Hyrax Metadata Ordering

Summary and Evaluation of Proposed Solutions and Solution Recommendation

By the Hyrax Metadata Ordering Working Group (Julie Hardesty, Ryan Johnson, Sadie Roosa, Ryan Wick, Jen Young)

In answer to a question posed regarding ordering multiple fields of the same metadata in Hyrax (creators, for example), locally implemented solutions were offered from the following institutions/working groups:

Durham University (Sufia)
MODS and RDF Descriptive Metadata Subgroup (working group of Samvera Metadata Interest Group)
Penn State University (Sufia)
University of Michigan (Hyrax)
University of Virginia (Sufia 7.4)

See https://groups.google.com/forum/#!topic/samvera-tech/g1oa5B2B0Wk for the original conversation.

Criteria for Ordered Things

The following criteria seem to hold true for the need to order metadata fields describing a digital object:

- Order needs to be applied to the data, not just in a particular view of the data, such as the web application
- The back end storing the data makes a difference in ordering options (Fedora cannot apply an order to properties stored on objects vs Postgres can store an ordered array)

Ordered Things as Objects[1]

Durham, University of Michigan, and UVA create objects for the creator/contributor examples reviewed but do not make use of those objects for ordering purposes. The MODS and RDF Descriptive Metadata Subgroup proposed a solution to only order things when they are objects that can be URI-referenced. The ordering in that case occurs with an additional property containing a delimited text list of those URI values.

Durham wrapped the text values for names with an order number, storing that as the property value in Fedora on that creator or contributor object, and that is then parsed by code to determine display order.

Benefit

- If the order is restricted to things that are objects themselves (and reference-able via URI) then ordering could be kept in sync when changes are made to those objects

Challenge

- None of the implementations have used those objects for actual ordering yet

[1] Some institutions have mentioned a hypothetical use of the hydra-pcdm gem, which contains a function called "ordered members". It is unclear what the status of this solution is, or whether it can be its own self-contained solution to ordering. Nurax shows the ability to order files attached to a Work and this is likely using PCDM but it does not seem as feasible to use PCDM ordering for multiple descriptive properties on a single Work.

- Each Work will need its own descriptive objects to be ordered, raising the likelihood of redundant objects, making updates harder to apply and maintenance more difficult over time

Additional Property with Ordered Things as Text List

All other institutions except Durham provided solutions that involved an additional property with a delimited text list of values to provide order. Except for the MODS to RDF Subgroup's proposed solution, those values were not URIs but were a delimited list of names for the creator examples reviewed. This property is in addition to each creator being its own property as well.

Benefit

- Easy to implement
- Easy to review and understand

Challenge

- Ordered lists that are text strings can become out of sync with values of properties when they are updated or removed, either through normal web form updates or bulk edit changes (regardless of being URI or text values)
- Requires programming to regenerate the delimited text list when any change is made, making this somewhat of a hack to work within the context of RDF and not applicable for data views outside of the web application (batch ingest, Rails console, etc)

Valkyrie

Valkyrie is new middleware between Hyrax and different back ends that will allow Hyrax to work with storage repositories other than Fedora. While no proposed solutions involved Valkyrie, ordered properties are not currently supported by Fedora as a metadata adapter so allowing for other repository back ends that can manage ordered fields (such as Postgres) is potentially helpful. Figgy from Princeton currently has Postgres in place as its digital repository and manages ordered fields using arrays through Postgres. Hyrax 3 is currently scoped to implement Valkyrie to provide for different storage repository options.

Solution Recommendation

The working group does not have a consensus for a recommendation. Some feel that for Hyrax 2, there is no workable solution since doing things the RDF way with Fedora means that there isn't an order or hierarchy involved - RDF is about the relationships between objects and not the order of the data describing the objects. Reality says that order is sometimes necessary. It would seem the best option to achieve order with RDF in Hyrax 2 is to apply a programmatic solution that will use a delimited string from an additional property to indicate the order of a specific set of fields. This will likely only work within the Hyrax web application and will not be applicable in any other views of the data (Rails console, exports). Programming changes will also be required to ensure any change to those fields regenerates that delimited list property. This can be done but is not a good solution to live in the Hyrax codebase. Recommended solutions along these lines can be suggested in the Samvera Community Knowledge Base to help those customizing a local implementation of Hyrax 2.

If Hyrax 3 implements Valkyrie, this might indicate a Hyrax codebase solution is possible but it would depend on Valkyrie supporting ordered properties across multiple metadata adapters and Hyrax choosing or requiring default adapters like Postgres with these capabilities.