Harvesting Quality: Evaluating Metadata for Digital Collections

Paromita Biswas

Hunter Library, Western Carolina University, NC 28723

Digital Library Projects and Quality Control

Context: Digital library projects often deal with multiple metadata creators, new formats, and dynamic metadata standards.

Result: Achieving accuracy and consistency in metadata records can be challenging.

Remedy: Digital repositories rely on periodic sampling or peer review of metadata records as a means of quality control.

Poster: Harvesting tools can provide a quick and effective means of running quality checks on digital projects' metadata.

Materials and Methods:

- WorldCat Digital Collection Gateway tool: Frequency Analysis feature
- Hunter Library's harvested digital collection: *Picturing Appalachia*
- Dublin Core metadata map for the digital collection
- Hunter Library's metadata schema

Hunter Library's digital collections are described using Dublin Core in Contentdm. The library has recently started harvesting its collections into WorldCat via the Gateway tool. The Frequency Analysis feature in Gateway showed usage of "contributor" field not being in compliance with the library's metadata schema.
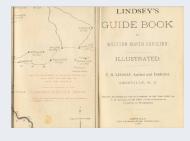
Harvested Collection: *Picturing Appalachia*

Metadata Map in Gateway

Frequency Analysis Tool: Contributor

An item in *Picturing Appalachia*

The item record in Dublin Core in Contentdm

The printing company is recorded as "contributor" in the item record; however the library's metadata schema specified only editors and illustrators as "contributors."

Results and Conclusions

- Updating of metadata schema: Redefinition of "contributor" field in metadata schema to incorporate contributors' plethora of roles.
- Updating schema has promoted consistency in recording this field. Possibly also benefitted users searching by multiple names related to an item.
- FA feature can be used to run quality checks on all fields:

Should "Sherill's Art Store" and "Southern Post Card Co. (Asheville, N.C.)" –both corporate bodies– been recorded as "contributors" instead of "creators"?

For more information email:
pbiswas@email.wcu.edu