Structured Metadata Spaces

Thomas D. Wason, IMS Project, GEM
David Wiley, Brigham Young University, DLE Group

1. Introduction

This paper will present the concepts of a **metadata space** as it relates to cataloging and discovery. A space has multiple dimensions; in the case of resource metadata, these **are descriptive dimensions**. We will explain the needs for orthogonal descriptive dimensions, and present a method for achieving maximally efficient, independent dimensions using **semantic structures** realized in **structured metadata**. A specific example of this system as developed in the IEEE Learning Technology Standards Committee (LTSC P1484) Learning Object Metadata (LOM) will be presented. The LOM is the collaborative work of many organizations including ADL, AICC, ARIADNE, GESTALT, and IMS. The scope of the concepts presented in this paper encompasses general concepts of metadata systems.

2. The Problems

In order to understand the requirements for a metadata system, the nature of the problems such a system addresses needs to be understood. The user community encompasses a wide variety of needs, not all of which can be predicted by the cataloging community. The user may be a teacher searching for individual resources or complete courses. It may also be a student looking for references for a paper or doing research for a project. (A data management system may be considered a special form of user.) Metadata can be considered a system that supports communications between two very diverse user communities. As a **communications system** it must have a well-defined vocabulary and syntax to support a wide variety of semantic needs. There is inherent variability in the use of a communication system: interpretation of the language may vary with individuals and/or may drift over time. This variability creates a certain lack of clarity in the communications, or "fuzziness." Additionally, the purposes filled by the system may change, causing the system itself to change to fulfill these new requirements.

2.1. Communication

Bibliographic metadata can be considered to be a system of communication between the **cataloging community** and the **user community**. The cataloging community ranges from the professional cataloger who describes a large number of resources on a daily basis to the casual creator of metadata who may want to provide some meaningful labels for a web page. Together, the community of catalogers creates a corpus of metadata.

There is a strong need for a common language that has adequate richness but will not fall apart due to the noise inherent in the system. Metadata is a poor personal communications system, as the cataloger cannot know exactly for what purpose every user will want a resource. Additionally, the user cannot ask for information that does not exist. As the communication becomes more uni-directional, the requirement for unambiguous messages increases. The need for a common, robust communications system is clear, but what are its characteristics? The purpose of the system, as with normal human language, is to enable the communication of meaning. Metadata should describe resources in ways that will be useful to the user. The user communicates with the cataloger in an abstract fashion: the user asks the cataloger to describe the resource in the same manner that the user does. There is a communications expectation created by the user, but the communications channel is restricted. Formative feedback from the user to the cataloger is diffuse, either through complaints or from lack of use of the resources. An effective metadata communications system must have some mechanism for providing this feedback. *An effective metadata system should provide good correspondence between the description of the resource by the cataloger and the strategies of the searcher.*

2.2. Description and Discovery

Discovery is the process of finding a previously unknown resource that satisfies a particular need. The discovery process itself may serve to sharpen the definition of that need. The definitions that underlie the description of a resource may not reveal the purpose of the resource. The searcher has a purpose for his or her discovery process. Some purposes can be clearly articulated in terms of the information needed. The definitions of the terms of description should be as free of implied purpose as possible. *In an efficient metadata system, the resource description should be as objective as possible, but provide enough information for a user to determine if the resource is appropriate for his or her purposes.*

2.3. Fuzziness

Ask ten catalogers to provide descriptions of a resource using an unrestricted vocabulary. Ask ten searchers to find the resource, again, using an unrestricted vocabulary. A perfect system would have the resource descriptions provided by all ten catalogers found in the searches by all ten searchers. This is highly unlikely, however, because the catalogers will differ in their descriptions of the resource and the searchers will differ in their search strategies and terminology. The lack of correspondence between the user community and the cataloging community creates a system with noise or **fuzziness**. Fuzziness can be considered from a statistical standpoint: it is the variability in repeatedly assigning the same value to the same resource over repeated trials. The same person may not produce exactly the same metadata for a resource now that he or she might have a month earlier. Descriptions that overuse multiple terms in a single field may create another kind of fuzziness.

2.4. Change

Metadata systems change. The fields used, the definitions of the fields, and the vocabularies used have changed and will change again. Some changes will be evolutionary, some revolutionary. *A metadata system must do more than contend with changes, it must manage them.*

3. Metadata Space

Metadata spaces provide an analogy for thinking about, describing, and creating effective metadata systems. The following is adapted from Wason (Wason, 1998).

3.1. Precision, Resolution, and Repeatability

Precision is the degree of fidelity with which something can be represented. **Resolution** is the ability to differentiate between two similar items. **Repeatability** is the ability to have the same resource described the same way on two or more occasions.

If one has a green square, it is possible to describe that *color* from a fixed vocabulary. If the vocabulary contains only *colored*, *black*, and *white* as terms, then one can describe the square as "colored". This is a low degree of fidelity, however. The **precision** is low. Next introduce a blue octagon to the collection. The vocabulary does not permit the two items to be separated in a search; the **resolution** is low. Now suppose two different people search for the green square. Both will probably pick "colored". The **repeatability** is high.

Continuing with the green square example, what if we change the color palette: black, white, red, orange, yellow, green, blue-green, blue, and violet. We have increased the potential precision of the descriptive field. Again, we have people independently categorize the square and the octagon. For the square, most will pick "green", but a few will pick "blue-green" or "yellow". For the octagon, most will pick "blue", but some will pick "blue-green", some "violet". The larger vocabulary has increased precision and resolution at the cost of lower repeatability. The variability among all of the catalogers has a *probability distribution*. This is the *fuzziness* referred to above. The searcher selecting a blue-green shape will get back a mix of shapes. The searchers also have some variability in their personal color interpretations. Thus the searchers have a probability distribution. Skilled catalogers may have a smaller distribution than searchers, but beyond a certain point the increased cataloging precision has no value in the meaningful resolution among instances. The variability of the searchers overwhelms the precision. As a consequence, the probability of the searcher discovering appropriate materials while excluding intrusions has a finite limit.

Different fields can support different size vocabularies. Well-structured vocabularies, including taxonomies, may support many more levels of precision, particularly if both catalogers and searchers agree upon the terms and definitions. But eventually the fuzziness overwhelms the precision such that increased vocabularies provide no increase in performance, instead creating a decrease in performance as the level of an increase in intrusions due to a reduced capability to provide meaningful resolution among items. At a certain point it is simply not possible to resolve the differences between two very similarly colored polygons based on the description in this field (the "color" metadata). It is necessary to turn to some other feature to discover the differences between these polygons.

Attempts to reduce system fuzziness usually focus on controlling the vocabularies used by the different communities. If the catalogers and searchers can agree on a common vocabulary with consistent definitions, fuzziness is reduced. Restricted vocabularies have narrower **resolution** and **precision**. The two concepts are closely related. When an online search produces a large number of hits that are poor matches for the users needs, it is frequently due to failures in resolution and/or precision in the metadata. *An efficient metadata system should reduce fuzziness without unnecessarily reducing resolution or precision.*

3.2. Dimensions

Differentiation is the ability to separate two items that appear identical along one dimension by using another dimension. If resources are described from different standpoints with smaller vocabularies, the chances of agreement on terms in any given field are higher, and the possibility of discovering a resource is greater. For our polygons above, if we use shape as a descriptor it is possible to reliably differentiate the polygons by the number of sides. Figure 1 demonstrates the manner in which different dimensions are combined to form a space.

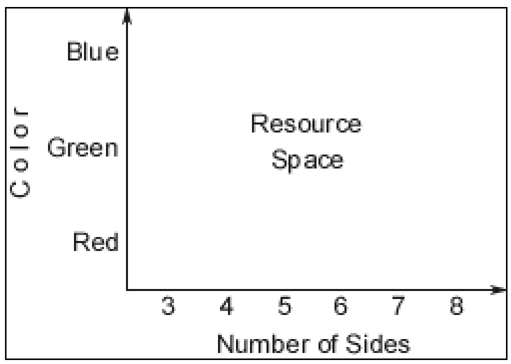

Figure 1. Space defined by multiple dimensions.

The number of sides is independent of the colors of the polygons. We can say that they are **orthogonal** dimensions. If we know the number of sides of a polygon, then we may know the relationship between its size and area. They are not completely dependent, nor are they orthogonal. The most efficient method for describing a space is with orthogonal vectors or dimensions. If the vectors are not independent, then more must be used to fully define the space. *An efficient metadata system strives to have as nearly independent dimensions as possible.*

3.3. Metadata Space Concepts

Here we present the underlying concepts of a **metadata space**, a geometrical analogy that will allow us to speak, albeit indirectly, about the properties of a metadata system.

3.3.1. Points, Lines and Spaces

We can consider resource discovery a process that explores geometrical space. The potential for unique identification of a resource is proportionate to the number of available *dimensions*. The more dimensions available, the more unique possibilities exist (e.g., a 10 x 10 x 10 cube has 1,000 unique addresses while a flat 10 x 10 grid has only 100). Thus, spaces can easily differentiate more items than lines, and points don't differentiate with respect to any sort of continua. The more uniquely defined an object is relative to other objects, the more easily it can be discovered and retrieved.

3.3.1.1. Dimensions

Dimensions are analogous to the resource metadata **fields**. A dimension is a measurement in a direction. The vocabulary terms used in the metadata fields are the metrics. For example, a ruler is able to measure along a line in some sort of standardized units, such as millimeters. Metadata can be analyzed according to its dimensions.

3.3.1.2. Metrics

Specific points along a dimension are defined by metrics. 1cm is a specific location relative to 0. The metrics that specify locations along the metadata field dimensions are not always ordinal. For example, learning style does not have an inherent order of terms. If there is no particular order to the dimension, then how is one to locate an object along it? Restricting the number of points on the dimension simplifies the task. For metadata fields, this means using finite, known vocabularies. If all resources are defined with a common set of terms, and if there are enough terms and dimensions to differentiate the resource well, then powerful discovery methods yielding small, focused sets of results are possible.

3.3.1.3. Points

Imagine that someone asked where a particular rock was, and you responded by saying "Oak." If you are standing in a field with a single oak tree in it, one can presume the rock is near the oak tree. The oak tree is a single point. If you don't know its context, it doesn't provide much information. Each word in a document can be considered a point of information. For example, if one searches for all content with the word "the", the return will be huge, but will provide little specificity. Sorting by selecting various points (terms) for filtering may allow one to access the largest number of resource items, but it provides no specificity, as the words are not metrics for any dimension. As Paul Saffo said, “It’s the context, stupid” (Saffo, 1994).

3.3.1.4. Lines

A line has a single dimension and can be resolved into smaller pieces or segments. One method of resolving a dimension is through hierarchies. Each finer level of granularity further resolves the dimension into finer levels, thus creating a hierarchy. A library can be considered to be one long shelf of books. The shelf is broken down into major sections that are further subdivided into minor sections, and so on. Hierarchies are logical structures for resolving dimensions. Hierarchies are useful methods of storing and retrieving information. The location on the "line" can be defined with increasing precision until the object is uniquely located. The Library of Congress cataloging system is a hierarchy.

Hierarchies have limitations. For instance, a book on armaments of World War II could be located under history (WWII) or under technology (armaments). A person wanting to locate information on WW-II armaments must decide how the "line" is resolved, and follow that hierarchy to locate the book. Libraries solve this problem through cataloging systems that cross-reference though subject descriptors.

3.3.1.5. Spaces

Resources can also in described by multiple dimensions. For example, a polygon may have a certain number of sides and be of a certain color. The number of sides can be varied independently from the color. Independent dimensions are **orthogonal**.

Obviously, the more orthogonal the dimensions, the more potentially differentiable positions are possible in the resulting space. The more dimensions one uses to define a space, the more likely one is to find that they are not truly independent. *An effective metadata system will utilize the right number of dimensions.*

3.3.1.6. Sub-Spaces

Orthogonal metadata fields serve the process of discovery. One may choose not to use all of the available metadata fields. For instance, one may search using only the fields of Subject, Content Type and Learning Level and a few key words to be used against the Abstract. As all of the other fields are not specified, such as Creator (e.g., author), the search takes place against all creators. Selecting values for some of the dimensions while ignoring others creates a sub-space. Thus, having fields that are not used by the searcher is not a penalty as the objects in the space are still accessible. If there are a large number of "hits", the searcher may choose to include additional fields to refine the search.

4. Definition Management

Stable definitions that can be commonly agreed upon are the basis of interoperability. Both fields and vocabularies need consistent definitions. Here we will address the creation of definitions for dimensions to create as nearly orthogonal dimensions as possible. Vocabulary definition is not addressed in this paper.

4.1. Semantic Structures

The following is adapted from Wiley (Wiley, 1999a, 1999b). A semantic structure is an organization that represents *meaning*. For example, an English sentence is a semantic structure. Consider the sentence structure: subject - verb – object. Before objecting by saying something like "the words have meaning, not the structure", consider two examples of the previous structure: (1) Dog bites man. (2) Man bites dog.

Obviously the words have meaning. But, two sentences containing the same words in the same structure can have very different meanings - depending on *where* in the structure the words appear. How do you know which position means what? Systems that use semantic structures (like languages) must supply a grammar, or way of interpreting meaning based on structure, in order to be useful.

You are asked to build a house out of blocks and given your choice of two different sets of blocks. The first set only contains five pieces: three walls, a wall with a door, and a roof. The second set are like Lincoln Logs, i.e., several different small pieces which can be hooked together by lining up the ends in a certain way. Consider the answers to the following questions: With which set could you build a house the quickest? With which set could you most easily add a new wing or garage if the need arose? Clearly, the Lincoln Logs allows you the highest degree of individuality and flexibility, as well as the greatest possibility for future expansion.

The LOM metadata standard is written like a set of Lincoln Logs. There is a dictionary of terms, like a bag of differently shaped logs, there is a grammar, like the rules for stacking the logs, and there is a schema, like a blueprint (one possible way of organizing the logs).

The LOM metadata standard contains a Master schema, or listing of many possible cabin features like walls, doors, windows, hot tubs, covered porches, tennis courts, etc. The schema is a hierarchical structure created from the elements in the dictionary. The dictionary is a list of elements with brief, non-contextualized definitions. We shall use the IEEE LOM *preliminary* draft metadata standard for examples.

There is a term in the LOM metadata dictionary called "person," meaning, "a specific human being." As in the dog bites man example, the term "person" itself has some meaning. However, it gets greater meaning by being placed in the following context:

Meta-metadata.Contribute.Person

The context, or structure, tells us that this particular person is the "specific human being who contributed to the metadata." In another position in the metadata structure, such as:

LifeCycle.Contribute.Person

Person now means the " specific human being who contributed to this version of the resource."

One may ask, "if the LOM is recommending a specific metadata schema, then why doesn't it just provide three walls, one wall with a door, and a roof instead of forcing us to build a metadata implementation from small, recombinant terms organized in semantic structures?" While the five-piece approach is the fastest, it is also a dead end. If we want to leave room for ourselves to improve the LOM metadata house in the future we need to build it from little blocks. There are a number of benefits to be reaped by utilizing a semantic structure approach to metadata -- many more than the extensibility issue addressed above.

An organization may extend the Master metadata set in some way. If that extension is done using terms from the LOM dictionary, a search agent that has never before encountered your particular metadata set could still understand it. How? Because it knows the meaning of the dictionary terms and the rule by which they are structured, it can infer meaning from a previously unencountered metadata schema. When you saw the example Meta-metadata.Contribute.Person above, if you knew the meanings of the words meta-metadata, contribute, and person, you probably had a pretty good idea what it meant even before we explained it.

You may extend the metadata set using terms outside the IMS dictionary, or structures other than those in the grammar. As long as you provide a dictionary and grammar that the search agent can access in order to learn the "meanings" of your terms and structure (link types), the agent could again infer meaning from the structure and interpret your metadata accordingly. The degree to which the elements in the dictionary have well-differentiated definitions will be reflected in the orthogonality of the dimensions. A major reason for using semantic structures is to manage the dimension (field) definitions.

4.2. Evolution

Evolution is not just about change. It is about change based on an underlying stability. Evolution has three main factors: change, stability, and selection.

The process of evolution may be summarized:

- An evolving system does *not* have a **goal**.
- It *may* have a **purpose**.
- Evolution produces an improvement in quality according to some **metric**.
- Adherence to the **metric** controls the **purpose**.

A brief summary on current thinking in evolution provides a useful model. The following is adapted from Wason (Wason, 1997). In the paleontological record, species appear to change in steps, rather than in a slow, continuous process. This step model is called "saltatory change".

In the natural, real-world process of evolution, one theory (e.g., Gould, et al.) states that evolution does not occur in the main population, but occurs in pioneer populations that splinter off from the main population. The pioneer populations are fairly isolated from the main population for some period. There may or may not be a connection for genetic communication between the main and pioneer populations.

A pioneer population will frequently occupy some *niche*. This may be by choice or due to external factors (e.g., a landslide blocks a mountain pass). The pioneer population is small, so the probability that a mutation will *persist* is increased relative to that in the main population. In addition, the pioneer population probably occupies an environment that differs from that of the main population. Mutations are of great value if they provide a survival advantage. If the pioneer population becomes re-connected to the main population, the pioneer population may serve to infect the main genetic pool with variant information that has been hammered out on the hard anvil of local survival. Or, in a more sinister fashion, the pioneer population may be so superior to the main population that it replaces it. This may occur through direct conflict or by more successfully competing for resources.

How does this all relate to the management of metadata? There are to be well-defined populations or communities of practice that may create changes in their own systems of metadata. These changes may propagate within the population. These populations interact with other populations in a well-controlled manner to select the best quality components for a continually improving system of metadata. Quality is defined as that which best fills the needs of the relevant population.

The metadata system is intended to support the entire evolutionary process in a state of control. The relationship between the main and pioneer populations will be illustrated in the IEEE LOM implementation, summarized below.

5. Relationship between Semantic Structures and Metadata

5.1. Structured Metadata

The metadata encoding, or binding, should reflect the definition structure. In this way the structure of the metadata directly reflects the semantic structure that defines the fields. Thus, it is interpretable by applications that have the semantic structure syntax and dictionary available. This supports both interoperability and extensibility.

5.2. Interoperability

Interoperability relies on the accessibility of stable definitions that are generated by a common method. Systems of structured metadata that use a well-defined binding, or method of encoding, that directly reflect a common semantic structure or method can be interpreted relative to each other. This provides the ability to map from one system into another, or create a crosswalk between the two. The two metadata systems do not need to be identical; they need to have non-conflicting dictionaries with as many common terms as possible, and adhere to a common syntax.

6. Filtering Model

The following is adapted from Wiley (Wiley, 1999c, 1999d). Most search interfaces do not take advantage of the multidimensional nature of metadata, providing nothing more than an open-ended, full-text search of the available metadata. And many of those search engines that do provide multidimensional resolution provide it a very unnatural way: they force simultaneous multidimensional resolution. For example, let's say that Jill is shopping for a house online. She may be able to search dimensions like location, number of bedrooms, square footage, and price range. The unnaturalness of the search comes from the way she is forced to make these many decisions at once. In the real world, most people probably go through an iterative process in which the resolution of the various dimensions (metadata fields) is spread out temporally. In Jill's case, perhaps she would decide first on a city and then view what is available. After observing current prices on homes in that area, she may next select a price range and see what is available within it. Each time she selects a value for a given metadata field, she reduces the metadata space in one dimension to a subregion of values, creating increasingly smaller and more relevant subsets. Four points merit special attention here.

First, Jill is creating a path by traversing a series of options. If she hits a dead end, instead of completely recreating the search she simply traverses up the path to the previous fork in the road, and heads another direction. All of the pertinent search information, i.e., the set of values creating the path to the fork where the ineffective decision was made, is preserved.

Second, a search that unfolds over time in this manner provides a greater opportunity for multiple metadata repositories to provide meaningful information to the searcher. For example, if Bill goes to the local mall to purchase a pair of shoes, he may not know exactly what he is looking for. Bill would probably not approach the information desk and say, ‘I would like a pair of white and blue high top shoes.’ He might ask the person at the information desk what stores sell athletic shoes. Once at a sporting goods store, he would ask the clerk where athletic shoes are on display. Bill would then look for a pair of shoes that appeals to him. If he doesn’t find a pair, he simply moves on to another store. Once he finds a pair, he checks the price to see whether or not he can afford them.

Third, inherent in this last example is the notion that the discovery process is only partially a search process, as opposed to the retrieval process, which is solely a search process. The **discovery** process is always part search, part browse. Once Bill reaches the athletic shoe section of the store, i.e., once he has collapsed the metadata space to a sufficiently small and relevant subspace, he stops searching *per se* and begins browsing.

Finally, also inherent in the shoe example is the potential economic characteristic of the iterative filtering model. When Bill arrives in the shoe section of the store, he is surrounded by a number of items relevant to his interest that he did not initially intend to purchase. Once he enters the subregion he is browsing not only shoes, but also other relevant resources such as shoelaces, T-shirts, shorts, water bottles, etc. This replicates point-of-purchase and other successful in-store advertising models.

7. Implementation

What organizations are using structured metadata spaces? There is a combined effort by a number of educational organizations to create a common educational metadata standard, and common method of managing metadata definitions.

7.1. Organizations

The IEEE Learning Technology Standard Committee (P1484) [Ref] is working to create a standard for Learning Object Metadata (LOM) using metadata spaces and semantic structures. Organizations participating in the effort include:

- ADL (Advanced Distributed Learning Initiative of the U.S. Dept. of Defense),
- AICC, (Aviation Industries CBT Committee)
- ARIADNE (Alliance of Remote Instructional Authoring and Distribution Networks for Europe),
- GESTALT (Getting Educational Systems Talking Across Leading-Edge Technologies), and
- IMS (IMS Project)

7.2. Audience

The IEEE LTSC LOM is designed to serve educators, learners, coordinators, and providers. Educators included teachers, tutors and trainers. Learners include K-12 students, home school students, college and university students, employees, military members, retirees, and other life-long learners. Coordinators include colleges and universities, public and private K-12 schools, the military, departments of labor, parents delivering home schooling, assessment specialists, and learners who are self managing. Providers include anyone that creates or edits materials such as professors, teachers, publishers, and students. The purpose of metadata is to help this audience discover and assess the appropriateness of resources.

7.3. Duplication

Duplication means that there are some fields that duplicate the information in other fields. The metadata space model encourages the creation of independent fields, or dimensions of the space. However, there is also a need to provide for the evolution of metadata. The model of evolution presented above features a main, stable population and one or more pioneer populations that differ from the main population. These pioneer populations in metadata may wish to create new fields and new vocabularies and taxonomies. How are stable metadata fields and evolving metadata fields managed together?

An overdetermined system is used to support evolution. IEEE LOM uses *Classification* to allow specialized taxonomies and vocabularies that may duplicate the field definitions elsewhere in the metadata hierarchy. For example, the IEEE LOM has a top-level element, called a "category", of *Educational*. This has a definition of "Educational or pedagogic features of the resource". *Educational* contains the following fields:

Educational

Typical Age Range

Difficulty

Difficulty is considered by many not to be independent of *Typical Age Range*. There is no standard taxonomy of difficulty and there is no standard method for defining an age range. *Educational Difficulty* and *Educational Typical Age Range* are considered by this community to be too important to be omitted. There is a desire to move slowly toward some standard taxonomy.

The IEEE LOM takes a two-pronged approach: (1) it created the two simple fields above, with limited controlled vocabularies, (2) it created an additional structure under a category element of *Classification*. The controlled vocabularies of the predefined fields are quite simple: *Typical Age Range* as values of one or two integers, and *Difficulty* has integer values of 0 to 4, with 0 being "Very Easy", and 4 being "Very difficult." These vocabularies are inadequate, but provide a consistent starting point.

Classification is defined as "Description of a characteristic of the resource by entries in classifications." This is a structure that does not directly create a specific definition. Instead, it contains a sub-element of *Purpose* to specify the type of classification. It has an unrestricted vocabulary including: *Discipline, Idea, Prerequisite, Educational Objective, Accessibility Restrictions, Educational Level, Skill Level,* and *Security Level*.

Classification has the structure:

- Classification (unordered list)
  - Purpose (Single)
  - Description (unordered list of LangStrings)
  - Keywords (unordered list of LangStrings)
  - Taxonpath (unordered list)

Source (single)

Taxon (ordered list)

Instead of the completely defined definition that is created by other parts of the semantic structure, this uses a dependent "clause" (i.e., *Purpose*) to create the full definition.

This is effectively an extension without requiring an extension to be created; new classification systems can readily be added, such as security or skill level, without the need to extend the metadata structure. More than one instance of *Classification* and its structure can be used (an unordered list). Referring to the evolutionary model above, *Classification* becomes a potential mechanism for the "pioneer" efforts in using new taxonomies and vocabularies, while the controlled vocabularies elsewhere in the structure provide stability.

The IEEE LOM standard should be submitted for vote by the publication of this paper; it is slated for a 19 July 1999 submission. This standard will be available for use by anyone. The organizations listed above will all be subscribing to this standard, providing a truly international standard for educational metadata that will support future growth.

8. Conclusion

Thinking about metadata as a space having structure provides a model for creating and using metadata fields. Orthogonal metadata fields provide the most efficient use of that space. Semantic structures provide a method of managing the definitions of the fields. Semantic structures provide a basis for interoperability among metadata systems, as interoperability is largely a matter of determining equivalence between metadata fields in different systems. Metadata encoding that captures the semantic structures helps to convey the field definitions among applications.

References

ADL: Advanced Distributed Learning Initiative, http://www.adlnet.org

AICC: Aviation Industries CBT Committee, http://aicc.org

ARIADNE: http://ariadne.unil.ch/

DC: Dublin Core, http://purl.org/DC

IMS: Instructional Management Systems, http://www.imsproject.org

IEEE (Institute of Electrical and Electronic Engineers) LTSC (Learning Technology Standards Committee, P1484), http://ltsc.ieee.org/wg12/

GEM: Gateway to Educational Materials, http://gem.syr.edu/

GESTALT: Getting Educational Systems Talking Across Leading-Edge Technologies, http://www.fdgroup.co.uk/gestalt/metadata.html

GILS: Government Information Locator Service, http://fiat.gslis.utexas.edu/~ssoy/pubs/l391d2a.htm

MARC: MAchine Readable Cataloging, http://lcweb.loc.gov/marc/marbi/dp/dp86.html

Saffo, Paul. (1994). *It's the context, stupid* [On-line]. Available: http://www.saffo.org/contextstupid.html

Wason, Thomas D. (1998). *Metadata spaces* [On-line]. Available: http://www.imsproject.org/technical/metadata/History/did054.html

Wason, Thomas D. (1997). *An evolutionary model for TIERS* [On-line]. Available: http://www.imsproject.org/technical/metadata/Tiers/did159.html

Wiley, David A. (1999a). *Why semantic structures?* [On-line]. Available: http://ims.byu.edu/semantics.html

Wiley, David A. (1999b). *My, what an intelligent tool you have!* [On-line]. Available: http://ims.byu.edu/intelligent.html

Wiley, David A. (1999c). *An intelligent method for searching metadata spaces* [On-line]. Available: http://wiley.byu.edu/search/if-search.pdf

Wiley, David A. (1999d). *Metadata mall: Two arguments for the IF-search technique* [On-line]. Available: http://wiley.byu.edu/search/mm.html