Works cited

Barroso, I., Hartmann, N., & Ribeiro, C. (2015). Metadata Crosswalk for a Museum Collection in a Thematic Digital Library. Journal of Library Metadata, 15(1), 36–49.

Bigelow, I., & Sparling, A. (2022). Control or Chaos: Embracing Change and Harnessing Innovation in an Ecosystem of Shared Bibliographic Data. JLIS.It, Italian Journal of Library, Archives & Information Science, 13(1), 67–85.

Bountouri, L., & Gergatsoulis, M. (2009). Interoperability Between Archival and Bibliographic Metadata: An EAD to MODS Crosswalk. Journal of Library Metadata, 9(1–2), 98–133.

Castelli, L., Felicetti, A., & Proietti, F. (2021). Heritage Science and Cultural Heritage: Standards and tools for establishing cross-domain data interoperability. International Journal on Digital Libraries, 22(3), 279–287.

Chen, Y.-N. (2015). A RDF-based approach to metadata crosswalk for semantic interoperability at the data element level. Library Hi Tech, 33(2), 175–194.

Zeng, M.L. & Qin, J. (2022). Metadata (3rd Edition). ALA Neal-Schulman.

A No-Nonsense Guide to Thinking About **Interoperability**

What is it?

Why bother with it?

By Rosemarie Fettig

For LBSC770

1 December 2022

>What is interoperability?

Well, if you're picking up this zine I would assume that you already know the answer to this, BUT like many terms in the library/information science sphere, "interoperability" gets used by a lot of different people in a lot of different ways.

For the purposes of this conversation, we're going to cast a wide net and say that interoperability is ***"the capability of different information systems to communicate with one another"*** (Zeng and Qin 359). Here, we'll specify "information systems" to mean "metadata structures." This can include everything from data models to element sets to encoding schemes – it really depends on what your goal is, what you're working with, and a whole bunch of other factors, some of which we'll do our best to touch on here.

If we want to get more specific, we can say that "Interoperability is succeeded when a set of processes ensures that systems manage their information sources in a way that supports the reuse and exchange of data from various sources inside and outside a particular system" (Bountouri 98-9).

>Concluding Thoughts

If you take away any ideas from this zine, I hope that the following are among them:

- Interoperability is a mindset, an approach, a process, and a goal – it is not a state of being!
    - A collection and its metadata cannot be interoperable in a vacuum – you must always establish what you want your metadata to be interoperable *with*
- You don't have to reinvent the wheel!
    - This advice applies to most metadata projects, but when it comes to interoperability and crosswalking, it will often be more efficient to start by making connections with what you already have. Don't assume a total transformation is necessary!
- Interoperable doesn't mean "matches 1:1 with other standards/schema/etc" it means "able to crosswalk or make connections with other metadata sets"

Very few collections will be interoperable without some kind of intervention for each instance of connection!

>Cross-domain crosswalking considerations

Interoperability often becomes part of the conversation when projects seek to connect collections and their corresponding metadata over domains that often have their own metadata standards and practices. A common example is integrating museum, library, and archive collections under a unified search portal.

Barroso et al. (2015) approached a project like this by focusing on the intended output, "keeping in mind the intended views for its potential users and leaving room for the addition of new records" (38) – this is the goal-oriented process that we previously discussed in action. Keeping the user experience at the forefront made it more straightforward to do things like select a subset of museum metadata field that "could provide an appropriate representation of a museum object in the digital library" (39) – emphasizing that not every crosswalk will need to count for every metadata filed in the source record.

>Why does it matter?

Again, if you're this deep in it you probably have a reason for wanting to bestow interoperability upon your metadata (and hopefully it's a more concrete reason than "because Tim Berners-Lee said I have to!" because otherwise you're just going to end up tied in all kinds of knots), BUT as we'll see, interoperability is fundamentally a goal-oriented process, and so it helps to keep those goals in mind.

In libraries, archives, museums, and other information services or cultural heritage institutions, interoperability is most often pursued with the goal of enabling cross-collection searching: the ability to use the same discovery portal for multiple collections, often collections of different domains or collections housed in different organizations.

A related but not-quite-equivalent goal is that of reaching the fullest capabilities of Linked Data. In order for Linked Data and the Semantic Web to work as intended, our information systems need to be able to "talk" to each other, and in order for that to happen, they need to be speaking the same language, or at least on the same page (if we try to continue this metaphor it will fall apart, so let's not). Interoperability brings us closer to being on that page.

"Without achieving interoperability of semantic mapping, application of metadata in the retrieval of networked resources while maintaining optimal recall and precision, will be inefficacious." (Park 74)

>*What is a crosswalk*?

Inevitably when you start talking about interoperability, you'll eventually run into the concept of crosswalks as well. Sometimes this turns up in conjunction with "mapping," and some folks will say that *mapping* is the process of making connections between metadata elements or fields belonging to different schemas, while a *crosswalk* is a visual representation of that process (usually in the form of a chart or table).

However, we've gotten to a point where crosswalk has become a verb unto itself, and is now used mostly interchangeably with mapping.The collapse of these two terms into each other serves as a demonstration of the fact that crosswalking is usually concerned with making direct connections between elements in order to transform records from one schema into another, to the point of using translation terminology like "source" and "target":

"most metadata crosswalks are still focussed on mapping equivalent lexical definitions of metadata elements between source and target metadata standards." (Chen 176)

>*Proposed Methodology*

This methodology for crosswalking, proposed by Chen (2015), is an RDF-based approach that attempts to go beyond establishing lexical equivalency between metadata standards. It consists of five steps:

- Step 1: identification of the multiple objects and their relationships embedded in source metadata elements
- Step 2: selection of the adopted objects and their metadata elements from the source standard – not all are necessary!
- Step 3: identification of the semantic and hierarchical relationships between source and target elements.
- Step 4: identification of the granular and syntactic relationships between source and target elements.
- Step 5: mapping of all equivalent lexical elements from source metadata into the target with reference to RDF-based relationships

This process allows us to maintain "the contextual relationships embedded in multiple objects and their corresponding metadata elements" (Chen 190), and because it largely operates at the data element level (within the semantic layer noted previously), a high degree of granularity is possible.

>Consideration 4: Levels of interoperability

Distinct from the stage of the project, it's also important to consider the level at which you need or want your metadata to be interoperable, within the overarching hierarchical structure of metadata and collections arrangement. These layers can be expressed as follows, in order of decreasing scope (adapted from Zeng and Qin):

- System: interoperability at this layer involves issues of data presentation and compatibilities between hardware and operating systems. System interoperability concerns are largely covered by protocol standards such as those established by W3C.
- Syntactic: interoperability at this layer involves issues of data language, encoding, decoding, and representation of data. Adherence to standards such as XML and RDF fall under this layer.
- Structural: as you might guess, this layer deals with data structures, including frameworks, data models, and schemas. This is where the conceptual and intellectual organization of metadata falls – things like the Dublin Core Conceptual Model and BIBFRAME.
- Semantic: this layer deals with the context of data, or the "communication of coherent purpose" (358). Put more simply, it encompasses the correspondence (or lack thereof) between classification schemes, taxonomies, thesauri, controlled vocabularies, and so on.

To give an example, if you are trying to make the metadata of two different collections interoperable, and they both adhere to difference RDF-based standards, then you already have achieved syntactic interoperability and your efforts will need to focus on structural or syntactic interoperability.

>Consideration 1: Goals

Whether you are a) trying to make sure the metadata you are creating will be interoperable in the future or b) staring down the barrel of a crosswalk and praying that you can connect the dots, the most important thing you should always keep in mind is your end goal. Trying to achieve interoperability for interoperability's sake is not actually an achievable goal because you haven't established what you're trying to make sure your records are interoperable with.

Unlike with adherence to other standards (like "well-formed" XML documents), interoperability is a two-way street, so you need to know which direction you're going before you head out.
This can be articulated more concretely as undertaking "transition planning," which consist of a need for community wide planning, standardization, and transparent communication." (Bigelow and Sparling 81)

>Consideration 2: Source Data

Just as important as knowing where you're going is having a solid understanding of where you're coming from – this means knowing the allllllll ins and outs of your source metadata and the information objects to which it corresponds. Furthermore, it means being able to pinpoint the most essential parts of the metadata so that you can ensure they survive the transformation more or less intact. While our impulse as information professionals is usually going to be to say that every aspect of a metadata record is essential, it's important to remember that this assessment is based on the goals that we discussed previously. The full record may not be strictly necessary for the fulfillment of those goals!

>Consideration 3: Stage of Project

Your approach to making your metadata interoperable can and should be dependent on what stage your digital collections project is in – an established collection undergoing a platform migration is going to have different needs than a collection in the earliest stages of development. One might assume that early interventions are essential for ensuring interoperability, and that attempting to render a large, established collection's metadata interoperable would be an insurmountable task. This is absolutely not the case! It just requires a slightly different mindset and deliberate planning. One way of determining the stage of a project is according to the following lifecycle (adapted from Zeng and Qin):

- Schema development: the bare-bones beginning, where you're trying to establish standards, element sets, application profiles, and so forth
- Record generation: the approximate middle, at this point you either have some records or are in the process of creating new ones, so you may be either transforming records or establishing guidelines for record creation
- Repository operation: relatively late stage, if you've already got a fully functioning repository then you're here. For this stage, interoperability might look like focusing on mapping existing values to enable cross collection searching, rather than complete overhauls or transformations.