

**Vocabulary Task Force**  
**September 1 and 2, 2010**  
8:00 a.m. to 2:30 p.m./Eastern Time  
Location: TBD  
Washington, DC

---

**Instructions and Questions for Panelists**

---

Judith Sparrow  
US Dept of Health & Human Services  
Office of the National Coordinator for Health IT  
330 C Street, SW, Suite 1104  
Washington, DC 20201  
Judy Sparrow [judy.sparrow@hhs.gov](mailto:judy.sparrow@hhs.gov)  
Jamie Ferguson [Jamie.ferguson@kp.org](mailto:Jamie.ferguson@kp.org)  
HIT Standards Committee Clinical Operations Workgroups – Task Force on Vocabulary  
September 1-2, 2010

**Re: Panel 1, Measure Developers – Value Set Creators**

Thank you for this opportunity to provide testimony to the ONC Standards Committee on the subject of value sets and subsets for Panel 1/Measure Developers – Value Set Creators.

My name is Dr. Nikolay Lipskiy. I am a Centers for Disease Control and Prevention (CDC) Health Scientist and acting CDC Health IT Standards and Interoperability Lead. During the last five years I have worked for CDC, which includes my assignment as the BioSense subject matter expert on electronic laboratory reporting and two most recent years during which I have worked on the assignment on leading CDC tasks on Health IT standards. I am also contributing as the technical expert to the WHO/CDC Public Health Informatics Collaborating Center.

My name is Dr. Sundak Ganesan. I am a Northrop Grumman contractor working at CDC as Vocabulary Specialist Lead in the CDC Vocabulary & Messaging team. I am also a co-lead for Public Health Vocabulary & Messaging Community of Practice. I am a physician by training and have been working in the field of medical and public health informatics for the past 10 years focusing specifically on the vocabulary and messaging standards. I have been creating and managing the value sets in CDC vocabulary server PHIN VADS for the past 6 years. I am very actively involved in the SDO vocabulary activities as well as with the implementation of standard vocabulary.

As it was mentioned in my testimony in March 2010, CDC has a strong commitment to support the development and deployment of standards-based public health information systems and for fostering the use and exchange of consistent information among public health partners. CDC acts as one of the leading government public health agencies that

are involved in the development of content exchange and vocabulary standards on the international and national public health arenas.

**1. What are the requirements for a centralized infrastructure to implement “one-stop shopping” for obtaining value sets, subsets, and vocabularies for meaningful use?**

We support a vision of the Vocabulary Task Force on development of the “one-stop shopping” infrastructure for obtaining value sets, subsets, and vocabularies for meaningful use. Furthermore, CDC had initiated this approach for public health vocabulary through Public Health Information Network (PHIN) Vocabulary Access and Distribution System (VADS) since 2003. We believe that following requirements will support a development of the “one-stop shopping” infrastructure for obtaining value sets, subsets, and vocabularies for meaningful use:

- Existence and acceptance by ONC and providers of national HIT standards that cover an exchange of standard vocabularies
- Inclusion of a national vocabulary metadata repository as the object of the Federal Health IT Architecture Model
- Assignment of vocabulary stewards for groups of vocabularies or domains of clinical and population care knowledge
- ONC leadership in a resolution of issue with SDO licensing fees, which is a barrier in national exchange of standard vocabulary
- Harmonization of standards for data exchange between clinical care (NHIN) and public health (PHIN)
- Development of HIT certification rules that will be applicable not only to clinical care provider but also to HIEs and Public Health (PH)
- Existence of a support team for management of the “one-stop-shop” infrastructure and working with implementers and SDOs
- Utilization of a web design for viewing vocabulary, accessing it and distributing it.
- Utilization of efficient tools for: a) integration of vocabulary into computer applications, and b) adding quickly new data to the system (i.e., Universal Authoring Network, UAF)
- Implementation of a subscription mechanism for updates and use the server pull mechanism (i.e., the RSS feed that can be pinged for users to find out if there is an update of their vocabulary; programmatic API access to inspect status dates).

**2. Which requirements or functionalities are urgent, i.e., absolutely required to support “meaningful use”? Which would be most useful immediately? What would be a staged approach over time to get to the desired end state?**

We believe that following requirements are urgent:

- Approval of messaging guides for all meaningful use Stage 1 objectives
- Messaging specifications should be constructed in a format that combine two approaches: 1) fostering implementation of national requirements (i.e., utilizing a national messaging specification for the meaningful use objective) and 2) provide opportunity for optionality that is based on standard data elements (i.e.

Stage 1 PH reporting objectives allow states to make their own specific reporting requirements. Even though states may incorporate their own additional data elements, those elements still should be based on standard vocabulary)

- Attention to governance of health care vocabularies and terminology

A staged approach for functionalities should provide:

- Logic continuation of HIT processes that will lead to improvement of quality of life and positive health impact
- Moving from optional functionality to required functionality
- Raising HIT certification benchmarks
- Progression from text format of HL7 segments to a structured format
- Adding national and international standard vocabulary requirements for structured data elements (i.e. progression from a requirement on a submission of laboratory results in structured format at Stage 1 to submission them with SNOMED codes at Stage 2)
- Progression from local to standard data elements (i.e., progression from local codes for laboratory procedures to LOINC codes)
- Progression from requirements on use of HL7 version 2.x to version 3.x. on latest stages of meaningful use
- Progression towards establishment and recognition by ONC of a terminology for pre-aggregated sets of vocabulary. Those pre-aggregated sets should provide capabilities for a selection of up-to-date sets of value sets for a specific purpose (i.e., Immunization, Cancer Reporting) or those that are children of the node in a tree of concepts. An example of a such development is a recently developed by IHE Shared Value Sets (SVS) profile (IHE SVS concepts: **Expanded Value Set** – a set of concept representations that were in effect at a specific time for a particular version of a Value Set vocabulary; **Intensional Value Set** – a set of concepts that is specified in terms of the “intension” of use, for example “all concepts that are children of this node in a tree of concepts”; **Extensional Value Set** – a set of concepts that is specified in terms of a list of concepts)

Also, a staged approach should be based on a clear understanding of an existing technical environment, resources and implications of added requirements. For example, currently most public health laboratories use the HL7 version 2.3.1. When a Final rule for Stage 1 added a PH laboratory reporting PH requirement for using only the HL7 version 2.5.1 it added a technical burden to public health. Many states have expressed their concerns about their capability to switch to the HL7 version 2.5.1 within the required next 6-12 months.

### **3. Where are you using value sets and subsets? For what domains? How many value sets and subsets?**

CDC uses value sets for development and harmonization of PH systems (biosurveillance and response, health status and disease management, population-based research etc). Value sets developed by CDC are used primarily to support the HL7 message (V2.x, V3) and CDA implementation guides that were developed for Electronic Laboratory Reporting (ELR), Immunization, Vaccine Adverse Event Reporting System (VAERS), Public Health Case Reporting, Case Notification,

Healthcare Associated Infections, Antibiotic Use and Resistance Surveillance, BioSense, Non-infectious conditions (Lead Poisoning), Chronic conditions and Cancer.

The PHIN VADS has already published the value sets associated with such population health meaningful use measures as the ELR to Public Health (HL7, version 2.5.1) and Immunization (HL7, version 2.5.1).

Currently the PHIN VADS has 557 value sets supporting 60 HL7 and CDA message implementation guides. The PHIN VADS hosts both the intrinsic and extrinsic value sets. The PHIN VADS has a robust mechanism to host current and the previous versions of value sets and messaging guide vocabulary views.

In addition to the hosting of value sets, PHIN VADS also hosts 149 code systems (92 HL7 code systems, Clinical Vaccine Names (CVX), Manufacturers of Vaccine (MVX) and Healthcare Service Location, and developed by CDC codes for Race and Ethnicity).

The main purpose of PHIN VADS is to distribute the value sets. PHIN VADS does not allow the users to download the code systems such as LOINC and SNOMED CT and expect the users to download from the SDO or official distribution source.

#### **4. In your experience with creating, disseminating, updating and/or using value sets, subsets, and entire vocabularies, what works and what does not work?**

##### **a) Creation of Value Sets:**

Collaborating with the stakeholders, Standard Development Organizations (SDO) and implementers during the process of developing messaging guide and value sets helps in the adoption of standard vocabulary. CDC vocabulary and messaging team provides support to the vocabulary and messaging community of practice program (VMCoP).

The biggest challenge in creating value sets is the timeliness of getting the new standard vocabulary concepts from SDO. For example, it takes at least 3 to 6 months to get the new concept codes from code systems like SNOMED. It is difficult for CDC to develop or update the value sets. If SDOs do not provide the concept codes in a timely manner, especially during an outbreak scenario, then CDC (PHIN VADS) assigns temporary codes for concepts.

##### **Distribution of value sets:**

CDC programs and implementers found it easier to download a specific version of the value set or vocabulary view from such single for PH location as PHIN VADS. Most of the HL7 implementation guides require at least 25 to 30 value sets that were developed from various coding systems. It works when PHIN VADS groups all value sets that are associated with an implementation guide as “Vocabulary Views”. Vocabulary Views have a versioning process, which is similar to a value sets versioning. The PHIN VADS value set download includes all metadata that are needed for implementing HL7 message coded data types (CE, CWE). It makes a downloading process easier for implementers. Many implementers have not adopted the SNOMED numeric codes. They are still using the SNOMED alphanumeric codes. PHIN VADS include the value set concepts with the preferred

concept codes as well as the alternate codes. It makes it easier for the implementers to make a transition from alternate to preferred concept codes.

b) Value set updates:

Maintenance of value sets at a central location makes it easier to obtain the various versions (past, current and future) of value sets. Implementers and CDC programs can access the previous, current and future versions of value sets and vocabulary views from PHIN VADS as they become available. This makes it easier for implementers to implement a particular release of the HL7 messaging or CDA implementation guide.

An e-mail notification process is not an optimal solution for notifying implementers about the updates in a timely manner. We believe that a subscription mechanism (e.g. RSS Feeds) would work better for getting updates to vocabulary views or value sets.

**(5) What human resources does it take to implement and manage value sets, subsets, and entire vocabularies? Informaticists? Clinicians? IT people? How are you organized?**

The CDC Vocabulary and Messaging team manages the PHIN VADS content as well as supports development and implementation of value sets. This team works very closely with SDOs, CDC programs and their Informatics staff and various communities of practice that provides communication with stakeholders and implementers. The following roles play important parts in the development and management of value sets:

- (a) Facilitator / Informaticist: Facilitates the communication between the CDC Program Subject Matter Experts (SME) and Informatics activities. . This should be an informaticist who has knowledge in the domains of medicine, public health, and informatics.
- (b) Vocabulary & Messaging Business Analyst: Translates the CDC programs or Public Health Departments use cases into messaging and vocabulary requirements.
- (c) HL7 Messaging Analyst: Analyzes the use cases and provides messaging solutions in the form of HL7 messaging or CDA implementation guides.
- (d) Vocabulary Specialist (Value Set Developer): Works closely with the HL7 Messaging analyst and the CDC program Subject Matter Experts during the process of developing value sets. Vocabulary specialists standardize the value set concepts based on Whitehouse E-Gov Consolidated Health Informatics (CHI) domain recommendations, HITSP and Meaningful Use standards.
- (e) Vocabulary Steward: Provides guidelines in the development and management of value sets, especially governance of value sets and the selection of standards. The Vocabulary Steward works very closely with SDO and keeps track of the new concept request. The Vocabulary Steward facilitates the interaction between the CDC programs, subject matter experts (value set requestor), and the SDO especially for getting the new concepts from SDO. The Vocabulary Steward reviews the value sets and administers the content of the vocabulary server (PHIN VADS).

Vocabulary Stewards and vocabulary specialists are responsible for updating the SDO vocabulary of the vocabulary server.

(f) Software Developer: Develops and maintains the vocabulary server web user interface, vocabulary authoring software as well as web services (API). Software developer works closely with the business analyst regarding the requirements. PHIN VADS software developers participate actively on HL7 CTS2 and IHE SVS activities.

(g) Database Architect and Administrator: Develops and maintains the data model of the vocabulary server. The DBA is primarily responsible for managing the vocabulary authoring, staging and production database. The PHIN VADS Database architect also participates on HL7 CTS 2 and IHE SVS activities.

(h) Tester: PHIN VADS tester tests the applications as well as the content. Value set content is also tested by the messaging analyst, CDC program SME, and vocabulary specialist.

(i) Project Manager: Manages the CDC vocabulary and messaging team, and plans future releases of PHIN VADS. The Project manager and vocabulary steward co-ordinate outreach activities (communication and training).

(j) Program Manager: Monitors the CDC Vocabulary & Messaging team activities and future development of the vocabulary server. The Project manager and program manager co-ordinate integration of the CDC vocabulary server with other applications or tools such as PHIN Messaging Quality Framework, which validates HL7 messages.

**(6) What national resources and services could be leveraged to reduce the level of effort required for local implementations? What is the irreducible minimum of local work at an implementation site, or within an organization or system?**

(a) National Resources & Services:

- National support team for implementing the messaging and vocabulary standards would be a very useful resource for implementers.
- Vocabulary mapping tools that may assist in the mapping of local vocabulary to standard vocabulary.
- Training regarding SDO vocabulary, tools, and web services (CTS2), which would allow the implementers to adopt standard vocabulary. Lack of training about the SDO vocabulary may cause semantically incorrect mappings as well as mappings with less specificity.
- National and regional PHIN infrastructure that supports implementation of messaging and vocabulary standards. CDC PHIN had the regional PHIN coordinators to support the implementation of PHIN messaging implementation guides.

(b) Local resources that are needed for standard vocabulary implementation

- It is crucial to provide the tools and resources for mapping the local vocabulary to standard vocabulary.
- Local staff needs to be trained to integrate the vocabulary server with their applications.
- Funds to support the vocabulary server, resources, training etc.

**(7) What is your maintenance process? How do you manage updates?**

PHIN VADS uses Universal Authoring Framework (UAF) to manage the vocabulary present in PHIN VADS. UAF is currently available within CDC intranet and plan on making this available outside CDC intranet later this year which allows vocabulary authors outside CDC intranet to manage the value sets.

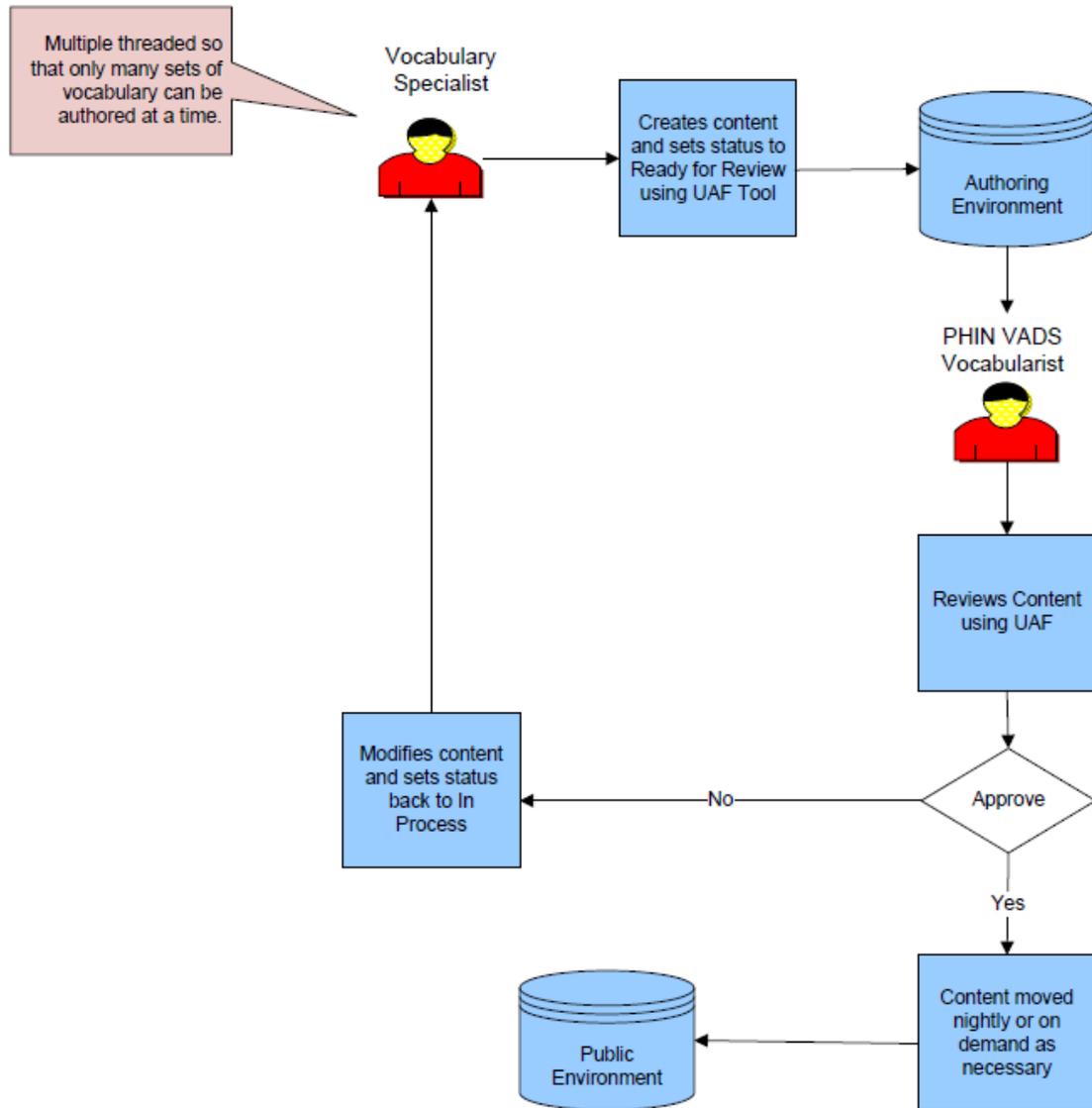


Figure 1 - PHIN VADS value sets authoring using UAF

Figure 2 illustrates the vocabulary governance in PHIN VADS which includes importing the code systems from SDO as well as management of value sets.

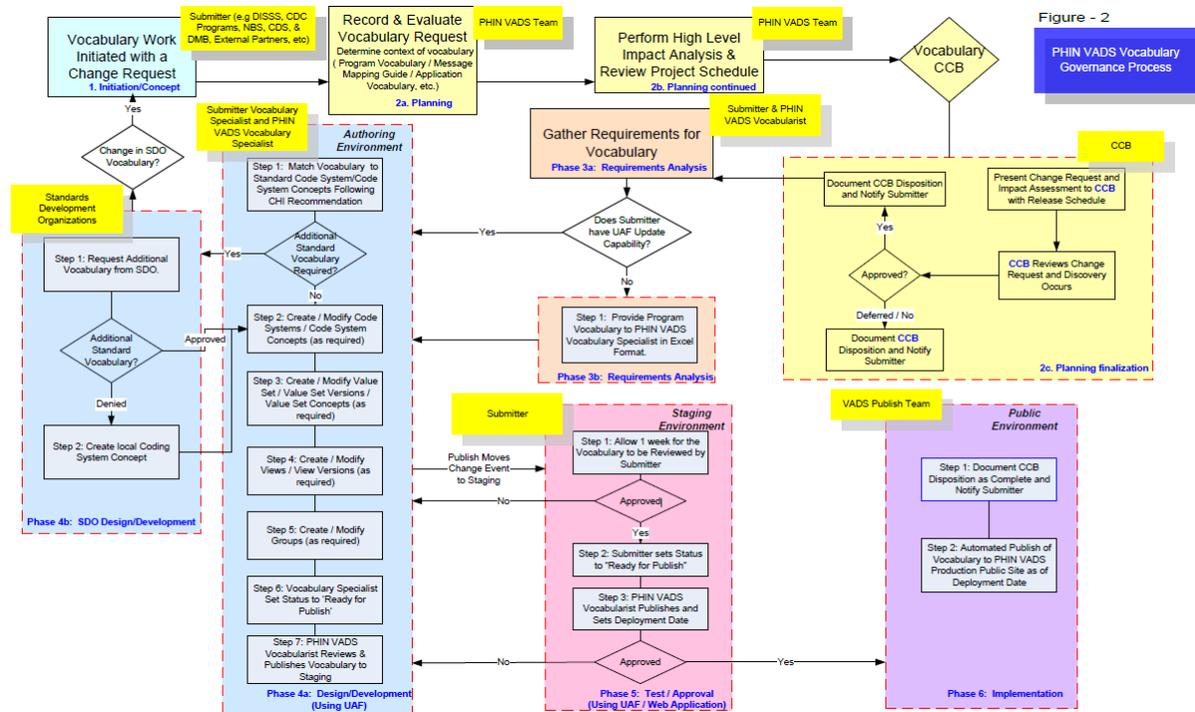


Figure 2 – PHIN VADS Vocabulary Governance Process

## (8) What metadata do you maintain and how do you maintain versioning?

PHIN VADS has adopted all the value set and code system metadata based on HL7 vocabulary technical committee recommendations and CTS2 specifications.

### a) Metadata:

Microorganism value set ([hyperlink](#)) example illustrates the value set metadata.

\* PHIN VADS Data Model:

[http://phinvads.cdc.gov/vads/WebHelp/PHIN\\_VADS.htm#The\\_PHIN\\_Vocabulary\\_Model.htm](http://phinvads.cdc.gov/vads/WebHelp/PHIN_VADS.htm#The_PHIN_Vocabulary_Model.htm)

\*Code system representation in PHIN VADS:

<http://phinvads.cdc.gov/vads/DownloadCodeSystemRepresentation.action>

b) Versioning:

PHIN VADS maintains the past, current and future versions of a value set.

Please see the Figure 3 below illustrating the various versions of Microorganism value set based on SNOMED CT code system updates.

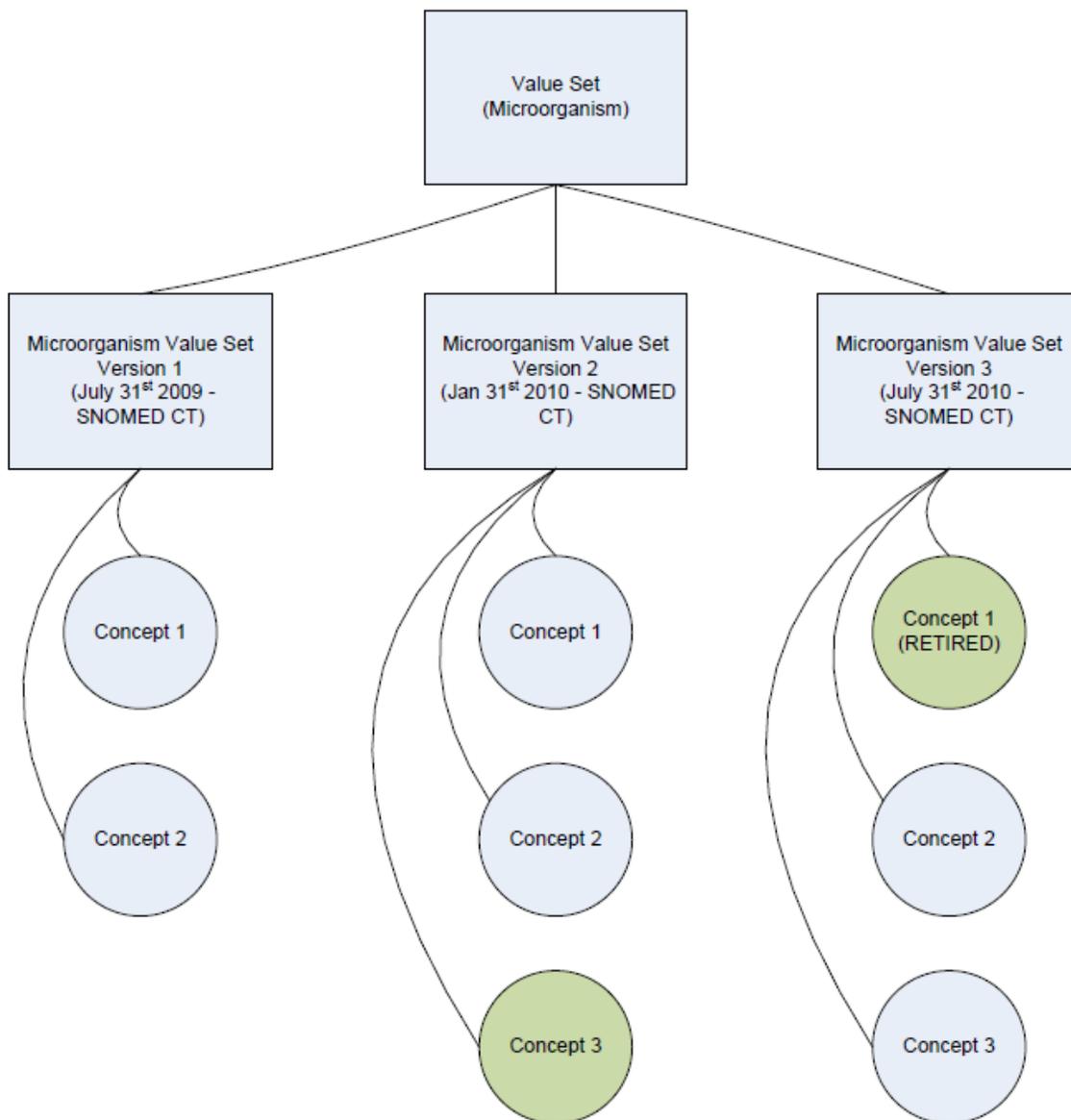


Figure 3 – PHIN VADS Value Set Versioning

- (9) **Is there a difference between versioning for clinical documentation vs. versioning for reported measures, i.e., when do you go live with a change in the EHR vs. when do you use the new version for measures?**

N/A

- (10) **How do you manage versioning in clinical decision support vs. changes in value sets?**

N/A

- (11) **How does an application know which value set is for which purpose? How is the specific context for a value set maintained at the message data element level of specificity? How is the English language intent of the value set context documented and maintained?**

- (a) **How does an application know which value set is for which purpose?**

Applications have information models to link data elements that are present in the information model to the value sets. Application information or logical model provides the context and the purpose of the value set. Many applications cache the value sets for quicker retrieval locally. Also, using vocabulary web services (API / CTS2), they may query a vocabulary server periodically for updates.

- (b) **How is the specific context for a value set maintained at the message data element level of specificity? Program driven requirements communicated through IG.**

HL7 messaging or CDA implementation guides provide a context, and define data elements and value sets. They also provide HL7 segment information in which the data element and the value set concepts would be passed in HL7 message.

- (c) **How is the English language intent of the value set context documented and maintained?**

All the public health messaging implementation guides are based upon the US-English language. PHIN VADS imports only the US-English synonyms from SDO vocabularies like SNOMED. PHIN VADS can be enhanced to support the language context (Spanish version of SNOMED). PHIN VADS supports the Unicode UTF-8 format, as many of the medical vocabulary has Latin characters (SNOMED terms) and special symbols (units of measure).

- (12) **What are lessons learned about web links vs. storage of the vocabulary or other artifact in a physical repository?**

Web links are easily accessible but it is not an optimal method to distribute the various versions of value sets. Web links do not help the implementers in mapping their local vocabulary to standard vocabulary.

Benefits of vocabulary servers (including the PHIN VADS server):

- Allow implementers to search a concept within a value set and/or an implementation guide, which makes easier a mapping task
- They provide synonyms, which facilitate the mapping of a local vocabulary to a standard vocabulary.
- They provide a context or hierarchy for the value set concepts. It allows the implementers a semantically correct mapping of their own vocabularies to standard vocabularies
- They handle very well a versioning of a single value set and groups of value sets (vocabulary views).
- Interface engines may call the vocabulary servers programmatically, retrieve value set concepts and display the standard concepts using their interface.
- They facilitate an adoption of standard vocabulary as well as help the implementers maintain their mapping to standard vocabulary
- Interface engine applications can also call the vocabulary servers to validate the coded concepts present in a HL7 message or CDA
- They can host multiple versions of value sets that allow implementers to have access to the current, past and the future value sets.

(13) **How do you manage distribution of updates to multiple sites?**

The PHIN VADS application and content has been integrated with various other public health applications, including the PHIN Messaging Subscription System, EpiInfo, PHIN Message Quality Framework, etc.

Public health surveillance applications can either integrate PHIN VADS data with their systems by storing the data they need locally or by calling-the PHIN VADS web service.

Some public health applications (i.e., EpiInfo and Outbreak Management System) may only include the PHIN VADS vocabulary content in their applications and expect users to update vocabulary directly from PHIN VADS. In addition to the messaging vocabulary, PHIN VADS hosts application- specific vocabularies. Also, it provides bookmarks (URL) for the application- specific vocabularies.

(14) **Where is local customization appropriate and how much customization is acceptable?**

PHIN VADS value set concepts have the standard concept name from SDO vocabulary as well as the display name (CDC preferred designation) for all the value set concepts. This allows the CDC programs and public health departments provide a display name that is appropriate for the end users. All the HL7 implementation guides recommend sending only the standard concept name in the HL7 message. The display names facilitate the end users while mapping the local concepts to standard vocabulary. The SDO concept names that come in the HL7 message and the CDC preferred display name can be retrieved from the vocabulary server while visualizing the HL7 message data through the data analysis tool.

PHIN VADS value set concept names do not contain any HL7 restricted characters (^ & \~|). It would be beneficial to have those characters removed from the concept names developed by SDOs.

**(15) How do you manage distribution of updates with local variations and optionality? Unique subsets? Local mappings?**

PHIN VADS provides web services as well as web user interface for value sets. The primary CDC installation of PHINVADS does not hold any local variations value sets. However, implementers can have instances of PHIN VADS at their sites and manage value sets that are local to them. PHIN VADS also provides the vocabulary authoring tool, Universal Authoring Framework (UMF), which allows implementers to author value sets. The primary CDC installation of PHIN VADS hosts value sets that are applicable for most of state public health departments. Sometimes public health departments or implementers may need to constrain or extend their value sets. In this situation implementers take a responsibility for managing their local versions of value sets.

**(16) What has to be local in an EHR implementation vs. what can be external in a vocabulary repository?**

All standard vocabulary, including the value sets, subsets and code systems, can be external in a vocabulary server or repository. A subscription system, which may provide an alert regarding the updates, can be present in a vocabulary repository. Vocabulary repositories can provide mapping between standard vocabularies (e.g., HL7 and SNOMED CT specimen concepts). This capability will facilitate adoption of standards that were recommended in meaningful use.

EHR needs to have mapping of local to standard vocabulary either in their vocabulary server or HL7 interface engine. EHR may also need to preserve display names that were added locally.

**(17) What functions are required that users have not yet appreciated?**

Implementers have not begun a broad adoption of web services (CTS2) and the vocabulary solutions that were created for managing various value set versions.

Some HL7 interface engine applications can provide interface for mapping a local vocabulary to standard vocabulary. It may simplify an adoption of standard vocabularies.

The PHIN VADS has been designed with a capability for integration with HL7 interface applications, knowledge management systems and public health surveillance applications through web services (CTS2).

A subscription mechanism for value sets could be made simple via RSS feeds which would allow the implementers to update their value sets and HL7 interface engine vocabulary mappings.

Validation of the vocabulary in HL7 messages can be performed using vocabulary servers like PHIN VADS. This type of validations may improve quality of HL7 messages.