# Dublin Core: An Overview of the Metadata Seminar

*by Maria Okonska*

**The problem.** The rapid growth of the World Wide Web has provided vast access to digital information and created many new opportunities while bringing problems previously unknown to the library community. Among these challenges encountered are the lack of rules or shared standards, inconsistencies in recording and indexing information, and the ephemerality of Web publications. These are only some of the barriers to the access to the Internet as the future unlimited, universal library.

**The solution:** "Metadata" standards. Formal resource description communities (such as museums, libraries, government organizations and commercial agencies) have increasingly come to agree that an interdisciplinary, international consensus is necessary to build Web descriptive "metadata" standards to facilitate discovery of electronic resources. Along with the recognition of the Web metadata as an area of public concern came increased research, resulting in a large body of publications, several international conferences, and marked technological advances. The development of Extensive Markup Language (XML) and Resource Description Framework (RDF) and such international descriptive standards as "Dublin Core" are all significant aspects of the digital era. While these developments have enormous potential, the speed at which they are appearing, along with new technologies and terminology, add more confusion than clarity to the present state-of-the-art. In this context, metadata, especially for the World Wide Web, has become a much-talked-about buzzword. Intended to enable users and various constituencies to describe, organize and manage resources (objects, documents, images, databases), it is understood differently by different professional communities. Though there is no clear definition of this term, it is essential that we realize the importance of different types of metadata in the development of networked information systems.

**The Dublin Core** is a metadata element set composed of 15 broad elements that can be used to describe a wide variety of information resources on the Internet for the purpose of cross-disciplinary resource discovery. The building of an interdisciplinary, international standard around a core element is its central feature. It was designed to support the creation of simple, but generally usable resource descriptions that will satisfy the needs of many users and communities, and to provide a shared semantic standard to allow users that operate under different rules and/or standards to be able to exchange metadata readily. (The 15 elements of the Dublin Core are: title; author or creator; subjects and keywords; description; publisher; other contributor; date; resource type; format; resource identifier; source; language; relation; coverage; rights.)

Recognizing the need to provide a foundation for understanding this new world of metadata, the OCLC Institute has initiated regular two-and-a-half day seminars/workshops in Dublin, Ohio, on the Dublin Core approach. I recently participated in the March 28–31, 1999, session on "Using Metadata for Knowledge Management." The seminar provided an overview on using Dublin Core. More specifically, its purpose was to:

- present an overview of networked information;
- define metadata and its role in knowledge management;
- define Dublin Core and its current status;
- apply Dublin Core to a variety of knowledge management task; and,
- assess the future of metadata and Dublin Core.

The seminar consisted of lectures, structured lab exercises, and small project teams. To evaluate the resource description system in terms of design and functionality, including data input, storage, retrieval and display, the participants learned how to use the OCLC CORC system. (CORC is a research project exploring the cooperative creation and sharing of metadata by libraries and other agencies.)

The seminar provided the opportunity for information, practical examples, and guidance to understand: metadata, its importance and why it differs from traditional cataloging; and the Dublin Core, its meaning for resource discovery, and its current state and future developments.

***continued on page 36***

## Suggested readings

*Introduction to Metadata: Pathways to Digital Information,* edited by Murtha Baca (Los Angeles, Calif.: Getty Information Institute, 1998, 41 p., bibliography) examines metadata from various viewpoints, focusing on the cultural legacy of information metadata systems. Useful for anyone interested in access to electronic information, the book features the issues that rise in creating metadata and reviews a number of initiatives already under active consideration.

The Dublin Core Metadata Initiative (http://purl.org/dc) assesses the Dublin Core, a worldwide metadata standard, in terms of its clarity, extensibility, and global, interdisciplinary nature.

Norman Oder's "Cataloging the Net: Can We Do It?" (*Library Journal* 123:16 (1998): 47) discusses methods to manage information on the Web and improve access for patrons by librarians. The crucial question is: Can the traditional cataloging or a 15-element set Dublin Core metadata describe Web resources properly?

"Show me the metadata!" (September 1998, Ed Summers, ed., online at http://libstaff.lib.odu.edu/~esummers/meta/meta.html) finally allows us to understand what metadata is. Summers presents a short history of various metadata initiatives, such as the Text Encoding Initiative, Encoded Archival Description, Geospatial Metadata, and Dublin Core.

The meeting signified the importance of the issues, and the dedication of the OCLC Institute and many other organizations and groups to identify, formulate and attract the world-wide attention to the problems and dilemmas associated with Web descriptive metadata standards. No doubt, one of the notable outcomes of the Dublin Core is the building of a cross-disciplinary consensus around a core element. However, with millions of users and potential applications for Web metadata, the creation of the most appropriate structure and content of metadata, despite all the efforts made, seems problematic. The Web metadata is still in its infancy and will certainly continue to evolve. But this issue is crucially important for everyone, from librarians to museum professionals to those who intend to make information available on the World Wide Web.

The information presented at the Dublin workshop helped to increase our understanding and sensitivity to the issues discussed. It is hoped that meetings like this will stimulate continued investigation in this growing area of interest.

***Maria Okonska*** *(mokonska.brooklaw @pcm.brroklaw.edu) is a Cataloger in the Brooklyn Law School Library, Brooklyn, New York.*

## Glossary

**Formal Resource Description Communities,** in the context of the Dublin Core Metadata, refers to every information professional, from museum specialists to librarians, archivists, to those who intend to make information available on a network. They provide digital versions of actual collection content as well as formal, structured description of Web resources that they feel are suitable for their communities.

**Metadata** describes an information resource. "Meta" comes from a Greek word that denotes something of a higher or more fundamental nature. Metadata, then, is data about data. It is the Internet-age term for information that librarians traditionally have put into catalogs, and it most commonly refers to descriptive information about Web resources.

The **Resource Description Framework** (RDF), produced as a part of the World Wide Web Consortium's Metadata Activity, is a metadata application of XML (Extensive Markup Language). It is seen by many experts as the successor of HTML and the future language of the Web. The RDF will provide a flexible architecture for managing diverse application-specific metadata packets that can be processed by machines.

**Hyper Text Markup Language** (HTML) is the language in which WWW documents are written.

**Standard Generalized Markup Language** (SGML) is an ISO (International Organization for Standardization) standard that was first used by publishing industry for defining, specifying, and creating digital documents that can be delivered, displayed, linked, and manipulated in a system-independent manner.

**XML** (Extensive Markup Language) is a simplified subset of SGML that is designed specifically for use with the World Wide Web and that provides for more sophisticated data structuring and validation than does HTML.

The **World Wide Web** (WWW or W3) is a browsing and searching system. It allows exploration of the seemingly unlimited worldwide digital "web" of information.