

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

## **HIT Standards Committee Clinical Operations Workgroups – Task Force on Vocabulary**

### **Re: Panel 3/Best Practices & Lessons Learned on Value Sets Creation and Distribution Questions:**

- 1) What vocabulary subset or value set creation and distribution services do you provide?
- 2) Who uses your services and what is the level of use?
- 3) What, if any, additional services and capabilities are in active development?
- 4) If applicable, what process is used to establish and revise any subsets or value sets that you distribute?
- 5) Based on your experience, what advice would you offer regarding best practices and pitfalls to avoid?

Thank you for this opportunity to provide testimony to the ONC Standards Committee on the subject of value sets and subsets for Panel 3/Best Practices & Lessons Learned.

My name is Dr. Nikolay Lipskiy. I am the Centers for Disease Control and Prevention (CDC) Health Scientist and acting CDC Health IT Standards and Interoperability Lead. I have more than 30 years of experience working as a clinical epidemiologist, informatician and health scientist. During the last five years I have worked for CDC, which includes my assignment as the BioSense subject matter expert on electronic laboratory records 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 Collaborative Center.

For decades, CDC has engaged in the development of expertise, information, and tools that help people and communities protect their health. CDC has a strong commitment to support the development and deployment of standard-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.

### **CDC's experience and history regarding value sets management**

I would like to highlight a few of the CDC's historical activities as well as its active participation with various standard recommendation workgroups (Consolidated Health Informatics ,CHI; Health Information Technology Standards Panel, HITSP) regarding the development, management and distribution of value sets:

- First value set was published in 1995 (15 years ago) using HL7 Version 2.x implementation guide for Immunization by CDC National Immunization Program (NIP).
  - HL7 Vocabulary Technical Committee incorporated the CDC's NIP Value sets as part of HL7 v 2.x tables (e.g. [CVX](#) – Clinical Vaccine Names, [MVX](#) – Vaccine Manufacturers)
- CDC closely worked with the HL7 Vocabulary Technical Committee to define the value sets that would be needed for Electronic Lab Reporting (ELR) such as, “Flags “and “[Specimen Source](#)” .
- CDC actively participated in the development of HL7 Common Terminology Services (CTS) specification and follows the HL7 Vocabulary Technical Committee's recommendation regarding Value Sets metadata, binding and its usage in the HL7 messages.
- In 2004, CDC developed a web-based enterprise vocabulary system called the Public Health Information Network Vocabulary Access and Distribution System ([PHIN VADS](#)) for accessing, searching, and distributing value sets used within the PHIN.
  - CDC Vocabulary and Messaging team manages the PHIN VADS application and content. Several vocabulary and messaging experts from this team actively participates in various Standard Development Organization (SDO) and standard recommendation workgroup (CHI, HITSP) activities.
- In 2006, the CDC created Public Health Vocabulary & Messaging Community of Practice ([VMCoP](#)) to facilitate the dialogue and problem solving around vocabulary and messaging development, usage, and distribution within the public health arena. The Vocabulary and Messaging Community of Practice also provides an [online forum](#) to share knowledge, expand their professional skills, and develop solutions to solving common public health vocabulary and messaging challenges.
- In March 2007, the CDC actively participated in the Value Set Summit for addressing the issues related to value set management and to harmonize the value set metadata between the following organizations / applications: (a) HL7 CTS / [value sets and domain binding document](#) (b) CDC [PHIN VADS](#) (c) National Cancer Institute (NCI) - [EVS](#) (d) Mayo Clinic – [LexGrid](#) (e) UK [Cancer Grid](#)

- Between April and June 2009, the CDC participated as a member of the HITSP Data Architecture Tiger Team to assist in the development of HITSP C80 value set metadata and the content.

Through the **PHIN Certification Program**, CDC has gained a certain experience to verify that state and local public health departments have the capability to electronically exchange public health data across jurisdictional lines based on national standards of interoperability and data exchange. The PHIN Certification process uses program specific requirements for the evaluation of message structure and content, as well as security and data integrity. Components of this program include the following:

- Development and maintenance of PHIN Requirements (providing national standards for electronic exchange of public health data);
- Development and maintenance of PHIN Certification Criteria (measuring compliance of state and local partners in meeting PHIN Requirements);
- Refining of PHIN Data Exchange Standards (PHIN Message Mapping Guides) development and implementation processes across CDC;
- Programmatic partnerships within CDC;
- Technical Assistance to 62 Public Health Emergency Preparedness Cooperative Agreement Grantees (state and local partners) to achieve PHIN Requirements; and,
- Objective assessment of 62 Public Health Emergency Preparedness Cooperative Agreement Grantees (state and local partners) ability to meet PHIN Certification Criteria. (see at: <http://www.cdc.gov/nceh/programs-projects.html#cert>)

I would like to preface my answers to the specific questions that the Vocabulary Task Force (TF) has posed, with some general observations and comments that are based on my experience.

1. I believe that the Public Health WG should be added to the Vocabulary TF. Furthermore, I believe that CDC and state/local public health representatives should be included in HIT committees and advisory boards, so that population health needs are better addressed from the outset.
2. I agree with the Vocabulary TF regarding needs to promote the development of effective “marketplaces” of easily identified and obtained value sets that are described in a uniform way. I want to suggest that adding and implementing the “group of value sets” term to other Vocabulary TF terminology may improve the efficiency of those “marketplaces”.
3. I believe that one of the most efficient solutions for the task outlined above is the building of Health IT industry-level value sets repositories.
4. Finally, my vision is that the management of value set repositories should be closely tied with the task of development business cases, which, in the perspective of Public Health and Population Health, may be resolved through the involvement of federal and state agencies.

Now I would like to respond to specific Vocabulary TF questions.

## 1) What vocabulary subset or value set creation and distribution services do you provide?

CDC distributes value sets primarily through vocabulary server PHIN VADS. Public health community uses PHIN VADS to get the value sets associated with the various HL7 implementation guides based on HL7 2.x, V3 and CDA.

PHIN VADS can be accessed using one of the following methods which have been described in detail below:

- (A) PHIN VADS Web Browser.
- (B) PHIN VADS Web Service – System to System exchange of vocabulary data.
- (C) PHIN VADS Lite – Local instance of VADS web services and database.

PHIN VADS contains and distributes several unique value sets based on HITSP and CHI recommended code systems as well as CDC developed code systems adopted by HL7 SDO such as Race & Ethnicity, Vaccine names (CVX) and Healthcare Service Delivery Locations. PHIN VADS value set groups and value set views are unique CDC categorizations and grouping of value sets. There are very successful results in the development and distribution of value sets at the program/project level (i.e., related to Immunization, Public Health Case Notification, etc...)

CDC vocabulary and messaging team do not publish any vocabulary subsets in PHIN VADS, but provides the vocabulary subsets to various CDC programs and public health community via e-mail or forum in order to assist the public health programs in the development of value sets (E.g. Microorganism hierarchy of SNOMED CT).

My response to this question also requires the addition of two CDC definitions, which we recently submitted to IHE as an extension to Enhanced Shared Value Set (SVS) profile. First, to address the needs of our partners, we have created a new vocabulary object called **Value Set Group** for categorizing the value sets by subject or domain. PHIN Vocabulary Groups are used mainly for indexing the huge number of value sets published in PHIN VADS.

### Examples of Value Set Groups:

- a) PHIN VADS *Demographics value set group* contains all of the value sets related to demographics such as Countries, Counties, Cities, Sex, etc.
- b) PHIN VADS *Laboratory value set group* will contain all of the value sets related to laboratory tests and results such as specimen or sample type, Hepatitis Laboratory Tests, Microorganism, etc.

Second, we have also created another vocabulary object “**Value Set View**” for categorizing the value sets by HL7 messaging or CDA implementation guides. E.g., Tuberculosis, Influenza H1N1 Case Notification. A view can also be referred to as the **Implementation Guide vocabulary**. Views can also be used for navigating the huge number of value sets published in PHIN VADS.

### Examples of Value Set Views:

- (a) Tuberculosis Message Mapping Guide Vocabulary View
- (b) Healthcare Associated Infection CDA Implementation Guide Vocabulary View

**(A) PHIN VADS Web Browser:**

PHIN VADS web browser is the most commonly used method for browsing, searching and downloading the public health value sets. This has been in use since 2004 and can be accessed at <http://phinvads.cdc.gov> .

Currently in PHIN VADS we have:

- i. Number of **Value Sets** (E.g. Microorganism, Rubella Specimen Type): 533
  - Number of **Value Set Groups** (E.g. Demographics, Lab): 14
  - Number of **Value Set Views** (E.g. Tuberculosis Case Notification, Measles): 59
  - Number of **Value Set Concepts** (E.g. Bacillus Anthracis): 1.8 million
- ii. Number of **Code Systems** (E.g. LOINC, SNOMED, CVX) : 130

Approximately 100 code systems belong to HL7 V2.x and V3 standard. These would include CDC developed code systems that are needed to support various public health data exchanges.

- Number of **Code System Concepts** (E.g. Anthrax): 1 million

**(B) VADS Service – System-to-System Exchange of Vocabulary Data (PHIN VADS web services)**

- Provides a collection of methods that can be invoked by any software application to access the data directly without using the PHIN VADS Web Browser.
- Can be accessed from any internet connection.
- Can be incorporated into other applications.
- Accesses the PHIN VADS database used by the PHIN VADS Web Browser.
- Comprised of three major components
  - *Data Transfer Objects (DTO's)* – Encapsulate parameters that are sent to the service and results that are returned from the service. The DTO's contain domain model objects, and other information, relating to the service call.
  - *Domain Model* – Representations of the vocabulary model that define the vocabulary objects and their constituent properties.
  - *Service Interface* – Defines the methods that can be invoked and the DTO's that are sent to and returned from those methods.
- Contains a link to the developer toolkit that contains libraries for integrating with the VADS web service in 3 popular languages (Java, .Net and PHP).
- E.g. [NEDSS Messaging Subscription Service \(MSS\) application](#) includes PHIN VADS vocabulary server and Orion Rhapsody that will help validate HL7 messages and assist mapping between local and standard terminology.

**(C) Implementation of a Local Instance of the PHIN VADS Web Service and Database-VADS Lite**

- Allows any individual or organization to implement a local instance of the VADS Service and Database.
- The VADS Lite deployment is available for Oracle and Microsoft SQL Server.
- This is the same Service and Database that are used by the PHIN VADS Web Browser.
- VADS Lite is not deployed with a User Interface
- The deployment model requires the users to download, install, and maintain a local installation on their hardware

**2) Who uses your services and what is the level of use?**

PHIN VADS users include:

- CDC programs
- State and local health departments
- Healthcare Providers including labs
- Value set developers (outside of CDC)
- Value set implementers
- SDO (for purposes of adding new value sets or data modeling)
- Standard harmonization and promotion organizations
- Researchers
- EMR and public health application vendors

Users have an option to download PHIN VADS value sets, groups of value sets in either Excel or text format. A second option for use is by using PHIN VADS Lite. This option is used by PHIN Messaging Subscription Services (MSS) to distribute to their state sites for vocabulary validation.

**3) What, if any, additional services and capabilities are in active development?**

1. The PHIN VADS Vocabulary Team has started CDC-wide collaborative efforts on assessment of vocabulary-related PH capacities for addressing needs of the meaningful use objectives and maintenance and distribution of respective groups of value sets. This task includes but not limited to a gathering of business requirements from CDC programs and our partners, filling out gaps in standard code systems, binding them into structures of HL7 messages etc.

2. Our ongoing task is the improvement of PHIN VADS capabilities for supporting the management of public health threats and emergency situations. As the example, PHIN VADS has a success story of a vocabulary support of the influenza H1N1 pandemics. We were initiators of a submission to IHTSDO of new SNOMED codes for reporting of cases. However, because of the urgent situation and based on consultations with CDC

influenza specialists, PHIN VADS had released their own value sets several months prior to the official IHTSDO release.

3. CDC informaticians, along with the PHIN VADS Vocabulary Team, are working on an information model for the PH Case reporting. The conceptual approach for this model based on the assumption that any event of public health interest, from infectious diseases to behavior surveillance, can be built within similar interoperable semantic structures. As an example, the proposed meaningful use measures, such as syndromic surveillance and laboratory reporting and immunization, may be presented as the instances of the hierarchy of concepts that will support the information model for PH case reporting.

4. PHIN VADS works on development and implementation the Universal Authoring Framework (UAF). UAF is the distributed terminology development environment, which provides:

- A web-based thin client distributed environment that allows for vocabulary authoring and publishing
- Versioning Capability for Vocabulary Objects
- Historical tracking of database updates
- Capability for multiple users to author data simultaneously via the Change Event
- Capability for vocabulary updates via the web services

5. PHIN VADS participates with Integrating the Healthcare Exchange (IHE) on initiatives as it relates to Sharing Value Sets. As the outcome of this collaboration we expect that a) PHIN VADS will be better connected to PH value sets developers and b) users will have easier access to latest sets, including groups of value sets supporting measures for population health and clinical decision support.

**4) If applicable, what process is used to establish and revise any subsets or value sets that you distribute?**

Stakeholders mentioned in the response to question #1 would be involved in the creation and maintenance of value set (governance) using the Change Event Model. This Model allows:

- Locking of vocabulary objects during vocabulary authoring
- Providing a historical record of all changes made to an object over time
- Supporting collaborative authoring managed through governance
- Supporting change management
- Processing of data between systems and database models and storing changes

**5) Based on your experience, what advice would you offer regarding best practices and pitfalls to avoid?**

1. We found that a governance of the PH vocabulary metadata repository, PHIN VADS, within CDC is very beneficial for all our partners.

2. I agree with suggestions of vendors from the previous meeting that details of what specific information needs to be communicated in each element of the information model are best known by the developer of the specific requirements. With the context of a population health in mind, I believe that public health agencies should play a primary role in this process.
3. I suggest for adding the “group of value sets” term to the Vocabulary TF terminology. It will help in a vocabulary support of use cases based on the meaningful use objectives.
4. Adding interoperability specifications (i.e., HITSP, IHE) to the ONC IFR standards for vocabularies and data exchange to accelerate implementation of the Health IT standards.
5. CDC programs and partners are urging on immediate attention to development and distribution of common for clinical care and population care value sets that are related to the preventive tasks. Examples of those value sets are those related to behavior risk factors (i.e., drug abuse, alcohol use), risk factors during pregnancy, family health history etc.
6. Another request from our partners is in bringing to Vocabulary TF attention that at least secondary diagnoses ICD-9/10 codes or SNOMED codes in addition to primary codes should be included in the electronic data exchange.
7. SNOMED codes for reporting of laboratory results must be added to the Stage 1 of meaningful use.
8. Engagement of the implementers in early stages of the development of the value sets and HL7 messaging/CDA implementation was mutually beneficial to the VADS team and implementers
9. One of the key conditions to PHIN VADS’ success is providing support to the programs and organizations that receive the coded data via HL7 messages or CDA documents (E.g. Value Set version history, Code System version history, Translation and Mapping between vocabularies)
10. Discussion of the value set implementation issues with vocabulary SDO’s (e.g. Post-coordinated SNOMED concepts in HL7 v2.x messages and CDA ) is very beneficial to PHIN VADS development
11. One of the vocabulary metadata repository functions should be providing value set authoring tools to create the value sets. Also, we found that authoring tools need to have all the HITSP / CHI recommended code systems as well as allow the users to create the local codes as needed.
12. Another vocabulary metadata repository function should be providing vocabulary mapping tools that would help the implementers map the local to standard vocabulary.
13. Establishment of subscription mechanisms to inform users regarding upcoming updates and new value sets (i.e., through the Vocabulary & Messaging Community of Practice, posts on the PHIN VADS website and webinars) should be one of highest priorities of a vocabulary metadata repository.
  
14. PHIN VADS team provides training on vocabulary basics as well as advanced topics including vocabulary server’s web services, which is helpful for users.

Nikolay Lipskiy, MD, DrPH  
Health IT Standards and Interoperability Lead,  
Centers for Disease Control and Prevention, CDC  
Atlanta, GA

03/17/2010