A comparison of the Two Traditions of Metadata Development
Kathleen Burnett, Kwong Bor Ng, Soyeon Park

The two traditional methods of Metadata Development as describe in the article are bibliographic control approach and the data management approach. These two approaches have evolved over time due to the rise of digital information stored in cyberspace and help facilitate metadata standards which include the identification, location, retrieval, manipulation and use of the information. So what is metadata?

From my understanding of what metadata is, metadata is simply defined as ‘data about data’. And to a certain extent and depending on the perspective which you define what metadata is, ‘data about data’ is a legitimate and accurate definition. Looking at the various definition of metadata Kathleen Burnett, Kwong Bor Ng, Soyeon Park (1999), define metadata as “structured data about data” whereas NISO Press, 2004 describes metadata as “data about data or information about information.” But according to Gill, 2000, “information about information” is often “incorrectly generalized” by the reader or author as the meaning of metadata, therefore, ‘data about data’ would be more concise and accurate for this paper. This begs the question of exactly the purpose of metadata and metadata standards. In theory, “metadata provides an effective mechanism for identifying and locating data that is relevant to a particular user.” (Kathleen Burnett, Kwong Bor Ng, Soyeon Park, 1999). The elements within the metadata standards may be divided into two categories: intrinsic (related to resource identification and discovery) and extrinsic (related administration and other nonbibliographic data). While some metadata formats vary in terms of intrinsic and extrinsic elements, some elements are more suitable for resource description and identification whereas others may help resource access and navigation. However, the issue still remains as to which metadata format to use and whether a universal or unified “grand scheme” that is rich in attribute should be employed, eliminating the need in choosing a specific metadata format.

In the past years, we have seen a rise of conflicting standards as an increasingly large amount of information becomes available electronically. This has become a growing problem as retrieving information becomes more complex. In an article by Jessica Milstead and Susan Feldman (1999), the author uses the term “Mercury” as an example of the documents the user would retrieve through the use of
the World Wide Web, which includes a mixture of “the planet Mercury, the element Mercury, the Greek God Mercury, and articles from the San Jose Mercury-News”. It would also disregard any non-textual documents such as images, audio, video, etc. To date, there is no single international standard for metadata. Some may argue that each metadata scheme for handling digital information is necessary in its own right as they all have different levels of “complexity and richness”, (Kathleen Burnett, Kwong Bor Ng, Soyeon Park, 1999) suitable for different purposes. Where others may argue that we have too many competing proposals of metadata standards such as Jessica Milstead and Susan Feldman, (1999), who agrees that each metadata has its own merit and would be “nicely applicable to the materials of the grouping proposing it,” however, in an information management environment, “we have run smack up against the "no one is in charge of the Internet" wall-and no obvious answers are in sight.”

A comparison of six metadata standards based on their scope and impact on other metadata schemes was demonstrated by Kathleen Burnett, Kwong Bor Ng, and Soyeon Park, (1999). Base on these six metadata formats, there is to a certain degree of overlap between them. According to Jessica Milstead and Susan Feldman (1999), generally, most metadata format elements are common. Elements such as author, title, date of creation and modification and identifier can be seen in most formats as a means of resource discovery and identification attribute. Therefore, “if like information is treated alike, the information is more accessible.” (Jessica Milstead & Susan Feldman, 1999) However, it is important to remember that each metadata format serve different purposes and audiences. “Complementary schemes can be used to describe the same resource for multiple purposes and to serve a number of user groups.” (NISO Press, 2004) A relatively simple format such as The Dublin Core format contains 15 metadata elements focusing on simplicity in order to describe electronic documents compared to a more complex format such as the Text Encoding Initiative (TEI), a set of tags and rules defined in Standard Generalized Markup Language (SGML) syntax that describes the structure and elements of the document. The Dublin Core format is arguable the most used format for metadata standard due to its simplicity and is widely used for either cataloguing or retrieving information from the Internet. The Dublin Core focuses on resource discovery, dismissing the need for extrinsic elements such as availability or access restriction. It was the collaboration of “group of librarians, information professionals, and subject specialists through an ongoing consensus-building process that has included seven international workshops
to date and an active mailing list” (Tony Gill, 2000) to come up and define the 15 elements which the Dublin Core consist of. Part of the advantage of Dublin Core is its ability to refine an element known as qualifiers. For example, the Dublin Core element ‘Date’ can be used with the refinement qualifier ‘created’ or ‘copyright’ to narrow the meaning of the element and hence a better understanding of the element.

This brings me back to my previous question for the need of a unified or universal ‘grand scheme’. If there were to be one grand scheme which would eliminate the need for the other metadata formats currently available, the Dublin Core Format at its very basic level would be it. The focus on simplicity and conciseness would favor most common users at the same time allowing for specialized groups the ability to extend the list to suit their needs. Expanding the Dublin Core formation to include more extrinsic elements such as ‘system requirement’ etc. would add to the attractiveness of the format and would prove more ideal for usage. “The primary concern of metadata is to discover appropriate methods for modeling various classes of information objects in the networked and distributed information environment.” Therefore, a ‘grand scheme’ would be inappropriate as the various metadata format serve different purposes for different audiences. Ultimately, the quantity of information available will continue to grow, simultaneously new technology, standards and practices are constantly advancing the application of metadata.
Reference

Last accessed on the World Wide Web on 29/11/05
http://www.getty.edu/research/conducting_research/standards/intrometadata/metadata.html

Last accessed on the World Wide Web on 29/11/05
http://www.niso.org/standards/resources/UnderstandingMetadata.pdf