January 2005

International Conference on Dublin Core and Metadata Application 2004

Mary J. Wu
Roger Williams University, mwu@rwu.edu

International Conference on Dublin Core and Metadata Applications 2004

By

Mary Wu

Introduction

The International Conference on Dublin Core and Metadata Applications took place on October 11-14, 2004 in Shanghai, China. It was the fourth in a series of expanded Dublin Core conferences that included a conference, tutorials and workshops. Prior to the first in this series of conferences held in Tokyo in 2001, eight Dublin Core workshop series had been held in various North American and European countries starting in 1995. These workshops and conferences have provided a forum where researchers and practitioners can exchange new ideas and demonstrate the development of metadata standards and applications from practical aspects.

The Dublin Core Metadata Initiative (DCMI), an open and international association, has been the driving force in the community of people who devote themselves to making information resources easier to find. Following the development and the adoption of the Dublin Core Metadata Element Set, other metadata standards and initiatives have emerged. Having recognized the growth of activities in the Chinese metadata community as well as the growing contributions and importance of metadata developments in the country, the Dublin Core Metadata Initiative chose China as its host for the 2004 conference. The Chinese people were delighted to have the opportunity to showcase their experiences and achievements. Meanwhile, they also took the occasion to draw international attention to the Chinese endeavor in developing metadata for organizing and

discovering its abundant cultural and historical resources, which otherwise would remain hidden from the rest of the world.

Conference Organization and Location

The conference was hosted by the Shanghai Library and sponsored by the Library of Chinese Academy of Sciences, the National Science and Technology Library of China, and the Dublin Core Metadata Initiative. The conference organization consisted of eight committees. Among them, the Conference Organizing Committee was chaired by the Director of Shanghai Library Jianzhong Wu, and the Program Committee was co-chaired by Wei Liu (Shanghai Library) and Thomas Baker (Germany). Other committees included the Workshop Committee chaired by Makx Dekkers from DCMI and the Tutorial Committee chaired by Diane Hillmam from Cornell University. The coordinators of the conference were Wei Liu and Makx Dekkers.

Shanghai, the host city of the conference, lies in central-eastern China with a total area of 2,448 square miles and a population of more than 13 million. A coastal city exposed to the East China Sea, Shanghai is one of the most booming urban areas in the world. It is an important center of finance, economy, trade and a seaport for both China and Asia-Pacific regions. The city, started as a fishing village in the 11th century, rapidly grew into an industrial and financial metropolis when the British along with the Americans and French came into Shanghai as a result of foreign trade after the Opium Wars in the 19th century. In 1949, the Communist Chinese government took over the control of Shanghai as well as the businesses left behind by the foreigners. After the cultural and economic turmoil generated by the Cultural Revolution from 1966 to 1976, the open door policy of

the Chinese Government made Shanghai once again a prosperous economic zone attracting more and more investors and entrepreneurs from all over the world. Shanghai Library, the venue of the conference, was first founded in 1952. The original library, an elegant and beautiful European-style building, was too old and too small to house a fast growing institution. In 1996, the new library, built in the center of Shanghai, was opened. With a total floor space of 83,000 square meters, the two tower buildings of 11 and 24 stories hold 48.5 million volumes in both Chinese and Western languages. A library with an auditorium of 842 seats and plenty of small rooms for group discussion proved to be a suitable conference place. Being the symbolic cultural center of the city, the library, equipped with modern multimedia technology, provided all the required support for speeches and presentations. Shanghai Library's digital project, which digitizes and stores the country's largest gramophone record collection of Chinese local operas and dramas as well as other documents of historical significance, pertains to the theme of the conference—metadata across languages and cultures.

Coference Schedules and Special Events

The conference was convened by the Managing Director of DCMI Makx Dekker and followed by one of the keynote speakers Xiaolin Zhang, the Director of Library of Chinese Academy of Sciences. Mr. Dekker highlighted the mission, scope and organizational structure of DCMI while Mr. Zhang described the background and development of the Chinese Digital Library Standards project. During the conference, 166 participants from 25 countries and regions heard 10 full and 23 short papers delivered at the 3 Plenary Paper Sessions:

- Metadata Frameworks
- Information Retrieval
- Managing & Harvesting Metadata

and four Short Paper Sessions:

- Vocabularies and Application Profiles
- Models
- Case Studies
- Tools and Methods

Four tutorials were also offered on the four consecutive mornings to provide an introduction to metadata for non-experts.

Conference Implications

Interoperability was one of the most prevailing topics at the conference. It is the key to ensure the accessibility and sustainability of any effective digital library. The diversified Internet environment makes it impossible for a single metadata element set to meet the needs of all domains and purposes. Yet, it is impractical to develop metadata sets application by application to fulfill the domain specification. Therefore, the Dublin Core community has made significant efforts in developing methods to realize semantic interoperability of metadata frameworks across domains and cultures. One of the approaches presented at the conference was to store its multiple metadata vocabularies in a Web-based DCMI schema registry which operates as an authoritative index database controlled by the Usage Board of DCMI. Once approved by the Board, a new term is assigned a Uniform Resource Identifier accessible via the Internet by both humans and

machines. The terms in English can be translated into non-English languages for adoption by local communities. The ability provided by the registry database to unambiguously define metadata terms is a precondition for developing semantic maps and interoperability services across multiple standards[1].

In addition to the Metadata Registry, the creation of an expressive, extensible and domain-specific metadata schema language is also crucial for interoperability. When a large number of domains adopt the employment of metadata, various domain-specific metadata terms need to be defined. For example, if two or more different domains use the same metadata term and each domain has its own term for property and value types as well as qualifiers used to enhance the meanings of the term, one has to refer to their domain specific definitions.

As metadata standards, such as Dublin Core, are used freely by application designers for a variety of purposes, the Application Profiles were developed by DCMI as another crucial method of documenting Metadata usages for sharing metadata models within communities of interest. Application Profiles are used primarily as a form of documentation for harmonization of different metadata practices. For instance, they would document what subset of available terms is selected, which elements are mandatory, and whether terms from multiple sources, e.g., Dublin Core elements together with elements from IEEE/LOM standard or with elements defined in house for application-specific purposes, are used. It is hoped that machine-processable forms of such Application Profiles will ultimately provide a basis for automating metadata interoperability such as semantic crosswalks and format conversions[2].

Other issues concerning effective information retrieval are the managing and harvesting of metadata. Institutions in the business of managing ever-increasing numbers of digital objects are facing challenges including selection of a metadata framework for resource discovery, automatic generation of metadata, metadata maintenance and quality control. Papers presented at the conference suggested that using the most appropriate international metadata standard in any given circumstance is essential, whereas it must be able to be mapped to Dublin Core at the common resource discovery layer. Creating metadata for every purpose is almost impossible within budgetary constraints. However, some new metadata can be generated by automatically converting from existing sources. During the process of converting or harvesting, problems with missing data, incorrect data, confusing data, or insufficient data can arise. In order to provide a reasonable level of quality for metadata, an automatic technique called "safe transforms" is designed to enhance the information present in the original metadata with no risk of degradation[3].

Conclusion

We have done a great job in making information retrieval easier by improving the interoperability of metadata across domains and cultures. As technology rapidly progresses, the ever increasing expectation of end users requires more intelligent and resource imbedded metadata with the capacity of capturing context, mass contribution and mass customization. In order to build the future of metadata, the DCMI community will expend its efforts by bringing together experts from various backgrounds. Solving technical issues related to managing and maintaining the Metadata Schema Registry,

improving metadata frameworks, and strengthening the cooperation among working groups will remain priorities for DCMI for the year of 2005.

For further information about the conference and DCMI visit:

http://dc2004.library.sh.cn/english/ and http://dublincore.org/

Mary Wu is a Catalog/Database Management Librarian at the Roger Williams University Library in Rhode Island. She may be reached at mwu@rwu.edu.

[1] Baker, T. (2004) "Maintaining a vocabulary: practices, policies, and models around Dublin Core", DC-2004: Proceedings of the International Conference on Dublin Core and Metadata Applications, pp. 15-16

[2] Baker, T. (2004) "Maintaining a vocabulary: practices, policies, and models around Dublin Core", DC-2004: Proceedings of the International Conference on Dublin Core and Metadata Applications, pp. 17-18

[3] Hillmann, D. (2004) "Improving metadata quality: augmentation and recombination", DC-2004: Proceedings of the International Conference on Dublin Core and Metadata Applications, pp. 62-63