CCDH Internal Monthly Report - 2021-02-17

2021-02-17

Community Development Update

1. **Progress over the last month**
   a. End-to-end Requirements Analysis (related [ticket](#))
      i. Draft timeline and description of deliverable (in progress),
      ii. Continue iterating to improve on CRDC data life cycle diagram - (from planning doc) to drive the E2E assessment.
   b. Documentation - aggregating documentation is in progress
      i. Fed Oversight team is consulting with NCI for the final SOP for releasing docs. May be a hybrid approach.
   c. Reimagining office hours:
      i. New community session will occur in April (after the 14th), details are in development:
         1. 10 min intro
         2. 20 min Q&A
      ii. ~~Nicole may present on Mondo at the inaugural meeting?~~
         https://mondo.monarchinitiative.org/
         1. Part of an ongoing terminology discussion with Lyuba on how we handle the various tools we employ (will wait a few months until we have a better understanding of cancer landscape and options) - terminology team needs to meet and discuss w/ Chris Mungall
            ToDo: check on status of LBL team
         2. Could use office hours to discuss terminology options; could talk about value sets, CaDSR
         3. Formal presentation of DMH model as an idea for office hours
         4. E2E report presentation. E.g., discuss file format for data transfers across different parts of the system
   d. Drafting Support Service and Engagement Plan
   e. Website updates needed (Jim, Brian and Jen):
      i. https://github.com/cancerDHC/community-development/issues/37
      ii. https://github.com/cancerDHC/community-development/issues/43
      iii. https://github.com/cancerDHC/community-development/issues/44
         Notes:
         ToDo: CD, T&DQ ws to work on a roadmap for deployment for functional components: data model navigation, terminology services, etc. Plans for implementation.
2. **Timeline check:** what's the status of your work compared to our project milestones?
   a. End-to-end Requirements Analysis - *anticipated completion date is end of March*

b. Documentation - aggregating documentation is in progress
      i.   Is there an update on where we'll host the documentation? GitHub?
  c. Reimagining office hours - announcement went out.
  d. Drafting Support Service and Engagement Plan - on pace for on time delivery
3. **Work for the upcoming months**
  a. Continue with tasks above

2021-02-10

Data Model Harmonization Update

1. **Progress over the last month**
  a. Meeting with ISB / CDA personnel on 1/29 to discuss helpdesk topic regarding Administrative Subdomain entities
  b. Work on remaining Administrative subdomain entities (Program, Project, Case, Study) continues
      i.   Incongruence across existing node models as reported by ISB / CDA
      ii. Examination of instance data in PDC wrt crossover with GDC c=
  c. Initial proposal for [data model change log](#) presented to DMH group on 2/1
      i.   Feedback welcome!
          1. E.g., Look at how CaDSR group handles changes and what sort of mechanism they use for collecting changes in upstream processes.
      ii. Suggestion to share details about the existence of this log to all nodes during the All Hands on Feb 22nd and 23rd.
  d. Introductory meeting with CTDC on 2/2
      i.   Follow-up meeting plans being discussed

  Notes:

  From Melissa & CCDH team:
  ● We planned to have a follow up meeting to review the Pediatric Commons model for trials, and invite a few additional people from our respective teams;
  ● We discussed mapping the needs for trial attributes across nodes, and then prioritizing these for addition to CRDC-H and otherwise integrating the Pediatric Commons Trial model within CRDC-H;
  ● We were going to do a deeper dive on the MATCH model;
  ● We were going to get CRFs from relevant sources to compare to all of the above.
  ● MH: I think it is very timely to do this alignment since there isn't too much yet in CTDC and as such, the prior work on the Pediatric Commons can be utilized for this next phase of CRDC-H with requirements from the above sources.

  From Cristina Russo:
  To complement Melissa's notes, below are my notes and action items.

ACTION ITEMS:
- CTDC to share roadmap of the datasets with CCDH to understand priorities
- CCDH/ Sam to give a demo on harmonization done in pediatric clinical trials
- John/David to provide CCDH with the data model used for the MATCH arms in the MVP
- Mark to provide CCDH with eCRFs for the Biobank dataset

Meeting notes:
- CTDC presented current preliminary plans for CTDC and expected datasets coming from multiple sources (high-level info in PPT)
- CTDC stated the wish to explore a universal model and inquired on alignment with CRDC-H efforts. However, CRDC-H's recent focus is in CDA, so CCDH will support this project as consultants, reviewers, and SMEs, and both teams will work on an iterative way. One of the goals for the model is to support interoperability across the nodes, and CCDH explained this is better done prospectively.
- CCDH wants to see what CTDC's current model covers and also see the CRFs of upcoming data sets to understand how they align. It would also be helpful to have a list of data sources and transformations, and of potential desirable requirements for the "perfect" harmonized model. CTDC has asked that owners of the data sets provide examples to CCDH.

2. **Status (compared to our timeline)**
   a. Generally tracking well, but <mark>need to discuss some prioritization of efforts</mark>, especially wrt bringing in new models (e.g. IDC) and consulting.

3. **Next steps, work for the upcoming month in the context of workstream or cross-cutting plans**
   a. Complete work on remaining Administrative entities
   b. Work on Clinical subdomain entities identified by CDA as priority elements for query (including diagnostic and therapeutic attributes)
   c. Continue work on tooling to support automation of model development workflow
   d. Continue regular tech meetings with CDA
   e. Continue discussion with CTDC around data model for clinical trials data

---

2021-02-03
Tools & Data Quality Workstream Update
1. **Progress over the last month**
   a. Reviewed NCIt + OBO integration workflow with Dazhi and Harold

i.   Identified some improvements to make to align with current approaches
                    1.   e.g. integrate SSSOM mapping format
            ii.   Discussions on plans for further development of pilot
                  harmonization/mapping tools
                    1.   Define interaction with CRDC-H and terminology server, in terms of
                         searching value sets, validation, etc.
                    2.   Develop existing prototype into one or more services that could be
                         plugged into different stages of the data lifecycle
                            a.   e.g., integrated as a data source for Ptolemy, or instead
                                 incorporated into a non-graphical workflow
2. **Status (compared to our timeline)**
    a.   Aiming to develop exploratory tools into a pilot implementation this quarter.
3. **Next steps, work for the upcoming month in the context of workstream or cross-cutting
   plans**
    a.   Work through concrete plans for interaction between CRDC-H, terminology
         servers, and harmonization tool with DMH & Terminology groups

---

2021-01-27
Terminology & Ontology Ecosystem Workstream update.
    1. **Progress over the last month**
        a.   Analyzed GDC and PDC emunrated terms and permissible values
        b.   Working on constructing a CCDH ontology for ICD-O terms which are used in
             several fields in the GDC and PDC data dictionary
        c.   Extracting the TCCM minimal representation of terms from several ontologies
             (NCIt, Mondo, HPO)
        d.   Discussed CCDH model and terminology in FHIR at HL7 BR&R
        e.   Preliminary ontoserver review
    2. **Status (compared to our timeline)**
        a.   Behind schedule (Q4 milestone)
    3. **Next steps, work for the upcoming month in the context of workstream or cross-cutting
       plans**
        a.   Finish the CCDH ontology for ICD-O terms
        b.   Bring up a testing service for concept descriptions
        c.   Complete the Q4 milestone

2021-01-20
Meeting was not held in observance of the Federal Holiday to Celebrate the Presidential
Inauguration.