ƒ i ® s † m ¤ ñ d @ ¥
PEER-REVIEWED JOURNAL ON THE INTERNET

Examining MARC Records as Artifacts That Reflect Metadata Utilization Decisions

by William E. Moen

The Institute of Museum and Library Services (IMLS) awarded a National Leadership Grant to the Texas Center for Digital Knowledge at the University of North Texas to investigate the coding of information and metadata utilization in a large set of machine–readable catalog (MARC) records. The project, *Examining Present Practices to Inform Future Metadata Use: An Empirical Analysis of MARC Content Designation Utilization*, is investigating the extent of catalogers' use of MARC 21, the markup language used by catalogers worldwide to create catalog records. We have named this research study, the MARC Content Designation Utilization (MCDU) Project. Information about the project is available at: http://www.mcdu.unt.edu.

The creation of metadata to represent information objects (whether digital or analog) results in records that reflect a set of decisions by metadata creators (human or machine). These records can be considered artifacts of the metadata creation process. Dictionary definitions of the term "artifact" commonly describe an artifact as something created by humans for some practical or utilitarian purpose. We assert that metadata records can be characterized as an artifact of the metadata creation process. This allows us to understand or infer various things about the decisions made in the course of metadata creation. Therefore, we further assert that investigating these artifacts is important to understand significant features of the metadata creation enterprise.

Bibliographic records are instances of metadata records. Library catalogers have a long history of concepts and practices related to the creation of bibliographic records that describe items in a library's collection. The records are put into library catalogs to assist users in identifying, finding, selecting, and accessing relevant information objects. MARC bibliographic records encode bibliographic data, data created by both catalogers and systems.

Library catalogers apply rules and standards such as the *Anglo–American Cataloguing Rules* and local cataloging policies in the process of creating the bibliographic data. They then encode those data into a MARC record using input templates provided by a cataloging module in the integrated library system. In addition, the cataloging module may insert data in the MARC record automatically (*i.e.*, machine–generated metadata). The resulting MARC bibliographic record can be viewed as an artifact of the entire cataloging enterprise, an enterprise that encompasses decisions by library professionals, their organizations, and the automated systems they use. Through methodologies developed for the MCDU Project, we are able to interrogate the artifact and through that interrogation, identify and reveal aspects of the cataloging enterprise related to metadata utilization.

The overarching research question addressed by the MCDU Project is: What is the extent of catalogers' use of content designation available in MARC 21? A previous analysis on a smaller set of records indicated that library catalogers commonly use only a small set of the available elements in a rich metadata scheme (Moen and Benardino, 2003). In the MCDU Project, we are systematically examining over 56 million metadata records from OCLC's WorldCat bibliographic database to arrive at conclusions about utilization in a particular metadata creation practice, namely library cataloging. The research is underway, and a first set of results have been compiled and are available on the project Web site.

Empirical investigations into metadata records as artifacts of the metadata creation process can provide insights into the process. Further, operational decisions about utilizing metadata schemes can be informed by such empirical investigations. In addition, being able to characterize repositories of metadata records using measures of metadata utilization can assist aggregators when working with such repositories. ▪M

About the author

Dr. William Moen has been a faculty member at the School of Library and Information Sciences, University of North Texas, since 1996. He currently serves as the Interim Director of the Texas Center for Digital Knowledge, and Associate Professor in the School. Dr. Moen was

the recipient of the 2005 Frederick G. Kilgour Award for Research in Library and Information Technology.

Acknowledgements

Reference

W.E. Moen and P. Benardino, 2003. "Assessing Metadata Utilization: An Analysis of MARC Content Designation Use," In: *2003 Dublin Core Conference: Supporting Communities of Discourse and Practice — Metadata Research and Application*, Seattle, Wash., at http://www.unt.edu/wmoen/publications/MARCPaper_Final2003.pdf, accessed 5 June 2006.

Editorial history