Strengthening Metadata Management for Governments and Formalizing OCL Governance

Submitted by Jonathan Payne (Open Concept Lab) on January 19, 2018 - 7:57pm
Last revised by Web Producer on June 21, 2018 - 3:09pm.
Proposal Status: Awarded--Pending Funding

Executive Summary

This proposal is being submitted on behalf of the Open Concept Lab community, a consortium of partners being represented, for this proposal, by Apelon, Cooper/Smith , OpenMRS, and Regenstrief Institute. The Open Concept Lab (OCL) is an existing "global good" consisting of an open-source terminology management system (distributed under MPLv2 with a healthcare disclaimer) to help collaboratively manage, publish and use metadata in the cloud alongside the global community. Imagine GitHub for indicators, terminology, and metadata-- a one-stop shop to access international standards, create and publish your own definitions, or browse country and global indicators with mappings to the diagnoses, procedures, or other data definitions used to calculate them.

We propose to use this award to achieve two objectives:

- Implement a sustainable governance model for the OCL software and community, establish ties to new software development resources, and strengthen community engagement. These are important steps to prepare for significant expected growth over the next two years to ensure mechanisms are in place to concretize community ownership, sustain development, and support our users.
- Make significant improvements to OCL's usability and functionality specifically for governments. Governments are a key stakeholder of OCL and these improvements will allow government administrators and partners to more easily perform collaborative management and publication of nationally endorsed information standards, adopt standards across a wide range of use cases and levels of expertise, harmonize data collection and reporting approaches across partners, and enable electronic information exchange. Government use cases and requirements will be informed by the Kuunika Project in Malawi with Cooper/Smith as a technical partner and the Ethiopia Data Use Partnership, both which are actively advancing their terminology management strategies.

Most countries struggle with the quantity and fragmentation of data collection tools and reporting requirements within the health sector. For example, at the facility-level in Malawi, health workers (HWs) record patient transactions and manually aggregate data into government registers and reports, as well as those required by donors. The sheer volume of data that HWs must manage and report does not only place an excessive burden on an already stretched workforce, but also results in poor data quality and sub-optimal reporting rates. Often, the requested information is duplicative as the data elements are not standardized. Managers and policy makers struggle to identify the right data and reconcile discrepancies across systems and reporting streams, leading to insufficient data use. A mapping of the data elements and indicators used across data collection forms, registers, and reports would allow governments to harmonize definitions, reducing duplicative reporting and enhancing data use.

By adopting sharable definitions for data and indicators, we empower stakeholders to exchange data effectively and make interoperability possible. For data use to be institutionalized among low and middle income countries (LMICs), there are several key obstacles that must be overcome. Among those obstacles is the limited access to already-defined indicators and terminologies, such as reference terminologies (WHO ICD-10, LOINC, HL7) or country-defined content such as Health Management Information System (HMIS) indicators or subsets representing domain-specific priorities. A second key obstacle is the absence of appropriate and accessible tools to support the metadata management and publication needs of resource-constrained governments or organizations working in the development sector. While the initial use cases for OCL were focused on health, there is every reason for other sectors to leverage the platform.

Like other global goods, OCL is both open-source software and a community. However, OCL has two other distinct elements: a hosted web presence and content. The hosted web presence is a global instance called the OpenHIE Metadata Clearinghouse that is a central service for organizations to publish public metadata--such as indicator definitions or clinical concepts--for their stakeholders and the international development community. Stakeholders may still implement a local instance of OCL to supplement the hosted service. The content is a growing set of reference vocabularies, country and donor indicators, and other definitions from donors, countries, universities, and implementers throughout the world.

Consortium Team

OCL is supported and used by a community of health, informatics, and technical partners. Past funding for OCL has supported a consortium of partners -- including Partners In Health, Columbia University, Vital Wave, Thought Works, Regenstrief Institute, and Apelon -- to contribute to software development and terminology curation. For this project, the consortium will be comprised of:

Apelon Inc. - Apelon is an industry leader in clinical informatics focused on accelerating eHealth through data standardization and interoperability. Apelon provides professional services, contributes to standards development and thought leadership through HL7, AMIA, ISO, among other forums, and supports the open source Apelon Distributed Terminology Service (Apelon DTS), which serves as a critical service in many health information exchange implementations.

Cooper/Smith - Cooper/Smith is a technical assistance organization that uses hard data to increase the effectiveness & efficiency of development programs worldwide. We have experience leading full-cycle, large-scale data initiatives including requirements collection, defining metrics, tool and system development, piloting and scale-up, and applied used of results. We have deep knowledge of strategic planning, evidenced-based resource allocation, and DHIS2 roll-out and training (specifically DATIM – PEPFAR instance of DHIS2). Our staff have served in many roles: researchers, economists, strategic information specialists, policy advisers, software developers, and clinical coordinators.

OpenMRS Inc. - The OpenMRS community is currently pursuing the design and development of a new user interface for OCL that will serve as the technical foundation for the user interface-focused software development activities proposed in this grant. We will leverage pre-existing OpenMRS community calls to coordinate and advance the design and development of this work.

Regenstrief Institute - The Regenstrief Institute has a long history of expertise on health informatics and will provide strategic and technical inputs into many of the proposed work streams. In addition, Regenstrief Institute supports the OpenHIE community, which will provide access to a broad community of experts and potential implementers and serve as a meeting point for this project via the extant Terminology Services subcommunity.

The Open Concept Lab LLC was established in 2017 to serve as a temporary legal home for OCL while an appropriate long-term structure is established, which is a key objective of 2018 for the OCL community. The OCL LLC is an important step in ensuring community ownership of OCL longer-term.

Project Description

There are several activities that we propose for inclusion in this project, centered around two key objectives:

1. Implement sustainable governance, build ties development resources, and strengthen community engagement - These are important steps to prepare for significant expected growth over the next two years to concretize community ownership, sustain development, and support our users. A recently established OCL Leadership Committee will help to ensure that the established mechanisms amplify our mission and set the stage for future growth. Activities to include:

- Coordinate the OCL Leadership Committee, a recently created multi-organization team, to recommend and support

establishment of an appropriate and sustainable long-term governance model, ensuring that OCL remains "community-owned"

- Coordinate development of an OCL roadmap built around community input (leveraging shared resource if available)
- Publicly launch OpenHIE Metadata Clearinghouse, the central, cloud-hosted instance of OCL designed to serve as a key source of shared vocabulary within health information exchanges and significantly lower the barriers to harmonizing data and indicator definitions across governments, donors, and projects
- Coordinate periodic community calls with public notes, leveraging the OpenHIE Terminology Services call, especially for topics related to the OpenHIE Metadata Clearinghouse, and adapting into separate calls (e.g. design, developer, leadership), as community needs evolve. Will include support for an OpenMRS Early Adopters Team with a clear timeline for wider testing and adoption of the OpenMRS-OCL Subscription Module and advising on needed support materials and processes.
- Overhaul OCL public web presence to better communicate mission, governance model, community ownership, content licensing, support processes (e.g. new virtual help desk system), and engagement opportunities and to provide easy access to documentation and training materials
- Improve user and technical documentation and support materials (leveraging shared resource if available). This may include implementing auto-generation of technical reference documentation from the code.

2. Enhance OCL for Ministry of Health use case - Improve support for Ministry of Health administrators and their partners to collaboratively manage, publish, harmonize, and use their national health data dictionaries.

It is important to note that the OpenMRS community is pursuing the development of an " OCL for OpenMRS" user interface (currently in the design phase). The "OCL for OpenMRS" user interface will serve as the technical foundation for the user interface-focused software development activities proposed in this grant. We will leverage pre-existing OpenMRS community calls to coordinate and advance the design and development of this work.

- Implement a modern and interactive web interface for navigating published metadata, visualizing relationships, and highlighting duplicates, that is more intuitive to government workers who may not have a background in terminology. The new interface will sit on top of the existing OCL API, replace portions of the existing OCL web user interface.
- Improve dictionary curation workflows in OCL, specifically to support the needs of a district or regional health officer. The key improvement will be to add support for content proposal, review, and verification processes. This may also include support for predefined attributes/properties of a definition based on resource type.
- Expand metadata search to custom attributes, with improved sorting and filtering (e.g. fuzzy search and facets).
- Support for "visual diffs" between resources and resource versions.
- Simplified and more flexible imports, exports, and search downloads.

Use Cases, User Stories and Activities

There are three primary target audiences we consider for terminology management services:

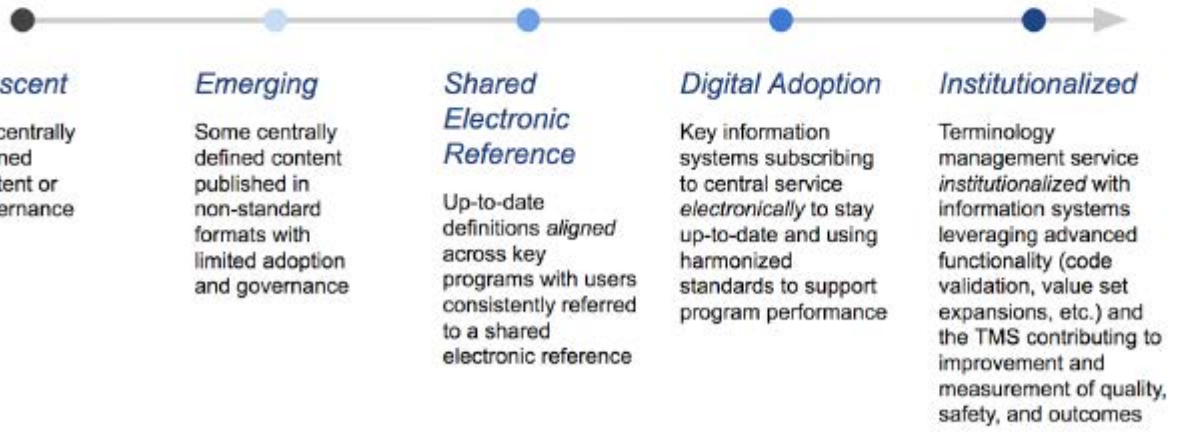
- Governments, who want to collaboratively manage and publish nationally endorsed information standards, harmonize data collection and reporting approaches across partners, and enable electronic information exchange
- Terminology consumers, especially implementers of point of service systems like OpenMRS, that want to define a locally relevant subset of terms, propose edits or new terms, and subscribe to it to fetch changes as they are made available

- Donors and other metadata publishers, such as PEPFAR, who want to simplify management of indicator definitions, maximize uptake of published standards across their partners, and improve quality and timeliness of electronic reporting

The latter two use cases (terminology consumers and metadata publishers) are progressing through separate activities:

- The OpenMRS community is launching the OCL Subscription Module later this year (undergoing final testing now), which is a key part of its transition to OCL as a primary terminology distribution service. A key component of this is the Columbia International eHealth Laboratory interface terminology, a popular vocabulary within the OpenMRS community that also uses OCL as a key distribution service.
- The OpenMRS community is also supporting the design and definition of a new user interface for OCL, as discussed previously in this proposal.
- PEPFAR is incorporating OCL into the DATIM ecosystem to host PEPFAR indicators and to automate transformations of country-level results data into standardized indicator representations. It is expected that DATIM will be using OCL to facilitate data transformations in ten countries by the end of 2018, with more countries expected in subsequent years.
- In addition, the OpenHIE community established the OpenHIE Metadata Clearinghouse, hosted on OCL, as a global resource for the publication of publicly available metadata.

The focus of the software development for this award will be on the third use case: Making the tool easier for government workers that may have limited experience with terminology to use for management, publication, harmonization and adoption of metadata within the health system. An appropriate and accessible tool that facilitated these processes would accelerate a country's progression to higher levels of terminology management sophistication. Five levels of sophistication of terminology management were described in a maturity model drafted in collaboration with the Ethiopia Data Use Partnership and are shown below. Maturity levels vary by country, however the experience of the authors indicates that helping countries transition from an "Emerging" terminology management level, in which health information standards begin to be managed and published centrally, to the "Shared Electronic Reference" stage is of particular importance. (Note that these maturity levels are still evolving, especially to allow for integration into the HIS Interoperability Maturity Model developed by MEASURE Evaluation.)



Specifically, following are a list of activities on the critical path to advance from "Emerging" to "Shared Electronic Reference" that we want to directly facilitate:

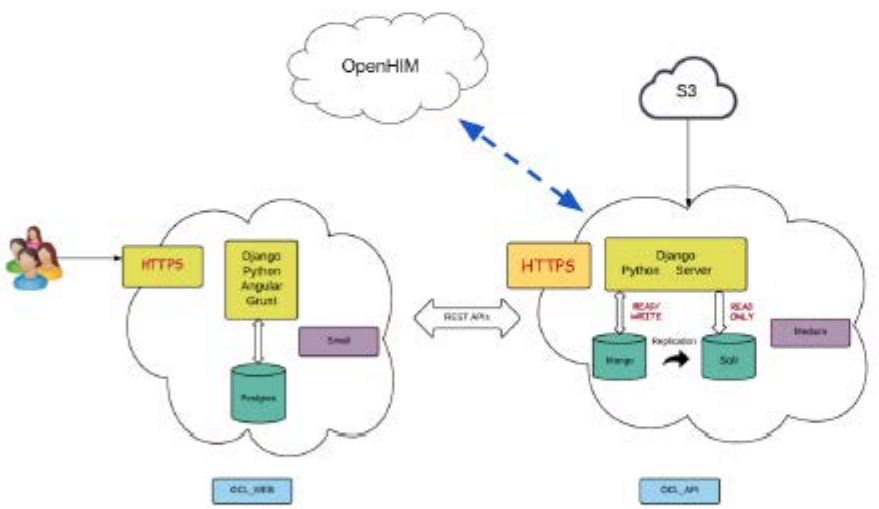
- Digital publication of priority nationally endorsed health information standards mapped to international reference vocabularies. Early endorsed standards typically include HMIS indicators, disease classifications, drugs and supplies lists, and insurance claims.

- Modeling and publication of data element definitions from data collection instruments used within the health system, again, mapped to international reference vocabularies.
- Harmonization and reconciliation of data collection and reporting requirements across programs, partners, and tools facilitated by an electronic tool that highlights duplicates and close matches.
- Providing an electronic service to the health system for management, publication, and distribution of information standards to health system in a variety of formats. This allows for expansion into other domains.
- Making published information standards available as a foundational service within a national digital health architecture, allowing for integrations and metadata subscriptions.

Digital Health Technologies

OCL currently consists of an OCL API Core that can standalone and a thin web user interface to simplify access. OCL has also implemented OpenHIM to support custom presentations and synchronization of metadata between information systems. OCL is cloud-hosted on the JetStream environment, provided by OpenMRS Inc and Regenstrief Institute. The diagram below shows the current architecture of OCL and technologies utilized.

TECHNICAL ARCHITECTURE - CURRENT



- S3 cloud storage is used to cache exports of versioned repositories
- Redis (in-memory database) used on OCLAPI to handle cron jobs for asynchronous processes
- Docker-based deploymet on JetStream cloud
- Bamboo used for continuous integration

Software development activities described in this award will consist of the following:

- OCL API Core -- OCL utilizes a REST API first model, meaning that 100% of functionality is exposed through the API, and the web interface simply uses the API. Some of the required change, including the metadata proposal/review process, expanded search capabilities, and the ability to evaluate a diff between resources will require changes to the API. Note that this does not involve a significant refactoring of code.
- OCL UI 2.0 -- The user interface improvements will be built using ReactJS, which will connect to the OCL API. This will also leverage work being done by the OpenMRS on the "OCL for OpenMRS" user interface. Additional technology decisions and requirements will be coordinated along with the OpenMRS community.

Community Feedback

We plan to host at least one platform-specific call and to leverage existing forums to the extent possible.

- OpenHIE Terminology Services Forum - Monthly call hosted by the OpenHIE community that will serve as a primary forum for discussions related to the OpenHIE Metadata Clearinghouse, in addition to other community topics. This is a key forum for public engagement on the use of OCL.
- OpenMRS Design Calls - Bi-weekly calls hosted by the OpenMRS community that we will join as appropriate to coordinate design and development of the new OCL UI.
- OCL Design and Dev Calls - Initially, OCL will host a bi-weekly or monthly open call to address community design and developer topics. We will adapt the frequency and focus of calls as the community's needs evolve.

We will host other project-specific meetings as needed.

Global Good Maturity Model Self-Assessment

The OCL Global Goods Maturity Model assessment can be found in this spreadsheet:
OCL Global Goods Maturity Model Assessment

Workplan, Project Deliverables & Timeline

Assuming start date of Summer 2018, we propose these milestones:

- Fall 2018 - Implemented Community Processes that will be ongoing throughout the project
- Summer 2019 - Established governance model and production deployment of new functionality and user interface to support government metadata management

Supporting Documents: global_goods_budget_narrative_ocl_2018-03-23.docx