

RESEARCH DESIGN LITERATURE REVIEW

MAKING SENSE OF METADATA:

A POTENTIAL APPLICATION OF SENSE-MAKING METHODOLOGY

by

Jon Mason
education.au limited
jmason@educationau.edu.au

CITATION AND COPYRIGHT INFORMATION:

Cite as: Mason, J. (2005, May). Making Sense of Metadata: A Potential Application of Sense-Making Methodology. Paper presented at a non-divisional workshop held at the meeting of the International Communication Association, New York City.

© Jon Mason (2005).

1. THE PHENOMENON

The focus of interest is the diversity of definition, interpretation, usage, and application of metadata in contexts where the World Wide Web is used to facilitate sense-making and learning. Problems identified with this interest involve conceptualization and development of shared conceptual models, semantics, terminology, politics, the technological environment, technical design and capability, and standardization (Duval, et al, 2002; Gorman, 2004).

2. THE DISCOURSE COMMUNITIES

There are two primary discourse communities (emerging from e-learning standards development and digital library research and development). While historically distinct, these communities share some common interest (managing information resources on the Internet). See, for example, Friesen, Mason, and Ward (2002) and/or McLean and Lynch (2004).

3. BROAD BRUSH STROKE PICTURE OF THE PHENOMENA AS IT IS NOW UNDERSTOOD

Metadata is often simplistically defined and its value is often debated. While a number of standard definitions are now in common use (and geared toward practical implementation) there remain significant differences in understanding and purpose. One recent definition agreed to by members of various standardization communities is as follows:

“Metadata is commonly defined as “data about data”. More generally, metadata is information about a resource, either physical or digital ... [in learning, education, and training contexts] metadata refers to information about resources used in the context of learning, education and training. Metadata often consists of characteristics and properties of a resource such as title and author. Metadata may include objective information such as the size of a file, classifications such as the subject or topic, and assertions such as the appropriate audience for the resource. Metadata appears in many forms. It can be presented as part of a resource, such as the title on a book, or be stored in a separate file.” (Metadata FAQ, 2005)

My own working definition is as follows: *metadata is information about things, events, and services that is typically organised in a conventional, structured way*

that helps identify and describe these things, events, and services. Elsewhere, I have discussed the relationship between context and metadata (Mason, 2004).

4. WHY RESEARCH/PRACTICE RELATING TO THIS PHENOMENA IS NOT "ENOUGH" AS THINGS STAND NOW

In tracking the histories of e-learning there are a number of commentators who observe that we are only now beginning to move beyond the "cottage industry phase" (McLean and Lynch, 2004; Blinco, *et al*, 2004). The field of e-learning can be characterized as an emergent one, triggered largely by technological enablers and the intersections of a number of distinct fields of practice: computer-based training, human computer interaction, information science, computer science, publishing, game theory, and distance education, to name a few. During this cottage industry phase (which could be characterized as the twelve years or so beginning in the early 1990s) there has been a proliferation of new jargon, some of which has stabilized or become sticky ("e-learning" itself an example), while many terms appearing as having either been appropriated or recycled from other domains (terms like: architecture, framework, services, ontology, objects, consumption). Terms like "metadata" are a little different as they are relatively new.

While one of the functions of metadata is to facilitate information-seeking and discovery, this is not its only function. A useful metaphor from the physical world is a bottle of wine – where the wine can be characterized as the *content*, the bottle the *container*, and the label as the *metadata*. Just like wine, much of the purpose-built e-learning content needs to be placed in a container for distribution and easy access. Most of us also generally like to know *what it is* we are drinking and so labels are very important for conveying this. If it is to be easily shared and re-used then e-learning content generally needs to be contained in modular chunks (just like wine bottles and standard cartons) and it needs to be labelled carefully for identification and potential uses. In the transport and management of e-learning content, however, there is also metadata about the 'bottle' that remains invisible largely to most users.

However, such metaphors from the physical world have limited use. Despite the scientific view that the physical world is made up of atoms it is generally easy to distinguish between wine bottle, wine, and labels. In the digital world, where all things are pretty much made up of bits and bytes there are an increasing range of contexts where *content is in the eye of the beholder*. As Shabajee (2002) suggests, "Any object, physical or virtual, could be described and discussed in possibly limitless ways". In an effort toward simplifying this complexity, Seely Brown has elegantly said: "What do we know that we didn't know ten years ago? That learning and knowledge are the result of multiple, intertwining forces: *content, context, and community*." (Seely Brown, in Ruggles, & Holtshouse, 1999: ix)

Another metaphor, the *history of money formats*, can be reasonably used to point toward a similar history – the *history of information formats*. Over the centuries human societies have evolved a series of increasingly sophisticated and flexible ways of handling and exchanging money. Most of us have now experienced monetary transactions involving coin, paper, plastic, and bits and bytes. What money *is*, is becoming increasingly abstract.

The forms in which data and information find expression are evolving in a similar way. While librarians have been cataloguing books and journals for centuries, retailers have been stocking their stores with bar-coded items for only about two decades. More recently, supply-chain specialists have started using technologies like Radio Frequency Identification to efficiently distribute and track physical

objects of all kinds. In each of these examples, as it is with e-learning, objects (digital or otherwise, large or small) need to be identified and described. Age-old questions such as *What, Who, When, and Where* – and sometimes *Why, How, and If* – are the kinds of questions that metadata addresses.

Investments made by the communities of practice harnessing new technology in innovative ways involved in developing standards are sufficiently large that 'orthodoxies' have also emerged during this cottage industry phase. These orthodoxies seem to exhibit some degree of reflective inertia, now that first generation international standards have been developed. While many members of these orthodoxies would argue that the inertia is really an essential state of stability required for the industry to develop there is also some effort toward looking forward to next generation metadata standards.

5. REASON SENSE-MAKING METHODOLOGY INFORMED RESEARCH/ PRACTICE MIGHT BE USEFUL

The formal processes involved in the international standardization of metadata schemata, such as the Dublin Core (2003) and the IEEE LOM (2002) reveal a high degree of sustained commitment over periods in excess of five years by quite different communities of practice. However, while a surface analysis of these schemata shows strong similarities, and sophisticated journal articles have been written concerning shared principles (Duval, *et al*), the conceptual models are sufficiently different in their detail and encodings to cause major issues for adopters and implementers. It is possible that an informed sense-making study may point to ways in which a "meta" understanding of these models can be properly articulated.

Innovative and successful search engines like Google™ also use metadata, but in novel ways – leveraging the relationships between information and patterns of information use. The success of Google™ has tended to stimulate debates between those who understand the importance of information management practices such as cataloguing and those who just want to find stuff using the latest, easy technology. These debates reveal a range of interpretations as to what metadata actually is. It seems likely that a sense-making approach to researching these interpretations and usages may inform the standardization communities better than they currently are.

6. POTENTIAL RESEARCH QUESTION

A number of questions arise in relation to metadata and its use, particularly in educational and training contexts:

- (i) Why is information science not a core subject in K-12 education now that a global information infrastructure is now in place? Reason for the question: it would seem that developing an understanding of information relationships is becoming an important foundation for a contemporary education, possibly as fundamental as reading and writing.
- (ii) How might the evolution of information and data formats be best described in order to show a growing importance in the role and function of metadata?

REFERENCES

Blinco, K., Mason, J., McLean, N. and Wilson, S. (2004). *Trends and Issues in E-Learning Infrastructure Development* – A White Paper for *alt-i-lab 2004* prepared on behalf of DEST (Australia) and JISC-CETIS (UK).

<http://www.msglobal.org/altilab/altilab2004/Alttilab04-Trends-Issues.pdf>

Duval, E., Hodgins, W., Sutton, S., & Weibel, S. (2002). *Metadata Principles and Practicalities*, *D-Lib Magazine*, Vol 8(4).

<http://www.dlib.org/dlib/april02/weibel/04weibel.html>

Gorman, G.E., (Ed.) (2004). Metadata Applications and Management – International Yearbook of Library and Information Management, 2003-2004. Facet Publishing: London.

Friesen, N., Mason, J., & Ward, N. (2002). *Building Educational Metadata Application Profiles*, Proceedings, Dublin Core Metadata Initiative annual conference, Florence, Italy. <http://www.bncf.net/dc2002/program/ft/paper7.pdf>

Mason, J. (2004). *Context and Metadata for Learning, Education, and Training*, in R. McGreal (ed.) Online Education Using Learning Objects, RoutledgeFalmer, London.

McLean, N. & Lynch, C. (2004). *Interoperability between Information and Learning Environments – Bridging the Gaps*, A Joint White Paper on behalf of the IMS Global Learning Consortium and the Coalition for Networked Information http://www.imsglobal.org/digitalrepositories/CNIandIMS_2004.pdf

Metadata FAQ (2005). <http://www.cetis.ac.uk/metadatabfaq/>

Ruggles, R. and Holtshouse, D. (eds.) (1999). The Knowledge Advantage: 14 Visionaries Speak on Leveraging Knowledge for Marketplace Success, Capstone: Dover, NH, USA

Shabajee, P. (2002). *Primary Multimedia Objects and 'Educational Metadata' – A Fundamental Dilemma for Developers of Multimedia Archives*, D-Lib Magazine, Vol 8(6) <http://www.dlib.org/dlib/june02/shabajee/06shabajee.html>