Metadata: What's all the fuss about?

By Jody Perkins, Metadata Librarian, Miami University Libraries

So what is it? Metadata is structured data that tells you something about a resource. All OPAC catalog records contain metadata. However, what we are commonly referring to when we talk about metadata is data about digital resources contained in nonMARC formats, typically XML (extensible markup language) or XML like formats.

Why is it important? It's fundamentally about access, discovery and preservation – making the greatest number of materials available to the largest number of people over the longest period of time. Metadata facilitates discovery in the same way that traditional cataloging does – by supporting collocation (bringing similar items together), disambiguation (supplying distinguishing details), and providing context and access or location information (typically a link to a digital object in this case).

Use of open architectures and related standards supports accessibility by eliminating the technical barriers of traditional proprietary database bound collections. Making digital collections available in open standard or open standard compatible formats, facilitates cross-domain, cross-system searching and harvesting. There is one caveat however, the metadata that accompanies these collections needs to be developed with an eye toward interoperability. If you want your collections to have increased access and longevity you can't create your metadata in a vacuum.

Metadata also supports the preservation of digital resources. Without metadata that includes technical details about a resource (e.g. file formats, hardware and software requirements etc.) it may not be possible to maintain a resource over time through multiple system migrations and changes in technology.

What's the deal with standards? Like anything that requires a certain level of compatibility, metadata is standards driven. There are four layers of standards that support metadata: data structure or schema standards (e.g. Dublin Core, VRA Core), data communication or encoding standards (e.g. MARC, XML), data content standards (e.g. AACR2, VRA-CCO) and data value standards (e.g. LCSH, AAT). Data structure and data communication standards are not necessarily mutually exclusive and may be combined into a single standard (e.g. MARC, EAD and TEI).

The variety of metadata standards reflect the variety of resource formats and communities of practice that create and use metadata. The ultimate goal of all standards creating bodies is interoperability. In the case of metadata it often requires some kind of compromise between the needs of specific types of resources or implementations and the more universal requirements needed to achieve interoperability.

One standard that is an exception to this is Dublin Core, which has come to be a nearly universal standard. It is often the bench mark standard that provides the

Useful Metadata Links

• Dublin Core (DC) - http://www.dublincore.org • Encoded Archival Description (EAD) http://www.loc.gov/ead/

• Extensible Markup Language (XML) http://www.w3.org/XML/

• Text Encoding Initiative (TEI) http://www.tei-c.org

Metadata continued...

required level of interoperability which others either adopt, refine and extend or map to in one way or another. It can function directly as either a base schema or as a format for universal exchange.

Why do I need to know about it? Developments in metadata are changing the way we create, access and use resources. Even if you are not actively involved in building digital collections or creating metadata, as a librarian it's important to have some understanding of such trends in order to make the most of digital resources - not simply to access them but also to harvest, deconstruct, reconstruct, reorganize, reuse and create new services around them.