

Abstract

This article highlights community and stakeholder mobilization initiatives in the library and heritage sectors that help in the transition to the next generation of metadata. We draw from the Next Generation of Metadata round table discussions organized by OCLC Research in March 2021. In these discussions, we saw next generation metadata mobilization taking place along two trajectories: (1) transforming and publishing institutionally sourced metadata and (2) improving metadata already in the supply chain. The article provides context and scope from the round table conversations and highlights national initiatives taking place in both mobilization areas. The article then discusses the challenge of managing at multiple scales, as efforts of local, national and global scale gear up to connect with each other.

Introduction

OCLC Research facilitated eight virtual round table discussions¹ with metadata experts on the topic of transitioning to the next generation of metadata, throughout the month of March 2021. In total, 86 participants from 71 different institutions – mostly university libraries, bibliographic agencies and national libraries – contributed to the sessions. Sessions were held in six different European languages so that participants could comfortably contribute to the discussion.

The main discussion question, “*How do we make the transition to the next generation of metadata happen at the **right scale** and in a **sustainable manner**, building an **interconnected ecosystem**, not a garden of silos?*” implies the need for intentional coordinated action towards achieving a shared vision of this future ecosystem. In the succeeding sections, we first provide the context of these conversations and then we present coordination and collaboration approaches pertinent to achieving transition.

Context

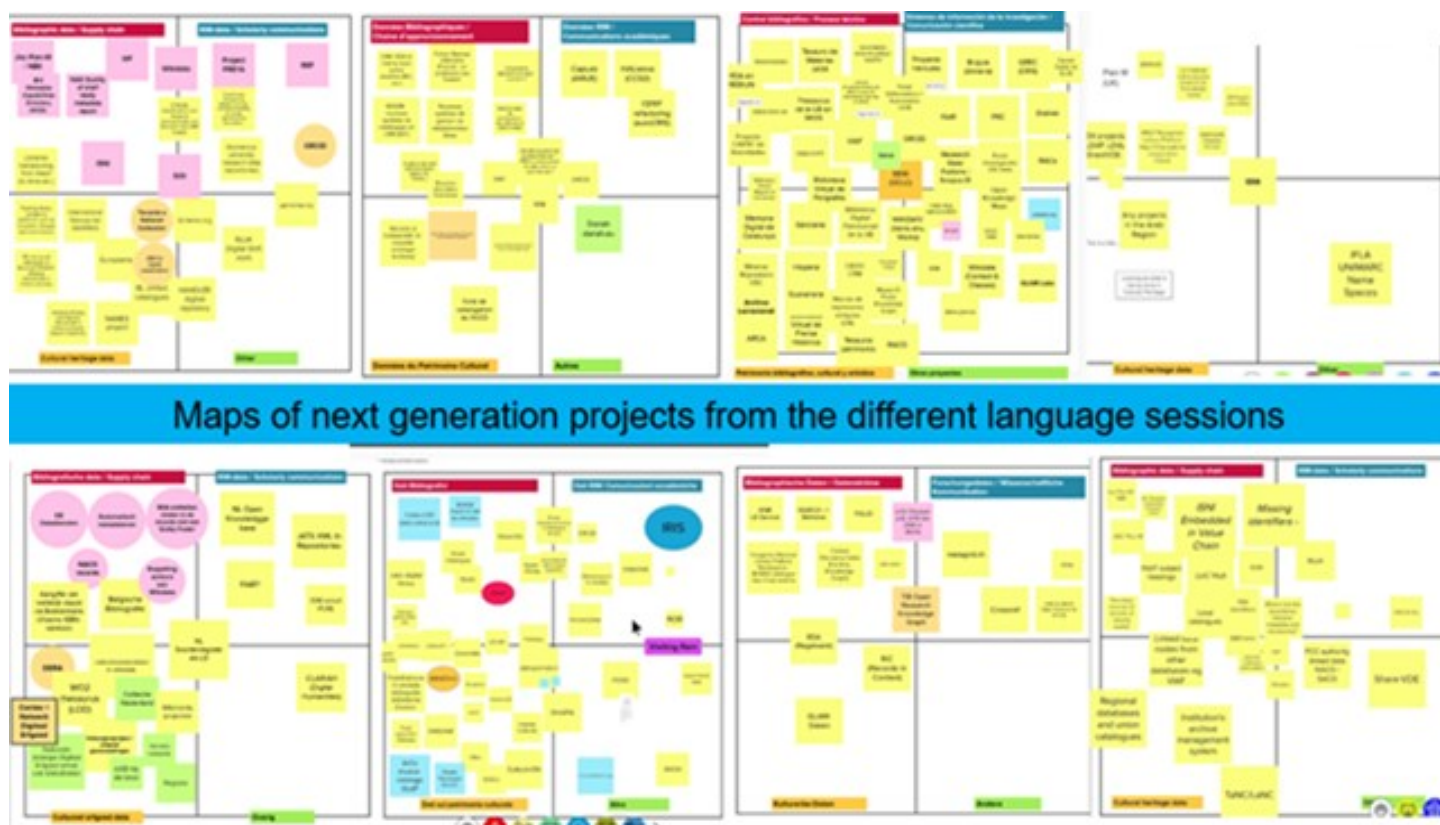
Karen Smith-Yoshimura’s report, “Transitioning to the Next Generation of Metadata”² served as background reading and inspiration for the sessions. This report consolidates six years (2015-2020) of discussions on topics covered by the OCLC Research Library Partnership Metadata Managers Focus Group³ about the future of metadata and the changes affecting their core areas of work: Creating and managing metadata. The changes taking place include the transition to linked data and the adoption of identifiers such as URIs (Universal Resource Identifiers), the move of metadata creation upstream in the data supply chain, and the evolution of metadata as a service. These changes open opportunities for libraries to engage in new areas where metadata is becoming key, such as Research Data Management (RDM) and Research Information Management (RIM). These changes are also key to achieving important library goals, such as support for multilingualism – which in turn, is closely connected to EDI (Equity, Diversity and Inclusion) principles.⁴

Working from the premise that the “what” and the “why” of the future of metadata operations are well understood, the round table sessions aimed to address the practical concern of the “how to” and “with whom” – a concern aptly articulated by Jane Daniels, moderator of the CILIP Conference 2020 session where Karen presented her report.⁵

“I have read the report right through, in fact I have read quite a lot of it twice now. And every time I read it, I think – yes, I can see the direction we should be going in. I suppose the question is, how do we mobilize, probably at an international scale, to make sure that what you are listing, or what you are putting in the report actually comes about”.

To get the conversation started we asked attendees to map next generation metadata projects they were aware of on a 2x2 matrix characterizing three different application areas: 1) bibliographic data and the supply chain, 2) cultural heritage data and 3) metadata to support RIM and scholarly communications.

The collage of the eight maps that came out of the round tables shows a large number of cultural heritage data projects and bibliographic data projects and few RIM projects, reflecting the focus and expertise of the attendees. Clusters of post-it notes naming persistent identifiers (PIDs) – such as ISNI, ORCID, VIAF, DOI – featured prominently in all the maps, showing their value to participants.



The conversation then took deep dives into some of the initiatives, projects and services in order to analyse what it takes to make the transition to next generation metadata **scalable and sustainable** and to build an **interoperable and interconnected ecosystem**.

In the next section we present coordination and collaboration approaches that stood out because of their intent to mobilize communities and stakeholders in bringing about this transition.

Approaches geared toward mobilizing stakeholders and communities

Transforming and publishing metadata

From the round table sessions, we learned that libraries are deeply invested in transforming and publishing their authority files (both the name authorities and the subject headings), in order to leverage them. We heard that open government data policies and national interoperability frameworks are strong drivers for this focus. National libraries have been pioneers in the field, publishing their national bibliographies and authority files as linked data in the course of the past decade. Whilst many libraries utilize the national authority file maintained by their country's national library or a large international authority files - such as the Library of Congress Name Authority File (LC/NACO Authority File) – others maintain their own local authority files. Many institutions with major cultural heritage or academic collections are also publishing or planning to publish their local authority files and bibliographic data as linked data. They mostly act individually. As a result, their approaches and implementations tend to differ, leading to a range of idiosyncratic solutions and much duplication of effort. This problem was identified during the round tables. Innovative approaches to create opportunity for more concerted effort are still sparse, however, two initiatives were discussed - one in France and one in the Netherlands – that provide interesting models for such concerted effort at a country level.

In France the collaboration between the two main bibliographic agencies – the [Bibliothèque nationale de France \(BnF\)](#), and the Bibliographic Agency for Higher Education (Agence bibliographique de l'enseignement supérieur, Abes) – is seen as a breakthrough needed to lead the bibliographic transition in France. BnF and Abes were enticed to collaborate because of the harder delineation between metadata production and release on the one hand and metadata management and (re-) use on the other, instigated by the open government data policy. The aim of the collaboration is to maximize the benefits of centralization, normalization, and efficiency at production and publication time. The two agencies decided to build and populate the French National Entity File (NEF) and to produce bibliographic data following agreed standards and practices, based on RDA and IFLA models. To involve the French library and archive community they launched the national Bibliographic Transition Program,⁶ which, in their own words, “*aims to support all stakeholders in France through the drastic changes induced by such a major normative and technological evolution.*” The program encompasses the massive processing, refactoring, and publication of the French legacy heritage data, the country-wide adoption of the new French cataloguing code that is in development, consultation with providers of library systems and software solutions, and professional development for French cataloguers and metadata experts to help them make the transition.

In the Netherlands, the Dutch Network for Digital Heritage (NDE)⁷ leads the bibliographic and cultural heritage data transition. It brings together the main players from the library, academic, media, museum, archive and architecture sectors with the aim to jointly implement the national strategy for digital heritage on behalf of the Ministry of Education, Culture and Science. The Dutch Ministry seeks alignment with the European Interoperability Framework (EIF),⁸ which promotes interoperability of digital public services at EU levels. As a result, NDE developed the Digital Heritage Reference Architecture (DERA)⁹ which, importantly, requires the application of linked data principles for the publication and linking of heritage data. The remit and ambition of the NDE are broader than just leading the bibliographic transition and address the full digital strategy for the heritage sector. On the other hand, its practical impact is limited to the voluntary adoption of best practices and guidelines, as it does not have a centrally coordinated approach or roadmap for the transition. However, the network is influential and successful in raising awareness, knowledge sharing, training professionals and deploying digital heritage coaches in its effort to mobilize the Dutch heritage community and its stakeholders.

The French and Dutch examples are very different in approach and scope, but both are efforts to mobilize the community towards a more unified approach to next generation metadata. They both showcase EU policies' impact on collaborative arrangements in the library and cultural heritage sector in Europe. The impact is obviously not the same in all EU member countries even if the guiding principle: “*digital-by-default, cross-border-by-default and open-by-default,*” is a shared one. The Italian round table session for example, mentioned Italian “Guidelines for Semantic Interoperability through Linked Open Data”¹⁰ produced as early as 2013 and inspired by the EIF. However, Italian participants observed a lack of coordination between initiatives within the country and a fragmented cultural heritage landscape resulting in a variety of independent projects and a lack of common standards.

Improving the metadata in the supply chain

Round table participants reported how libraries are intentionally feeding external systems - e.g., the university's research portal, the ORCID-database, Wikidata, etc. - with their authority data to achieve critical mass. They are also systematically embedding persistent identifiers (PIDs) throughout the supply chain to rationalize it. This is particularly true for libraries and bibliographic agencies that act as registration agencies for identifiers, such as the International Standard Name Identifier (ISNI) for example, and who operate in the context of their national bibliography and legal deposit tasks.

More coordinated approaches that go beyond individual initiatives and institutional scale, are currently taking place in the UK to achieve efficiencies at scale. These were discussed during the English round table discussions. One such effort concerns the embedding of PIDs within the supply chain to get them out to where cataloguers can pick them up. To this end, the British Library is collaborating with publishers to encourage the adoption of ISNIs upstream. As a result, publishers are starting to use them in their ONIX feeds. Also, cataloguing agencies in the UK are being supplied with ISNIs so that they can embed them in the metadata at source, in the cataloguing-in-publication (CIP) metadata, that they supply to libraries in the UK. Efforts are also underway to work together with OCLC to systematically match ISNI entries against VIAF entries, and to provide a reconciliation file to enable OCLC to update the VIAF with the most recent ISNI. These identifiers could then be fed through to the Library of Congress, who can use these to update their LC/NACO Authority File. As one of the participants said:

“With 6 million files to update, this is a perfect example of a leading edge dynamic next generation metadata initiative that will have to overcome the considerable challenge of scalability for it to succeed at a global level.”

In this context, the UK-HE-wide approach to funnel authority data in a coordinated way to the LC/NACO Authority File was also mentioned.

Finally, Plan-M¹¹ was briefly mentioned. It is an initiative, facilitated by Jisc, to rethink the way that metadata for academic and specialist libraries is created, sold, licensed, shared, and re-used in the UK. It is a conversation between libraries, suppliers and intermediary organisations to streamline the metadata marketplace in the UK so that it is more coherent, transparent, robust and sustainable.

Managing multiple scales

It was fascinating to hear about these country-initiatives on the one hand and to hear about the challenges faced by individual libraries striving to effectively move to the next level of metadata management, on the other. For the latter, often the most convenient and attractive way to overcome the issues they face, is to connect to national or global scaled efforts. Of the latter, two were most mentioned:

- LC/NACO Authority File, one of the largest international authority files maintained cooperatively, and which is massively being enriched with relevant identifiers.
- OCLC's Shared Entity Management Infrastructure, which is being built at global scale to provide a centralized infrastructure for managing shared library data about entities such as person names and bibliographic works.

These global components were also considered important to link up with for the country-initiatives. The challenge then becomes one of managing multiple scales and global interoperability. Andrew Pace, Executive Director, Technical Research at OCLC, described the landscape as a “*semantic continuum*” – portraying the challenge of managing multiple scales as one of “*bridging the effort between the short tail and the long tail*”, in other words, between scaled effort and localized domain and collection expertise. To that end, libraries need to “*balance large, shared, homogenous collections and data, while accounting for a myriad of de-centralized and heterogeneous collections*”.

Conclusion

The round table conversations gave participants an interesting glimpse into the relief formation of the newly emerging landscape of the next generation of metadata. Unsurprisingly perhaps, but noteworthy nevertheless, national scale programmes - in France, the Netherlands and the UK - stood out as most intentional in mobilizing community, from the many different projects and initiatives discussed during the round tables. Participants in the Italian-, German- and Spanish-speaking sessions expressed a strong desire to move their next generation metadata efforts forward in a more concerted way. They clearly saw the benefits of cooperating more strongly at a national level and, also internationally, especially when cooperating with those further ahead. They expressed the wish to see OCLC help them organize more discussions like this at the country level and to play an active role in this moment of transition. This message did not fall on deaf ears, and at OCLC we are planning next steps with the community to keep the momentum going.

References

- ¹OCLC Research Discussion Series on the Next Generation of Metadata (23 February – 13 April 2021): <https://www.oclc.org/go/en/events/next-generation-of-metadata/Register.html>
- ² Smith-Yoshimura, K. 2020. *Transitioning to the Next Generation of Metadata*. Dublin, OH: OCLC Research. <https://doi.org/10.25333/rqgd-b343>
- ³ The OCLC Research Library Partnership (RLP) Metadata Managers Focus Group homepage <https://www.oclc.org/research/areas/data-science/metadata-managers.html>
- ⁴ OCLC Research. Equity, Diversity, and Inclusion Initiatives (viewed 7 June 2021). <https://www.oclc.org/research/areas/community-catalysts/edi.html>
- ⁵ CILIP Conference 2020, session 'Transitioning to the Next Generation of Metadata' (19 November 2020): <https://pheedloop.com/cilipconf20/site/sessions/?event=cilipconf20§ion=38349&id=HxkkpJ>
- ⁶ The Bibliographic Transition in France webpage (in English): <https://www.transition-bibliographique.fr/enjeux/bibliographic-transition-in-france/>
- ⁷ The Dutch Network for Digital heritage homepage: <https://netwerkdigitaalerfgoed.nl/>
- ⁸ Annex 2 to the Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions 'European Interoperability Framework – Implementation Strategy'. COM(2017). English version. http://eur-lex.europa.eu/resource.html?uri=cellar%3A2c2f2554-0faf-11e7-8a35-01aa75ed71a1.0017.02/DOC_3&format=PDF
- ⁹ Digitaal Erfgoed Referentie Architectuur webpage (in Dutch): <https://netwerkdigitaalerfgoed.nl/activiteiten/dera/>