5-1-2007

The Dilemma of Access: Describing Open Access Journals with MARC and [Other] Metadata Schemes, Summary of a Presentation by Monica Berger and Gloria Rohman

Beth Evans
Brooklyn College

Follow this and additional works at: http://academicworks.cuny.edu/ulj

Part of the Library and Information Science Commons

The Dilemma of Access: Describing Open Access Journals with MARC and [Other] Metadata Schemes, Summary of a Presentation by Monica Berger, New York City College of Technology, CUNY and Gloria Rohman, New York University

Monica Berger and Gloria Rohmann bring to the open access discussion the librarian's point of view, both from the standpoint of the cataloger (Berger) and of the position of the public access librarian (Rohmann).

The speakers define what content is likely to be found in open access journals. Some OA journals offer published material, in its final form. The publisher may offer all content, for all years of the publication, freely. Some, as in the case of Highwire, may embargo selected years. Other journals include pre-print material, or articles that continue to see the light of day as they go through two or more stages of revision. Journals that publish pre-prints are often hybrid journals that contain finished articles as well. Pre-prints force us as librarians to look at the issue of versioning, or, determining which is the correct or final or desired version of a work. Unpublished material, such as what might appear on an author's web site but not in a journal, (e.g., an unpublished conference paper or the invisible college, as Rohmann calls it, i.e., the informal communication between scholars, fill out the arena of open access, scholarly communication. Typically, though, in the open access discussion, the content is offered through a commercial vendor and select parts are offered freely.

When authors archive their own material, in an ideal world, they will adhere to the standard, the open archives initiative protocol for metadata harvesting to enhance findability. The Sherpa Project (http://www.sherpa.ac.uk/) was instrumental in achieving this goal – of using OAI-PMH -- for a number of institutions developing local repositories. OAIster (http://www.oaister.org/), a union catalog of digital resources, makes use of the same protocol and includes a significant number of non-article materials.

The important consideration for both technical and public services librarians is to learn whether "open access content will find its way into the OPAC and other bibliographic tools and [whether] bibliographic utilities [will] continue to function successfully as unique catalogs… [W]ill all this Web-born content bypass our world?" (Berger) Literature is falling from the hands and control of librarians. The Directory of Open Access Journals (DOAJ) is a good example of how portions of the available literature are not passing through the hands of librarians before coming into the public eye. Many titles are missing from the DOAJ. Librarians are not putting locally produced e-title publications into OCLC. The catalog itself may be endangered. Nevertheless, a good sign is that the Library of Congress is using

MARCXML, which is a version of XML that very neatly maps to MARC. Other schemes include MODS, which is essentially a smaller version of the same thing, but it is not equivalent to MARC, and MADS, or, metadata authority description scheme. Another good sign is OCLC's loading bibliographic references into Google. These will lead people back to the OPAC. Google Scholar is useful in that it clusters things and link resolvers will lead users to the library's catalog. The time may come to reverse the role of the library with the OPAC. The goal will be to map from an indexing and abstracting service to your holdings rather than mapping your holdings to your abstracting services. Consider the example of Cornell. CU did a project where they took XML metadata and mapped it back to the catalog. The University of Illinois developed software that lets translates MARC to different metadata schema and back. Currently, libraries seem happy enough to get MARC records tape-loaded or dumped and not concern themselves too much with quality.

Another channel for finding open access material is the DOI ▢ digital object identifier. The problem with DOI is that not many titles are registered with DOI, and these are a minority of those journals, particularly outside the biomed or physics and math areas. DOI establishes a persistent link to a digital object, and provides a container which can accommodate any existing identifier. Existing identifier meaning ISBN, ISSN, or simply some kind of abbreviated name of the journal and the date of the issue. DOIs are no good unless you can resolve them to actually get an article. And DOI resolution is provided by the first part of that URL that you see http://doi.org. The journal URL can change, but the DOI stays the same.

FRBRization is a cataloging process where items that it actually considers are grouped. FRBR was discussed by IFLA in 1997. The whole point behind it is to show relationships and hierarchy. FRBRization has four levels of specificity, the item, manifestation, expression and work. When an item is the most concrete, the work is then abstracted. FRBR will improve how catalogs work. It's going to provide this hierarchical, easy to digest display. We can't catalog individual titles at that kind of vast level. And then we have different physical and intellectual versions of same article. We want to connect the different levels of the entities, ---- the supplements, and show title change history.

As technology develops, consolidation will reduce the number of standards. Interoperability will remain key. The current mix of standards don't necessarily match each other, but they match to each other at least decently so you can translate things one to the other.

Rohmann's concluding comments sum up the discussion. She sees the OPAC developing in number of different ways, FRBR emerging as important and greater interoperability coming with XML. ERMs, electronic resource management systems, will become integrated with interlibrary systems. Standards and adding identifiers

will become more cohesive. Most importantly, though, user behaviors on the Web will change the way librarians work.