



BIBFRAME and Open Metadata Sharing

Solving the Gordian Knot of Interoperability

Ian Bigelow
Head, Cataloguing Strategies
University of Alberta Library



Context: Open Metadata and Cataloguing Practice

Since the very formation of online catalogues, resource description standards and practice have been built on supporting interchange and open sharing of data.¹

- Development and ongoing use of MARC for interchange
- Universal Bibliographic Control
- International Cataloguing Principles
- Work of the PCC

So much of the infrastructure for resource description, from policy to practices hinges on the ability to openly share bibliographic metadata.

But this is now being challenged by organizations looking to monetize bibliographic data and/or where the open sharing of data is seen as a threat to their business models.

1. Deschamps, D., Farnel, S., & Bigelow, I. (in press). From open metadata practice to open metadata policy: Developing an Open Metadata Sharing Framework. In *Open Movements: Recognizing challenges and building connections*, ACRL



Context: Open Metadata and Cataloguing Practice

International Coalition of Library Consortia (ICOLC) and American Library Association (ALA) Joint Statement on the Metadata Rights of Libraries:

“Given how important it is for libraries to use, re-use, and share such metadata freely, current industry trends are troubling. Limited choice in affordable metadata service options, along with restrictions imposed on metadata use, re-use, and sharing across collaborations, platforms, system types, and providers, are impeding libraries’ rights and ability to innovate and collaborate to serve their users effectively.”¹

1. International Coalition of Library Consortia (ICOLC). “Joint Statement on the Metadata Rights of Libraries.” March 21, 2023, 1-5.
https://icolc.net/system/files/Joint%20ALA%20Core%20%26%20ICOLC%20Statement%20on%20Metadata%20Rights%20of%20Libraries_3.21.23_0.pdf.



Context: Open Metadata and Cataloguing Practice

Not surprisingly, the issue of metadata re-use has given rise to further research in the area. For instance, in *No One “Owns” That Metadata*, the authors argue that:

“both locally and collectively, we must encourage the open sharing of member institutions’ library-created metadata through development of aligned policies, and supporting these policies through vendor negotiation practices, including requests to contractually acknowledge library metadata ownership. OCLC and other community-based providers of metadata are encouraged to permit broad sharing of metadata, and clarify policies and contract language that may interfere with the open distribution of metadata.”¹

This captures three key areas that are worth further consideration - copyright, agreements, and vendor business models.

1. Courtney, Kyle, Kathleen DeLaurenti, Matthew Kopel, and Katie Zimmerman. "No One “Owns” That Metadata, Copyright, and the Problems with [Library] Vendor Agreements," 2024.



The Knots to Untangle: Copyright

What does the research indicate for copyright of bibliographic metadata?

“The vast majority of metadata is not copyrightable. Insofar as metadata contain any “thin” copyrightable material, they should be released under a Creative Commons 1.0 Universal (CC0 1.0) Public Domain Dedication (CC0) to ensure the widest possible distribution to support research, aid in the dissemination of knowledge, and promote innovation.”¹

“as with Canada, barring a legal decision on whether library metadata may or may not be protected by copyright, libraries wanting to pursue open metadata sharing can utilize open licences to waive certain or all copyright protections that they may or may not hold over their metadata.”²

So ... the majority of bibliographic metadata is likely not copyrightable, but law varies by region and institutions may want to create open metadata policies to be safe, and to assert that others are welcome to use it in a given way.

1. Courtney, Kyle, Kathleen DeLaurenti, Matthew Kopel, and Katie Zimmerman. "No One “Owns” That Metadata, Copyright, and the Problems with [Library] Vendor Agreements," 2024.

2. Deschamps, d., Farnel, S., & Bigelow, I. (in press). From open metadata practice to open metadata policy: Developing and Open Metadata Sharing Framework. In Open Movements: Recognizing challenges and building connections, ACRL



The Knots to Untangle: Agreements

“For records created by publishers or vendors, contractual restrictions represent a more limited risk of breach, which may be mitigated via negotiation and direct communication with relevant vendors.”¹

This is an important area that can be difficult to outline in terms of specific cases given how it ties to contractual agreements that may differ from institutions to institution. That said, at UAL there has been an ongoing effort to disentangle ourselves from agreements that place limitations on the re-use of bibliographic metadata.

1. Courtney, Kyle, Kathleen DeLaurenti, Matthew Kopel, and Katie Zimmerman. "No One "Owns" That Metadata, Copyright, and the Problems with [Library] Vendor Agreements," 2024.



The Knots to Untangle: Business Models

- Advocacy and engagement is needed here, and hopefully vendors and other agencies working with bibliographic metadata that rely on more restrictive access will be open to change.
- Vendors/Agencies in question tend to rely on the openness/availability of metadata from libraries that create it
- Evolving technologies, such as linked data and AI are reliant on available open data and networked environments, where restrictions on re-use pose both technical challenges and risks. Not changing business models accordingly may end up isolating vendors that don't adapt.



Open Metadata at UAL

- Discussions on how to approach the sharing of UAL metadata began in a more systematic way in 2015
- Based on a very cursory scan of what other libraries and institutions were doing at the time and conversations with staff from the University's Copyright Office, we planned a page within our website that would include various metadata records sets and a statement along the following lines: "The University of Alberta Library believes that the vast majority of metadata is not copyrightable. However, in the case of discrete record sets, a Creative Commons Attribution-Non-Commercial ShareAlike 4.0 License will apply"¹
- 2016 - first few record sets were made available
- 2022 - Worked with partners such as Share-VDE and PCC on respective open metadata policies
- 2023 - UAL Open Metadata Sharing Framework²

1. University of Alberta Library, "University of Alberta Library Data Collection," accessed June 17, 2024, <https://borealisdata.ca/dataverse/UAL>

2. University of Alberta Library, "Open Metadata Sharing Framework," accessed August 25, 2006, <https://library.ualberta.ca/about/policies/open-metadata-sharing-framework>



Context: UAL Open Metadata Sharing Framework

“University of Alberta (UAL) strives to encourage open access to data and support free sharing of information. UAL is committed to ensuring our metadata is as open as possible to support and enhance the discovery of library and cultural heritage resources and permit users to use, reuse, remix, and repurpose our metadata through clearly communicated open licences.”

“UAL’s formalization of open metadata, or metadata sharing, through this framework is in keeping with long-standing library practices of shared, collaborative, and community metadata work.”

“This framework aligns with UAL’s mission: to foster knowledge creation, enable its dissemination, and ensure its preservation, as well as UAL’s long-standing support of and commitment to open access and open data.”

...

“UAL’s Open Metadata/Metadata Sharing Framework is a living document and will be reviewed at regular intervals.”

University of Alberta Library, “Open Metadata Sharing Framework,” accessed August 25, 2006,
<https://library.ualberta.ca/about/policies/open-metadata-sharing-framework>



UAL Catalogue Open Metadata Set Policy

UAL resource descriptions in the UALOCMS are dedicated to the public domain, indicated by the CC0 dedication mark. Any data originally created from other sources with restrictions preventing release under CC0 will not be added to the UALOCMS, but may be made available as a separate open data set under the appropriate licence for re-use (ODC-By for instance). In addition, select UAL catalogue data will have restrictions on re-use that prevent re-use entirely, and these descriptions will not be made available for re-use.

UAL is actively working with partners on reducing reliance on any metadata for the description of our collections that is not openly reusable.



Metadata in the UAL Catalogue comes from a variety of sources, ranging from:

- Descriptions originally created by staff at UAL
 - Identified by MARC 040 \$a = UAB or AEU or CaAEU
 - These will be included in the UALOCMS
- Descriptions originating from other sources significantly modified in some way by UAL
 - Identified by MARC 040 \$d = UAB or AEU or CaAEU
 - In these cases if the 040\$a is from a known source with limitations on re-use, the appropriate licence will be indicated, or the description will be removed from the UALOCMS as appropriate.
- Descriptions from other sources without significant modification by UAL
 - Identified by library codes in 040 \$a other than those for UAL, and without UAL 040 \$d
 - In these cases if the 040\$a is from a known source with limitations on re-use, the appropriate licence will be indicated, or the description will be removed from the open data set as appropriate. Otherwise, these will be included in the UALOCMS.



So...this means, out of all records in the UAL Catalogue (7,420,934):

1. All records in the catalogue created by UAL (365,831)
 - a. Probably the safest set to publish, but more limited
 - b. If all libraries did this and they were organized in a central pool, even this level of open sharing could have a massive impact
2. All records in the catalogue minus ones with a 040\$a from a known institution that doesn't support CC0 (6,926,919)
 - a. This number highlights that the majority of records are not created by those agencies that don't support open re-use. Most are created by a variety of individual libraries.
 - b. But, what does this approach mean in terms of modifications to records, and for agencies with undisclosed/unknown open metadata policies?



What about records modified by an agency that doesn't support CC0?

- MARC and the 040\$d do not provide sufficient detail to determine which modifying agency made any given changes or updates. It only indicates that said agency made “any correction to a record, including cataloging, content designation, or keying changes, but excluding the addition of holdings symbols.”¹
 - Surely this isn't sufficient to claim any ownership of the original metadata. It isn't sufficient to identify which institution made which change.
 - Where records are continually re-used and updated, there are also often multiple 040\$d to consider, and it is impossible to determine who changed what.
- So, unless the 040\$a is from an agency that specifies more limited re-use then consider that
 - Generally bibliographic metadata does not fall under copyright
 - No claim is being made about other policies for the re-use or restriction of re-use of the metadata as originally created
 - Therefore, why not include it in an open metadata set?

1. Library of Congress. (n.d.). MARC standards. Network Development and MARC Standards Office. 040 - Cataloging Source (NR). Retrieved August 25, 2026, from <https://www.loc.gov/marc/bibliographic/bd040.html>



What about the BIBFRAME context though?

With bf the structure of triples/quads, as well as the developing use of AdminMetadata, identifying the creators and derivations of individual entities will be much more likely. So further attention will be needed for which metadata elements have been created by which agencies, as well as those that have been derived. This has benefits in terms of better tracking, but also poses technical challenges for the proliferation of AdminMetadata.

With bf and a linked open data ecosystem though, what are the implications for the use of entities that are not open for re-use?

- Use when necessary applying appropriate attribution and try to limit redistribution? In such a networked environment this could be seen as a significant risk.
- Use the open metadata and strip entities that fall under restrictive licenses
- Do not use at all and go to the source, or alternative open metadata pools

In short, this creates a “use at your own risk” kind of landscape for more closed metadata, and those agencies without open metadata policies could become closed islands



But that premise has a problem, as it assumes that there is a good way to find all the more open data created by libraries through sustainable workflows.

- There are promising developments in terms of open metadata pools (Blue Core, Share-VDE)
- But just looking at the UAL Catalogue, there are 6,561,088 records that are created by other libraries and agencies that so far as we can determine have no assertion of any metadata re-use policy. How do we make the re-use policies for these records more transparent, and how do we leverage them for current workflows?
- Wouldn't it be nice if we had a registry of open metadata policies that:
 - For all institution codes, list the policy for re-use
 - Perhaps where none is provided by the library, default to CC0?
 - Would make it easier for LSP vendors (or individual libraries) to create combined search targets
 - Would allow for record contribution or harvesting of open data sources for larger open metadata pools



Framing open metadata sharing in a bf environment as a Gordian knot type challenge

- Restrictive metadata sharing policies will be challenging as libraries move to linked open data and library data is connected to the open web
- Linked data will allow for clearer, more granular identification for who has created and/or modified specific data elements and entities
- Significant technical development is going into open metadata pools, yet the majority of bibliographic metadata has not been leveraged in part because it hasn't been asserted as open.
- An open metadata registry could be just the right tool to cut this Gordian knot!



Thank You!

Ian Bigelow
Head, Cataloguing Strategies
University of Alberta Library
bigelow@ualberta.ca