



OHDSI Goals for 2023

Pillar #1: Standardized vocabularies

- Opportunity: Increase transparency and maturity with vocabulary development and evaluation process
- Proposed solutions:
 - Conduct landscape assessment to understand community needs
 - Develop code of conduct and developer guidelines
 - Disseminate vocabulary process and end-user documentation and roadmap
 - Establish centralized development infrastructure
 - Create standardized test development
 - Build vocabulary version release distribution service





-
- The treemap visualization displays the distribution of 100 medical terminologies, categorized into four main groups: Observation (blue), Procedure (purple), Condition (red), and Device (orange). The size of each rectangle represents the relative frequency of the terminology. The largest terminologies are RxNorm Extension, SPL, GRR, RxNorm, BDPM, NCCD, Gemscript, VA Product, JMDC, NDFRT, GPI, DA_France, Multilex, AMT, KDC, NDC, DPD, dm+d, AMIS, RxNorm Extension, SNOMED, ICD10CM, ICD10CN, ICD10, MedDRA, CIEL, SNOMED, Read, MedDRA, SNOMED, GRR, ISBT, dm+d, SNOMED, dm+d, and Gemscript. The treemap also includes labels for 'Standard' and 'Non-standard' terminologies.



OHDSI Vocabulary Team



Alex



Oleg



Vlad



Varvara



Masha



Mik



Dmitry



Christian



Irina



Anna



Maria



Timur



Dmitry



What we did in 2022

- Regularly updated >25 external vocabularies (such as SNOMED or RxNorm)
- Created/maintained our own mappings (incl. manual) for more than 20 vocabularies (such as NDC)
- Added 31 new vocabulary in 2020-22
- Addressed 37 issues, bugs and requests on GitHub in 2022
- Made 9 releases in 2022, which were downloaded >48,000 times



What we did in 2022

- Regularly updated >25 external vocabularies (such as SNOMED or RxNorm)

But our documentation lags

- Created/maintained our own mappings (incl. manual) for more than 20 vocabularies (such as NDC)

But mapping coverage is not 100%

- Added 31 new vocabulary in 2020-22

But more to add

- Addressed 37 issues, bugs and requests on GitHub in 2022

But there are 169 outstanding issues

- Made 9 releases in 2022, which were downloaded 48,000 times

But we don't know what you do with them, and which problems are the most serious



Landscape assessment of community needs

1. Community survey

We are looking for feedback from:

- ETLers
- Developers
- Researchers

and others!

2. Interviews and focus group discussions with the working groups



What we will ask about

- Which vocabularies you use in ETL, research or development
- Problems you encountered with Vocabularies completeness and correctness
- Problems you encountered with Vocabularies recency and updates
- What you like to see improved

What standard and source vocabularies do you use or have in your source data? Do you have vocabularies that are not in the OHDSI Vocabularies?

Have you encountered missing mappings to standard concepts? Wrong mappings or domain assignment?

Have you had problems with Vocabularies download from Athena or upload into your database?

Have you had problems with delayed Vocabularies release or when doing research on multiple Vocabularies versions?

What is needed to improve your confidence in Vocabularies content and processes?

Are Vocabularies intuitive to use?



What we will do with it

- Which vocabularies you use in ETL, research or development → • Determine how to allocate the resources across the vocabularies to prioritize more important content
- Problems you encountered with Vocabularies completeness and correctness → • Prioritize process improvement activities
- Problems you encountered with Vocabularies recency and updates
- What you like to see improved → • Establish a better way for community contribution
- Publish the report




Survey

Takes ~ 10 mins

Second survey for those
who have a data source
converted to OMOP CDM

Vocabulary landscape assessment

This survey aims to collect your perspectives on the current state of the OHDSI Standardized Vocabularies. We will ask you about the vocabularies you use and the challenges you encountered. The survey will take approximately 10 minutes. If you have the data converted to OMOP CDM we will ask you to provide the feedback on each of the data sources at your disposal at the end of the survey.

aostroplets@gmail.com [Switch account](#) 

*** Required**

Email *

Your email

Name *

Your answer

Organization *

Your answer

Next

Page 1 of 9

Clear form



Please help us and fill the survey!

bit.ly/3iTnyco



Due February 23rd