

Metadata Support Survey Results Community Message

At the request of the DSpace Committers and Developers, the DSpace Community Advisory Team (<https://wiki.duraspace.org/display/cmtgpp/DSpace+Community+Advisory+Team>) (DCAT) has begun an effort to build a community consensus on improving the metadata support in future DSpace releases. Last month this process began with a community survey to collect feedback on what improvements organizations would like to see relative to metadata support in DSpace.

Survey Result Highlights

- * 84% of respondents use DSpace or plan to use DSpace
- * 86% of respondents indicated that Dublin Core was the most relevant metadata standard for their organization
- * In addition to Dublin Core, respondents saw a need to support ingest and export for the following standards, in order of priority: Dublin Core, MODS, PREMIS, Open URL ContextObject, ETD-MS
- * 72% of respondents indicated it is a priority to add metadata authority controls/vocabularies to the data model
- * 62% of respondents indicated it is a priority to have enhanced metadata for Communities, Collections and/or individual Files (bitstreams)
- * 56% of respondents indicated it is a priority to update to update the Qualified Dublin Core registry to the latest standards of the DCMI.
- * 54% of respondents indicated would not have objections to prohibiting changes/deletions that would break compliance with the standardized default Dublin Core metadata schema

The complete survey results are available here: INSERT LINK

Where do we go from here?

From the survey it is clear that improvements to metadata support in DSpace could help solve some common challenges for many organizations. Many will agree that there is no one size fits all repository software. DSpace is the most popular repository software, used by over 1,000 organizations for a reason – it is relatively easy to get a repository up and running. Because of this the DSpace developers have to be careful that adding functionality it is not at the cost of making it more difficult for the vast majority of DSpace use cases. With that in mind, DCAT has proposed exploring the following priorities with the DSpace Committers/Developers and the user community:

- 1) Add metadata authority controls/vocabularies to the data model. Since there is an existing add-on for controlled vocabularies (<https://wiki.duraspace.org/display/DSpace/Authority+Control+of+Metadata+Values>), DCAT interprets this to mean that we should open up the rights to use a controlled vocabulary and possibly link it from an external source. Some examples would be the National Agricultural Library (linked open data) and the National Library of Medicine (subject based).
- 2) Update the Qualified Dublin Core registry to the current DCMI standards, adopting the newer [DC Terms namespace](http://dublincore.org/documents/dces/) (<http://dublincore.org/documents/dces/>) in the dc namespace. Based on the community survey, DCAT also recommends that the default configuration should separate out standardized DC metadata, administrative metadata and local customizations into distinct metadata schemas. Repository administrators would be prevented from modifying the standardized DC metadata which would effectively break compliance. Instead, local modifications would be accomplished in a separate metadata schema.
- 3) Enhanced the metadata available for Community, Collections and Files (bitstreams). Unlike Items, Collections, Communities and Files (bitstreams) currently do not have Dublin Core metadata associated with them. The community has expressed an interest in making the available descriptions more granular.
- 4) Simplify and make local customizations more accessible by bringing more functionality into the user interface or metadata registry. This could also include a utility to import/export metadata to a schema other than DC. Some community members expressed an interest in exposing the RDF triples. In addition, DCAT recommends a comprehensive tutorial on how/where to perform specific metadata schema tasks be developed.
- 5) Explore the implications of offering hierarchal metadata in DSpace and whether it makes sense, given the DSpace/Fedora integration work. Hierarchal metadata would allow for relationships between community/collection/bitstream metadata. Because Fedora metadata is not stored in the database, as it is in DSpace, but in files, the relational aspect functionality in Fedora is more flexible. It is hoped that the DSpace/Fedora integration work would take advantage of Fedora's flexibility. DCAT proposes that the community explore the implications/reasonableness of a separate effort on hierarchal metadata in addition to the DSpace/Fedora integration.

How can you help?

If any of the above priorities are important to you or your organization, we need your help. In order to move forward we plan to put together a project team from the user community for each priority. If there are no volunteers, there will likely be no meaningful movement forward. Please contact Valorie Hollister at vhollister@duraspacespace.org if you are interested in contributing your time and effort.