

# Making the most of metadata with Metadata 2020

Patricia Feeney, Crossref and Metadata2020

CSE Annual Meeting  
April 2018



---

---

---

---

---

---

---

---

## What is Metadata 2020?

Metadata 2020 is a collaboration that advocates richer, connected, and reusable, open metadata for all research outputs, which will advance scholarly pursuits for the benefit of society.



---

---

---

---

---

---

---

---



## RICHER

Richer metadata fuels discoverability and innovation.



---

---

---

---

---

---

---

---



**CONNECTED**

Connected metadata bridges the gaps between systems and communities.

METADATA 2020

---

---

---

---

---

---

---

---



**REUSABLE**

Reusable, open metadata eliminates duplication of effort.

METADATA 2020

---

---

---

---

---

---

---

---

**“The goal? To demonstrate why richer metadata should be the scholarly community's top priority, how we can all evaluate ourselves and improve, and what can be achieved when we work harder, and work together.”**

- Ginny Hendricks, *Instigator*, Metadata 2020

METADATA 2020

---

---

---

---

---

---

---

---

We already have a lot of standards...

Source: <http://jennriley.com/metadatamap/> CC BY-NC-SA  
 Jenn Riley, Associate Dean, Digital Initiatives, [McGill University Library](#)

METADATA 2020

---

---

---

---

---

---

---

---

---

---

---

---

Who is involved?

- A team of 19 advisors initiated and helped launch
- A core team of 12 individuals drive structural next steps
- Over **70 individuals** in community working groups:  
 Publisher; Librarian; Service Provider / Platforms and Tools; Data Publisher and Repositories; Funders
- So far around 40 individuals signed up to projects

METADATA 2020

---

---

---

---

---

---

---

---

---

---

---

---

Community Groups

- **Publishers;** Chairs: Terri Teleen and Duncan Campbell, Wiley
- **Librarians;** Chair: Juliane Schneider, Harvard Catalyst
- **Service Providers & Platforms and Tools;**  
 Chair: Marianne Calilhanna, Cengage Publishing Services
- **Data Publishers and Repositories;**  
 Chair: John Chodacki, California Digital Library
- **Researchers;** Chair: Cameron Neylon, Curtin
- **Funders;** Chair: Ross Mounce, Arcadia Fund

METADATA 2020

---

---

---

---

---

---

---

---

---

---

---

---

# Defining the issues

METADATA 2020

---

---

---

---

---

---

---

---

## Publisher Challenges

- Metadata management is **enormously costly** for publishers
- The community **lacks an effective metadata distribution** model
- The author community does not yet understand the rationale for supplying full metadata and see deposit as an inconvenience
- Metadata **versioning is a problem**; metadata goes through a number of manipulations which create challenges around version of record
- Vendor submission systems have certain requirements that **publishers cannot control**
- Attaching appropriate metadata to previously published material would be a huge challenge

METADATA 2020

---

---

---

---

---

---

---

---

## Publisher Opportunities

- Potential for **collaboration to define a consistent vocabulary** around metadata for researchers
- Potential to work with submission system platform providers to facilitate more **efficient systems**
- **Develop business cases** for improved metadata that directly results in optimized discoverability; and examples of, or pilots for incentive structure for optimized metadata internal to organizations
- Thoroughly **research customer needs** regarding information required
- **Map metadata**, and collectively work with service providers to define interoperability requirements

METADATA 2020

---

---

---

---

---

---

---

---

## Service Provider/Platforms and Tools

### Challenges

- Metadata creators make **assumptions that don't travel** with the data
- **Lack of consistency of metadata schema**, requirements, and entry; and inconsistency in dates and special characters
- **Records can disappear** without transparency as to where or why
- **No followed community standards about metadata vocabulary**
- Inadequate understanding of the **context of metadata** and its uses lead to improper choices of schema
- Metadata is often **out-of-date** and updates are not supplied by the metadata creators

METADATA  
2020

---

---

---

---

---

---

---

---

---

---

## Service Provider/Platforms and Tools

### Opportunities

- Collaborate with other communities to form and distribute **best practice guidelines**
- Collaborate with other communities to **map metadata**, build awareness about metadata uses, and detect inefficiencies
- Build **awareness of metadata** uses among other communities through use cases, metadata mapping, and community engagement
- Develop and integrate **metadata evaluation tools** for metadata creators in consultation with others in schol comms
- Develop and integrate **new metadata tools** to increase interoperability across systems and avoid duplication of metadata entry

METADATA  
2020

---

---

---

---

---

---

---

---

---

---

## Researcher Challenges

- **Duplication of metadata entry efforts** across multiple systems
- Researchers **don't understand the ramifications of their actions**; entry in submission systems are mapped directly into metadata records
- Institutions require deposit of research and data in repositories which further increases **workload**
- There is **little apparent incentive** to spend extra time to complete the metadata
- There are many frustrations with submission systems - why you can't drag and drop information into relevant fields, **why information cannot be stored for use for subsequent submissions** etc.

METADATA  
2020

---

---

---

---

---

---

---

---

---

---

## Researcher Opportunities

- **Researchers respond well to evaluation systems**, and so will likely be responsive to metadata completeness systems upon submission
- **ORCID** is being adopted
- Researchers are well-positioned to respond to **mandated funder metadata requirements**
- **Researchers respond well to evidence.** Strengthening the story between complete metadata and discoverability, will likely result in more complete metadata
- Help **reimagine a submission system** that is designed for maximal information gathering and interoperability

METADATA  
2020

---

---

---

---

---

---

---

---

## Funder Challenges

- Many funders are **not even aware of 'metadata'**
- Funders are **focused on funding not on schol comms** (publishing, discoverability, citations, etc.)
- Funders are not communications-focused. **Not concerned with communicating stories around their funding**
- Many funders in the Research areas are run by individuals or groups that are traditionalists and **lack technical skills or knowledge**
- Even progressive funders are still focused on policy changes (OA, data sharing, etc.) and not on technical infrastructure issue

METADATA  
2020

---

---

---

---

---

---

---

---

## Funder Opportunities

- Using **data visualizations** as a way of bridging tech novices
- Make sure to draw parallels/connections between enacting new policies and the required metadata
- Funders love **metrics**. Show that citations go up, etc.
- With the advent of Chronos and Funder collaborations, we see more communication within funders and this could mean more possibility of them **educating each other**.
- **Publishers are willing to collect Funder metadata.** So there is an opening for funders to join the dialogue.

METADATA  
2020

---

---

---

---

---

---

---

---

### Data Publisher and Repositories Challenges

- **Repository managers feel caught** between the need for metadata and the need for researchers to deposit in the repositories
- **Lack of awareness** by researchers for the need for metadata
- **Deficient understanding** by researchers about the journey of metadata, and how individual objects are carried by their corresponding metadata
- In some cases, data are fed from instruments and the feeds are set on auto-pilot. This leads to **metadata that is not customized for each dataset**

METADATA  
2020

---

---

---

---

---

---

---

---

### Data Publisher and Repositories Opportunities

- Development and **integration of metadata evaluation tools**
- Collaborate with other communities to **synchronize vocabulary** and evolve messaging for researchers surrounding metadata
- Collaborate with other communities to **develop best practices** and principles to share with repositories and data publishers, so that metadata requirements are consistent
- **Map the metadata lifecycle** in a way that is easily demonstrated to others in the community

METADATA  
2020

---

---

---

---

---

---

---

---

### Librarian Challenges

- There are a variety of different uses of the word 'metadata' and surrounding **vocabulary**. Researchers have built their own systems
- Libraries are overly focused on **legacy systems** and workflows that prevent new efforts getting very far
- Financially, there is a lot invested in proprietary systems or **siloed metadata standards**
- There are **major interoperability challenges**, resulting in a lot of duplication of data entry for use with different systems
- There is an ongoing challenge of extracting enough metadata for discovery but not overburdening the metadata providers

METADATA  
2020

---

---

---

---

---

---

---

---

### Librarian Opportunities

- Librarians have extensive knowledge so are in an excellent place to **repurpose metadata** expertise elsewhere in the scholarly communications chain
- Librarians can **collaborate with researchers** and publishers to create a common metadata vocabulary and consistent communication plan
- A **more widespread adoption of ORCID** in libraries would help
- Librarians and researchers will likely be receptive to use cases of easy systems implementation for heightened **interoperability** (e.g. ORCID)
- The community responds well to principles and **best practice outlines** so can contribute and respond to best practice guidelines

METADATA  
2020

---

---

---

---

---

---

---

---

### Making progress

METADATA  
2020

---

---

---

---

---

---

---

---

### 2018 Projects

- Identified common themes among groups
- 6 projects kicking off in 2018
- All projects cross community groups

METADATA  
2020

---

---

---

---

---

---

---

---

### 2018 Projects

- Researcher communications
- Metadata recommendations and element mappings
- Defining the terms we use about metadata
- Incentives for improving metadata quality
- Shared best practices and principles
- Metadata evaluation and guidance

METADATA  
2020

---

---

---

---

---

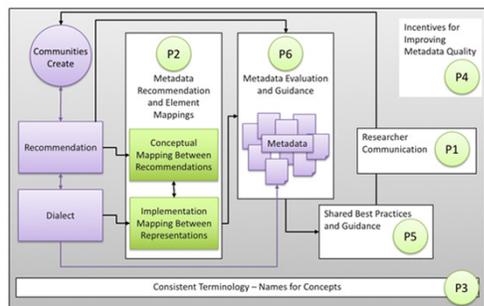
---

---

---

### Metadata 2020 Thread

Supporting the entire metadata life cycle



METADATA  
2020

---

---

---

---

---

---

---

---

### Can you help?

- Contribute to Metadata 2020 [projects](#)! Email Clare Dean at [cdean@metadata2020.org](mailto:cdean@metadata2020.org) for details, or sign up [here](#).
- Help promote our efforts to the wider community through your organizations, word of mouth, and social media
- Find us on @Metadata2020 Twitter, Facebook, LinkedIn, and at [metadata2020.org](http://metadata2020.org)

METADATA  
2020

---

---

---

---

---

---

---

---

**Thank you!**

Metadata2020.org  
@metadata2020  
info@metadata2020.org

METADATA  
2020

---

---

---

---

---

---

---

---