement. All of this takes time and staff resources, which can be a challenge to institutions that may be lacking in one or both.

To understand the current status of rights information described in digital collections, this paper shares the key findings from the data analysis of metadata values in the Rights fields contributed to the Illinois Digital Heritage Hub (IDHH)–the Illinois State service hub for the DPLA. This analysis included how digital collection owners describe copyright statuses for their digital items and how that information is presented in the metadata in the local system and made available to service providers in the <dc:rights> field. The findings of the data analysis were then used to determine if and how the current rights metadata from Illinois contributors could be mapped to the standardized rights statements from RightsStatements.org.

The data analysis revealed several common issues such as incomplete values and inaccurate use of the Rights field. Recommendations are given for how issues like those can be avoided and remediated. The article concludes with forward steps that the IDHH is undertaking to provide contributing institutions with copyright information and training, and how standardized rights statements can be utilized effectively in digital collections and in the digital environment.
Literature review

For digital collection owners, the first challenge to describing rights of the digital collection and items associated with the collection is understanding what the rights statement should describe, the digitized or physical resource. The Bridgeman decision is regarded as the case that defined the relationship of copyright to digitized objects (Bridgeman Art Library v. Corel Corp., 1999). The court case determined that slavish copying (such as photographing or scanning) of an object was not a sufficiently creative endeavor and as such a new copyright could not be applied to a slavishly-copied and digitized version of an object. The Bridgeman decision, therefore, informs how copyright is discussed throughout this article—that copyright is for the original object and not the digitized surrogate—and copyright descriptions will be addressed from this perspective.

One of the greatest challenges of describing copyright is identifying the copyright date of an object. Copyright does not accrue based on the date of digitization of the underlying work, but rather based on the date that the object was originally created (for unpublished works) and the date of publication (for published materials). For many archival materials, which remain unpublished works, the copyright would expire 70 years after the death of the author. So, a more appropriate information to include, rather than the date of digitization, would be the author’s name and (if deceased) the date of the death of the author, if known. And, in many instances, it is helpful to include both the date of publication and the date of the author’s death (Hirtle, Hudson, & Kenyon, 2009). Or, for unpublished works, the appropriate date to include would be the date that the work was created and not the date of publication (Hirtle et al., 2009).

As evidenced above, copyright date is hard to determine accurately, and since death dates of authors are rarely easy to determine, it may sometimes be impossible (Dickson, 2010). As an alternative to including complete and accurate rights metadata, such as including the date of publication of the work (if any), and the death date of the author (if known), many institutions might choose to have general copyright or fair use information, or use a boilerplate copyright claim, such as “© 2018 institution name” which may be borne out of a desire to convey the year of digitization. However, according to Mazzone (2006) and Sims (2017), a note like this often amounts to copyfraud—“claiming falsely a copyright in a public domain work” (p. 1028)—as the library does not own the copyright on the materials included in the online collection since making a “slavish copy” of a work in digital form does not create a separate copyright in the digital work, as per the Bridgeman decision (Sims, 2017). And, omitting copyright data altogether is also misleading to users who
may think that indicates that the items are (by default) owned by the library or in the public domain. What is the best practice in these instances? To use a standardized rights statement that indicates either that: (1) copyright status is unknown or (2) no copyright determination has been made by the institution. In any event, this will alert the user that the library is uncertain about the copyright status of the work and will give the user an opportunity to do some further rights determinations on their own.

One of the objectives laid out in the *Recommendations for Standardized International Rights Statements* was for the rights statements to be both human and machine readable, which is a significant improvement from traditional text based rights statements, that are only human readable (International Rights Statements Working Group, 2015). Another objective addressed was for the standardized statements to effectively communicate copyright and use restrictions to users that also improve the reuse of the digital item—therefore the language needed to be clear so that users can interact with digital items appropriately. The International Rights Statements Working Group (2015) acknowledges that determining copyright statuses can be complex, especially when compounded with legal uncertainty. Such legal uncertainty and complex copyright statements can lead to unclear language, or as Ballinger, Karl, & Chiu et al. (2017) found in an assessment of the Penn State Library CONTENTdm collections, boilerplate rights statements are applied when the library wants to make an effort to provide consistent legal language around the rights in their collections. However, these boilerplate rights statements are full of copyright information (e.g., general fair use information or source citation guidance) that is not specific to the collection to which the rights statement is affixed, offering little useful information to researchers regarding specific rights as they relate to individual items.

Another issue that can potentially be addressed by using standardized rights statements, is the aforementioned copyfraud. If, in the boilerplate example above, all items in a collection are given an in-copyright status but not all items are actually in copyright, those items with false assertions of copyright are in copyfraud. Standardized rights statements can certainly be used inaccurately by metadata practitioners—the statements are, after all, only a controlled vocabulary, and can be used however the practitioner chooses—but in order to implement those statements correctly into a collection, an assessment of the rights in that collection, and the items therein, should be performed to determine accurate copyright statuses to avoid falling into the trap of copyfraud.

While using standardized rights statements may not necessarily solve inaccurate copyright or outright copyfraud statements problems, they do
build a framework to work within as well as provide legally accurate language to describe rights held in and over a digital object. Schlosser (2009) pointed out that there were widely-accepted guidelines for the responsibility of libraries to provide copyright information for photocopying, but that there was no such consensus for the necessity or content of copyright statements in digital collections. While there continues to be no explicit guidelines for the responsibility of libraries to provide copyright information for digital collections, the availability of standardized rights statements has begun a national conversation among cultural heritage institutions regarding the necessity of accurately describing copyright in their digital collections.

### Analysis of Values In `<dc:rights>` Contributed To The Illinois Digital Heritage Hub

#### Data collection and methods

Data for this study was collected from the metadata contributed to IDHH. As the Illinois DPLA service hub, the IDHH harvests metadata from contributing institutions via the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH)\(^3\) in the Dublin Core Terms\(^4\) format mapped from partner institutions’ local digital assets management systems. At the time of the data collection (January 2017), a total of 203,895 metadata records from 338 collections contributed from 120 institutions were harvested when the IDHH prepared its first ingestion to DPLA. As a state-wide service hub, the IDHH has metadata from many different types of cultural heritage institutions from across the state of Illinois. Table 1 shows the types of contributing institutions in IDHH at the time of the data collection.

| Institution type             | No. of institutions |
|------------------------------|---------------------|
| Public libraries             | 63                  |
| Academic & research libraries| 43                  |
| Museums                      | 5                   |
| Historical societies         | 4                   |
| Special libraries            | 3                   |
| Archives                     | 1                   |
| State library                | 1                   |

At the time of the analysis, no contributors to the IDHH were utilizing standardized rights statements or licenses, and all `<dc:rights>` field values were free-text. However, since the analysis, over 10% of all records harvested by the IDHH include Standardized Rights Statements.

After the initial harvest, IDHH minimally enhances the contributing institutions’ metadata with information required by the DPLA, including
institution and collection names if the information is not already present. It should be noted that Rights information is not altered in any way during this process. Once the metadata enhancement process is completed, DPLA harvests the metadata from the IDHH’s OAI server.

The DPLA Metadata Application Profile (DPLA MAP) defines the \textless dc:rights \textgreater field as a required field, therefore all metadata contributed to the IDHH includes at least one \textless dc:rights \textgreater field. The \textless dc:rights \textgreater field is one of the elements whose values are harvested from contributing institutions without any alteration. For this study, values used in \textless dc:rights \textgreater were collected from the DPLA OAI Aggregation Tool developed by the North Carolina Digital Heritage Center. The tool showed all values contained in the \textless dc:rights \textgreater field for each collection with the number of items associated with each value. All values (1220 in total) were compiled in a spreadsheet with the collection and contributing institution’s names. Using Excel spreadsheets, all duplicate values used as rights statements were removed; this process did not remove or correct any spelling, grammatical, or spacing differences, resulting in 938 unique values.

The data analysis was conducted with the 938 unique values collected from the harvested \textless dc:rights \textgreater field in two ways: first, an analysis was performed on the values in the \textless dc:rights \textgreater field provided to IDHH; second, local metadata fields were studied to understand how they were mapped to \textless dc:rights \textgreater. This two-way data analysis informed researchers of how collection owners perceived and described rights, and how that information is added in metadata fields in both their native environments and made available to service providers. Analysis of values used for the \textless dc:rights \textgreater field was also used to determine whether there was a corresponding standardized rights statement from RightsStatements.org for each value found in the \textless dc:rights \textgreater field. For the data analysis, spreadsheets were used to organize findings, such as whether or not copyright was described, what kinds of non-rights information was described, and a possible equivalent standardized rights statement when current rights information described copyright.

While a full qualitative analysis would determine accuracy, and whether or not the copyright was accurate to the item being described (e.g., whether an item was in copyfraud, whether the copyright date was based on the original or digitization date, etc.), the researchers determined that that was out of scope for this study since there were clear limitations of understanding and identifying rights information for each item contributed to IDHH. It would be impossible to perform a qualitative analysis of the rights statements because the variables needed to assess quality (e.g., whether a risk assessment was performed, whether there was access to a copyright professional, what the rights metadata workflow was, etc.) are nearly infinite.
Analysis of `<dc:rights>` values

Values used in the `<dc:rights>` field

Collection owners used the `<dc:rights>` field not only for a rights statement or copyright information, but for many different types of information about the digital or physical object (see Figure 1). When analyzing the 938 unique values in the `<dc:rights>` field, many included more than one information type value (e.g., copyright and contact type values) in a single `<dc:rights>` field, resulting in a total of 1172 type values amongst the 938 unique values. According to the analysis results of values used for the `<dc:rights>` field, only 289 (25%) values described the copyright status of the item, including who owns the copyright(s), and/or whether the item is in the public domain or in copyright. The other 651 values did not include any copyright information: Among them, 312 values (27%) had personal names, and only 34 of those included a statement that declared that the name was of the copyright holder (e.g., “Copyright held by Margaret Schotz.”). The remaining 278 values that included personal names did not have any associated contextual information (e.g., “Michael O’Daugherty”) that described whether the name belonged to the copyright holder, owner, or a donor of the item (or any combination of those).

The `<dc:rights>` field was also used for contact information for the holding institution (275 occurrences, 23%) usually including email address,
phone number and/or the street address of the institution or department that owns the physical collection or a digitization unit. Funding information for the digitization or project of which the digital object is a part (62 occurrences, 5%), condition of use and permissions (76 occurrences, 6%), and reproduction information (48 occurrences, 4%) also were frequently included in the `<dc:rights>` field. 110 values (9%) described other information, including general copyright and fair use, related materials, location of the physical item, and the date digitization permissions were granted.

Multiple values in `<dc:rights>`

In some cases, the `<dc:rights>` field contained more than one type of information. Of the 938 values used for the `<dc:rights>` field, 224 (24%) values had more than one information type in a single field, which might confuse users as to how they may use the digital item. The most commonly paired values were copyright and contact information, such as:

“All rights held by 0000. Please contact the 0000 Library for permission to reproduce, distribute, or use this image. 555-123-4567, or librarian@0000lib.org.”

As well as condition of use and contact information, such as:

“This image may be used freely, with attribution, for research, study and educational purposes. For permission to publish, distribute, or use this image for any other purpose, please contact 0000 at 555-123-4567.”

Although the majority of collections (298 collections, 88.2%) had one `<dc:rights>` field in their metadata, the data analysis also showed that 40 of the 338 collections (11.2%) had more than one `<dc:rights>` field (see Table 2a,b). There are 23 collections with two `<dc:rights>` fields, 15 collections with three, and two collections have four `<dc:rights>` fields. In such

| Local label | Dublin core mapping | OAI metadata |
|-------------|---------------------|--------------|
| (a) Two instances of the `<dc:rights>` field in a single metadata that contain two different values | Rights dc:rights | This material may be protected by U.S. Copyright Law (Title 17 U.S. Code). This resource is provided for educational purposes only. Please contact: library@0000.org for permission to use this resource. |
| GiftBy | dc:rights | This project has been funded in whole or in part with oooo, under Contract No. xxx-xx-1-1111 and xxx-xx-2-2222. Also, funding for this grant was awarded by the 000. |
| (b) Three instances of the `<dc:rights>` field in another metadata, two fields include near duplicate information | Rights dc:rights | This material may be protected by U.S. Copyright Law (Title 17 U.S. Code). May not be reproduced without permission from the 0000 History Center. |
| Copyright dc:rights | Notice: This material may be protected by U.S. Copyright Law (Title 17 U.S. Code). May not be reproduced without permission from the 0000 History Center. |
| GiftBy | dc:rights | Jacqueline Smith. |
a case, usually one \(<\text{dc:rights}>\) field was used for copyright information and the other for additional information such as funding or donor information. Since all IDHH contributing institutions (at the time of analysis) used CONTENTdm as their digital assets management system, it is assumed that collection owners mapped multiple local fields to the \(<\text{dc:rights}>\) field. This can create difficulties for service providers trying to identify the correct \(<\text{dc:rights}>\) to use for displaying accurate copyright information.

**Human error**

The analysis also revealed that a simple human error can cause incomplete or possibly incorrect rights information in the metadata. For example, a typo, incomplete sentence, wording variation, and/or punctuation can result in inconsistencies in the rights statement.

\[
\text{Notice: This material may be protected by U.S. Copyright Law (Title 17 U.S. Code). May not be reproduced without permission from the OOO Historical Society.}
\]

The two rights values above are from the same collection and are identical in the first half, but due to a human input error, users may not fully understand the permissions on the item if they were to come across the incomplete second rights value. Errors like these can be alleviated by using a standardized rights statement in URI format, which can auto-populate a rights field with the full rights statement.

**Analysis of rights values in local environments**

The necessary next step of the data analysis was a closer examination of the difference between local metadata and harvested metadata, e.g., how those values in the \(<\text{dc:rights}>\) field were represented (with different labels) in the local environment to understand the original context. As discussed above, more than 69% of values used for the \(<\text{dc:rights}>\) field did not describe actual rights information. The field is used for other information, including personal names and dates without any contextual information, contact, or funding information. Additionally, there were many items that included more than one \(<\text{dc:rights}>\) field, as shown in Table 3, or with more than one value in the single \(<\text{dc:rights}>\) field. Since
the data used for the analysis were mapped from local system metadata to the Dublin Core Terms format for service providers, it is critical to understand where those values came from in the local environment.

Local metadata and mapping analysis showed that collection owners use various labels for the rights information, and mapped assorted local fields to the `<dc:rights>` field. The most commonly used local field labels for rights information were `<Rights>`, `<Rights Statement>`, and/or `<Copyright>`. In addition, collection owners mapped other local fields to `<dc:rights>`, such as `<GiftBy>` or `<Other>`. Table 2a,b show two examples in which multiple `<dc:rights>` fields were used in two collections. Table 2a shows two values mapped to `<dc:rights>` from two local fields in one record; one value contains copyright information, while the second contains funding information and describes no rights. Table 2b shows three values that were mapped to `<dc:rights>` from three different local fields in one record. The first field does address copyright to an extent (albeit, not to a particularly helpful degree), however, the second repeats this information but adds “Notice”: at the beginning, but this does not change the meaning of the rights statement. It is unclear, from a service provider’s point of view, which is the preferred rights information to use, or the benefit to using one over the other because the values are nearly identical. The third field does not describe copyright in any way, and because of this, that field should be mapped to a field other than `<dc:rights>`.

Table 3 shows the local labels for each value collected for the analysis (1220) regardless of whether the value was duplicated, and how those values were mapped to `<dc:rights>`. Values were sometimes exactly duplicated in a record, but mapped to `<dc:rights>` from two separate local fields (such as Rights and Copyright) or nearly exact duplicates, as in Table 2b.

| Local field label       | No. of `<dc:rights>` fields |
|-------------------------|-----------------------------|
| Copyright               | 121                         |
| GiftBy                  | 427                         |
| None                    | 5                           |
| Other                   | 4                           |
| Rights                  | 604                         |
| Rights on item          | 16                          |
| Rights statement        | 31                          |
| Rights; copyright       | 10                          |
| Rights-access rights    | 1                           |
| Rights statement; rights| 1                           |
Possible applications of standardized rights statements

In order to support using standardized rights statements, values used in `<dc:rights>` fields were compared with rights statements available in RightsStatement.org. Based solely on the information that is currently available from contributors, regardless of accuracy of the copyright statement (as in the case of copyfraud statements), seven of the twelve standardized rights statements were identified to be equivalent to the free-text rights values used in the IDHH metadata as shown in Table 4.

Among the 265 rights values that addressed copyright (whether the item was or was not in copyright), 219 values (66%) claimed that the items described were in copyright and described no educational or noncommercial use exceptions. The following is an example of common terminology that was used to describe this:

Copyright University Library. To request reproductions or inquire about permissions, contact University Library, 1200 University Ave., City, IL 60000, (123) 456-7890

For a statement like this, where there was no use-exceptions, an IN COPYRIGHT⁸ rights statement would be the most appropriate.

A total of 46 values (15.2%) described that the copyright was undetermined. Therefore the closest equivalent statement was considered COPYRIGHT UNDETERMINED⁹. These values did not state whether the object was in copyright or not, but did state that the object may be in copyright, which implies that the copyright has not been determined. However, it is possible that the intent of the original statement may be more inline with COPYRIGHT NOT EVALUATED, though without contacting each institution using this type of statement it is not entirely certain how these declarations are intended to be read. The following is an example of how this was often described:

Notice: This material may be protected by U. S. Copyright Law (Title 17 U. S. Code). May not be reproduced without permission from City Public Library.

The two following examples show how use exceptions (educational and noncommercial) were described in rights values. These state that the object is in copyright, but that the digital object can be used for educational and

| Equivalent standardized rights statements | No. of IDHH rights statement |
|-------------------------------------------|-----------------------------|
| In Copyright                              | 219                         |
| In Copyright – Educational Use Permitted  | 18                          |
| In Copyright – Non-Commercial Use Permitted| 5                           |
| Copyright Undetermined                    | 46                          |
| No Known Copyright                        | 5                           |
| No Copyright – United States              | 18                          |
| No Copyright – Contractual Restrictions   | 2                           |

Table 4. Equivalent standardized rights statements from the RightsStatements.org for 313 IDHH rights statements.
noncommercial endeavors. The first example is the equivalent of an IN COPYRIGHT - EDUCATIONAL USE PERMITTED\(^\text{10}\) while the second example is the equivalent of IN COPYRIGHT - NON-COMMERCIAL USE PERMITTED\(^\text{11}\).

Copyright 1985, the Village Museum. The images in the Village Museum collections may be viewed, downloaded, and printed for educational use, but commercial use is prohibited without permission. For more information, contact the Village Museum at (123) 456-7890, or info@villagemuseum.org.

The Town Public Library holds the copyright for this image. It may be downloaded for personal use, but may not be published or redistributed, in whole or in part, without written permission from Town Public Library. You may contact Town Public Library at 1600 Main St., Town, IL or at (456) 789-0123.

There were three instances where an item was described as having no known copyright restrictions. These values were somewhat unclear as they described an object as having no known copyright, but that there were possible other restrictions, though those restrictions (rights, reproductions, permissions) were not identified. Items out of copyright but with use restrictions could most closely equate to NO COPYRIGHT - OTHER KNOWN LEGAL RESTRICTIONS\(^\text{12}\) or NO COPYRIGHT - CONTRACTUAL RESTRICTIONS\(^\text{13}\).

No known U. S. Copyright restrictions; other restrictions may apply. Please credit the City Historical Society when citing or reproducing this item.

The three cases in which an item was no longer in copyright, but still had noncommercial restrictions were items that were in the public domain, and had been digitized by Google for HathiTrust. The rights metadata for these objects included a link to the Access and Use Policies page on the HathiTrust website. The equivalent standardized statement is NO COPYRIGHT - NON-COMMERCIAL USE ONLY\(^\text{14}\).

Rights fields that described the item as public domain (or not in copyright) often had the shortest description, because there were no restrictions on use, there was little to be described. These values occurred 20 times, and had an equivalent standardized statement of NO COPYRIGHT - UNITED STATES\(^\text{15}\). Because the Standardized Rights Statements are intended for international use, a work that is no longer in copyright is designated as NO COPYRIGHT - UNITED STATES for items that function under the jurisdiction of the United States, which is different from a public domain demarcation because the laws of public domain vary by jurisdiction. The two following examples were the most commonly found:

Public domain.
Not in copyright. Contact archives@university.edu for more information
Discussions

Analysis of the rights values used in the digital collections contributed to the IDHH showed three notable characteristics:

1. Harvested `<dc:rights>` fields include more than rights information. As found in the data analysis, 69% of the rights values contain information about the donor of the item, funding sources, location of the institutions, or contact information of the department or institution that created and published the collection. In many cases this extraneous information was included while the actual rights statement was not in the `<dc:rights>` field. Collection owners and metadata creators often want to add this information in the metadata, and add it into a customized local field label, such as `<Contact Information>` or `<GiftBy>` for their respective users, however, since there is no proper element for these types of information in Dublin Core (or Terms), they simply map the information to the `<dc:rights>` field due to the limitations of the Dublin Core schema.

2. Mappings from local metadata to OAI metadata requires careful consideration. Collection owners and metadata creators want to make their local metadata available for aggregators as much as possible. However, not all local metadata elements need to be mapped or made available in the OAI metadata, especially when the local labels imply that those fields are not important to users, such as `<Other>` or `<None>`. In addition, the mapping only transfers values used in the field, not the local field labels that contain contextual information. If the donor has the copyright for the item, then that information should be included in the proper local metadata field with an appropriate label, such as “*John Doe (donor) holds the copyright for the item.*” and mapped to the `<dcterms:rightsHolder>` field, instead of using a specific label but mapping to `<dc:rights>` where the meaning will be lost when that metadata is harvested by an aggregator (Han, Cho, Cole, & Jackson, 2009). If the donor does not own the copyright and the information is not useful for discovery, then collection owners and metadata creators should consider keeping the information in other local fields with proper explanation, such as `<dc:description>` *John Doe donated this item to the institution.* `<dc:description>`, or moving the metadata into a hidden field not visible to users that is also not harvestable by aggregators.

3. Understanding what kind of information should be used for `<dc:rights>` can be improved significantly. The data analysis results revealed that values used in the `<dc:rights>` field varied from having all-inclusive descriptions of the rights held in and over a digital object, to contact information or a single date without contextual information,
such as whether the date referred to copyright. In some cases, while there was a rights statement, it was unclear whether an item was in copyright or not, due to the vagueness of the language used to describe rights information. Currently, the most common values (33%) in the \texttt{<dc:rights>} field describe who owns the physical resources or the collection as a whole, and 29% of values describe contact information for the holding institution. This information does not inform users of how they can utilize the digital items. While the specific ownership information is important in the local environment, it is generally more useful for users in aggregators’ environments to know how they can use and engage with the digital resources in a broader sense. As more and more users of local digital collections come from the web or aggregators’ sites, collection owners and metadata creators should identify and provide an appropriate standardized rights statement for users on the web, in addition to information about the ownership of the physical item that is important for users in their local environment.

**Moving towards standardized rights statements**

The IDHH works in three different ways to promote accurate use of the \texttt{<dc:rights>} field and standardized rights statement. The IDHH developed and deployed the Illinois Metadata Best Practices in January 2017 as a first step, which included recommendations for how standardized rights statements could be utilized, as well as what information may be useful to users who want to engage with the digital objects.\textsuperscript{16} The Best Practices defined what the Rights field should be used for, what established standardized statements can be used in the field, and ways to avoid common issues such as input errors or mappings, through thoughtful application of rights statements.

As a supplement to the Metadata Best Practices, a series of ongoing metadata training workshops are being taught. The workshops review the best practices laid out by the IDHH, and include how to best use the Rights field, as well as introduce the concept of utilizing standardized rights statements. The information gathered in the analysis on how contributing institutions were currently utilizing the Rights field helped the authors develop training that includes what information can and should be included, as well as what can be avoided. Additionally, the trainings offered a way to provide an introduction to the Standardized Rights Statements and their use in digital collections.

The workshops open up conversations with partners about the need to evaluate current rights information for accuracy and clarity, and the benefits of implementing standardized rights statements. All workshops include
time for questions and answers so that participants can engage directly and discuss the importance of accurately describing the rights of the digital item. Further webinars on copyright and utilizing standardized rights statements were also developed with partners throughout the state of Illinois. These workshops are invaluable because the IDHH is not able to alter rights metadata, therefore building educational tools to equip contributors with accurate copyright knowledge is key.

In conjunction with metadata workshops, IDHH provides metadata assessments for all collections. The assessment evaluates the metadata fields used in a collection for accuracy and consistency. This includes an evaluation of the rights field to determine if the field is being used correctly, whether that value includes accurate information (e.g., whether the information describes copyright or not), and whether there is a possible RightsStatements.org match to the current statement. The assessments offer a starting point for collection owners that may be uncertain of where to begin with remediating rights information currently in the metadata. Useful resources are included with the assessments, such as the Copyright Term and the Public Domain chart\(^1\)7, with the hope that the collection owners are able to make appropriate copyright determinations and changes as necessary. With the metadata best practices document, webinars, and assessments, four institutions within the IDHH have already begun implementation of statements from RightsStatements.org for 38 digital collections, and have replaced the vague and/or inaccurate language in the `<dc:rights>` field with clear and concise language as of January 2018.

A rights statement for a digital object is a tool that connects users to objects. If the rights statement informs users with clear and correct information regarding whether there are or aren’t any restrictions for use, users can then utilize the object accordingly. As standardized rights statements—such as those available through RightsStatements.org—can improve communications between cultural heritage institutions and users, collection owners and metadata creators should work together with copyright specialists to identify the most appropriate rights statements for their digital collections in order to better connect their digital collections to users.

Notes

1. http://rightsstatements.org/files/151002recommendations_for_standardized_international_rights_statements.pdf
2. http://rightsstatements.org/en/
3. https://www.openarchives.org/pmh/
4. http://dublincore.org/documents/dcmi-terms/
5. https://ildplmetadatawkgrp.wordpress.com/2016/04/29/metadata-for-digital-collections-contributed-to-the-illinois-digital-heritage-hub/
References

Bridgeman Art Library v. Corel Corp. (1999). 36 F. Supp. 2d 191 (S.D.N.Y. 1999).
Ballinger, L., Karl, B., & Chiu, A. (2017). Providing quality rights metadata for digital collections through RightsStatements.Org. Pennsylvania Libraries: Research & Practice, 5(2), 144–158.
Dickson, M. (2010). Due diligence, futile effort: Copyright and the digitization of the Thomas E. Watson Papers. The American Archivist, 73(2), 626–636.
Han, M.-J., Cho, C., Cole, T. W., & Jackson, A. S. (2009). Metadata for special collections in CONTENTdm: How to improve interoperability of unique fields through OAI-PMH. Journal of Library Metadata, 9(3-4), 213–238.
Hirtle, P., Hudson, E., & Kenyon, A. (2009). Copyright and cultural institutions: Guidelines for digitization for U.S. libraries, archives, and museums. Retrieved from http://hdl.handle.net/1813/14142
International Rights Statements Working Group (2015). Rightsstatements.org White Paper: Recommendations for standardized international rights statements. Retrieved from http://rightsstatements.org/files/160208recommendations_for_standardized_international_rights_statements_v1.1.pdf
Mazzone, J. (2006). Copyfraud. New York University Law Review, 81, 1026–1100.
Schlosser, M. (2009). Unless otherwise Indicated. College & Research Libraries, 70(4), 371–385.
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.