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

Background

Testimony from this two-day hearing will help the Vocabulary Task Force formulate recommendations to the HIT Standards Committee and the National Coordinator on what would constitute the right set of requirements for vocabulary infrastructure that would provide an ideal "one stop shop," and which of these requirements have the highest priority, given the drive to achieve meaningful use. If you have any questions, please contact Jamie Ferguson, Chair, or Betsy Humphreys, Co-Chair, Jamie.ferguson@kp.org or Betsy.Humphreys@nih.hhs.gov

The September hearing is a follow-on to a previous hearing the Vocabulary Task Force held on March 23, 2010. The March 2010 hearing focused on general questions about national governance of terminologies, value sets and subsets, related to Meaningful Use. A theme in those presentations and discussions, especially with the implementers and users of EHR technology, was a desire for "one-stop shopping" and this was reflected in the subsequent recommendations of HITSC. (See attached recommendations).

Format of Presentation:

The Task Force respectfully requests that panelists limit their prepared remarks to **five (5) minutes**. This will allow the Task Force to ask questions of the panelists and allow every presenter time to present his or her remarks. We have found that this creates a conversation for a full understanding of the issue. You may submit as much detailed written testimony as you would like. Written testimony received by the requested deadline will be reviewed in detail by Workgroup members before the hearing. PowerPoints will not be needed.

Pre-Presentation Questions/Themes:
The questions below represent areas the Task Force intends to explore at the hearing. Please feel free to use them in preparing your oral and written testimony; the Task Force recognizes that certain questions may not apply to all presenters.

The Task Force respectfully requests panelists to provide written testimony by **August 26, 2010**. Please submit the testimony to Judy Sparrow at Judy.sparrow@hhs.gov

Presenter Biography
In addition, the Task Force requests that all presenters provide a short bio for inclusion in the meeting materials. Please send your short bios to Judy Sparrow, judy.sparrow@hhs.gov

THEMES/QUESTIONS

INTRODUCTION

Columbia University and NewYork-Presbyterian Hospital have employed a central terminology service since 1989. We use a knowledge-based approach to organize the terms in the data systems of the medical center. Terms are organized into a semantic network that classifies items into useful groups and defines attributes for each term and relationships among terms. This knowledge is used to define terms, to map terms from one system to another (e.g., from the laboratory system to the clinical information system), to group terms (e.g., so that when a clinician queries for bloood potassium, all of data stored under the many laboratory codes for potassium appear in the clinician's display), and to manage terms (e.g., when a new term is added, the system checks whether a similar term has already been defined). The system consists of a terminology structure, a primary knowledge base with all the terms (over 100,000), a set of tools to manage those terms, and a set of runtime routines that deliver definitions, print names, and traversal of the terminology hierarchy with microsecond-level response time. Our clinical applications can query the terminology directly, or they can receive a code table that is automatically generated by the knowledge base. The terms include local terms that were defined in the institution and the appropriate subsets of national standard terminologies (e.g., LOINC, ICD9, SNOMED, UMLS). Over the past two decades, we have swapped clinical information system applications (e.g., laboratory systems) with completely different code sets but without the clinical users ever noticing the change because of the terminology mapping.

Over the past 15 years I have been a user of clinical vocabulary services, either as a graduate student at Columbia University where I used the Medical Entities Dictionary and UMLS to assist data mining in a clinical database, eventually building a vocabulary-based data viewer we still use today at Columbia, or as a system developer and project director at Intermountain Healthcare where I was responsible for creating clinical documentation templates and point-of-care reporting tools that used the Health Data Dictionary, or in my current role as Director of Clinical Databases at NewYork Presbyterian Hospital where I am responsible for extracting data out of our system for various regulatory reporting, decision support, clinical research and quality improvement initiatives. From this experience, in multiple roles across two institutions both using robust but different clinical vocabularies, I have come to appreciate the difficulty in controlling vocabularies as data are collected or tools are developed, the value of a terminology in facilitating data retrieval, and the frustration and regret when either data are not coded properly or when retrieval tools do not make use of the available terminologies.

Overall questions

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

The first requirement would be value sets that are defined according to the meaningful use criteria. One of the challenges in using vocabularies is that terminologies need to be comprehensive, yet in a comprehensive form they are almost unmanageable to end users. Many of the meaningful use criteria require a percentage of use, which makes it likely that at first the more common demographic data, vital signs, problems, medications and allergies will be documented in structured form, and need codes. A centralized infrastructure should be able to identify which of the terms are or should be more commonly used, and be able to report the more common terms. For example, in the Medical Entities Dictionary (MED) at NewYork Presbyterian Hospital/Columbia University Medical Center (NYP/CUMC), there are 22,565 terms under the concept of ICD9 Disease, but only 8,342 terms were used in 2009, 1,470 terms represent 80% of all usage, and only 155 terms represent half of all usage. Since meaningful use does not necessarily mean complete use, it

would be helpful for the centralized infrastructure to be able to give some idea of what the most common terms actually are.

In addition, there needs to be an understanding that even with a centralized infrastructure there is a need for local terminology expertise for management of the local variations of the vocabulary.

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?

Defined value sets are more urgent, since they influence how easily the vocabularies can be implemented in collecting data; i.e., they are needed as data entry concepts are defined. Maintenance is important, but is more relevant over time as semantic drift occurs.

<u>Detailed Questions</u>

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

We use specific terminologies in our larger dictionary to match terms according to external requirements. For example, we use LOINC to define our terms relative to Department of Health terms for reporting requirements, Other subsets used are ICD9 and CPT, as well as parts of the UMLS.

We also identify subsets within our data for research and quality queries based on their use. For example, we will often look at what terms are actually used in our database vs. available to be used when writing retrospective queries to extract the data.

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

NYP/CUMC has a central system developed for and dedicated to managing institutional code sets from multiple clinical departments (including laboratory, radiology, pharmacy, documentation, etc). NYP has gained certain

benefits from centralizing the terminology resource, including consistency across systems, and streamlining of knowledge management (maintenance can be kept in one place).

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?

NYP terminology team consists of 3 dedicated FTEs. Informaticists with clinical experience (MDs, nurses) are probably most appropriate for managing the specifics of clinical code sets and associated knowledge. Mapping of local codes to standards often requires detailed domain knowledge that must be sought from clinicians from the applicable departments. In that respect, fostering an environment of open communication with clinical departments is important.

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?

No matter what national resources are available, the requirement for local terminology resources can often be underestimated. The correct use of standard code sets at any local site will require an active process by dedicated personnel.

The effort at local sites could be facilitated by making code selection more global. For example, if a certain clinical laboratory instrument is used at 1000 sites, then the standard code could potentially be chosen once by the instrument maker instead of 1000 times at each site. Although this kind of code "globalization" may make sense in some cases, there will always be local factors that affect final code choices.

<u>Detailed Questions, continued</u>

7. What is your maintenance process? How do you manage updates?

There are different types of maintenance. Some maintenance comes from a change in the underlying systems, such as a change of the lab system from one

vendor to another. Other maintenance addresses semantic drift. For changing underlying systems, a lot of time is spent mapping the concepts in the new system to the classes in the vocabulary. Where concepts do not exist, they must be created. We try to prioritize the concepts according to their expected use, and use standard vocabularies (e.g., LOINC) to map terms where possible. This is still a labor-intensive process, though. Fortunately we have been able to swap out back-end systems without a clear change for the end users during such system migrations.

For changes that occur during semantic drift, we follow a standard process of adding terms in the semantic network. Once the terms are identified (usually because the interface engine notes a new term from a source system is not mapped), a new term is created, and the interface message is re-processed. A centralized infrastructure will need to either allow local updates or be able to propagate out changes, as not all sites would be able to support the manual process of adding terms.

8. What metadata do you maintain and how do you maintain versioning?

We maintain structured multiple hierarchical relationships, diseases tested for, mappings, links to external resources, measured chemicals, specimens, etc. Versioning is accomplished by adding and deactivating terms in the hierarchy of the classes.

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?

Yes. Data are checked for consistency more as they are stored, so versioning follows the vocabulary maintenance process. With reporting, however, it is possible though not recommended to write queries that include the actual codes used rather than classes. This creates queries that are unaware of the changes in the vocabulary, and can lead to incorrect queries over time. Right now we have no standard testing to identify whether queries are written correctly, or certification of query quality. We train the query developers, and hope that the principles are followed.

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

This is similar to the above answer. If the classes are used rather than the leaf-node terms, it works fine. Otherwise we have problems in maintenance.

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?

If the value sets are represented by classes in the MED hierarchy, we maintain context by slot number and hierarchical relationships. Otherwise, the context is not maintained. As a result, it is imperative that the structure support multiple hierarchies that can define the value sets, and that creating such hierarchies is easy.

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

We have not used links to external data sets directly, but have included the external vocabularies as slot values in our MED. For some applications, we have linked out to specific information sources where the vocabularies of the information sources are mapped to the MED. Our experience in these cases was that the websites and access protocols of the external knowledge sources need to be maintained as much as the vocabularies. Whenever an external source upgrades its site or how access is structured, we need to update our application as well. These changes seemed more frequent than the semantic changes, actually.

13. How do you manage distribution of updates to multiple sites?

We use a central terminology service, that propagates changes out nightly to the various sources. Content is disseminated by different modalities to diferrent downstream systems, but the content derives from the same central data.

14. Where is local customization appropriate and how much customization is acceptable?

Local customization must be allowed, based on local differences in modeling and use. One simple example was an experience I had in building growth

charts at Intermountain. I had thought that height was height, and weight was weight. And it was, for many clinicians. But the nutrition specialists at the children's hospital were much more particular about what heights and weights should be considered in making their clinical decisions. Without local customization of the vocabulary that would account for these differences, the applications that use the data must include complex specifiers such as location, user type, etc., to differentiate appropriately. This should be maintained in a vocabulary rather than the downstream query.

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

In the semantic network which allows multiple hierarchies, the local variations are included in the network which is then maintained separately. To only include the local information, the terminology would be only selected among that hierarchy. For other hierarchies, the local terms just sit alongside the other terms. We had an example of this in our lab terms, where the data from a lab in one community hospital had a different reference range than the data from the academic medical center. Both terms for a serum potassium test existed under one node of "Serum Potassium Test", but the sub-terms defined their source. And both terms were in the central vocabulary.

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

This is architecture dependent. In our experience, the master library of terminology data and mappings is best maintained in a vocabulary repository, where terminology experts can manage it independently. Then, as connectivity and archtecture allows, the EHR either connects in real time or models subsets of the data as it requires.

17. What functions are required that users have not yet appreciated?

Initially, the users don't appreciate the value of using the classes to define the terms rather than the actual codes stored in the database, especially for building queries. We have since built tools that show the density of the terms in the database, and then the users (developers) can see how one class concept represents multiple terms, or they can select the appropriate level of concept in

the hierarchy. That seems to be the best way to get them to appreciate the classes, though not all users want to learn at first.

We're not exactly sure what other functions users have or haven't appreciated, but perhaps the challenge of getting terminology right may require more active effort than is globally understood.

Attachments:
Vocabulary Task Force Member List
Vocabulary Recommendations to HIT Standards Committee