Scheme Versioning in the Semantic Web

Joseph T. Tennis

SUMMARY. This paper describes a conceptual framework and methodology for managing scheme versioning for the Semantic Web. The first part of the paper introduces the concept of vocabulary encoding schemes, distinguished from metadata schemas, and discusses the characteristics of changes in schemes. The paper then presents a proposal to use a value record–similar to a term record in thesaurus management techniques–to manage scheme versioning challenges for the Semantic Web. The conclusion identifies future research directions. doi:10.1300/J104v43n03_06 *[Article copies available for a fee from The Haworth Document Delivery Service: 1-800-HAWORTH. E-mail address: <docdelivery@haworthpress.com> Website: <http://www.HaworthPress.com> © 2007 by The Haworth Press, Inc. All rights reserved.]*

KEYWORDS. Indexing, classification, versioning, controlled vocabulary management, indexing language management

STRUCTURES IN THE SEMANTIC WEB

The Semantic Web is a collection of structures that work together (or interoperate). These structures constrain the range of possible interpretations in order to enable the transfer of meaning across the Web.

Schemas and schemes work together to constrain meaning. Schemas are the total set of assertions that can be made about a resource. A resource is anything that is addressable on the Web. For example, a surrogate for a person is a resource as much as a Web page is a resource. Both of these are addressable by URIs–Uniform Resource Identifiers (Connolly, 2006). The Dublin Core Element Set Version 1.1 (ISO Standard 15836-2003) is a particular type of schema. See Table 1.

1. Title
2. Creator
3. Subject
4. Description
5. Publisher
6. Contributor
7. Date
8. Type
9. Format
10. Identifier
11. Source
12. Language
13. Relation
14. Coverage
15. Rights

TABLE 1. Fifteen Elements Dublin Core Element Set–An Example of a Schema

Schemas work with schemes. Schemes are the range of values that can be provided for an assertion about a resource (date-time format, authority list, controlled vocabulary, etc.). There are two types of schemes: Vocabulary encoding schemes and Syntax encoding schemes. Vocabulary encoding schemes indicate that the value is a term from an indexing language, such as the value "China–History" from the Library of Congress Subject Headings. Syntax encoding schemes indicate that the value is a string formatted in accordance with a formal notation, such as the ISO 8601 date-time format, "2000-01-01" as the standard expression of a date. An example of the value taken from Vocabulary encoding scheme is given in Table 2 below.

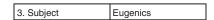| 3. Subject	Eugenics

TABLE 2. Value Provided for the Subject Elements in the DCES

This value, *eugenics*, came from a vocabulary encoding scheme. It came from the Library of Congress Subject Headings (LCSH). We could have taken the same subject from a different vocabulary encoding scheme, the Dewey Decimal Classification, for example. That might look like this, in Table 3 below:

3. Subject	176

TABLE 3. Value Provided for the Subjet Element of the DCES

As we can see from these examples, schemas allow us to say that a resource has an attribute (a subject), and a scheme allows us to make explicit what that subject is (the value of that attribute). However, we can also see another level of complexity that arises from this act of metadata creation.

We must now say where this value comes from. Furthermore, it is not enough to say that the value in Table 3 came from the Dewy Decimal Classification (DDC), but we also have to say what edition of the DDC. In so doing, we cite a place that one can reference this value. Furthermore, by declaring one edition instead of another, we also call into question the potential and demonstrable change in meaning from edition to edition, such that notation 176 (hereafter all occurrences of three digits together, or three digits followed by a decimal are DDC notations) in DDC means something different in different editions. The same can be said of terms in the relative index, and their position in the schedules. Table 4 provides an example how eugenics has changed in discipline and other characteristics in DDC.

Eugenics		100	300	500	600
Year	*Edition and relative index entry*				
1911	**7th**		575.6		
1942	**14th rev. enl.**				
	Eugenic method crimol.		364.3018		
	crime prevention		364.42		
	evolution			575.1	
	hygiene				613.94
	mental psychology	136.3			

Eugenics		100	300	500	600
Year	***Edition and relative index entry***				
1958	**16th**			613.94	
	Eugenic practices crime prevention		364.42		
	Hygiene				613.94
		[301.323 officially killed]	[575.1 officially killed]		
2003	**22nd**	363.92			
	crime prevention		363.4		
	ethics	176			
	health				613.94
	social services		363.92		
	sterilization services		363.97		

TABLE 4. Eugenics in Four DDC Schedules

In this table we see eugenics contextualized in disciplines, and refined with additional words. We can speculate on the many reasons why eugenics has changed over the years. DDC does provide us with some general reasons for change (OCLC, 2006). We can interpret, in this case, that different aspects of eugenics have surfaced over the years, the term has been used in different ways in the literature (and DDC wants to reflect that change both in scholarship and in viewpoint), and they want to reduce bias. To reflect these differences, the entry for eugenics in the relative index points to many different places in schedules–and in some cases, no longer points to places it once did.

In Semantic Web applications using DCES, we might use many versions of a scheme. For example, in a hundred years of creating metadata for resources we are likely to go through as many versions of DDC as we have in the past hundred years. We would then see instances of metadata from DDC 22 through DDC 37, and it is also possible, with digitization projects, to see resources with metadata back to DDC 7 online. Further, DDC is not the only scheme that goes through revisions. Every vocabulary encoding scheme that is kept up to date is revised. They are revised so they can maintain their purpose: to retrieve documents; or to couch it in terms of bibliographic control: to find, collocate, and identify resources.

If schemes are used to find, collocate, and identify items in the Semantic Web applications, like they are in catalogues, then we must be

sensitive to the versioning of these schemes, and the changes affecting the values from these schemes. In some domains, schemes change even more rapidly than DDC, and in a distributed networked environment, managing the semantics of these changes is vital to the functioning and utility of these schemes. To that end, scheme versioning is a management activity that requires a conceptual framework and a methodology. It is important that this conceptual framework and methodology are in place before more metadata is created. As we have learned for electronic records systems, the conceptual models must be built into the metadata for the system to fulfill its purpose (Duranti, Eastwood, and MacNeil, 2002).

This paper looks at the conceptual framework and one consequent methodology in order to aid interoperability on the Semantic Web synchronically and diachronically–at one point in time, but also through time.

METHODOLOGIES AND CONCEPTUAL FRAMEWORKS FOR SCHEME CHANGE

The methodologies for scheme versioning management might be simple or they might need to be more robust. It is possible that noting a change is enough in some contexts. This depends on the purpose of the scheme, the system that employs the scheme, and the professionals and users engaged with the system. For example, if the purpose of the scheme were to represent the subject matter of a resource coextensively (completely and expressing every aspect of the subject), then professionals and users alike would benefit from a robust scheme version management methodology–one where differences and similarities between the old and the new scheme were made explicit. If the scheme were not semantically rich (not used to represent many aspects of the resources, for example), then a less robust scheme version management methodology might be all that was required.

Regardless of methodology, scheme version management requires a conceptual framework in order to understand (1) the phenomenon of change in schemes, (2) what characteristics of change need to be made explicit in a methodology to manage this change, (3) to serve the purpose of scheme viability through versioning. Scheme viability is an important consideration here. Schemes are built for particular purposes, and versioning should not change those purposes. If anything, the purposes of schemes should be strengthened through change. Managing

that change will ensure constant and consistent improvement. The follow section outlines the conceptual framework of scheme change.

CONCEPTUAL FRAMEWORK FOR SCHEME CHANGE

This section outlines the characteristics of scheme change in order to better represent the phenomenon of scheme change. We address change in relation to scheme viability at the end of this section.

Characteristics of Scheme Change

Scheme change occurs in three general categories: structural change, word-use change, and textual change. Structural change deals with the relationship structures in schemes and how editors alter them. Word-use change affects definitions, word forms, lead-in terms, etc. Though both structural changes and word-use changes are semantic, the latter do not explicitly affect relationship structures. Textual changes can affect both structural and word-use changes. Textual changes are changes in the interpretation and assignment of values to types of resources. The first two fall into the purview of the editors, while the third falls to both the editor and indexers. The following characteristics of scheme changes are adapted from Soergel (1974), Aitchison, Gilchrist, and Bawden (2000), Ranganathan (1967), and Beghtol (1986).

Structural Change

Structural changes affect a user's navigation through the scheme. Structural changes affect the semantics of a scheme because they change the relationships that obtain between values in that scheme. Structural change falls into five basic changes.[1] The five basic changes are:

- Addition of a new value
- Change in synonym structure (use eugenics to lead to both genetics and psychology)
- Change in equivalence structures (e.g., USE and/or USED FOR)
- Assignment of value to another group in the hierarchy
- Addition or elimination of associative relationship (e.g., RT).

The degree to which these changes affect indexing or classification is dependent on the purpose and structure of the scheme before the change.

That is, if a scheme is a thesaurus built on principles of mutual exclusivity (only one place for each concept–no overlap) then these changes are dramatic. If the scheme is not built on principles of mutual exclusivity, then navigation is hindered, but not confounded through these changes. In either case, it is desirable in a digital environment–in a Semantic Web–to track these characteristics of change in order to mange the meaning communicated through indexing and retrieval process.

Word-Use Change

The second type of change is word-use change. Word-use changes do not affect navigation through the structure. They are changes that preserve the structure of a scheme, while adding or replacing words. This may affect indexing practice, but it does not affect the scheme structurally. Word-use changes are:

- New word used as lead-in
- New synonyms added (replaced one for one, for example, *genetics* for *eugenics*)
- New preferred value added
- Change in definition of value

Like structural changes, combinations of these changes can occur. The effect of word-use on scheme versioning is powerful. In our example of eugenics, the lead-in terms, the synonyms, and the definition all affect the use of the value. Eugenics has been a concept that has affected a number of areas of science, social science, and philosophy. How words are used to present this concept affect the way it will be used by the indexer. One can also imagine a scenario where a value may be present in the scheme, but not used because of the definition. If this remains a constant in the use of the scheme, then this has ramifications for the structure. The value may disappear for example if its not used. So it is not structure alone that affects structural changes. Textual changes also affect structural changes, as well as word-use changes.

Textual Changes

Textual changes are changes in the relationships between texts and a version of the scheme. There are two primary types of textual changes. The first is textual warrant change and the second is the document-set change. Textual warrant is a term that is close to literary warrant–but

does not mean the same thing. Textual warrant is the combination of all texts (literature of the field, user studies, search logs, checklists, etc.) that would be used to create a value or relationship in a scheme. Soergel calls these sources and authorities (Soergel, 1974). Any change in this collection of texts results in a change of the evidence considered when managing the scheme, and hence managing changes to the scheme.

The second kind of textual change is the document-set change. In this case a set of documents has been indexed and given a value (for example, 575.6). This set will change as the scheme changes, and therefore shifting the representation power of the scheme. So the texts once classed under 575.6 are not the same kind of texts, because the relationship between the document set and the value has changed.[2]

Scheme Viability

Schemes are built for the purpose of information retrieval–to find, collocate, and identify resources. They do this by establishing a set of values and relationships between values. Changes to a scheme must strengthen this purpose, yet in the context of multiple versions of schemes, it is not always clear what values mean, and what relationships obtain between values. A clear methodology of tracking changes made to values in schemes enables a scheme to carry out its mission across various versions.

Methodology

How does one track these changes? In order to track changes, each value must be identified as an entity in relation to other values at a point in time. Also, each characteristic of change must be accounted for–so that structural, word-use, and textual changes can be made explicit. To create an explicit statement about values in this way is to create a value record, like a term record used in thesaurus construction and maintenance. The next section identifies the components of a value record for the purpose of tracking changes in schemes.

VALUE RECORDS

Scheme versioning, tracked through changes in values, can be managed with value records, an expansion on term records. Thesaurus management manuals and standards suggest the use of a term record to manage values (Soergel, 1974; Aitchison, Gilchrist, and Bawden, 2000;

ANSI/NISO, 2005; Anderson and Perez-Carballo, 2005a). An edited example of Soergel's term record is provided in Table 5 below.

01. Hierarchical Level	01. When terms are later sorted into hierarchies, based on BT and NT descriptors, each term will fall at a particular hierarchical level.
02. Type: DS, OP, NP, EL, CH	02. These codes indicate the current status of the main term in field 10: DS = descriptor (authorized term); OP = other preferred term (but not adopted as an authorized descriptor); NP = non-preferred term; EL = eliminate term; CH = change term information.
03. Subject Field	03. In order to find different terms indicating the same, or essentially the same concept, terms must be sorted conceptually. The subject field is the first large category for conceptual sorting. For faceted thesauri, these first level categories will be the main facets.
05. Notation	05. Later, when cards are sorted into final conceptual order, a notation can be assigned to maintain this order.
10. MT	10. This is the main term for this card. All the information on the card will relate to this term.
12. Standard Abbreviation	12. A standard abbreviation for a term is often helpful to indexers, who can use it to save time. Later, before an index is prepared for users, most abbreviations would be expanded to the full standard form. (Abbreviations can be the standard form when they are better known, as with acronyms such as "radar" and "Unesco.")
20. Variant Spellings	20. Variant spellings go here (as well as variant abbreviations).
30. Synonymous Terms (ST), including Equivalent Terms (ET)	30. Synonymous and equivalent terms go here.
40. Classification	
42. Category (CA)	40./42. This field can be used for finer categorization within the broad subject field, noted in field 03.
44. Broader Terms (BT)	44. Broader terms go here.
45. Narrower Terms	45. Narrower terms go here.
46. Related Terms	46. Related terms go here.
50. Translations	50. If the thesaurus is to be multilingual, than the equivalent terms in others languages go here.
60. Definition, Scope Note	60. A definition of the term, if needed, or a scope note explaining the usage of the term in the indexing language, goes here.
65. Sources/Authorities	65. Here is recorded the source of the term, or the authority for the definition/scope note.
70. Unspecified Relation (UN)	70. Any terms whose relationship to the main term has not yet been determined can go here.
81. Editor/Date	81. The name or initials of the thesaurus editor, plus the date, go here.

TABLE 5. Soergel's Term Record from Anderson and Perez-Carballo (2005b)

In this term record Soergel has given each area its own numerical notation. The 0x area uses numerical codes to place this term record in relation to others. 1x area identifies the main term. 2x variant spellings, 3x identifies types of equivalence relationships, 4x syndectic relationships (broader, narrower, and related terms), 5x translations if needed, 6x definition, scope note, and sources/authorities that give the term meaning,

7x provides a space for terms with unspecified relationships to the main term, and 8x, identifies the author and date of creation for this thesaurus. As can be seen from the comments on the right side of the table, this term record is used primarily for the creation of thesauri. Yet we could easily transfer these areas to other types of schemes besides thesauri, and identify values instead of terms. These features could be used as starting points for creating a value record that would account for characteristics of scheme change. We would make some additions. For example, Aitchison, Gilchrist, and Bawden (2000, 148-149) suggest an additional field: Frequency of Occurrence. This would identify how many times the value has been used. Another aid to creating a value record is the SKOS specification.

SKOS–Simple Knowledge Organisation System is a w3c area of work that develops specifications and standards to support the use of schemes in the Semantic Web (Miles, n.d.). SKOS creates these specifications in line with RDF (Miller, Swick, and Brickely, 2006) and OWL, two other w3c metadata recommendations. SKOS provides a set of assertions that can be interpreted as a value record. Table 6 outlines a selection of the SKOS assertions.

Types of Labels	
skos:prefLabel,	Preferred label for a concept
skos:altLabel,	Alternative label of a concept
skos:hiddenLabel,	Hidden label accessible to applications only for search
Types of Notes	
skos:definition	A complete explanation of the intended meaning of a concept
skos:scopeNote	Some, possibly partial, information about the intended meaning and/or use of a concept, which can be phrased as information about what is or isn't included within the meaning ('scope') of the concept.
skos:example	An example of the use of a concept
skos:historyNote	Instructions or useful information for users of the scheme, specifically relating to significant changes to the meaning/form/state of a concept.
skos:editorialNote	Information that is an aid to administrative housekeeping, such as reminders of editorial work still to be done, or warnings in the event that future editorial changes might be made.
skos:changeNote	Fine-grained changes to a concept, for the purposes of administration and maintenance.
Types of Relationships	
skos:semanticRelation	
skos:broader	Broader relationship
skos:narrower	Narrower relationship
skos:related	Associative relationship

TABLE 6. SKOS Assertions

SKOS provides some of the same assertions as Soergel, for example, the Types of Labels mimic his term record. Yet, in Table 6 we see more fields designed to manage change in a scheme. The Types of Notes assertions provide history notes, editorial notes, and change notes. These structures provide us with information on versioning. Thus the Soergel term record provides us information on relationships and definitions, the Types of Notes in SKOS provide us with the ability to create evidence of change. The combination of these two sets of assertion, these records, will allow us to make explicit relationship between values in different versions of schemes. These are the first steps in managing scheme versioning.

The next step for the effective use of a value record is to create explicit statement of these characteristics. The next section combines the Soergel and SKOS term records, with suggestions from Aitchison, Gilchrist, and Bawden to create a value record–a record that can account for scheme change.

Scheme Change Value Record

If we were to combine elements from the value records above, and make explicit in these records how to handle the characteristics of scheme change, we would be one step close to managing semantic change on the Semantic Web. Table 7 outlines what a value record would look like if it accounted for the characteristics of scheme change. The major changes are shown in italics.

01. Hierarchical Level	01. When values are later sorted into hierarchies, based on BT and NT descriptors, each value will fall at a particular hierarchical level.
02. Type: DS, OP, NP, EL, CH	02. These codes indicate the current status of the main value in field 10. For example in a thesaurus we might have: DS = descriptor (authorized term); OP = other preferred term (but not adopted as an authorized descriptor); NP = non-preferred term; EL = eliminate term; CH = change term information (which can be subdivided by versioning or not versioning)
03. Subject Field	03. In order to find different values indicating the same, or essentially the same concept, terms must be sorted conceptually. The subject field is the first large category for conceptual sorting. For faceted thesauri, these first level categories will be the main facets.
04. Classification	04 This field can be used for finer categorization within the broad subject field, noted in field 03.
05. Conceptual Notation	05. Later, when cards are sorted into final conceptual order, a notation can be assigned to maintain this order.

06. Scheme	*06. Scheme name and identifier (citation to the scheme or a URI).*
07. Version of the Scheme	*07. Version of the scheme. For example Edition 22 of the DDC.*
08. Date of Version of Scheme	*08. Date of Version of Scheme*
09. Accession Number	09. A unique number given to the value record when it is created.
10. Main Value	10. This is the main value (term) for this card. All the information on the card will relate to this value (term).
11. Version Number	*11. Version of the value (corresponding to the scheme, but attached here to each term). For example, 22 for edition 22 of DDC.*
12. Standard Abbreviation	12. A standard abbreviation for a value is often helpful to indexers, who can use it to save time. Later, before an index is prepared for users, most abbreviations would be expanded to the full standard form. (Abbreviations can be the standard form when they are better known, as with acronyms such as "radar" and "Unesco.")
15. First Addition Version Number	*15. Version number of value when term was added. (May be redundant to 11 until changes occur)*
16. Deletion Version Number	*16. Version number of value when it was deleted.*
18. Addition of Main Value	*18. Main value and version number of value (May be redundant with 10, until changes occur).*
19. Deletion of Main Value	*19. Main value and version number of value.*
20. Variant Spellings	20. Variant spellings go here (as well as variant abbreviations).
30. Synonymous Values	30. Synonymous and equivalent values go here.
31. Equivalent Values	31. Equivalent Values go here.
32. Addition of Synonyms	*32. Synonyms and version number of value.*
33. Deletion of Synonyms	*33. Synonyms and version number of value.*
34. Addition of Equivalence Structure	*34. Equivalent value and version number of the value.*
35. Deletion of Equivalence Structure	*35. Equivalent value and version number of the value*
36. Addition of Lead-in Values	
37. Deletion of Lead-in Values	
40. Broader Values	40. Broader values go here.
41. Narrower Values	41. Narrower values go here.
42. Related Values	42. Related values go here.
43. Addition of Broader Values	*43. Broader value and version number the value.*
44. Deletion of Broader Values	*44. Broader value and version number the value.*
45. Addition of Narrow Values	*45. Narrower value and version number of the value.*
46. Deletion of Narrow Values	*46. Narrower value and version number of the value.*
47. Addition of Related Values	*47. Related Value and version number of the value.*
48. Deletion of Related Values	*48. Related Value and version number of the value.*
50. Translations	50. If the scheme is to be multilingual, than the equivalent values in others languages go here.
60. Definition	*60. A definition of the value*
61. Scope Note	*61. Scope note explaining the usage of the value in the indexing language, goes here.*

62. History Note	62. Instructions or useful information for users of the scheme, specifically relating to significant changes to the meaning/form/state of a value.
63. Change Note	63. Narrative of changes to a value, for the purposes of administration and maintenance.
64. Editorial Note	64. Information that is an aid to administrative housekeeping, such as reminders of editorial work still to be done, or warnings in the event that future editorial changes might be made.
65. Sources/Authorities	65. Here is recorded the source of the term, or the authority for the definition/scope note.
66. Addition of Sources/Authorities	66. Sources/Authorities for the value and its version
67. Deletion of Sources/Authorities	67. Source/Authorities for the value and its version.
68. Addition of Definition	68. Definition and version number of the definition (from main value record)
69. Deletion of Definition	69. Definition and version number of the definition (from main value record)
70. Unspecified Relation (UN)	70. Any terms whose relationship to the main term has not yet been determined.
80. Editor	80. The name of the scheme editor
81. Record Creator	81. The name of the value record creator
86. Contact Information	86. Contact information for the scheme editor
87. Date Record Created	87. Date the record was created.
90. Information System(s)	90. Information system(s) that use this value (and its URI)
91. Frequency of Use	91. Number of times value has been used
92. Document-Set Date	92. Date Document-Set was created
93. Document-Set Resources	93. Set of documents indexed by the value. Provide the citations and if possible locators for these resources

TABLE 7. A Value Record

The Soergel term record has been rearranged slightly to bring the areas of the value record into a more meaningful sequence. As can be seen from Table 7 above, changes in schemes are operationalized as simple additions and deletions. In order to track the changes a scheme goes through, a value record should track the additions and deletions of values in relation to other values, word-use attached to the value, and its link to texts (both texts indexed and sources for values). The following section describes how structural, word-use, and textual change can be reflected in an example value record modeled off of Table 7 above.

The Value Record and Characteristics of Scheme Change

In order to make the term record meaningful to scheme version management, we first have to add a version number to each value in a scheme.

This appears as assertion 11 in the value record–Table 7. This allows us to track each value as a separate entity, but still tied with all other entities in the same version of the scheme (assertion 06 and 07 Scheme and Version of Scheme respectively).

Structural Change Reflected in the Value Record

All five types of structural changes are represented in the value record. These assertions should be repeatable as often as needed. To reflect the addition of a new value to a scheme, the editor can use assertion 15 Addition of Value and signify the version number, placing it in a sequence with other values. To change synonym structure addition and deletion happens in assertions 32 and 33 (Addition and Deletion of Synonyms). The same goes for equivalence structures in assertion 34 and 35, and associative relationships in assertions 47 and 48. In order to show a change in hierarchical grouping the value must be deleted from one set of broader and narrower relationships to a new set. This requires four types of assertions: addition and deletion from broader values, and addition and deletion of narrower values (assertions 43-46).

Word-Use Change

Changes in word-use are similar, in that they follow the addition/deletion format. So that new lead-in values, new synonyms, new preferred values, and changes in definitions. These additions and deletions are above in assertions 36-37, 32-33, 18-19, and 68-69 respectively.

Textual Change

Textual changes are reflected through sources and authorities and in document sets created at a particular point in time. For Textual Warrant change, assertions 66 and 67 (Addition and Deletion of Sources/Authorities) make explicit which sources and authorities were used for which version of the value. For Document-Set Changes, once the frequency of use is added (assertion 91), then it is possible to compare this number with set of documents (or resources) that are and have been indexed with this particular value. The set of documents is represented in assertion 93, Document-Set Resources, and it is dated in assertion 92, Document-Set Date.

Example: Value Record of Eugenics in DDC

In order to illustrate the addition and deletion aspects of the value record, we present examples using the value 613.94, from the DDC. The first, in Table 8, points to a hypothetical record for a value created for the 1942 14th revised and enlarged edition of the DDC. Table 9 that follows is a value record created for the 22nd edition.

01. Hierarchical Level	
02. Type: DS, OP, NP, EL, CH	DS
03. Subject Field	Medical Sciences
04. Classification	613.94
05. Conceptual Notation	613.94
06. Scheme	Decimal Classification
07. Version of the Scheme	14th Revised Enlarged Edition
08. Date of Version of Scheme	1942
09. Accession Number	
10. Main Value	613.94
11. Version Number	14 rev. enl.
12. Standard Abbreviation	N/A
15. First Addition Version Number	14 rev. enl.
16. Deletion Version Number	
18. Addition of Main Value	613.94
19. Deletion of Main Value	
20. Variant Spellings	
30. Synonymous Values	Stirpiculture. Eugenics.
31. Equivalent Values	
32. Addition of Synonyms	
33. Deletion of Synonyms	
34. Addition of Equivalence Structure	
35. Deletion of Equivalence Structure	
36. Addition of Lead-in Values	Eugenics (from Relativ Index 14th rev enl edition)
37. Deletion of Lead-in Values	
40. Broader Values	613.9 Hygiene of Offspring. Heredity.
41. Narrower Values	
42. Related Values	613.91 Congenital Defects of the Body
42. Related Values	613.92 Inherited Mental Disability
42. Related Values	613.93 Transmitted Disease
42. Related Values	Heredity in 575.1

43. Addition of Broader Values	
44. Deletion of Broader Values	
45. Addition of Narrow Values	
46. Deletion of Narrow Values	
47. Addition of Related Values	
48. Deletion of Related Values	
50. Translations	
60. Definition	
61. Scope Note	Stirpiculture. Eugenics. [editors might see more than what is in the schedule]
62. History Note	
63. Change Note	
64. Editorial Note	
65. Sources/Authorities	
66. Addition of Sources/Authorities	
67. Deletion of Sources/Authorities	
68. Addition of Definition	
69. Deletion of Definition	
70. Unspecified Relation (UN)	
80. Editor	Constantin J. Mazney, editor; Myron Warren Getchell. associate editor.
81. Record Creator	Joseph T. Tennis
86. Contact Information	
87. Date Record Created	2006-04-01
90. Information System(s)	
91. Frequency of Use	
92. Document-Set Date	
93. Document-Set Resources	

TABLE 8. Value Record for a Class14th Revised and Enlarged Edition of DDC

01. Hierarchical Level	
02. Type: DS, OP, NP, EL, CH	DS
03. Subject Field	Technology–Medicine and Health–Personal Health and Safety
04. Classification	613.94
05. Conceptual Notation	613.94
06. Scheme	Dewey Decimal Classification
07. Version of the Scheme	22nd Edition
08. Date of Version of Scheme	2003

09. Accession Number	
10. Main Value	613.94
11. Version Number	22
12. Standard Abbreviation	
15. First Addition Version Number	14 rev. enl.
16. Deletion Version Number	N/A
18. Addition of Main Value	613.94
19. Deletion of Main Value	
20. Variant Spellings	
30. Synonymous Values	Birth control and reproductive technology
31. Equivalent Values	363.96 22nd
32. Addition of Synonyms	
33. Deletion of Synonyms	Stirpiculture. Eugenics. 14th revised and enlarged edition
34. Addition of Equivalence Structure	
35. Deletion of Equivalence Structure	
36. Addition of Lead-in Values	Eugenics–heath (from Relative Index 22nd edition)
37. Deletion of Lead-in Values	Eugenics (from Relativ Index 14th rev enl edition)
40. Broader Values	Birth control, reproductive technology, sex hygiene
41. Narrower Values	613.942 Surgical. . .
41. Narrower Values	613.943 Chemical. . .
42. Related Values	
42. Related Values	
42. Related Values	
42. Related Values	
43. Addition of Broader Values	
44. Deletion of Broader Values	613.9 Hygiene of Offspring. Heredity.
45. Addition of Narrow Values	
46. Deletion of Narrow Values	613.91 Congenital Defects of the Body
46. Deletion of Narrow Values	613.92 Inherited Mental Disability
46. Deletion of Narrow Values	613.93 Transmitted Disease
47. Addition of Related Values	
48. Deletion of Related Values	Heredity in 575.1
50. Translations	
60. Definition	
61. Scope Note	Stirpiculture. Eugenics. [editors might see more than what is in the schedule]
62. History Note	
63. Change Note	
64. Editorial Note	
65. Sources/Authorities	

66. Addition of Sources/Authorities	
67. Deletion of Sources/Authorities	
68. Addition of Definition	
69. Deletion of Definition	
70. Unspecified Relation (UN)	
80. Editor	Mitchell, Joan S., Beall, Julianne, Martin, Giles., Matthews, Winton E., New, Gregory R.
81. Record Creator	Joseph T. Tennis
86. Contact Information	
87. Date Record Created	2006-04-01
90. Information System(s)	
91. Frequency of Use	
92. Document-Set Date	
93. Document-Set Resources	

TABLE 9. Value Record for 22nd Edition of DDC

In Table 9 we see how the value is changed, but retains its links back to the value and relationships in the earlier version (Edition 14 rev. enl.). We see what was deleted, and what was added. This is the first step in aiding semantics of the Semantic Web across versions of schemes. By making these changes explicit–both in this value record, and in conceptualizing the categories of change we put ourselves in a better position to manage the dynamic and quickly evolving world of the Semantic Web.

Future Work

The next stages of research and development in the Semantic Web will test these conceptualizations and methodologies, and then move them into machine-readable assertions about resources. Allowing editors to manipulate different versions of schemes for management and retrieval purposes.

CONCLUSION

The National Science Digital Library has embarked on a Registry project where they aim to store schemes for reuse and interoperability (Sutton and Hillmann, 2006). As more schemes are added to this regis-

try, and as different versions of those schemes are added to the registry, researchers and developers at NSDL will have to wrestle with scheme versioning. This registry must account for the structural, word-use, and textual changes schemes go through. This scheme registry will be a proof of concept application for scheme versioning in a Semantic Web environment. They are building metadata for schemes that move beyond SKOS–pushing the edge of what we know about contemporary scheme management. It is hoped that the value record, conceptual framework, and methodology developed in this paper will be an aid in their work, and others, toward developing a more robust and meaningful Semantic Web.

Schemes change, and if the Semantic Web is going to be a Web of meaning, then it must be a web of meaning that is dynamic, not static. If schemes form a major part of Semantic Web metadata, then we have to account for scheme versioning in ways that are both human and machine-readable. The value record presented above is a step toward making the elements of scheme versioning explicit so that we might construct management tools to aid in a dynamic and meaningful development of the next web, the Semantic Web.

NOTES

1. Other structural changes beyond these seven are combinations of two or more changes. For example, splitting one value into many is an act of adding new values, adding new preferred terms, or perhaps changing lead-in terms to preferred terms.

2. In bringing up this concept of a document-set and its shifting representation, I am invoking an analytical device similar to Melanie Feinberg's (2005), though not identical in use; they are similar in composition.

REFERENCES

Aitchison, J. Gilchrist, A, and Bawden, D. (2000). *Thesaurus construction and use: a practical manual*. London: Aslib.

Anderson, J. D. and Perez-Carballo, J. (2005a). *Information retrieval design*. DRAFT. http://www.scils.rutgers.edu/~carballo/ird2005.html.

Anderson, J. D. and Perez-Carballo, J. (2005b). Chapter 13: Vocabulary Management. In *Information retrieval design*. DRAFT. http://www.scils.rutgers.edu/~carballo/chapters/chapter13.htm.

ANSI/NISO. (2005). *Guidelines for the construction, format, and management of monolingual controlled vocabularies*. Bethesda, MA: National Information Standards Organization.

Connolly, D. 2006. *Uniform resource identifier (URI) activity statement*. http://www.
    w3.org/Addressing/Activity.
Beghtol, C. (1986). Bibliographic classification theory and text linguistics: aboutness
    analysis, intertextuality and the cognitive act of classifying documents. *Journal of
    Documentation* 42(2):84-113.
Duranti, L., Eastwood, T., and MacNeil, H. (2002). *Preservation of the integrity of
    electronic records*. Dordrecht: Kluwer.
Feinberg, M. (forthcoming). Expression of feminism in three classifications. In *Ad-
    vances in classification research*. Information Today Inc.: Medford, New Jersey.
Miles, A. (2006). *Simple Knowledge Organisation System (SKOS)*. http://www.w3.
    org/2004/02/skos/.
Herman, I. E. and Hendler, J. (eds.) (2006). *Web ontology language (OWL)*. http://
    www.w3.org/2004/OWL/.
Herman, I. Swick, R., and Brickley, D. (eds.) (2006). *Resource description framework*.
    http://www.w3.org/RDF/.
OCLC. (2006). Dewey Decimal Classification: New Features in Edition 22. http://
    www.oclc.org/dewey/versions/ddc22print/new_features.pdf.
Ranganathan, S. R. (1967). *Prolegomena to library classification*. 3rd edition. Bom-
    bay: Asia Publishing.
Soergel, D. (1974). *Indexing languages and thesauri: construction and maintenance*.
    Los Angeles: Melvile.
Sutton, S. and Hillmann, D. (2006). *NSDL Metadata Registry*. http://eg2.ischool.
    washington.edu/registry.