

Vocabulary Task Force Responses

Jacob M. Reider, MD

Member, EHRA Executive Committee
CMIO, Allscripts Healthcare Solutions, Inc.

Overview

In the responses that follow, I have done my best to represent the state of the EHR vendor community as a whole, rather than only one vendor's perspective. The narrative that follows is based on my understanding of the marketplace and its readiness/needs regarding value sets and vocabularies.

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?*

A Centralized infrastructure would require:

- **A Simple API.** The functions of the API could be defined well before implementation, allowing "real-world" implementations to be accelerated.
- **A fast, scalable technical.** The framework needs to address redundancy and disaster recovery instances so that the service can be reliable and responsive.
- **Predictable Update Schedules.** This is more important than "frequent" update schedules.
- **Well-defined processes.** An articulated, published process that allows for a quick turnaround (2 business days?) on inquiries regarding corrections, amendments, or concerns regarding the content of certain value sets and vocabularies. This allows providers, hospitals and Health IT vendors to know that such inquiries will be addressed promptly.
- **Availability of all necessary vocabularies and value sets.** The goal is to ensure that additional licensing processes and payments become unnecessary. The required vocabularies and value sets need to be licensed for the nation (by the nation). CPT is an obvious (and necessary) first step, but others such as Medcomp's Medcin, HLI's PFT and IMO's ProblemIT and ProcedureIT should be considered.
- **Processes and technical infrastructure to support versioning.**

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?*

- Urgent needs include **Value Sets for Quality Measures**. In the Final Rule, and associated documents posted to the CMS website, a series of value sets appear in a series of spreadsheets. While this is an acceptable first step, it is not sustainable, nor does it support any of the requirements above.
- Where vocabularies are referenced in Federal Rules regarding Meaningful Use, **mappings must be made available** from existing vocabularies to the referenced vocabularies. For example, in the Quality Measure spreadsheets, RxNorm and SNOMED-CT are referenced, but these vocabularies are not required in Stage 1. Eligible Providers and Hospitals thus require mappings from commercial medication databases to RxNorm, or mappings from ICD-9 to SONMED-CT where possible, potentially creating more work during an industry transition period that is already overwhelming.

Detailed Questions

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

- Most commercial EHRs make use of value sets inefficiently, since there has been no standard repository from which one could reliably retrieve such content. For example, an EHR might create a subset of most frequently used diagnoses, orders, prescriptions, or even examination phrases for a provider of specialty. An EHR might also leverage a value set of medications that represent ACE Inhibitors or allergens that contain Iodine. All-too-often, such sets are “hard coded” in database tables of an EHR and thus are difficult to maintain.

To answer the question of how many – really, this is impossible to measure. Hundreds at least. Thousands?

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

- **What works:** Coordinated, “central” management of value sets, vocabularies and subsets in large organizations – this includes using informatics expertise, deep engagement of all stakeholders, and constant revision.
- **What doesn’t work:** “Central” management cannot and will not happen in the majority of practice settings because the majority of practices have a very small number of providers with no IT infrastructure (or spare time) to invest in such activities. EHR systems that permit end-users to create custom vocabularies, subsets and value sets will run the risk of pleasing the practice now but burdening the practice later with the need for a manual mapping initiative to align their local terms with industry standards.

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?*

- Given the breadth of our market, the answer to this question is variable: for the larger EHR vendors, this is often done by a team of in-house clinician informaticists. Smaller EHR vendors often lack such resources, however, and may avoid managing these tasks internally by licensing as much content as possible or by purchasing services from one of the handful of companies who offer vocabulary services (3M, Apelon, IMO, HLI, etc.)
- Occasionally, small EHR vendors will also contract with their clients for creation and/or maintenance of subsets or value sets for a given specialty.

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?*

- To begin, we need value sets that align with quality measures. This will enable not just quality measures but will pave the way toward clinical decision support. It is THIS ENPOINT for which we must aim.
 - i. Quality measures are generally retrospective – “did X happen?”
 - ii. Conversely, clinical decision support is prospective (“for patients with (1) and not (2) then (3) must happen”). In each instance of (1) (2) and (3), it is imperative that we have standard value sets that represent the intent of the decision support rule (and the corresponding quality measures).

- If we achieve what I have just described, there will be much to celebrate. Over time, as quality measures/CDS rules propagate, we can enhance the breadth of uniform descriptors from a few hundred items to several thousand. I would not support a wholesale attempt to map local terms in the tens of thousands of implemented systems to industry standard vocabularies.

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

- Most vendors include updates to licensed vocabularies on a monthly or quarterly basis. Local vocabularies, lab compendia, etc., are much more static and may not be updated in a substantive manner for years and requires more resource investment on the part of the provider.

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

- For licensed vocabularies and value sets, the date and version number (if supplied) is maintained. For locally created content, the date/time, version, creator, and active/inactive identifiers are maintained.

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?*

- Generally, there are not universally applied best practices for versioning of vocabularies in the EHR vendor community. Some products use versioning well, with clinical documentation dovetailed with the measurement. Others do not, and this disconnect can cause quality measurement and clinical decision support to become “stale.”

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

- Again, there is wide variability in the marketplace largely due to the absence of a reliable trusted source of vocabularies and value sets. If versioning were to be managed “in the cloud,” vendors would rapidly embrace best practices here, and we could see an enablement of significantly improved quality measurement and clinical decision support capabilities.

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?*

- In general, most systems make no distinction between a value set for one purpose and a value set for another, and I would be concerned that an effort to create certain value sets for certain purposes would be problematic. Indeed, I question the necessity, for example, of having a “diabetes” value set for public health, another for clinical research, and yet another for clinical care. As we look carefully at the spectrum of value set sources, I can understand the wish to allow these various entities (CDC, USHIK, CDISC, NQF, LOINC, IHT-SDO, etc) to continue to create and maintain value sets that they believe best represent the needs of their domains, but without harmonization, we will continue to have chaos. How would an EHR know which “diabetes” value set to request from the ValueSetCloud?
- The question about context for value sets being maintained at the “message data element level” is unclear to me. Are we being asked about transactional systems (in which messages carry value set references)? Or are we being asked about EHR systems in which value set pointers are stored? Perhaps we can discuss this further during our meetings.
- English Language intent is managed well by only a fraction of systems in the marketplace today, and they all use some form of interface term system (locally developed or licensed), ideally with mappings to reference or administrative vocabularies. Such interface terms are in fact value sets that represent a collection of concepts using one word that is familiar to a provider.

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

- In general, web links require systems to be online and have reliable connectivity all of the time. If systems are to rely on cloud-based lookup functionality, we have found that redundant Internet connectivity is imperative, as is sufficient bandwidth to provide lookup with sub-second response. This generally works well in an enterprise setting that has made heavy investments in IT infrastructure. It does not, however, work well in rural settings. Rural settings or small practices with limited

IT infrastructure will require local replication of vocabularies or value sets, either in the EHR or on a local “node” of the vocabulary server (similar to the Internet’s DNS system).

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

- There are many methods used for this, from CD-ROM mailings to web services. This is often left to the provider’s preference.

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

- Local customization is a legacy industry solution that we will have to manage for a long time. Local terms will not go away, but we can mitigate the propagation of local terms by providing GOOD vocabularies, interface terms, and value sets going forward. With these three elements made available to the United States Health IT community as a public service, local terms will ultimately become unnecessary. Deliberate creation and management of local or custom value sets will also be unnecessary, but this does not mean that such value sets won’t exist in some form.

For example, Google has an internal representation of a “value set” that returns the most useful results FOR ME in my next search for a given term, a restaurant, etc. In the same way, EHR systems can and will track patterns and usage history so that predictions can be made regarding terms that may be most useful. These predictions could technically be considered “value sets” but would be managed by the system, not by a human or set of humans.

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

- Some systems **prohibit** updates with local variations, or vendors sometimes advise against it with the message that this should not be done, even though the technology may support it. In other cases, local variations are expected and tolerated, but they must then be normalized using a central mapping engine in an enterprise or community.

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

- As with #14 and #12 above, this varies with the level of local IT infrastructure (both technical and human). Ideally, local vocabularies would be minimized, or would be automatically maintained and updated in sync with external sources. Latency is the greatest concern, but as anyone who has used “Google suggest” or “Bing autocomplete” knows, latency can be minimized with appropriate technical infrastructure and adequate bandwidth.

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

- **Interface terms.** Users are unfortunately accustomed to the dysfunctional use of ICD-9 and CPT. Great interface terms will:
 - i. Prevent providers from having to learn ICD-10 – ever!
 - ii. Better capture clinical intent for diagnoses, procedures, allergies, findings and medical/surgical history
 - iii. Enable better quality measurement
 - iv. Facilitate clinical decision support