1
2
3
4
5
6
7
8
9
10

V3 Technical Editorial Services

For HL7 Contract Work Announcement
"V3 Technical Editor"

Vocabulary Editorial Assessment

8 June 2007

Ockham Information Services, LLC
303 Adams Street
Decatur, GA 30030

1

Vocabulary Editorial Assessment

Document Purpose

This document summarizes the initial findings of the V3 Technical Editing Project regarding Vocabulary documentation as identified in spring of 2007. These findings may be used by the document owner, the Vocabulary committee, to guide composition and revision. If the editing team's engagement is extended, the committee may work with the team to revise, refine, or extend these findings.

Assessment

Vocabulary was identified as being a high priority area for the editorial team to review because of its fundamental relationship to the data model. The *Introduction* to the vocabulary section of the ballot is concise and specific, and it serves its purpose well for a reader wishing to navigate the vocabulary topics.

However, most of the documentation for vocabulary is still in the conception stage. The lack of documentation in the vocabulary domain is a key gap in the V3 standard: vocabulary is a fundamental component of the architecture. Our work to this point has involved issue identification and introducing the user profile concept into the discussion. Due to the heavy commitments by the committee co-leads and suggested liaisons, the work has not progressed beyond the issue identification stage.

Recommendations

1. Bring existing ballot materials up-to-date

There are several issues outstanding in the current vocabularies listed in the ballot, and the HL7 National Library of Medicine project is currently in the process of identifying issues and recommendations for these vocabularies. A review of the current vocabulary repository should be on the schedule for editorial review after the initial clean up is complete.

In addition, the current ballot introduction contains text that is five years old, and there is a perception that it is out of date. While the introduction should be reviewed for methodological currency, the methodology may be in flux while new documents are created. If currency cannot be confirmed, then the introduction should contain a note informing the reader of its status.

2. Adopt a user model to specify document purposes

A user model consists of user profiles, user objectives, and use cases directing the user to specific documents for a given objective. The first step in developing such a model is defining the user profiles that will guide document development by ensuring coverage

and guiding author assumptions regarding reader knowledge and needs. We have tentatively identified the following set of user profiles, which the committee should review and confirm or modify.

- Methodologists: HL7 experts who define the core standards and processes upon which standards developers build new standards
- Vocabulary Developers: Methodologists who specialize in the definition of vocabulary content and processes
- Standards Developers: HL7 experts who define message types, structured documents, and other standards for use in the healthcare domain
- Voters: Stakeholders who have an interest in and desire to influence the creation of standards
- Software Implementers: Stakeholders who build software that supports HL7 standards
- Standards Implementers: Stakeholders who implement software that supports HL7 standards

User objectives are the sets of tasks the identified user profiles might be expected to want to perform. Cataloging and coordinating objectives and documents is a significant task: in the short term, profiles and objectives can be defined to support editorial decisions for specific documents.

3. Catalog documentation needs

If one half of the user model is the user definitions, the other half is the set of documents that will meet the users' needs. Vocabulary is a core component of the V3 architecture, but it suffers from a severe lack of documentation. We have identified a draft set of needs, to be confirmed by the committee, listed in appendix A. Each need should be identified with specific user profiles.

4. Collect, refine, edit, and ballot existing documents

- The *Value Set Binding Definition* is a deliverable of the HL7 NLM contract. It has undergone extensive discussion and revision and is ready for inclusion in the normative part of the vocabulary section of the ballot.
- The 16-Point document of good vocabulary practices
- The Vocabulary section of the *HL7 Definition Framework*

5. Create new documents

Document needs identified under recommendation 3 and not met by documents refined under recommendation 4 should be prioritized and created.

New documents may include the following:

- *Vocabulary Data Architecture*, including purpose and use, RIM structural elements, structure and use of CD data elements, approach for postcoordination, stewardship of HL7 and external vocabularies
- *HDF* section on vocabulary, including vocabulary development & maintenance processes and artifacts as well as vocabulary-related standards development processes and artifacts
- Implementer's guide, including document index and task-specific guidance

6. Leverage educational materials

The vocabulary technical committee currently provides two educational offerings developed by the vocabulary technical co-chairs. *Vocabulary 1* introduces the general topic of vocabulary usage within the HL7 family of standards and standard terminologies such as LOINC and SNOMED CT; *Vocabulary 2* goes into more detail about the use of vocabularies within HL7, tooling, creation and binding of value sets, and the work of the vocabulary technical committee. Incorporating the topics of the vocabulary tutorials into the informative part of the ballot will enhance members knowledge of the key role of terminology within HL7 products.

In addition to HL7-specific topics, the inclusion of descriptive materials for standard terminologies such as SNOMED and LOINC would give new HL7 members a starting point for the use of those vocabularies within HL7 artifacts. References for these assets should be included in the informative portion of the vocabulary ballot.

7. Structure ballot to reference existing documentation

The current introduction refers to three assets (the HL7 Vocabulary Domain Values Table, the External Domains Table, and the Coded Attributes Cross-reference Table). Only the domain values are included in the ballot package, and even they are not in the suggested table, but rather in 198 web pages. Clear reference to the provided data will help readers navigate effectively, as will provision of the other two assets.

8. Clarify relationship among vocabulary, the reference information model, and data types.

The RIM document is the closest thing to a data architecture in V3. A similar effort should be given to articulating representation of data elements. Such an effort would have two results: a formal specification of the relationships for methodologists, and a clear articulation of how to implement for end users.

Appendix: Identified Documentation Needs

Meta-Documentation Needs

- The purpose and use of vocabulary in the V3 specification - why vocabulary, how vocabulary relates to the RIM and other information models.
- Vocabulary as direct RIM documentation - many of the core RIM elements, such as data types, units of measure, message states, general message categories, roles, moods, etc. are, in part, documented within the vocabulary itself.
- The purpose and use of value sets, realms and other vocabulary specific artifacts

Structural Needs

- How vocabulary is used in the various CD data type derivatives. This includes code system identifiers, codes, original text, display name, translations, etc.
- How vocabulary content is represented in the HL7 V3 RIM (Access Tables, MIF, LexGrid, etc.)
- How value set definitions are represented inside and external to the HL7 RIM, including versions, post-coordinated constraints, etc.

Meta-Process Needs

- How new vocabulary elements are submitted to and adopted by the HL7 V3 RIM.
- How external terminology references are submitted to and adopted by the HL7 V3 RIM
- How realms are created, defined, updated and removed
- How realms interact with vocabulary, value set and other artifact definition, creation, etc.
- How value sets are submitted to HL7
- How value sets are changed, updated and removed
- How (and what) vocabulary components are included in ballotable RIM documents
- How vocabulary service components are defined, balloted, revised, retired and implemented (CTS and CTS 2)
- What are the rules for downloading, utilizing, extending, etc. vocabulary content
- How do revisions to external vocabularies interact with HL7 terminology references.
- How is vocabulary incorporated into other RIM documents (both human and computational process)?

Process Needs

- How and when a CD-derived data element acquires translations, codes where none are present, etc.
- How and when a CD-derived data element should be validated and what should occur in the case of exceptions.
- How and when value sets are bound to message definitions and instances (versions, realm specific issues
- CTS - how one invokes a terminology service.

Content Needs

- What is currently in the HL7 Terminology and how does one find it, download it and use it?
- Where does one find realm-specific content? How does one acquire it and use it?
- Where does one find external terminology references? How does one acquire it and use them?

Implementation Needs

- How to implement CD data types
- How to use vocabularies to populate data instances
- How to extend value sets and accept extensions
- How to handle changes & versioning to prevent collisions & failures

Representation of Information

- How is HL7 terminology content represented? (Access tables, MIF, etc)
- How is HL7 terminology bound to the rest of the RIM?
- How does HL7 terminology reference external terminological resources?
- How are value sets represented (including versions, external references and potential post-coordinated structures)?
- How are value sets bound to the various RIM artifacts?
- How are realms defined and represented?