

What Does Transaction Log Data Tell About Collection-Level Subject Access?

Oksana L. Zavalina

Assistant Professor

Department of Library and Information Sciences

College of Information, University of North Texas

Problems in subject access to information have been under investigation in the library and information science field for a long time. In the 1980s and 1990s, focusing on item-level information seeking behavior and information discovery in library catalogs and databases, researchers identified a range of problems the users experience with subject access, including quality and application of metadata. While a growing number of aggregations of digital collections are being built in the United States and abroad, there has been little consistency in application of collection-level metadata that describes collections as a whole across different aggregations. This is not surprising given that collection-level information seeking behavior has not been researched and the value of providing collection metadata has not been established and supported by the objective data. Does collection-level information seeking behavior differ from item-level information seeking behavior in aggregations? To what extent do the users explore aggregations by subject? How does collection-level metadata help the users of aggregations find digital objects and collections of objects through searching? How rich should collection-level metadata be to help meet the user information needs? Analysis of transaction log data has a lot to offer in answering these questions.

This poster reports results of the quantitative and qualitative study of a systematic sample of transaction log data that recorded user interactions with a large-scale aggregation of cultural heritage digital collections¹ over a period of one year. Analysis of user interactions with an aggregation yields a number of interesting findings. A high level of engagement with collection metadata records, with the total page views more than 4 times greater than item metadata page views, was observed. Search and browse at collection level were found to occur almost as often as at the item level. Both at collection and item level, browse is a prevailing type of user interactions with an aggregation of digital collections. Subject browse (topical and geographic) was used more often than other types of browse. Functional Requirements for Bibliographic References (FRBR) model set of entities was used as a framework for categorizing user searches in aggregation. Majority of collection-level search queries fell within FRBR Group 3 (or subject entities) categories – *object*, *concept*, and *place* – which allows to conclude that subject exploration through searching is an important function to be supported by aggregation. Some differences were observed between collection-level and item-level searching: significantly more *object*, *concept*, and *corporate body* searches and less *individual person*, *event* and *class of persons* searches in collection-level searches than in item-level searches. Analysis also shows that collection-level search queries are most often satisfied by *Description* and/or *Subjects* collection metadata fields and would fail to retrieve a significant proportion of collection records without controlled-vocabulary subject metadata (*Temporal Coverage*, *Geographic Coverage*, *Subjects*, and *Objects* fields), and free-text metadata (the

¹ Opening History (<http://imlsdcc.grainger.uiuc.edu/history/>) national-scope aggregation of digital collections with a focus on the United States history is developed by the IMLS-funded Digital Collections and Content research and demonstration project (<http://imlsdcc.grainger.uiuc.edu/about.asp>) at the University of Illinois at Urbana-Champaign.

Description field). The findings of this study extend the understanding of the information seeking behavior in aggregations and the value of collection-level subject metadata in the user interactions with aggregations of digital collections and indicate great importance of applying a variety of subject-bearing collection metadata fields in describing digital collections to facilitate subject access in aggregations.