Update on the OHDSI Standardized Vocabularies

05-March-2024

Alexander Davydov
Oleg Zhuk
Anna Ostropolets
Overview

1. Current status
   ○ Vocabularies improvement plan
   ○ New mode of operation
   ○ Roadmap and releases
   ○ New features in Athena

2. February release in detail
   ○ 3 overhauls
   ○ Vocabularies on maintenance

3. Community contribution
   ○ Submissions
   ○ Other ways of contribution

4. Q&A during the Vocabulary WG meeting
Who is talking to you

Alexander Davydov

Oleg Zhuk

Anna Ostropolets
The team behind us

Timur Vakhitov
Aliaksei Katyshou
Vlad Korsik
Maria Rogozhkina
Mikita Salavei
Varvara Savitskaya
Irina Zherka
Dmitry Buralkin
Tetiana Orlova
Tanya Skugarevskaya
Janice Cruz
Masha Khitrun
Current status on improvement plan and other activities
Vocabularies process improvement

In 2023, we launched an initiative to systematically improve vocabularies

- Alignment with the community
- Roadmap and release schedule
- Community contribution: pipeline & other forms
- New Quality Framework and QA system
- Common development infrastructure
- Vocabulary versioning
New mode of operation

• Alignment with the community
Systematic outreach through landscape assessment, forums, Vocab WG

• Focus on most important and painful points
Committee for prioritization, transparent and predictable roadmap, stable releases

• Scalability
Community contributions as a pathway to accommodate community needs and build a collaborative community
Roadmap and releases

- New roadmap timeline spent: 429/609 days (70%)
- Winter release v20240229 has just arrived to Athena
Vocabularies process improvement

We launched an initiative to systematically improve vocabularies

- Alignment with the community
- Roadmap and release schedule
- Community contribution: pipeline & other forms
- New Quality framework and QA system
- Common development infrastructure
- Vocabulary versioning through new Athena
Vocabulary versioning in Athena *(very soon…)*

- You can pick a version (release) when downloading
  - we’ll include all important releases since Sep, 2021
- You can generate delta when downloading: SQL script or diff tables
February release in detail
Why you should download this vocabulary release

• **More concepts:**
  - Refresh of SNOMED, MedDRA, ICD10PCS, ICD10CM, CVX, RxNorm and more

• **Better hierarchies:**
  - Improved LOINC - SNOMED hierarchy
  - de-novo constructed MedDRA - SNOMED hierarchy

• **More good mappings:**
  - ICD family refresh
  - community contributions
  - bug fixing

• **What you specifically asked for:**
  - We closed 41 GitHub issues and addressed many forum posts
1. Important bug fixes from the Regenstrief institute

1. We incorporated LOINC-SNOMED hierarchy from their strategic partnership and built more on top of that.

1. We deduplicated some concepts between SNOMED and LOINC

<table>
<thead>
<tr>
<th>LOINC</th>
<th>SNOMED</th>
<th>CT Heart</th>
<th>Non-standard + mapped</th>
</tr>
</thead>
<tbody>
<tr>
<td>58744-4</td>
<td>241547009</td>
<td>Cardiac CT</td>
<td>Standard</td>
</tr>
</tbody>
</table>
SNOMED: main changes

1. Improved domain assignment and its stability

2. De-standardized concepts:
   - Attribute
   - Location (except countries)
   - Social Context (except relatives, religion, occupation)
   - Physical Force and Physical Object (except Device domain)
   - Secondary neoplasms (now mapped to Cancer Modifiers)

Watch WG meeting
SNOMED: main changes

3. Helped you to build consistent cohorts by post-coordinating allergies and tests+results

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>165516004</td>
<td>Neutrophil count within reference range</td>
</tr>
<tr>
<td>30630007</td>
<td>Neutrophil count</td>
</tr>
<tr>
<td>281301001</td>
<td>Within reference range</td>
</tr>
</tbody>
</table>

Maps to value
SNOMED: other changes

4. Refreshed three editions:
   - 2023-07-31 SNOMED CT International
   - 2023-09-01 SNOMED CT US Edition
   - 2023-09-27 SNOMED CT UK Edition

5. Excluded SNOMED UK Drug Extension

6. Improved validity dates - now fully correspond to the source data

7. Deduplicated concepts inside SNOMED (with mappings)

Read release notes
MedDRA: main changes

1. We built a new hierarchy!
MedDRA-SNOMED CT Hierarchy Rules

- Complete semantic correspondence between MedDRA and SNOMED
  MedDRA above SNOMED
- Incomplete correspondence
  MedDRA position varies

Important!

If you use MedDRA-SNOMED hierarchy for concept sets, check descendants as not all of SNOMEDs may be included.
Example of complete semantic correspondence between MedDRA and SNOMED

<table>
<thead>
<tr>
<th>MedDRA High-level terms (HLT)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Is a</td>
</tr>
<tr>
<td>MedDRA Preferred terms (PT)</td>
</tr>
<tr>
<td>Is a</td>
</tr>
<tr>
<td>MedDRA Low-level terms (LLT)</td>
</tr>
<tr>
<td>SNOMED-CT</td>
</tr>
<tr>
<td>Is a</td>
</tr>
<tr>
<td>SNOMED-CT</td>
</tr>
<tr>
<td>Is a</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>35104833</th>
<th>Eosinophilic myocarditis</th>
<th>Condition</th>
<th>MedDRA</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>36312698</th>
<th>Colonoscopy</th>
<th>Procedure</th>
<th>MedDRA</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>4141815</th>
<th>Eosinophilic myocarditis</th>
<th>Condition</th>
<th>SNOMED</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>4249893</th>
<th>Colonoscopy</th>
<th>Procedure</th>
<th>SNOMED</th>
</tr>
</thead>
</table>

February release: MedDRA
Examples of incomplete semantic correspondence between MedDRA and SNOMED

<table>
<thead>
<tr>
<th>MedDRA ID</th>
<th>Description</th>
<th>Type</th>
<th>MedDRA (HLT)</th>
</tr>
</thead>
<tbody>
<tr>
<td>37503861</td>
<td>Joint therapeutic procedures</td>
<td>Procedure</td>
<td>MedDRA (HLT)</td>
</tr>
<tr>
<td>37520668</td>
<td>Tophus removal operation</td>
<td>Procedure</td>
<td>MedDRA (PT)</td>
</tr>
<tr>
<td>37604022</td>
<td>Non-site specific embolism and thrombosis</td>
<td>Condition</td>
<td>MedDRA (HLT)</td>
</tr>
<tr>
<td>36617339</td>
<td>Tumour thrombosis</td>
<td>Condition</td>
<td>MedDRA (PT)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SNOMED-CT ID</th>
<th>Description</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>4010109</td>
<td>Operation on joint</td>
<td>Procedure</td>
</tr>
<tr>
<td>4231363</td>
<td>Thrombosis</td>
<td>Condition</td>
</tr>
</tbody>
</table>
MedDRA: other changes

2. Refreshed MedDRA to the most recent version 26.1 (September 2023)

3. Added mappings to SNOMED for 11K PT and ~3K LLT terms

4. Started collecting mapping meta-data
ICD Family
ICD10, ICD10GM, ICD10CN, CIM10, KCD7 and ICD9CM

1. Developed new mapping reuse approach (Common Data Environment) to improve consistency of mappings for the whole ICD family

**Strict group** - concepts with the same or insignificantly different meaning
**Medium group** - concepts with close meaning
**Broad group** - groups by other ways of grouping (ICD categories)

<table>
<thead>
<tr>
<th>Resistance to antiviral drug(s)</th>
<th>ICD10</th>
<th>U84.2</th>
<th>{CIM10:U84.2, ICD10CM:Z16.33, ICD10:U84.2}</th>
</tr>
</thead>
<tbody>
<tr>
<td>Resistance to antiviral drug(s)</td>
<td>ICD10CM</td>
<td>Z16.33</td>
<td>{CIM10:U84.2, ICD10CM:Z16.33, ICD10:U84.2}</td>
</tr>
<tr>
<td>Resistance to antiviral drug(s)</td>
<td>CIM10</td>
<td>U84.2</td>
<td>{CIM10:U84.2, ICD10CM:Z16.33, ICD10:U84.2}</td>
</tr>
</tbody>
</table>
ICD Family: other changes

2. Based on most recent version of ICD10CM (April 2024)

3. Refreshed mappings with new SNOMED version for all ICD family concepts

4. Added 30K+ CIM10 concepts + translations (France)
ICD 10 PCS

1. Improved ICD10PCS - SNOMED hierarchy (removed 4K erroneous ‘Is a’ links)

2. Mapped all new drugs to RxNorm

3. Refreshed to the most recent version (April 2024)

4. Renamed XW0 section got renamed to show that it is Drug administration section

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>36802939</td>
<td>XW023 New Technology @ Anatomical Regions @ Introduction @ Muscle @ Percutaneous</td>
</tr>
<tr>
<td>1535226</td>
<td>XW043G New Technology @ Anatomical Regions @ Introduction @ Central Vein @ Percutaneous @ Drug administration</td>
</tr>
</tbody>
</table>
OMOP Extension

1. ~200 new concepts for antibody measurements, stage and scales (e.g., “Severity of alopecia tool”, “Itch Numeric Rating scale”)

1. Incorporated the Extracorporeal Life Support Organisation Maastricht Treaty to support the upcoming ECDM - Expansion for the Extracorporeal life support data
OMOP Genomic

- OMOP Genomic variants still contain somatic mutations relative to cancer. The rules of collecting OMOP Genomic concepts have changed in the following concept classes:
  - Structural variant (new name)
  - Gene Protein Variant (new name)
  - Gene Variant (new name)
  - Gene DNA Variant (new name)
  - Gene RNA Variant (new name)
Other changes

- Refreshed **RxNorm** (January 2024), **NDC/SPL** (February 2024), **CVX** (2023AB), **MeSH** (2023AB)
- **Improved mappings** for CPT4, Read, OPCS4, UK Biobank, OMOP Invest drug
- **SUS** is now distributed under an open-source license - thank you for your contribution!
- Other bug fixes

Read release notes
We try not to give you wrong mappings anymore

<table>
<thead>
<tr>
<th>concept_id</th>
<th>concept_name</th>
<th>relationship_id</th>
<th>target_concept_id</th>
<th>target_concept_name</th>
</tr>
</thead>
<tbody>
<tr>
<td>45019144</td>
<td>Lovastatin 20 MG Oral Tablet [Mevacor]</td>
<td>Maps to</td>
<td>42707418</td>
<td>24 HR lovastatin 60 MG Extended Release Oral Tablet [Altocor]</td>
</tr>
<tr>
<td>45314667</td>
<td>PV VINYL GLOVES</td>
<td>Maps to</td>
<td>4298923</td>
<td>Ibuprofen</td>
</tr>
<tr>
<td>45597198</td>
<td>Abnormal glucose tolerance test</td>
<td>Maps to</td>
<td>4149519</td>
<td>Glucose measurement</td>
</tr>
</tbody>
</table>
Community contribution in more detail
Community contribution: what and how

I found a mapping error!

I want to add my vocabulary!

My code is missing mapping!

More info: bit.ly/42qOscr
Contributions in Aug 2023 release

- Adding NHS Ethnicity codes and Admission and Discharge codes to the OMOP vocabulary
- ICD10CM: Limb vs. Lip in Carcinomas
- Adding New RxE Drugs

3 contributions: 2 simple and 1 complex use case

Contributions in Feb 2024 release

- To fix that ICD10CM "substance use with other disorders" mapped to Mood disorder
- ICD10CM: G92.0 Immune effector cell-associated neurotoxicity syndrome (ICANS) -> Maps to SNOMED 1230414002 Immune effector cell-associated neurotoxicity syndrome
- Sexual activity concepts in the incorrect domain of Condition. Should be Observation
- Neuropsychiatric Scores and Scales - Deduplication request [Psychiatry WG]
- SNOMED - RxNorm/RxNorm Extension deduplication
- OMOP Standard Concepts Deduplication Request
- Lovastatin 60 MG Maps to Lovastatin 20 MG

9 contributions: 5 simple, 3 medium complexity and 1 complex use case
<table>
<thead>
<tr>
<th>#</th>
<th>GitHub</th>
<th>What</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>850</td>
<td>ICD10CM-SNOMED <strong>change mapping</strong> for coma</td>
</tr>
<tr>
<td>2</td>
<td>856</td>
<td>ICD10CM-SNOMED <strong>change mapping</strong> for substance use</td>
</tr>
<tr>
<td>3</td>
<td>872</td>
<td>SNOMED <strong>change domain</strong> for sexual activity</td>
</tr>
<tr>
<td>4</td>
<td>920</td>
<td>ICD10CM-SNOMED <strong>change mapping</strong> for neurotoxicity</td>
</tr>
<tr>
<td>8</td>
<td>928</td>
<td>NDC <strong>change mapping</strong> for lovastatin</td>
</tr>
<tr>
<td>5</td>
<td>795</td>
<td>Psychiatry scales <strong>deduplication</strong></td>
</tr>
<tr>
<td>6</td>
<td>914</td>
<td>SNOMED-RxNorm <strong>deduplication</strong></td>
</tr>
<tr>
<td>7</td>
<td>915</td>
<td>Intensive care flowsheets <strong>deduplication</strong></td>
</tr>
<tr>
<td>9</td>
<td>896</td>
<td>Extracorporeal life support <strong>vocab</strong></td>
</tr>
</tbody>
</table>
### What got in and what did not

<table>
<thead>
<tr>
<th>#</th>
<th>GitHub</th>
<th>What</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>944</td>
<td>ICD10CM-SNOMED change mapping for carpal tunnel syndrome</td>
</tr>
<tr>
<td>11</td>
<td>933</td>
<td>Promote concepts for FNCLCC grades to standard</td>
</tr>
<tr>
<td>12</td>
<td>940</td>
<td>Change name for autoimmune polyendocrine syndrome type</td>
</tr>
<tr>
<td>13</td>
<td>851</td>
<td>Add new Type concept</td>
</tr>
<tr>
<td>14</td>
<td>846</td>
<td>Cancer modifier Grade deduplication</td>
</tr>
<tr>
<td>15</td>
<td>PR 925</td>
<td>Refresh ICDO3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>#</th>
<th>GitHub</th>
<th>What</th>
</tr>
</thead>
<tbody>
<tr>
<td>16</td>
<td>860</td>
<td>Create new route hierarchy</td>
</tr>
<tr>
<td>17</td>
<td>890</td>
<td>Refresh and decouple SNOMED Veterinary</td>
</tr>
<tr>
<td>18</td>
<td></td>
<td>Refresh EDI</td>
</tr>
<tr>
<td>19</td>
<td></td>
<td>Refresh CIEL</td>
</tr>
<tr>
<td>20</td>
<td></td>
<td>Refresh HemOnc</td>
</tr>
</tbody>
</table>
Contribution next steps

1. Complex contributions: we are prioritizing contributions to see what can be done + writing documentation for a new development schema setup.

1. Simple contributions: Stay tuned for a community-wide activity!
Summary

1. Please download the new release - it contains refreshes, improvements and fixes relevant to all of you!

1. We have multiple opportunities to get engaged (education, feedback and development) - come to the Vocab WG or shoot us an email!

Questions? Go now to Teams -> CDM WG -> Vocabulary WG to join the call