Chapter 3
Content Analysis of CONTENTdm Item Level Records and Their Aggregated Metadata in WorldCat

Iris Lee
American Museum of Natural History, USA

Mary Tyson
Whitney Museum of American Art, USA

ABSTRACT
This study examines the content in CONTENTdm item level records and their aggregated metadata in WorldCat. The research ascertains some amount of loss of contextualization in semantic meaning of digitized primary resources. Data was collected from non-probability sampling of CONTENTdm item records selected randomly from CONTENTdm’s Collection of Collections Web page. The same aggregated digitized resources were retrieved in WorldCat, and the data was recorded and analyzed against the CONTENTdm records. Evidence shows that the value of the metadata is altered or ambiguous when the local field name is divorced from the descriptive metadata. Contextual meaning is lost when metadata is shared. In fact, the study reveals that the de-contextualization problem is two-fold: some inconsistent mapping from aggregated metadata and lack of clarity in the metadata itself.

INTRODUCTION
The purpose of this study is to understand how interoperable metadata is shared and used through gathering evidence of problematic mapping. More illumination on issues that arise from using Dublin Core (DC) as the schema to describe archival collections has the potential to help managers and data providers. This study examines records made from OCLC’s digital content management system, CONTENTdm. We made a content analysis of 241 unique CONTENTdm collections. Data was...
gathered from one randomly sampled item-level record from each of the 241 collections to determine how metadata descriptions help or hinder their access on the World Wide Web via WorldCat.

The justification for this exploration is to evaluate how data providers customize and use DC elements to describe their digitized material. Because CONTENTdm items can also be easily published to WorldCat using Digital Collection Gateway, the resources become more visible to potential researchers. An important change is noted that the platform on which the metadata is presented is very different. It is important to understand how the descriptive information is read when mapped to another cataloging format, in this case, Machine-Readable Cataloging (MARC).

Tested metadata crosswalks are available for transforming DC records to MARC. However, the DC element set does not always meet data providers customized needs for special collection and archival descriptions. They then create semantically appropriate local field names. Even when crosswalks are applied, meaningful descriptors are dissociated, which render the metadata incomplete or ambiguous. Good metadata is shareable. Shareable means that a cultural object projects metadata that is aggregated in a meaningful and contextual way (Shreeves, Riley, & Milewicz, 2009). Descriptions for item-level records should be read and understood completely from one catalog to the next.

Libraries and archives seek to cast a wide net on the World Wide Web for increased accessibility of their primary source material. As Riley and Shepherd (2009) say, “The concept that records exist to be used is a core archival value” (p. 94). This study aims at observing the accessible view of CONTENTdm records aggregated in WorldCat. As Greenberg (2011) says now there is a concept of the macroscope: large views of things from potentially aggregating vast amounts of available data. This study points out where there are pitfalls in aggregating today and the importance of mending them.

**BACKGROUND AND LITERATURE REVIEW**

A full color giga-pixel photograph is a much richer visual experience than a high-contrast black and white copy. As Coyle (2005) says, “… metadata is not the world, it is how we see the world at some moment in time for some purpose” (p. 160). The full information in a record is a surrogate for the real thing, and as such, metadata is a surrogate for the real thing found online (Coyle, 2005). However, when at one time, a user would sit in the library and look up resources in the ILS, now many kinds of users, from many locations, with many different purposes, from people to crawlers, all access metadata through the World Wide Web.

Research on metadata spans interoperability to the implementation of standards with quality and shareability at the center. Howe (2011) said about quality that to have good metadata you must make good metadata meaning automatically generated metadata will not suffice. Metadata that is shareable means it can be mapped to schemas such as DC and others. Riley and Shepherd (2009), as well as Suleman (2001), describe shareable as a known concept distinct from local. It is with this shareable structure that metadata can be mashed up outside of relational mapping (included in commercial sources such as Google, Wikipedia, the Atlas Initiative).

**The Challenge of Metadata Creation**

What seems challenging for many metadata creators is to comprehend the value of using standards. Librarians have long been trained to adhere to cataloging rules and standard practice. The concept of metadata standards is to help search functionality work better. According to Reese (2009), the lack of national standard or shared best practices has contributed to a high level of inconsistency in the metadata currently being produced. This makes capturing subtle relationships expressed within the metadata difficult.
Related Content

Core Topics in Digital Library Education
[www.igi-global.com/chapter/core-topics-digital-library-education/19914?camid=4v1a](www.igi-global.com/chapter/core-topics-digital-library-education/19914?camid=4v1a)

Nested Partitions Properties for Spatial Content Image Retrieval
[www.igi-global.com/article/nested-partitions-properties-spatial-content/45736?camid=4v1a](www.igi-global.com/article/nested-partitions-properties-spatial-content/45736?camid=4v1a)

First Person Singular: A Digital Library Collection that Helps Second Language Learners Express Themselves
[www.igi-global.com/article/first-person-singular/39035?camid=4v1a](www.igi-global.com/article/first-person-singular/39035?camid=4v1a)

Digital Libraries in Africa: Evolution, Status, and Challenges
[www.igi-global.com/article/digital-libraries-africa/73645?camid=4v1a](www.igi-global.com/article/digital-libraries-africa/73645?camid=4v1a)