of options and constrained by limited resources, the challenge for libraries is to do the "next right thing." Each library will define that next step differently, but usually, the next right thing is not a discontinuous, high risk, high-cost leap. More often, the next change is marginal, and so is the one after that. The risks are lower and the costs more affordable. Cataloging Internet resources, metadata, PURLs — these represent some changes at the margin that offer significant immediate and long-term benefits.

Lest we watch the waves and ignore the tide, however, we should be open and alert to the continued application of computing technology to information management. Old ways of working are good ways of working only if they continue to make sense. So, too, with the precepts of cataloging.

Think back to a time before the Internet. For many of us, it wasn't that long ago. Consider how far we have come in such a short time. Yet the Internet is still nascent. We are in a time of unprecedented changes, many of which are directly related to recent advances in computing and telecommunications technologies and their rapid and widespread adoption. Libraries will not stay this flood of technology, but, with reasoned action, should expect to do more than just survive.

References

Dillon, *et al.* 1993. *Assessing Information on the Internet: Toward Providing Library Services for Computer-Mediated Communication.* Dublin, OH: OCLC.

Olson, Nancy B. 1995. *Cataloging Internet Resources: A Manual and Practical Guide.* Dublin, OH: OCLC. http://www.oclc.org/oclc/man/9256cat/toc.htm

The Dublin Core: A Simple Content Description Model for Electronic Resources

by Stuart Weibel

Stuart Weibel is senior research scientist in the OCLC Office of Research. He currently coordinates networked information research projects, including applications of World Wide Web technology and Internet protocol standardization efforts. He can be reached at http://purl.org/net/weibel

The term *metadata* simply means data about data. It is the term most often used in the Internet community for what has been known in the library community as cataloging data or resource description. The Dublin Core is a 15-element metadata element set intended to facilitate discovery of electronic resources. Originally conceived for author-generated description of Web resources, it has also attracted the attention of formal resource description communities such as museums and libraries.

The Dublin Core Workshop Series has gathered experts from the library world, the networking and digital library research communities and a variety of content specialties in a series of focused, invitational workshops. The building of an interdisciplinary, international consensus around a core element set is the central feature of the three-year evolution of the Dublin Core. The progress represents the emergent wisdom and collective experience of many stakeholders in the resource description arena. An open mailing list supports ongoing work.

The characteristics of the Dublin Core that distinguish it as a prominent candidate for description of electronic resources fall into several categories.

Simplicity

The Dublin Core is intended to be used by non-catalogers. It is expected that authors or Web-site maintainers unschooled in the cataloging arts should be able to use the Dublin Core for resource description, making their collections more visible to search engines and retrieval systems.

Most of the 15 elements have a commonly understood semantics that represents what might be described as a lowest common denominator for resource description (roughly equivalent to a catalog card). As such, the Dublin Core is not intended to replace richer description models such as AACR2/MARC cataloging, but rather to provide a core set of description elements that can be used by catalogers or non-catalogers for simple resource description.

Semantic Interoperability

In the Internet commons, disparate description models interfere with the ability to search across discipline boundaries. For example, libraries, museums and the geographic information systems community use different standards for resource description. This reflects the different description needs of these communities and the fact that such standards have evolved independently.

At a fine-grained description level, element sets are different because they must describe different things. Most writers seldom associate a cloud-cover attribute with their documents, but if you are describing satellite images of farmland, this is a

critical descriptor.

But most resources share a core set of attributes that are similar from one discipline to the next, but have different names simply because they have evolved independently and at different times. Promoting a commonly understood set of core descriptors will improve the prospects for cross-disciplinary search by unifying related attributes. For example, an *author* and a *creator* can be thought of as the same attribute for the purposes of resource discovery. The Dublin Core is intended to serve as this core element set.

International Consensus

Recognition of the international scope of resource discovery on the Web is critical to the development of an effective discovery infrastructure. The Dublin Core has benefited from active participation and promotion in many countries around the world.

Flexibility

Although initially motivated by the need for author-generated resource description, the Dublin Core has also attracted the attention of formal resource description communities. As the diversity and volume of Web resources increase, trusted intermediaries (such as museums and libraries) will achieve greater recognition as preferred sources of metadata for persistent resources.

The Dublin Core, in the hands of cataloging experts, is expected to provide an economical alternative to more elaborate description models such as full MARC cataloging. The Dublin Core includes sufficient flexibility to encode the additional structure and more elaborate semantics appropriate to such applications.

Metadata Modularity on the Web

The wide diversity of metadata needs on the Web requires an environment that supports the coexistence of many independently developed and maintained metadata packages. The Dublin Core is targeted specifically toward resource discovery, but one can imagine many functionally distinct packages that serve other goals (terms and conditions, archival management, administrative metadata and many others). For example, a *Terms and Conditions* metadata package would include elements that describe rights holders, cost of acquiring a resource, restrictions on reuse of the resource and related information.

Recognition of the desirability of this sort of modularity has guided the evolution of the Dublin Core since the Warwick Workshop and has been formalized as the *Warwick Framework*. The concepts articulated in this work have informed the ongoing development of a metadata architecture for the Web as well.

A Metadata Architecture for the Web

The World Wide Web Consortium (W3C) is the primary standards forum for the Web and has recently begun to focus on implementing an architecture for metadata for the Web. The Resource Description Framework (RDF) is evolving to support the many different metadata needs of vendors and information providers. Representatives of the Dublin Core effort are actively involved in the development of this archi-

tecture, bringing the digital library perspective to bear on this important component of the Web infrastructure.

Models for Deploying Dublin Core Description on the Web

The evolving RDF metadata architecture will support a variety of resource description models, each with implications for functionality and management.

1. Embedded Metadata

The easiest way of deploying metadata on the Web is by embedding it in HTML documents (using the META tag). Conventions exist to support inclusion of simple metadata in HTML versions 2.0 and above. The HTML 4.0 specification released in July includes additional attributes for the META tag that allow the qualifiers necessary for more complex implementations. The advantage of embedded metadata is that no additional system must be in place to use it; the metadata is integral to the resource and can be harvested by Web indexing agents.

2. Third Party Metadata

A model more familiar to the library community includes what is known in Web parlance as a third party label bureau, that is, an entity that collects and manages metadata records that refer to resources but are not embedded in the resource (a library catalog, for example). This model is important not only to libraries and museums, but also supports the development of agencies that might label resources according to age appropriateness or other acceptability criteria.

3. View Filter

A third model also involves management of records by a distinct entity, but not necessarily Dublin Core records per se. Managing a wide variety of data stores often involves reconciling very different description models. One approach to achieving interoperability in such an environment involves mapping many description schemas into a common set such as the Dublin Core, giving users a single query model.

Unsolved Problems and Future Directions

Much remains to be done to bring the Dublin Core to a state of sufficient maturity and stability to fulfill its promise as a foundation for resource discovery on the Net. The main thrusts of continued development include

1. Continued Refinement of Dublin Core Elements

The Dublin Core elements emerged from the collective judgment and experience of the many participants in the process to date. As deployment spreads, the evolution of the Dublin Core will reflect experience with the ambiguities, conflicts and deficiencies in the set. Standards of best practice will evolve in light of such experience.

2. User Education and Application Guides

The spread of a common set of resource description conventions depends in part on the availability of clear user guidelines. Such guidelines must be developed in many languages but with a common purpose and orientation.

3. Metadata Registries

The Warwick Framework describes the characteristics of an architecture for metadata that will allow independently

developed metadata element sets to co-exist. This implies that the consumers of metadata (either people or software agents) will need formal, online registries that describe the semantics, the structure and the transport syntax of a metadata element set. Thus, an application finding Dublin Core metadata associated with a collection of resources might access the Dublin Core Metadata Registry to better understand the characteristics of the metadata. Work on metadata registries is still in an embryonic stage, but as the functional specifications evolve, they will become a central part of the infrastructure necessary to develop and manage change for a metadata set such as the Dublin Core.

4. Tools

Tools for creating and managing Web-based metadata are evolving now. As the infrastructure evolves and standards become stable, these tools will become commonplace in authoring, site management and resource management applications.

5. Standardization

The development of the Dublin Core has been a voluntary effort on the part of many disparate stakeholders in resource description. As it becomes more widely deployed, standards of best practice must be formalized.

Places to Learn More

- **The Dublin Core Home Page**: The official home page of the Dublin Core includes links to the workshop reports, the reference description of the element set and to projects using the Dublin Core, and the evolution of the Warwick Framework. *http://purl.org/metadata/dublin_core*

- **The RDF Home Page**: The home page of the RDF working group, maintained by the World Wide Web Consortium (W3C), reports developments of the metadata architecture on the Web. *http://purl.org/w3c/rdf*

- **Crosswalks of Metadata Element Sets**: Maintained by Michael Day at the United Kingdom Office for Library and Information Networking (UKOLN), this site is an excellent source for crosswalks and element mapping for a variety of metadata sets, including the Dublin Core. *http://www.ukoln.ac.uk/metadata/interoperability*

- **The IFLA Metadata Home Page**: A good collection of links to various metadata activities and related resources. *http://www.nlc-bnc.ca/ifla/II/metadata.htm*

- **The UKOLN Metadata Page**: A summary of publications, projects and metadata resources in general and projects in the UK in particular. *http://www.ukoln.ac.uk/metadata*

The Dublin Core Elements
See *http://purl.org/metadata/dublin_core_elements* for further information.

1. TITLE
The name given to the resource by the CREATOR or PUBLISHER.

2. CREATOR
The person(s) or organization(s) primarily responsible for creating the intellectual content of the resource.

3. SUBJECT
The topic of the resource: keywords or phrases that describe the subject or content of the resource, including controlled vocabularies or classification schemes.

4. DESCRIPTIONS
A textual description of the content of the resource, including abstracts in the case of document-like objects or content descriptions in the case of visual resources.

5. PUBLISHER
The entity responsible for making the resource available in its present form, such as a publisher, a university department or a corporate entity.

6. CONTRIBUTOR
Person(s) or organization(s) in addition to those specified in the CREATOR element who have made significant intellectual contributions to the resource but whose contribution is secondary to the individuals or entities specified in the CREATOR element (for example, editors, transcribers and illustrators).

7. DATE
The date the resource was made available in its present form.

8. TYPE
The category of the resource, such as home page, novel, poem, working paper, technical report, essay, dictionary. It is expected that TYPE will be chosen from an enumerated list of types.

9. FORMAT
The data representation of the resource, such as text/html, ASCII, Postscript file, executable application or JPEG image.

10. IDENTIFIER
String or number used to uniquely identify the resource. Examples for networked resources include URLs and URNs (when implemented). Other globally unique identifiers, such as International Standard Book Numbers (ISBN) or other formal names, would also be candidates for this element.

11. SOURCE
The work, either print or electronic, from which this resource is derived, if applicable.

12. LANGUAGE
Language(s) of the intellectual content of the resource.

13. RELATION
Relationship to other resources. The intent of specifying this element is to provide a means to express relationships among resources that have formal relationships to others, but exist as discrete resources themselves.

14. COVERAGE
The spatial and temporal characteristic of the resource. Formal specification of COVERAGE is currently under development.

15. RIGHTS
The content of this element is intended to be a link (a URL or other suitable URI as appropriate) to a copyright notice, a rights-management statement or perhaps a service that would provide such information dynamically.