2007

People Profile: Sheila Bair

Editor
Cataloging/Metadata Librarian at The National Science Digital Library further defines metadata as “structured, standardized descriptions of resources, whether digital or physical, that aid in the discovery, retrieval and use of those resources.” Such information about information becomes increasingly valuable in a knowledge economy: the faster and more efficiently you can get the information you want or need, the less effort you waste — and the smarter and richer you become.

Exploding by the nanosecond, information threatens to overwhelm the frail bark of our capacity to order it. But information seekers are seldom cognizant of the metadata behind the database. For example:

- A book jacket image appears in the record because an ISBN was recorded in the metadata;
- A needed book can be borrowed from a library in Beijing because MARC records enable sharing of records in an international electronic union catalog;
- All the resources in a discipline or subject area can be perused because Library of Congress subject headings and classification have been added to the metadata;
- All the works by a favorite author can be instantly called up because a standard authorized heading was used for the author’s name.

Metadata effectuates connectivity, interoperability, searchability, accessibility, and findability.

**Types of Metadata**

Whatever its purpose, a database runs on quality, standardized metadata, which comes in a number of types. Descriptive metadata aids in the discovery, identification, evaluation, collocation, and selection of resources. Technical metadata describes information about creation and revision of digital objects, including resolution, compression, and pixel dimensions — information that may be needed later for migration. Structural metadata defines the relationships between multiple digital files. As it “relates the pieces of a compound object together,” it can synchronize audio with text or facilitate navigation through an eBook. Administrative metadata, finally, facilitates management of information resources and records information about provenance, history, ownership, and intellectual property rights.

**Components of Metadata**

We further characterize metadata in terms of three main components: syntax, semantics, and standards. As in language, metadata syntax, or encoding, defines the rules for construction of metadata “sentences.” Examples of syntax include Machine Readable Cataloging (MARC), an alphanumerical encoding that enables one to go online to determine a library’s holdings, and Extensible Markup Language (XML), a “human readable” or language-based encoding that allows Web publishing, electronic data exchange, and portable, reusable metadata. A feature of personal digital assistants, cellular phones, and automatic phone banking, XML will figure importantly in the library catalogs of the future. In semantics, by contrast, we find the meaning of semiotic markers — in a metadata scheme as in language where the word “chair” can refer to the piece of furniture or to the person presiding over a committee. Thus a metadata system requires a third and final component, standards, which fix meanings that would otherwise — as in actual language — be unixed, subjective, and contextual. Standards make possible the exchange of information by making metadata records compatible with each other and aiding interoperability between databases. There are standards for metadata element sets or schemes, element content, controlled vocabularies, and encoding.

**Metadata Schemes**

Because of the need for differences and levels of complexity in semantics for describing different types of resources, several different but standardized metadata schemes have been developed. The most common of those geared to specific disciplines and purposes include:

- Visual Resources Association Core (VRA), used for describing cultural objects and works of art;
- Encoded Archival Description (EAD), for describing archived collections;
- Text Encoding Initiative (TEI), which facilitates the description and marking up of texts; and, most prominent,

**Dublin Core**, an all-purpose metadata scheme that, used in its simple or qualified forms, can integrate many different formats, including maps, images, and texts. In its simplest form, Dublin Core is a “lowest common denominator” scheme that facilitates system-to-system operability.

The original purpose of the Dublin Core was to organize the Web. Back in 1995, it was thought that the Web could be organized like a library if Website creators would assign access points, descriptors, and subject headings to their content so that it could be located more easily. Website creators did not have the motivation to catalog their Websites, but museum curators, librarians, and visual arts librarians adopted the Dublin Core and were instrumen...