

University of California Libraries Metadata Sharing Policy

November 25, 2014

Revised April 23, 2015

Revised May 8, 2015

Introduction

In FY 2013-2014, Strategic Action Group 2 (SAG2): Access, Discovery & Infrastructure was charged by the Council of University Librarians to develop a policy regarding the sharing of metadata managed by all University of California (UC) Libraries, including the California Digital Library (CDL). In early 2014, SAG2 appointed the Metadata Policy Project Team to draft this policy [Appendix A].

The guiding principle of this policy is that metadata managed by an individual campus library or collectively managed by CDL should be shared as widely as possible in support of Goal 3 of the University of California Libraries Systemwide Plan and Priorities for FY 2014-2018 (<http://libraries.universityofcalifornia.edu/about/vision-and-priorities>), which is:

Maximize discovery of and access to information resources: Faculty and staff are adopting new techniques to identify, find, and use information resources. As user behaviors evolve, library services must adapt to provide better, faster, and broader access and support.

This policy applies to metadata that the UC Libraries intentionally share, including but not limited to:

- Bibliographic and authority data in local library systems
- Descriptive, administrative, and technical metadata in local digital asset management systems
- Archival finding aids
- User-contributed metadata
- Metadata in institutional repositories

Policy

All University of California (UC) Libraries, including the California Digital Library (CDL), are empowered to share metadata within constraints imposed by contractual agreements or by law. This includes metadata that is managed by an individual campus library or collectively managed by CDL.

For the purposes of this policy, metadata is defined as structured information that describes resources, allowing users to locate, retrieve, interpret, or manage those resources. Although the UC Libraries are not obligated to share their metadata, sharing is encouraged because the use and re-use of open metadata facilitates discovery, furthers knowledge and research, and provides public access to rich scholarly and cultural heritage content.

Guidelines for sharing metadata managed by the UC Libraries:

1. Libraries will place the fewest possible restrictions on the reuse of metadata they share. The preferred standard under which to share UC Libraries metadata is CC0, the Creative Commons Zero Public Domain Dedication.¹ UC Libraries may request appropriate attribution as the source of their shared metadata, to the extent such attribution is technologically feasible, in accordance with community norms.² The provision of a CC0 standard has the additional benefit of removing any uncertainty about reuse.
2. Metadata will be shared in accordance with the law:
 - a. UC Libraries will share metadata in ways consistent with contractual terms to which they have agreed³ and in a manner consistent with copyright law⁴ and privacy rights.⁵
 - b. UC Libraries will make reasonable efforts to retract shared metadata should such sharing be found in violation of contractual terms, copyright law, or privacy rights.⁶
3. If any campus wishes to opt out, that action would require approval by the Council of University Librarians, or their designated agent, who would make an evaluation based on the effect of such an action on the collective purpose of the UC Libraries and its relevant projects.
4. Metadata should be shared in formats that facilitate its reuse, such formats to be determined by the sharing library.
5. Those who contribute metadata to UC Libraries that is not covered by a purchase and/or contractual agreement must be made aware of the terms under which their metadata may be shared. An example of this would be user-supplied metadata, such as tags contributed by individuals for images.⁷ Libraries should provide a clear statement about how the metadata will be shared.
6. UC Libraries will take reasonable measures to keep shared metadata accurate and correct, updating metadata in a timely fashion when errors are observed or reported, and whenever required by legal obligations.⁸
7. UC Libraries should include their preferred URL in the metadata they share to direct users back to the resource being described. For example, they may choose to include a link directing users back to the point of aggregation or to the original contributor's site.

¹ A description of CC0 ("No rights reserved") can be found at <https://creativecommons.org/about/cc0>.

² In some circumstances campuses may wish to use a license such as CC BY (Creative Commons Attribution license, <http://creativecommons.org/licenses/by/4.0/>). Creative Commons, Open Data Commons, and other licenses generally require attribution and may impose other restrictions. Since these restrictions may make metadata reuse difficult or impracticable to implement, UC Libraries should think carefully about sharing metadata under terms more restrictive than CC0.

³ The campuses enter into a variety of contractual agreements with vendors regarding the use of the metadata provided by vendors. When sharing metadata, campuses should be conscious of these agreements and limit the sharing of metadata per any contractual obligations they have with vendors. Some examples are:

- A library purchases content notes under contractual obligations not to share those content notes with any other institution. If the library chooses to share its bibliographic records, they must remove these content notes from any copy of any record they share.
- A library acquires vendor records for systemwide use under contractual obligations not to share those records with any other institution. That library is therefore prohibited from sharing any of these records beyond UC.

⁴ Facts are not protected by copyright. This is generally true worldwide.

Most metadata (especially bibliographic metadata) will be facts, and therefore not eligible for copyright protection.

- Metadata that exhibits sufficient "original authorship" (i.e. creativity), will be eligible for copyright protection. Potential examples include abstracts and summaries.
- Expression of factual information, like a biographical essay, may be protected by copyright, but the underlying facts are not. Someone could write their own essay using all the same facts and it would not infringe the copyright in the original essay. However, someone copying the whole essay - the way it expresses those facts - likely would be. Arrangement that is standardized, such as in a form with common fields like birth date, full name, degrees obtained, etc., would usually not be protected by copyright.
- Short phrases are generally not copyrightable in the United States because they don't exhibit enough "authorship" (creativity), because of the potential effect on free speech, and because ideas and facts can't be copyrighted.

Databases and other collections of metadata, as compilations of facts, are different.

- In the United States databases of facts are copyrightable as compilations, but only to the extent that the selection and arrangement of facts exhibits sufficient creativity. Even then, the underlying facts in the database are not protected, and it generally would not be infringement to pull them out of the database and reuse them for something else without permission.
- Other countries offer different legal protection for databases.

Between the complication of the fact that the rules vary depending on the country a user is in and the potential for confusion about the what metadata is eligible for copyright protection, it is better to have a clear statement about reuse (e.g. CC0) than to assume people will know when reuse is allowed.

⁵ Privacy rights, which vary from state to state, are unlikely to be relevant to most metadata managed by UC Libraries. Unusual exceptions might include research or archival data with personally identifying information like social security numbers, health information, or student records.

⁶ Such efforts might include no longer making the metadata available and posting a notice about the problem. It generally would not include attempting to locate users who might have downloaded the metadata and contacting them.

⁷ In some cases UC Libraries may wish to share metadata for reuse that potentially contains copyrightable expression that is not owned by the University of California. Notifying the authors of such contributions about how the metadata will be shared is advised in order to prevent infringing any copyrights in such contributions. Such notifications may be posted in a location where contributors are most likely to see it, such as on web pages, submission forms, etc.

Sample statement:

Note: By contributing metadata (e.g. tags, descriptions, transcription) for this resource, you hereby declare the contributed metadata to be in the public domain and that it can be freely distributed, repurposed, and reused in any manner, with or without attribution, as deemed appropriate by the UC Libraries.

⁸ This guideline does not obligate the campuses to undertake any special efforts to update and replace metadata they may have shared unless the metadata is so flawed that it undermines its reuse by the receiving party. In which case, at the minimum, the campus is only obligated to notify the recipient of the metadata of the flaw. This guideline is meant only to encourage campuses to update and replace metadata, usually as an outgrowth of normal operations. For example, catalogers regularly perform bibliographic maintenance in OCLC; these updated OCLC records benefit the wider community with which they are shared.

Respectfully submitted,
Metadata Policy Project Team

Katie Fortney
Claudia Horning
Xiaoli Li
Louise Ratliff
Kathryn Stine
Adolfo Tarango

Appendix A

University of California Libraries

Metadata Policy Project Team

Summary and Background

The Metadata Policy Project Team (MPPT) is responsible for examining the primary issues surrounding the topic of shared metadata and developing a UC wide policy for shared metadata. The team is charged and managed by SAG 2. This is in response to the Council of University Librarians' (CoUL) priorities for 2014-2017 which states "Develop a policy regarding sharing of UC metadata."

Responsibilities

- Formulate a definition of metadata; considering whether it extends beyond bibliographic metadata to include additional content and/or format types
- Articulate the distinction between UC and contributor generated metadata
- Determine how widely/freely shared metadata should be distributed, especially with regards to commercial providers and other educational institutions
- Investigate the copyright/rights management issues surrounding metadata, especially that which exhibits authorship
- Identify potential types of sensitive metadata and make recommendations regarding access restrictions for this content
- Investigate the campus perspective on how shared metadata should direct users back to the original resource (e.g., is the preference to direct users back to the point of aggregation or to the original contributor site)
- Provide recommendations pertaining to the shared accessibility of contributor-generated metadata
- Determine the level of exposure for metadata as it relates to human and machine accessibility
- Address the issues relating to credit, compensation, and control of shared metadata
- Address remedies for metadata that is found to infringe copyright or other legal requirements
- Formulate UCs responsibilities as they pertain to metadata accuracy and the potential need for revision and/or correction of metadata
- Ensure the resulting shared metadata policy is aligned with other UC system and campus wide initiatives
- Articulate what organizations are empowered to share metadata and under what circumstances or conditions (i.e., CDL's role and responsibility in sharing metadata, individual campuses ability to share another campuses metadata, etc.)
- Recommend what group is responsible for monitoring the field and tracking issues that may impact the established policy
- Recommend what group is responsible for maintaining and updating the policy once established
- Recommend what licenses or agreements are required to implement the sharing of metadata

Expected Deliverables

The Metadata Policy Project Team will be responsible for drafting a UC wide policy for shared metadata that will be delivered to SAG 2, but will include consultation with other SAGs and interested stakeholders. The team will also be responsible for documenting its activities, including posting to a publicly accessible environment (e.g., Web, wiki) meeting agendas/minutes, charges, and recommendation/decision outcomes.

Communications

The Metadata Policy Project Team will communicate primarily through regular conference calls, e-mail correspondence, and collaborative Web tools. Decisions pertaining to the frequency of regular meetings will be left to the team, though with the aggressive timeline that has been assigned SAG 2 recommends that the team meet at least bi-monthly. The team's SAG 2 liaison will manage the communication to and from SAG 2.

Decision-making

The Metadata Policy Project Team is empowered to make recommendations relative to the issue of shared metadata to achieve the overall UC vision for this practice. The team will consult with SAG 2 if any financial or staffing resources are required to develop this policy. Final decision-making will come from the Coordinating Committee and CoUL.

Membership and Terms of Appointment

The Metadata Policy Project Team will be formed in consultation with the other SAGs and will consist of no more than six staff members throughout the UC system that have expertise in the following areas:

- OCLC metadata policy
- Purchased/licensed data (e.g., SCP program, vendor supplied records)
- Copyright/Rights management
- Bibliographic and non-bibliographic metadata creation
- User/contributor generated metadata

Other identified experts throughout the UC system, including IT staff, legal counsel, and faculty members, should be contacted and brought into discussions as needed.

A liaison from SAG 2 will also be assigned to this team.

Timeline

- Submit a draft proposal to SAG 2 for review and comment by May 5, 2014
 - Very soon after the scope of the metadata is defined, a quick check-in with CoUL should be scheduled to make sure we are meeting their goal

- This process should include a review cycle that includes input and vetting by the other SAGs
- Submit final proposal to SAG 2 by June 30, 2014

References and Suggested Resources

- HathiTrust Bibliographic Metadata Use Policy Recommendations prepared by Melissa Lohrey CDL Discover and Delivery Intern, May 12, 2013, with contributions from Kathryn Stine, HathiTrust Metadata Management System Project Manager and Metadata Analyst
- Principles on Open Bibliographic Data <http://openbiblio.net/principles/>
- OCLC Control Numbers, Lots of them; all public domain <http://hangingtogether.org/?p=3318>
- Metadata and Copyright | Peer to Peer Review <http://lj.libraryjournal.com/2013/02/opinion/peer-to-peer-review/metadata-and-copyright-peer-to-peer-review/>
- Copyright, Metadata, and Attribution <http://kcoyle.blogspot.com/2013/09/copyright-metadata-and-attribution.html>
- The Digital Public Library of America Policy Statement on Metadata <http://dp.la/info/wp-content/uploads/2013/04/DPLAMetadataPolicy.pdf>
- Europeana Data Exchange Agreement <http://pro.europeana.eu/documents/900548/8a403108-7050-407e-bd00-141c20082afd>
- OCLC offers an alternative Open Data Commons Attribution <http://www.oclc.org/worldcat/community/record-use/data-licensing.en.html>
- Harvard Library <http://openmetadata.lib.harvard.edu/>
- Arizona State University Library <http://repository.asu.edu/about/policies/metadata/>
- Syracuse University Library https://library.syr.edu/libcgi/cataloging/wiki/index.php/Metadata_policy