

# Metadata 2020: a community collaboration to advance metadata for scholarly communications



*Metadata 2020 is a collaboration of scholarly communications stakeholders working towards richer, connected, reusable, and open metadata for all research outputs. **Clare Dean** explains what the collaboration's goals are, what common problems and opportunities for progress have already been identified, and how these are being addressed by six main projects. These include improved communications with researchers on the importance and uses of metadata, more clearly defined and better understood terminologies around metadata, and the sharing of best practices, principles, and guidelines.*

[Metadata 2020](#) is a collaboration of publishers, librarians, service providers, providers of platforms and tools, data publishers, repositories, researchers, and funders. All have the shared mission of working towards richer, connected, reusable, and open metadata for all research outputs for the benefit of our society.

Right now there are so many problems with incomplete, inaccurate metadata and the interoperability of the systems that use it, that one year in we are still in the process of gathering a list of all those shared by our participants.

As Founder Ginny Hendricks noted in her [inaugural blog post](#), “there are huge gaps in the metadata that permeate throughout thousands of systems downstream, and we all suffer from mistyped, misplaced or just plain missing metadata. We initiated Metadata 2020 to bring together all the relevant parties from around the world, air the grievances, understand the barriers, and then to make it easier to reach and evaluate research outputs through better metadata”.

Without complete, open, connected, interoperable metadata, discoverability, use, and reuse is thwarted and the progress of research stymied. All metadata pretty much starts and ends with the researcher. If something is good for them, it's good for all of us.

## What are we doing about it?

Our collaboration is working to:

- Address interoperability challenges
- Create useful resources for the industry to navigate schemas and guidance
- Create consistent metadata terminology
- Provide guidance in the form of shared best practice and principles
- Equip individuals with the resources and guidance to help their organisations understand the need for infrastructure improvements
- Provide information about metadata, its use, and significance for researchers
- Help organisations to accommodate researcher metadata needs.

The logo for Metadata 2020 features the words "METADATA" and "2020" in a bold, blue, sans-serif font. The letters are slightly spaced out and have a subtle shadow effect, giving them a three-dimensional appearance. The text is centered on a bright yellow rectangular background.

# METADATA 2020

## A cross-community approach

Early on, the Metadata 2020 Advisory Board realised any effort to improve metadata for scholarly communications needed to involve stakeholders from across the community. While efforts have been made to improve metadata and metadata workflows within the librarian community, for example, we understood that if we were to address interoperability challenges, break down silos, and increase consistency in communication we needed to involve the full range of scholarly communications stakeholders.

The collaboration launched in September 2017, to an overwhelmingly positive reaction, and shortly afterwards participants met to define the core problems for their respective communities. From there we were able to identify common problems and opportunities for progress across the community groups, and formed six distinct cross-community projects, each launching in March 2018. Several months on, each has made significant progress thanks to the 130+ individuals who have volunteered their time and expertise. The six projects are:

### [1. Researcher communications](#)

While some researchers are well-positioned to understand the importance and uses of metadata, there are many who remain uninformed, which contributes to the incompleteness and inconsistency of metadata deposited as research is published. This project looks to gain a deeper understanding of researcher perceptions and needs around metadata, with a view to developing resources to assist them and to enable other communities to better serve them. This group is led by Alice Meadows of ORCID and Michelle Urberg of Proquest.

### [2. Metadata recommendations and element mapping](#)

Mapping between the recommended concepts and schemas is an important step towards a single recommendation that is consistent across communities. Jim Swainston of Emerald Group Publishing leads the effort to map between schemas.

This project also leads a sub-group of Metadata 2020 participants in the development of a metadata flow diagram, charting the flow of metadata in and out of types of organisations and systems throughout scholarly communications.

### [3. Defining the terms we use about metadata](#)

In the metadata space there is no agreement on what words like “property”, “term”, “concept”, “schema”, or “title” refer to. This project, led by T. Scott Plutchak of the University of Alabama at Birmingham, is working on developing a glossary of terms we use about metadata, to mitigate further confusion.

A sub-group has also been formed between this project and project #1 (Researcher communications), to convene focus groups and survey researchers in order to gain a more complete understanding of their metadata comprehension, needs, and uses, and how these may vary between disciplines.

#### [4. Incentives for improving metadata quality](#)

How can we motivate organisations and businesses in scholarly communications to improve their metadata? How do we support individuals to make the case for metadata being a strategic (not an operational) solution to decision-makers in their organisations? How might we elevate the importance of metadata to motivate publishers, service providers, and libraries to make the sometimes costly infrastructure changes to enhance the completeness, connectedness, openness, and reusability of metadata? Fiona Counsell of Taylor & Francis leads the effort to provide resources for individuals at organisations to help them make a case for the system changes needed to improve metadata quality.

#### [5. Shared best practice and principles](#)

This project is advancing quickly to build a set of high-level best practices for using metadata across the scholarly communication cycle, to facilitate interoperability and an easier exchange of information and data across stakeholders in the process, irrespective of chosen schema or standard. This effort is led by Jennifer Kemp of Crossref and Howard Ratner of CHORUS.

#### [6. Metadata evaluation and guidance](#)

Ted Habermann of the HDF Group rallies this group to examine current metadata evaluation tools and guidance around existing schemas, and identify areas where further guidance and tools are needed.

### Next steps

Metadata 2020 is in the process of seeking funding to expand our outreach to the researcher community, and to hold practical in-person meetings. This autumn, we will be convening in-person participant workshops to run focus groups and progress some of our more technical projects.

### An open invitation

Metadata 2020 welcomes anyone from scholarly communications who may be interested in contributing to this work. Whether you are a metadata expert or a novice with project management or communication skills, there is a place for your contribution. One of the key successes of the work so far has been the coming together of technical with non-technical people. We understand that we need to break down silos and work as a wider scholarly communications community to find solutions that work for everyone, with the ultimate goal of helping research be of benefit to society.

*If you're interested in getting involved or if you'd like to be kept updated with this work via an email newsletter, please contact Clare at [cdean@metadata2020.org](mailto:cdean@metadata2020.org).*

*Note: This article gives the views of the author, and not the position of the LSE Impact Blog, nor of the London School of Economics. Please review our [comments policy](#) if you have any concerns on posting a comment below.*

### About the author

**Clare Dean** is an independent scholarly communications outreach consultant, and manages the day-to-day running of Metadata 2020, alongside other projects for non-profit organisations. Most of her 13 years' experience has involved the development of academic journals through strategic marketing and communications. Originally from the UK, she now lives in Massachusetts, USA.