LIBRARY FUTURES

OCLC EMEARC WEBINAR SERIES
This webinar will be recorded. The recording will be published on the OCLC website.
Some tips with WebEx

- To go to the chat
- Send a chat message to All Panelists
- Ask the host for technical help
Welcome

Reinhard Altenhöner
EMEA Regional Council Executive Committee Delegate

#oclcemearc21
Reinhard Altenhöner
Deputy General Director
Staatsbibliothek Berlin, Germany
EMEARC Executive Committee 2020/2021

EMEA CHAIR
Jan Simane
Library Director
Kunsthistorisches Institut Max Planck Institut in Florenz Italy

EMEA VICE CHAIR
Sarah Hurter-Savie
Library Director
Nice Sophia Antipolis University France

Reinhard Altenhöner
Deputy Director General
Staatsbibliothek zu Berlin Germany

Shaikha Almehairi
Director of the Libraries Department
Department for Culture and Tourism Abu Dhabi

Adelaida Ferrer
Projects Coordinator
Universitat de Barcelona Spain

Stuart Hunt
Director, University Library & Collections Services/University Librarian
University of Reading United Kingdom

Beate Meinck
Head of the City Library
Library of Reutlingen Germany

Matthijs van Otegem
University Librarian
Erasmus University Rotterdam The Netherlands

Evi Tramantza
Director of Libraries and Archives
Anatolia College Greece

Anna Wołodko
Library Director
University of Warsaw Poland
Today’s speakers
Findings from the round table discussions
Transitioning report by Karen Smith-Yoshimura

oc.lc/nextgen-metadata-report

This report served as inspiration and background reading for the discussion series

Spanish translation of Karen’s report on the BNE website

#OCLCmetadataseries
The OCLC Research Series on Next Generation Metadata

OPENING PLENARY WEBINAR
Tuesday 23 February 2021, 15:00 (CET)

INTERACTIVE ROUND TABLE
During the first two weeks of March 2021.

CLOSING PLENARY WEBINAR
Tuesday 13 April 2021, 15:00 (CET)

#OCLCmetadata_series
Mapping exercise

• Bibliographic data / Supply chain

• Cultural heritage data

• RIM-data / Scholarly communications
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?
Participants and outputs

• **Broad interest**: representatives from 71 institutions (67% university libraries) - from 17 countries

• **Insightful reports**: 8 blog posts + translations on [HangingTogether](https://www.hangingtogether.org)
"It was very inspiring and interesting to hear how other experts from different backgrounds and projects deal with similar problems. I liked that it was a general conversation without a strict agenda which gave participates the opportunity to freely express their thoughts and the conversation to be lively and adaptable.”

[First English session]
Findings

1. Predominant topics

2. Barriers and needs during the transition

3. Expectations and questions about the Shared Entity Management Infrastructure
Predominant topics of the discussions
Maps of next generation projects from the different language sessions
Predominant topics

1. Investing in authorities and identifiers
2. Interlinking data
3. Managing multiple scales
Investing in authorities and identifiers

Context: Open Government Data

- strong focus on **transforming** and **publishing** library data
- is driving **collaboration** at **national** levels

*e.g.* [French National Bibliographic Transition Program](#)
Investing in authorities and identifiers

Aim: To achieve a critical mass of interoperable library data

• increasing the usefulness of authority data through reconciliation
• populating data sources and systems with library data
• embedding PIDs in the value-chain

  e.g. ISNI => VIAF => LC/NACO

#OCLCmetadata
Interlinking data

“When you refer to data that has already been defined by others, you relinquish control over that piece of information, and that can be a mental barrier against doing linked data the proper way. It is much safer to store and manage all the data in your own silo. But the moment you can let go of that, the world can become much richer than you could ever achieve on your own.”

[Dutch session]
Interlinking data

• About aggregators and discovery hubs and the value proposition of linked data

• Need for semantic interoperability: there is no unique model; reconcile across vocabularies/ontologies inside and outside library domain

• Opportunities:
  – Connecting between different languages
  – Connecting between different granularities
  – Automatic metadata generation/AI/machine-learning
“There are multiple, co-existing ‘right scales’ for next generation metadata that need to be interconnected and managed as a whole.”

[French session]
Andrew K. Pace
Executive Director for Technical Research, OCLC
What we learned: Project Passage & CONTENTdm Pilot

Project Passage
- Linked Data can exceed current library standards.
- Ontologies are vitally important.
- User interface enhancements are needed.
- A knowledge graph requires new interoperability and import functionality.
- Linked data diminishes the distinction between bibliographic and authority data.

oc.lc/passagereport
oc.lc/transform-linked-data
What we learned: Project Passage & CONTENTdm Pilot

Project Passage
• Linked Data can exceed current library standards.
• Ontologies are vitally important.
• User interface enhancements are needed.
• A knowledge graph requires new interoperability and import functionality.
• Linked data diminishes the distinction between bibliographic and authority data.

CONTENTdm Linked Data Prototype
• The value proposition succeeds.
• Shared data models are an excellent starting point.
• Decentralized domain expertise is needed for the transformation.
• Substantial resource commitments required, but benefits accrue long before the journey is complete.
• Working partnerships represent strength in numbers.
What we heard

- Publishing data on the web drives national collaboration, requiring
  - the presence persistent identifiers and the ability to create them in
    library workflows
  - the interfaces and ecosystem to create native linked data
    descriptions
- To achieve a critical mass of interoperable library data, libraries are
  - investing in authorities and identifiers
  - seeding the web with PIDs and library data
- For successful interlinking, libraries need to
  - find tools for broad reconciliation across vocabularies & ontologies
  - relinquish control and grow trust in data from external parties
  - reconcile vocabularies/ontologies inside and outside library domain
- Libraries will have to manage multiple scales
What we discovered: a semantic continuum
What we discovered: a semantic continuum

- Shared, homogeneous, and centralized entities...
- ...accounting for the reality of localized, heterogeneous, de-centralized collections
<table>
<thead>
<tr>
<th>Shared, homogeneous, and centralized entities…</th>
<th>…accounting for the reality of localized, heterogeneous, de-centralized collections</th>
</tr>
</thead>
<tbody>
<tr>
<td>Machine-matching, highly automated reconciliation…</td>
<td>…with tools for hand-matching, semi-automated reconciliation</td>
</tr>
</tbody>
</table>

What we discovered: a semantic continuum
What we discovered: a semantic continuum

- Shared, homogeneous, and centralized entities...
- Machine-matching, highly automated reconciliation...
- Well-accepted context: Persons & Works

...accounting for the reality of localized, heterogeneous, de-centralized collections
...with tools for hand-matching, semi-automated reconciliation
Granular context: About, Depicts, Annotations, Notes
What we discovered: a semantic continuum

| Shared, homogeneous, and centralized entities… | ...accounting for the reality of localized, heterogeneous, de-centralized collections |
| Machine-matching, highly automated reconciliation… | ...with tools for hand-matching, semi-automated reconciliation |
| Well-accepted context: Persons & Works | Granular context: About, Depicts, Annotations, Notes |
| Blurs the line between bibliographic & authority work | Blurs the line between object description & context description |
## What we discovered: a semantic continuum

| Shared, homogeneous, and centralized entities… | ...accounting for the reality of localized, heterogeneous, de-centralized collections |
| Machine-matching, highly automated reconciliation… | ...with tools for hand-matching, semi-automated reconciliation |
| Well-accepted context: Persons & Works | Granular context: About, Depicts, Annotations, Notes |
| Blurs the line between bibliographic & authority work | Blurs the line between object description & context description |

A large centralized infrastructure is needed. Custom applications and interfaces are needed.

**This will facilitate new Knowledge Work in libraries.**
A large centralized infrastructure is needed. Custom applications and interfaces are needed.

This will facilitate new Knowledge Work in libraries.

<table>
<thead>
<tr>
<th>Shared, homogeneous, centralized entities…</th>
<th>Machine-matching, highly automated reconciliation…</th>
</tr>
</thead>
<tbody>
<tr>
<td>Well-accepted context: Persons &amp; Works…</td>
<td>Hand-matching, semi-automated reconciliation…</td>
</tr>
<tr>
<td>Blurs the line between bibliographic &amp; authority work…</td>
<td>Blurs the line between object description &amp; context description…</td>
</tr>
</tbody>
</table>

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?

A large centralized infrastructure is needed. Custom applications and interfaces are needed.
Barriers and needs during the transition
Systems used in libraries are not ready

Current systems are not ready to support libraries in the transition:

- No integration of PIDS, local authorities, or links to/from external sources
- No integrated workflow, additional systems, duplication of effort, waste of time
Sustainability versus projects

At the basis of sustainability is **knowledge**. Sustainability is closely linked to:
- the purpose and objectives of projects
- the data quality
- the extent to which the end user benefits

"Let us get rid of the 'project' concept, but rather acknowledge that this is an ongoing effort which needs appropriate staffing, unlimited job positions, and sufficient financial resources! In this realm - at least of next generation metadata - we should no longer be working on a project basis."

[German session]
Knowledge sharing/professional development

• Next generation metadata is often a “stretch assignment” for staff, on top of everything else
• Having to learn new tools/systems (e.g. Wikidata) is not easily done. Learning curves are steep
• Regular training for staff in these areas (graph data, tools) is necessary but is not always available
• Looking to OCLC for leadership in this area
Rachel Frick
Executive Director, Research Library Partnership (RLP), OCLC
OCLC support and professional development

- **Practitioners** skilled to implement
- **Managers** understanding opportunities
- **Leaders** recognizing priority
OCLC Active Learning

- **WebJunction Course Catalog**
- **OCLC Research Library Partnership Metadata Managers Focus Group**
- **OCLC Community Center**
Expectations and questions around the Shared Entity Management Infrastructure
Expectations and questions

- OCLC's platform is expected to help connect or even integrate collections internally and externally

  “... this project is really going to transform the way we work. It will mean that we’re able to connect our metadata with our partner institutions, it will mean that medieval manuscripts will no longer be in this walled garden.”

  [First English session]

- A solution for multilinguality
- A platform for building national infrastructures
- Questions: Functionality, Data ownership, Governance, Business Model
John Chapman
Senior Product Manager, Metadata Services, OCLC
Shared Entity Management Infrastructure

Basics:

- Entity data published on the web as linked open data (CC-BY)
- Subscription access for editing and API use
- Shared system; OCLC staff responsible for some quality and maintenance tasks
Shared Entity Management Infrastructure

Goals:

• Address infrastructure needs identified by libraries
  • Stand behind entity URIs
  • Expand on “native” metadata management
  • Link library data to non-library data… and shared data to local data
• Operate sustainably, at scale, and complement other efforts
• Deliver December 2021
Shared Entity Management Infrastructure

What’s next?

• Exploring multilingual approaches
• Emphasizing quality
• Integration with other services and applications
Shared Entity Management Infrastructure

Let’s continue the work together!

- Determining cooperation and efficient use of resources
- Coexistence of shared environments and regional, national “domains”
CLOSING REMARKS
Implementing Sustainable Development Goals in my library
14 April, 14:00 – 15:00 CET

Register now: oc.lc/emearc21
Stay connected

oc.lc/mail

Updates, news and more