ABSTRACT

Objectives: To determine the potential value of importing user-contributed data from LibraryThing into a large academic library catalog, by analyzing the overlap between the two databases, and the quality and quantity of metadata available from each source.

Methods: A random integer generating program was used, to create a list of non-repeating indexes from 1 to 28,000,000; 20 records were selected from each of 15 randomly chosen intervals of 1,000,000 relative database size (as of 1 February 2008) to test LibraryThing. To determine if each record had a matching record in the MNCAT, a random sample of 80 records were selected from LibraryThing. To determine if each record had a matching record in the MNCAT, a random sample of 80 records was selected from LibraryThing. The characteristics of matched and non-matched records were examined. For this project, metadata analysis was carried out by the user, with the results being compared to the MNCAT database.

RESULTS

Comparison of subject metadata

- Format
  - LCSH/MeSH—controlled vocabulary consisting of a main term alone, or with a combination of subheadings identifying further topical, chronological, geographic, form or genre subdivisions.
  - LibraryThing tags—single free-text terms contributed by users.
- In order to make a meaningful comparison, LCSH/MeSH terms were “faceted”, i.e., broken down into their component parts and counted separately. Each unique facet was counted only once per record.
- Assessment of terms
  - The most popular tag from LibraryThing was compared with the LCSH/MeSH terms for the same record. 28 remaining titles lacked subject terms in one or both sources and so could not be compared.
- Subject metadata for literary works were compared.

- No attempt to evaluate accuracy from either source.

- Though 15 MNCAT records and 12 LibraryThing records had no subject descriptors, only one record had no subject metadata from either source.
- For 25 titles, the “most popular” LibraryThing tag roughly matched with the LCSH/MeSH heading for the same record. 27 titles did not have a corresponding close match. The 28 remaining titles lacked subject terms in one or both sources and so could not be compared.
- When “personal” tags are eliminated, LibraryThing has 584 tags, or an average of 7.3 per record. However, some personal tags might be useful to some users.

- Due to the small sample size, further research is needed to draw strong conclusions.

CONCLUSIONS

While overall LibraryThing is not a strong source of enhanced metadata for large academic libraries, for recent, popular materials and literary works the user-contributed metadata enhances and complements what is found in a traditional catalog record. For 22 records in sample were literary works. 28 LCSH terms were provided, or an average of 7.97 per record. 281 user tags operated on the LibraryThing records (13.32 per record). The most popular tag from LibraryThing was compared with the LCSH/MeSH terms for the same record. 28 remaining titles lacked subject terms in one or both sources and so could not be compared.

- Due to the small sample size, further research is needed to draw strong conclusions.

Examples of “Personal” LibraryThing Tags

- For additional information please contact Christine DeZelar-Tiedman Technical Services Department University of Minnesota Libraries ddez002@umn.edu