VirginiaTech Invent the Future | Digital Library and Archives | University Libraries
Search Addison
Off Campus Sign in
Ask a librarian
Search EJournals GO
Advanced Search
Electronic Journals
About DLA | Collections | Copyright | ETDs: Theses and Dissertations | EJournals | News | VT ImageBase

Virginia Libraries

Main Contact:
Lisa R. Varga vla.lisav@cox.net

January/February/March, 2004 **Volume 50, Number 1**

DLA Ejournal Home | VALib Home | Table of Contents for this issue | Search VALib and other ejournals

The Future of MARC: R.I.P. or Let Her Rip?

by Mary Finn

The 39th Annual Potomac Technical Processing Librarians Meeting—"The Future of MARC: R.I.P. or Let Her Rip?"—took place in Washington D.C. on October 31, 2003, with speakers Roy Tennant and Rebecca Guenther. Tennant manages eScholarship Web & Services Design for the California Digital Library of the University of California, while Guenther serves as Senior Networking and Standards Specialist for the Network Development and MARC Standards Office of the Library of Congress. Tennant's articles "MARC Must Die" and "MARC Exit Strategies," which appeared in *Library Journal* in late 2002, led to a national discussion of metadata issues. Guenther countered with a letter to the editor, "MARC: Not Dead Yet," published in LJ in January 2003.

At the meeting, Tennant argued that although the development of the Machine-Readable Cataloging record (MARC) was a "crowning achievement" of the library world, at its base it is an inventory control system. He feels that MARC is anachronistic and deficient, considering our bibliographic description needs today. MARC was introduced in the late 1960's, when computing power and storage were expensive, scarce, and unwieldy. Because of developments such as inexpensive and compact disk space, major changes in software systems and new technologies, we can do better today.

Tennant suggests that we develop new encoding standards. Rather than continue to adapt MARC, he believes we would be better off starting from scratch and devising a new system as we did over thirty years ago. We need more than a format or rules; we need an actual infrastructure that is so versatile it can deal with many different kinds of formats at once. "We

need to be capable of ingesting, if you will—importing, merging—different formats. We need an infrastructure that can take any arbitrary chunk of metadata and do something useful with it. We need extensibility because as time goes on, we are going to need to be able to extend it." He proposes that we use Functional Requirements for Bibliographic Records (FRBR) as a base from which to devise an encoding standard that provides us with more power and flexibility. The information science world is already moving to XML, which allows the hierarchical expression of information and uses an alphabetic rather than numeric syntax. An alphabetic syntax would make bibliographic records easier to navigate and metadata more understandable and accessible to a greater number of people. Further, metadata schemata that express hierarchical relationships (such as different editions or printings of the same title) would be of great use to library patrons.

Important also, said Tennant, is openness and transparency so that the new infrastructure will allow people at different institutions to create their own code and whatever tools they may need to do their jobs. The components of the system should be modular, providing the ability to plug in other components. Dealing with hierarchy is important. The flat record structure of MARC isn't appropriate for many objects. For instance, the table of contents of a book should be a higher object than the book itself in terms of coding and display. Granularity is the next important attribute of the system, necessary for coding and separating such information as last name, first name or date.

Tennant said that the Metadata Encoding and Transmission Schema (METS) enables us to do this very thing. METS can be used as a "wrapper for different metadata packages. But it can also encode the structure of a file or the components of an object so you can keep track of the fact that you have a Word version, a PDF version, and an XML version, know where they live, and be able to reference them."

Rebecca Guenther is one of the architects of METS. She agreed with Tennant on many points, but not on the supposition that MARC will die, either by murder or natural causes. As she noted, "In the environment of XML, it's easy to combine elements from different schemes." In MARC 21, the MARC relaters code list provides terms and codes to relate an author to a work, expressing that author's specific creative role. In Dublin Core (DC), that role can also be expressed. Guenther pointed out that both MARC and XML are evolving. "We need to take advantage of what XML has to offer and establish a standard MARC 21 representation in an XML structure, because XML is really the new language of the web...it's gaining a lot of ground and there are a lot of tools out there that use XML." Guenther said that MARC is going to XML too. MARC is possibly the earliest descriptive metadata standard, yet it remains viable—many national formats have converted to MARC 21. Many bibliographic utilities with widespread use have MARC coding underneath. MARC is also a standard communication format with predictable content, thus enabling a high degree of record sharing.

Investigating a more flexible syntax for the MARC element set, the Library of Congress created the Metadata Object Description Schema (MODS). Derived from MARC, MODS is a bibliographic element set that uses XML. It has the ability to provide rich descriptive metadata that is simpler than full MARC, but is especially useful for those complex digital library objects with many joined parts. Its level of descriptive ability is midway between the full MARC XML and the DC XML, which is sparser. In a transfer from MARC records to DC, a lot of information is lost. The element set of MODS is richer than DC's fifteen elements. It takes advantage of XML's ability to express hierarchy, and thus allows for very complex descriptive objects. MODS is not intended to be a MARC replacement. It was developed for electronic resources, and does have some elements that MARC doesn't, such as digital origin, which offers two values, digitally born or digitally reformatted. One can't transfer whole data from MODS to MARC and back again, but the formats are still closely aligned. Though XML is more hierarchical than MODS, the latter does have some hierarchy.

MARC is a number of things, including a markup language, in that it is a data element set, like HTML or XML. It has semantics or meaning, defined mainly in the MARC Standard and MARC documentation. There is content, primarily defined by other standards outside MARC. "And then there is structure, which is the syntax for communicating records, and that's often what people say needs to move or die, because the syntax was written in 1968." However, because MARC has been evolving and is so heavily used around the world, Guenther believes it is likely to survive.

Both Tennant and Guenther agreed that there will be transfer schemata offering communication between libraries and other organizations. These will likely be XML-encoded, since that is now the language of the web. Any such schema should be capable of containing intact packages of metadata, so that records from different schemata can live together. As Tennant said, "You may have a MARC record in there. Next to it you might have a Dublin Core record, next to it you might have an ONIX record, and so on." XML can "wrap those and keep them intact."

Mary Finn serves as Catalog Librarian at Virginia Tech. She can be reached at maryfinn@vt.edu.

This work is licensed under a Creative Commons Attribution-Noncommercial-Share Alike 3.0 United States License.

URL: http://scholar.lib.vt.edu/ejournals/VALib/v50_n1/finn.html
Last modified: 09/01/14 16:13:11 by KMB