

OCC FAQ

This Q&A might help in putting to speed those trying to understand the OCC.

Let me start by saying that the OpenMRS Concept Cooperative is a really hard thing to fully explain and for others to grasp, b/c IMO it represents a large amalgam of new ideas that haven't traditionally been placed together.

Question 1.

While the OCC doesn't intend to serve as a standard vocabulary, you (Paul) wrote on the wiki page that "commonly used concepts will themselves become de facto standards". **Isn't this as sort of conflicting?**

In the long run, the OCC is probably "consciously" avoiding being called a controlled 'standard' vocabulary set (while it is)? I think that what is different from other controlled vocabularies is that its contents will be "driven up" from implementation sites fostering a web 2.0 kind of collaborative control/standard (a standard/control that will evolve with the community). I will call the OCC, a community-controlled and standardized vocabulary.

From my perspective, it's one thing to build a framework that has a dedicated intention to become a standard, and another to take the hard work of the community and aggregate it in such a way as to see what evolves as common conventions. Having built a fair bit of vocabularies which drive system architectures (like the OpenMRS architecture), I'm convinced that there isn't a "right" or a "wrong" way to model concepts. There's likely however, a set of conventions to modeling concepts that derive the most value for the most people when it comes to clinical care and retrospective analysis. I have seen no real evidence of a consensus on this, and it's pretty clear as to why that is. The amount of inertia required to get developers thinking about how to model concepts in synonymous ways much less how they've chosen to represent medical facts within those metadata models is significant. The point of the OCC, therefore, is to aggregate the community's hard work and make it available to everyone else in such a way as to evolve common conventions and commonly used concepts. Why would we want to obscure the inevitability of what OCC evolves? I believe in an idea penned by James Surowiecki called the "wisdom of crowds", which basically points to the notion that a group of people that meet a certain set of requirements are inevitably smarter than a few content experts. Those characteristics are:

- Diversity of opinion: Each person should have private information even if it's just an eccentric interpretation of the known facts.
- Independence: People's opinions aren't determined by the opinions of those around them.
- Decentralization: People are able to specialize and draw on local knowledge.
- Aggregation: Some mechanism exists for turning private judgments into a collective decision.

Understanding these characteristics will give you some sense as to why we're building the OCC (what I consider to be the aggregator) in the way I've described. If you think about vocabulary developers in the OpenMRS community, they fit these first three requirements well. That is assuming that they don't participate in the OCC knowing that they are creating a standard. A good quote: "The irony of group wisdom is that it is only when a group is unaware of its intelligence that it can be effective." We want people to contribute and have independent thinking, drawing on their local knowledge. Hope I didn't lose you with this.

Question 2.

More than the concepts just being drawn from the OCC by intending implementations for forms design, lets not lose sight of the fact that the concepts being used by implementers are nothing compared, in volume, to the amount of data/concepts being entered by users (clinicians, etc). To paraphrase, using data from openmrs (clinical research questioning, decision support, etc) requires a lot more robustness in concept representation for user-entered data. Providing this for openmrs will (maybe... might) become inevitable as we intend to integrate with numerous other important systems eg dhis(who also have their own concept/data dictionary). Also, openmrs implementations might need to share forms and data (while maintaining meaning in form concepts and data concepts). This happens to be one of the goals of the UMLS - a comprehensive mapping system that allows disparate implementations for different/related purposes with sometimes different vocabularies to share data meaningfully (with minimal informational loss) while allowing structured representation of data that can be used for research, decision support. **What will be the relationship between the UMLS and the OCC? *If implementations utilize licensed vocabularies mapped within the UMLS (or directly from the licensee), when they share their concepts, does the OCC intend to leave the details of using these concepts with their licenses to each implementation?**

Your question about the UMLS is a good one. You're correct in understanding that the OCC will, by the product of its ability to aggregate concepts, create necessary mappings between OpenMRS implementations. This will serve (at some point) as a possible foundation to allow OpenMRS implementations to share information between systems using messaging protocols such as HL7. However, OCC's primary intention is to serve as a pragmatic "starting point" for those interested in populating their own OpenMRS implementation with a dictionary that meets their local needs. OpenMRS installs will not come with a "starter" vocabulary over the long run. I'm frankly somewhat hesitant whenever we share the current "starter set" with the community, knowing that I created it, and knowing that a lot of the decisions I made in how to model concepts are disputable. We as a community can do better than that. OpenMRS very soon will come with an ability to link up to this service and browse the OCC much like they would their local vocabulary. This functionality is beyond the scope of the UMLS. Additionally, the atoms of the UMLS metathesaurus by their very nature have disparate metadata models associated with their source origins. Not a good starting point for a practical OpenMRS implementation. So, while there are similarities in what they might look like on their surface, they are fundamentally different tools for different purposes.

Question 3.

Is there anybody using openmrs for a wide range of clinical data - eg primary care? I can imagine that the concept dictionary behind such will be really huge.

Yes. Yes, the dictionary can get large if not modeled optimally.

Question 4

What vocabularies are you drawing from? Are you using a defined UMLS subset?

We are not drawing from any reference terminologies, as they are not the foundation of our concept dictionary. Standard vocabularies such as LOINC and SNOMED, which were first built as a way to communicate through message protocols, are mapped to OpenMRS concepts. One of the fundamental choices we made early on is to not start from standard vocabularies as their metadata doesn't completely describe the information needed to drive system development. The OCC draws from the work of OpenMRS implementations, and so much as they map their concepts to LOINC and SNOMED, etc, then everyone else will potentially benefit from this work. Hope this makes sense.

Question 5

Can you share concept proposals within your implementation?

Concept proposals will not be shared immediately with the OCC central server. Once proposals are assimilated as either synonyms of a given concept or a new concept into the implementation's master dictionary, it will then be imported into the OCC.

Question 6

I think there is a need for reuse of modeling methodology - a kind of Openmrs forms collaborative. This will help intending implementers to get to speed with modeling data at their implementation site. What will be most important is not the forms themselves (since they should as close as possible copy their paper parent) but "what-has-been-learned-from-forms-design". **eh?**

I agree wholeheartedly. One of the early features we intend for OCC is to have the potential for threaded conversations about each concept. As you correctly realized, ultimately: the real end game for HIV care (and other high priority conditions in the developing world), will be forms that the clinical community evolves as standards for patient care, modeled in such a way that the subsequent data generated is re-usable to foster decision support and retrospective analyses. Today, there's significant rebellion against forms such as the IMAI, etc, b/c they've not evolved from the practical experiences of the clinical enterprise in the trenches. Perhaps our work can help to contribute towards that panacea.

Also see [Form Bank](#) in combo with the [Metadata Sharing Module](#)

Question 7

TBD