

A New World of Data

Populating the Semantic Web

by Karen Coyle

The world today is clearly not that of our library predecessors, of Melvil Dewey and Charles Ammi Cutter, not even of Seymour Lubetzky or Michael Gorman. The changes that have taken place since the introduction of the personal computer and the globalization of communication over the World Wide Web are huge, and they affect in particular anyone involved in knowledge re-



As information creation moves to the cloud, so will library services.

search and creation. Libraries have formed trust relationships with information providers—proof that linking does not have to be entirely uncontrolled or open. And finally, we are already seeing the advantages of moving discovery beyond the library catalog to other environments where the user is searching and interacting.

This evolution of library catalogs is like a dress rehearsal for moving library data from its storage silo in library systems and databases to the web of linked data.

Through the Semantic Web, library data will link to select other data sources in order to provide more value and services for users. Conversely, other users and resources will be able to link to library data, thus making library data discoverable from a variety of points in web space. As information creation moves to the cloud, so will library services, not because libraries create their own cloud but because there will be no separation between libraries and the web.

The Semantic Web

Like the World Wide Web, the Semantic Web is about linking, but it adds the linking of data, not just documents. It also changes the nature of the link: Whereas the link between documents has no meaning other than “link,” in the Semantic Web, the links themselves have a specific meaning. Consider this ci-

tation example: In a standard document, a citation is simply a number in the text and a bibliographic citation at the end of the page. You don’t know why the author is citing that work other than what you can glean from the surrounding text. Using the richly semantic links of the Semantic Web, you could characterize each citation with a meaning such as “cites as evidence,” “disagrees with,” or “extends.” (Those examples are from an actual Semantic Web vocabulary, CiTO, or Citation Typing Ontology.)

The Semantic Web will develop in two ways: First, by linking information that exists within documents, and second, by making the data itself accessible on the web. The ability to mark up information in documents could allow smarter access to that information than we get with keyword searching. For example, markup could identify the author of a document so that an author search could be done, something search engines do not provide today.

The second method of populating the Semantic Web is that of adding actual data sets to the web, as represented by the growing linked data cloud (see linkeddata.org). It is of most interest at the moment to libraries because the library catalog itself qualifies as data that can become part of the linked data cloud. ■

KAREN COYLE, library technology consultant, wrote *Linked Data Tools: Connecting on the Web* in the May/June issue of *Library Technology Reports*.

Linking and federating

What if you extrapolate from developments within library systems, such as federated searching, enhanced catalogs, and OpenURL, to the idea of libraries on the web? Through federated searching, we have learned how to create displays that combine different types of data, which our users navigate without great difficulty. Next, we know that we can enhance the user experience by linking out to select web-based resources. These resources may not be 100% reliable, but the risk is