2017-02-16 Metadata Management Meeting Notes

Date

16 Feb 2017

- 6:00pm UTC+01
- 5:00pm UTC±00
- 12:00pm UTC-05
- 11:00am UTC-06
- 9:00am UTC-08

Link to Meeting

https://meetings.webex.com/collabs/meetings/join?uuid=MBN9HHHWUB7P0RGHOGF388TIBA-2C6Z

Attendees

- Doreen Herold
- Lisa
- Joseph D. Olivarez
- Laurel Narizny
- Anne L. Highsmith
- Marc Johnson
- D. Bonner
- Lynn Whittenberger
- Peter Murray
- Charlotte Whitt
- Paul Hoffman
- Lisa Furubotten
- Christie Thomas
- Natascha Owens
- Alexandra

Discussion items

Item	Who	Notes	Action Items
Product Council update	Doreen	- Racine from Duke will fill in for Naun as liaison from the Product Council. - Naun will continue to serve on this SIG while at Harvard. - Metadata Management SIG roles have been filled: Alexandra will be our reporter to OLE. Doreen will back her up. Lisa will take minutes. Joseph Olivarez agreed to be the representative for the Metadata SIG on the Folio Forums Facilitator's Group.	☑
Discussing Knowledge Base and Metadata - Scope and Domain	All	The questions below came from Anne Highsmith. They are used here as a launching point for discussion.	☑

Relative to the stated initial goal of "Establish the groundwork for support of data formats beyond MARC … ," information that I have read or heard about FOLIO seems to state that any external record, such as MARC21 or DC or Bibframe, would be translated into a FOLIO-specific data storage format. Is that correct? Additional comment/question: Initial reaction is: why create yet another metadata format? It seems that it would make more sense to facilitate conversion between existing metadata formats rather than producing another one... So is that because there is no desire to build FOLIO on Marc, but there is nothing else as yet that can be used?		Marc Johnson comments: • The document serves three purposes 1) capture a high level summary of discussions around metadata and 2) encourage conversations in this area 3) get into a rough initial plan • We need to support MARC in FOLIO. • We want FOLIO to be able to support eventually resources that are not well expressed in MARC21, such scientific data sets • There needs to be a balance between describing resources without trying to cover all of MARC21 or Bibframe, for example. We don't want every part of FOLIO to have to understand all languages when and where it doesn't need to. We want the parts of the system that need to understand the formats to do so, but then to relieve the systems Doreen asked for reading recommendations that will help the group better understand this document. Anne - Is the intent to translate the records coming in to a single format? Paul - Some of the metadata that is going into FOLIO will need to be edited, some will not. Several possibilities of how FOLIO will work, depending on how data will be edited. Marc • Yes, we would like to translate aspects of the data coming in into something." Not sure how that will work, or how much will need to be translated. • The intention is to keep the original records. • Scenario: The "correct" representation will be the original, but editable. • Trying to consider all scenarios so we are not "backed into a corner" Peter **-** We could have apps that do not need to understand the underlying data format. For example, with circulation, that app does not need to understand MARC in its entirety. It only needs to understand basic descriptive info. Idea is to have circ apps that can be somewhat interchangeable, while allowing for description to Christy - What is the relationship between the source records and stored record? Peter - Distinct metadata formats would have distinct editors. Example: something in BibFrame would have a BibFrame editor, etc. What is "agreed upon between apps" is a high level of metadata. Sometimes this is expressed in Dublin Core. There would be translators in FOLIO that would help apps that understand different metadata formats do what they need to do. Some work was done on this over the summer, the output of which are two Google Documents with abundant comments (Kabalog Discussion and Descriptive Metadata Handling in FOLIO). He would like to clean up these documents to provide a further reference point for this SIG.	☑
Perhaps I'm being premature, but I look at the stated goals, outcomes, and deliverables in the context of existing record conversion, especially bibliographic /holdings/item records. In that context, I find the "physical item" depending from the "entitlement" to be problematic. It implies to me that at the point of conversion we'd need to provide data on how we're entitled to a given physical item; we don't have such data for hundreds of thousands of our physical items because they were bought while we were using a previous system and the purchase data didn't come with them into the current environment. Also, I would see Location connected to Physical Item rather than Entitlement.		Marc - Thinks that a new diagram, in development, that will be shown soon will clarify some of these questions. "Entitlement" is a word that implies a particular commonality between Electronic and Physical items owned by a library: users are "entitled" to use a print book or an electronic book. This is not intended to pull in acquisitions, licensing history. The definition and diagram will be refined. "Location" is an example of what might be a common piece of information among entitled materials. Both Electronic and Physical items have "locations". Entitlements have a common set of attributes. "Location" is one example of a common attribute for entitlements.	☑
On the other hand, I am somewhat reassured by the stated intent of maintaining the storage format of the original records because I don't know what we'd do with the mountain of management metadata in our existing holdings records, as an example. I have an image of them being converted piece by piece as standards develop.		Anne emphasized value of keeping original information in FOLIO when migrating. Marc - Value in distinguishing types of metadata	☑

The typical workflow at this library (Texas A&M) is to create or update a bibliographic record on OCLC, then export it into the local Voyager catalog. Could you describe how that workflow relates to your use of instance, both internal and external? Would the record on OCLC be the "external instance" and the copy of the record within FOLIO the "internal instance"? On the model, the "derived from" arrow goes from the internal instance to the external. Shouldn't there be two arrows, one in each direction? The answer, I guess, depends on whether I have correctly understood the definitions of external and internal instance. If I have, our library and I imagine many others derive records from the external instance and load into internal and almost never derive in the opposite direction.		Terms "external" and "internal" may be causing confusion. Key thing, the arrow and Internal Instance knows about the External Instance it was derived from. An internal instance "remembers" it was derived form an external one. Anne - Is Instance "a copy of a record"? Marc - Instance is meant in the BibFrame definition of the word, i.e. 1st edition of a book, but not the specific copy of a book. Lynn - How can linked data be woven into this diagram? Marc - Attributes can be expressed at the Work level, don't need to be re-expressed at the Instance level often. "Trickling down attributes." Lynn - How would we actually connect the FOLIO connect to an external authority record? How then, does that link to discovery?	☑
Could you please explain "reference cataloging."		Marc - "Reference cataloging" is the notion that the system remembers the external source of the original metadata. With that remembered by the system, that change can get pulled into the local copy of the record. Lynn - Instead of an internal record, would it make sense to only pull locally needed info and the system an call up the external info as needed? Why bring the whole record into your system when you can pull bits of it and "layer" the parts? Marc - Yes, they have been talking about it this way. Doreen - Example: URL can be institution specific and reside at the most granular level, whereas it is often given at the Instance level.	☑
Could you give some more detail about the relationship among the elements connected by the diamond, i.e. instance, package, platform, etc.? I think of packages as typically residing on platforms, for example, so I'd be interested in how you see the relationship.			☐
Regarding teh conceptual domain model, I would like to have clarification on what is meant my Instance - is this in a Bibframe sense or something else? If external /internal instances are sets of metadata either external to the institution or internal, I find the use of the word Instance in this context confusing: surely those sets of metadata would be more than relating to Instances only, but again it all depends what is meant here. If we also have Works, which would presumably be created as part of the FOLIO metadata format, then the external/internal metadata sets encompass more than Instances?			
Wrap up: Additional thoughts, comments, suggestions?	All	Marc and Peter will try to answer these questions online, on discuss.folio.org. Paul - Observation: We would want a distributed storage of bibliographic information, where different pieces can reference each other. Reminder: Link to discussion Knowledge Base and Metadata - Scope and Domain	☐