DCPAPERS
Proc. Int'l Conf. on Dublin Core and Metadata Applications 2018

Analysis of user-supplied metadata in a health sciences institutional repository

Poster

Joelen Pastva
Galter Health Sciences Library & Learning Center, Northwestern University, USA
joelen.pastva@northwestern.edu

Keywords: metadata; institutional repository; interface design; health sciences

Abstract

Launched in October, 2015 by the Galter Health Sciences Library, the DigitalHub repository is designed to capture and preserve the scholarly outputs of Northwestern Medicine. A major motivation to deposit in the repository is the possibility of improved citations and discovery of resources, especially for non-traditional materials such as poster presentations and teaching resources that are typically never made publicly accessible.

One of the largest barriers hampering discovery is a lack of descriptive metadata. DigitalHub was designed for ease of use for the depositor, requiring very minimal metadata in order to successfully deposit a resource. However, many optional descriptive metadata fields are also made available, some using auto-complete suggestions from controlled vocabularies wherever possible to encourage the consistent and detailed entry of descriptive information. Although the library can deposit materials on behalf of researchers, the repository is largely intended for the self-deposit of items by researchers. In an effort to improve the discoverability of resources deposited in DigitalHub, the Collection Management and Metadata Services department at Galter Library provides metadata enhancement services for all publicly accessible items. However, the library was curious to evaluate how users were approaching available metadata fields and accompanying instructions prior to the performance of enhancement operations.

In order to evaluate user-supplied metadata, an export was made of all of the metadata in DigitalHub for a 2.5 year period. Records previously enhanced by librarians, or records initially deposited by library staff were excluded from primary consideration. The metadata was then evaluated for completeness, choice of dropdown terms for resource type, inclusion of collaborators, use of controlled vocabulary fields, and any areas that indicated a clear misunderstanding of the intended use of the metadata field. This poster presents the preliminary findings of this analysis of user-supplied metadata.

Although all fields were used appropriately by depositors, over half of all optional metadata fields were left blank, with another 25% of optional fields underutilized. It was especially interesting to observe no use of the Contributor field, although depositors did often record multiple authors. 38% of depositors used a filename for a resource title, which is supplied by the repository by default upon deposit. Depositors were comfortable supplying their own keyword tags, but never utilized auto-suggested controlled vocabulary terms such as LCSH or MeSH for indexing. Despite a rich offering of nearly 160 resource types to accommodate different outputs, only 17 unique resource types were selected by depositors over 72 individual deposits.

It is hoped that the findings of this analysis will help guide future system and interface design decisions, cleanup activities, and library instruction activities. The lack of complete metadata supplied by depositors indicates the continued need for library metadata enhancement for improved discovery. There are also opportunities for the system to pre-populate fields that tend to

https://doi.org/10.23106/dcmi.952139061

be standardized across all records to improve the richness of resource description upon deposit. Ultimately the goal is to make the interface as usable and effective as possible to encourage depositors to supply an optimal amount of descriptive metadata upfront, and to continue using the repository in the future. These results should be of interest to repository managers that rely on users to supply initial descriptive metadata, especially for health sciences disciplines.

https://doi.org/10.23106/dcmi.952139061