



Update on the OHDSI Standardized Vocabularies

02 September 2025



Overview



Anna
Ostropolets



Masha
Khitrin



Melanie
Philofsky



Vlad
Korsik



Alexander
Davydov

Release
highlights

Deep
Dive

Community
in action

Community
contributions

Vocabul-
a-thon



Extended Vocabulary Team



Aliaksei Katyshou



Maria Rogozhkina



Mikita Salavei



Varvara Savitskaya



Tetiana Orlova



Polina Talapova



Dmitry Dymshyts



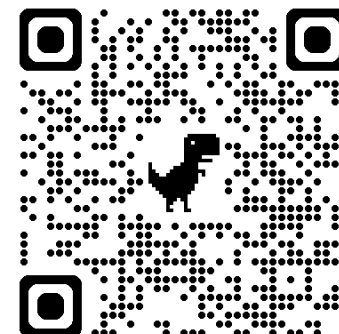
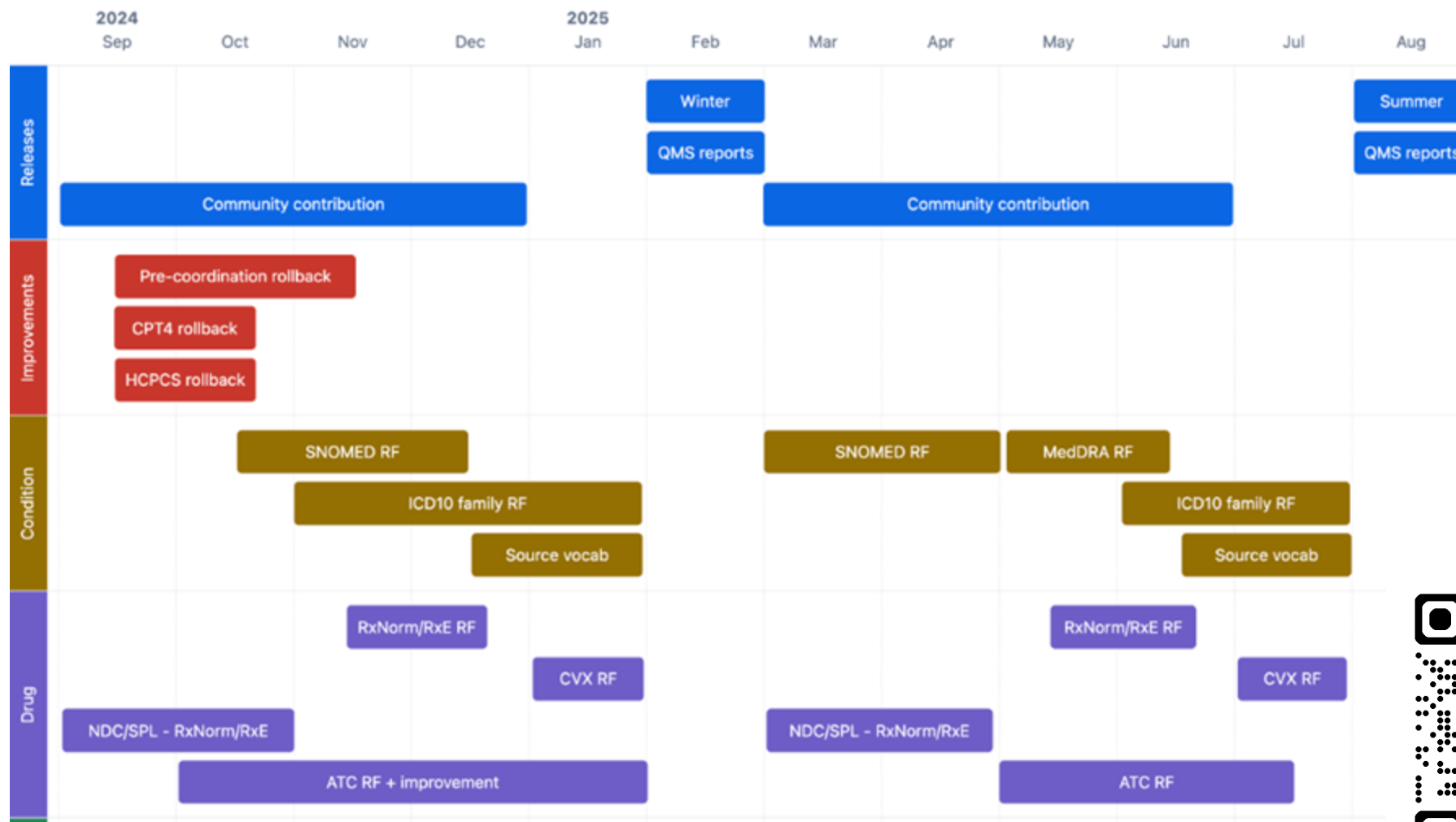
Alexander Davydov



Anton Tatur



Roadmap and releases: August release 2025





Why you need to download this release: new procedures, labs, conditions and drugs

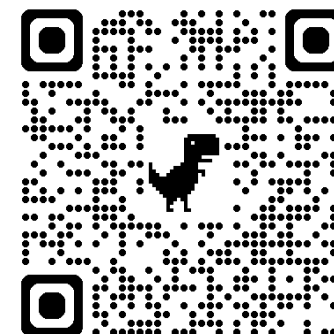
~1,000 new procedures in CPT4, HCPCS, ICD10PCS

2,000+ new lab tests in LOINC

50,000+ new source drug terms (incl. NDC, CVX and dm+d)



Detailed release notes:





Why you need to download this release: better mappings

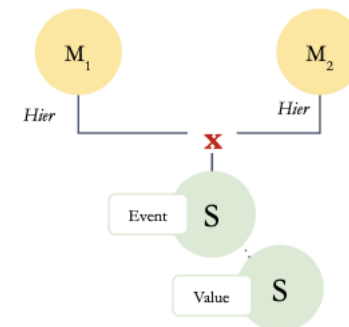
Source ICD10CM	Old target SNOMED	New target SNOMED
ICD9CM 205.32 Multiple myeloma in relapse ICD10CM C90.02 Multiple myeloma in relapse	437233 Multiple Myeloma	764229 Relapse multiple myeloma
ICD10CM F02.811 Dementia in other diseases classified elsewhere, unspecified severity, with agitation	444401 Delusion	37312035 Agitation due to dementia 374888 Dementia associated with another disease



Why you need to download this release: easier concept sets with expanded measurement, drug and procedure hierarchies

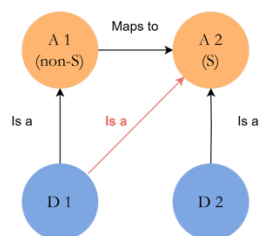
Expanded and more accurate hierarchies
for MedDRA – SNOMED, CPT4/HCPCS –
SNOMED, LOINC – SNOMED

Detection and correction of **orphan drugs**
(drugs not linked to their ingredients in the hierarchy)





Vocabularies machinery



Hierarchy improvement



Propagation of hierarchy



MaximMoinat opened on Jun 18

Describe the problem in content or desired feature

RxNorm Extension concept 44179872 - 'Albuterol 0.2 MG/ACTUAT Inhalant Solution' has no parent concept in the concept_ancestor table. This creates unexpected behavior when creating concept sets where e.g. albuterol ingredient and descendants are added.

How to find it

<https://athena.ohdsi.org/search-terms/terms/44179872/graph>

Expected adjustments

Make the concept a descendant of 36814286 - 'Albuterol 0.2 MG/ACTUAT' and related concepts.

TERM CONNECTIONS (1)



HIERARCHY

RELATED CONCEPTS

NUMBER OF PARENT LEVELS 10

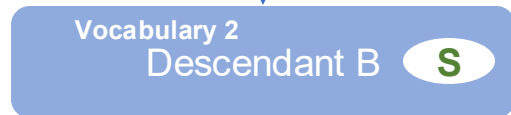
LEVEL OF DETAILS

MAXIMUM

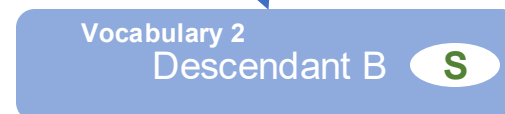
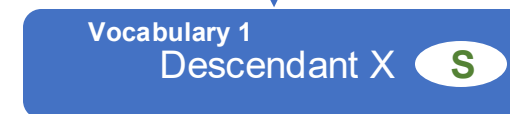
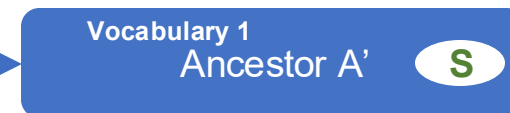
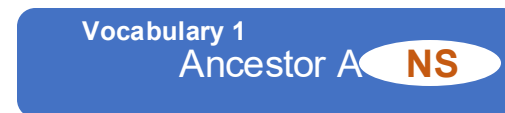
→ Albuterol 0.2 MG/ACTUAT Inhala... → 200 ACTUAT Albuterol 0.2 MG/AC...



Hierarchy changes between releases



Release 1



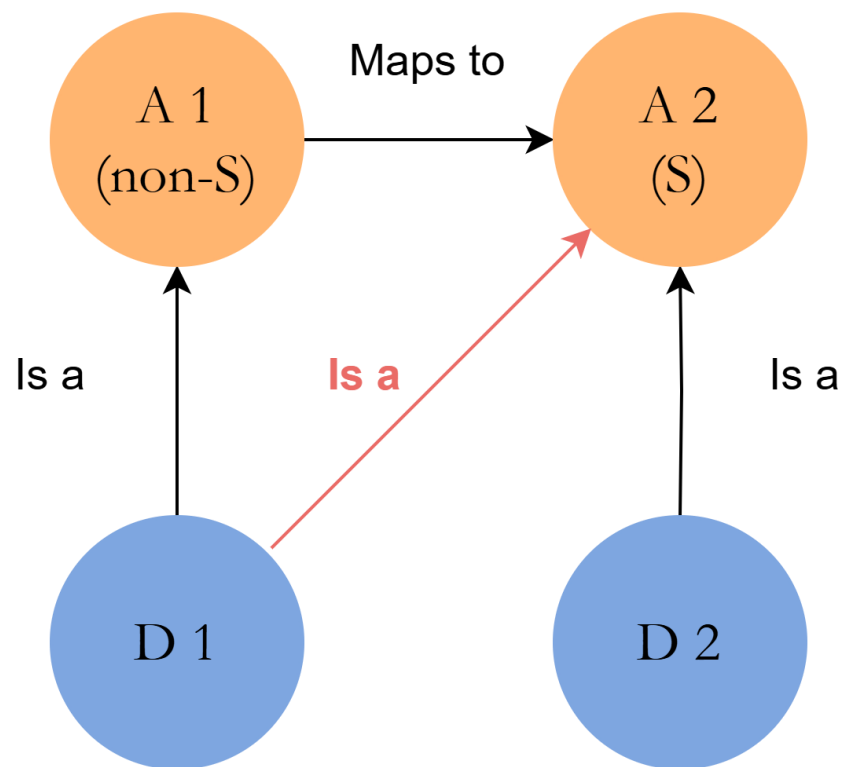
Orphaning



Release 2



Propagation of hierarchy – new function



- **Check** if mapping from A1 to A2 exists in mapping tables
- **Check** if relationship to-be-constructed is plausible (custom configurations +/- vocabulary-specific restrictors)
- **Store** facts in concept_relationship_stage
- **Log** changes in local schema for dry vocabulary runs and in dev schema while integrated.

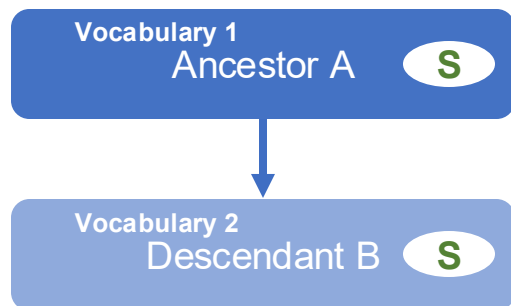
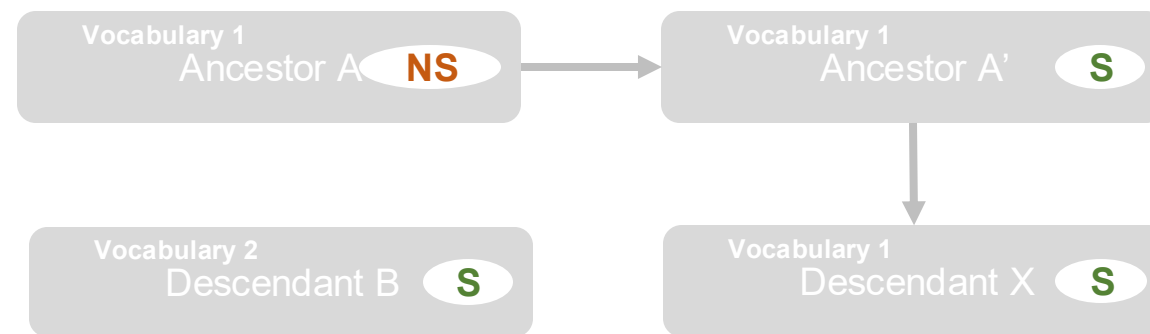
CPT4

Rx/RxE

HCPCS

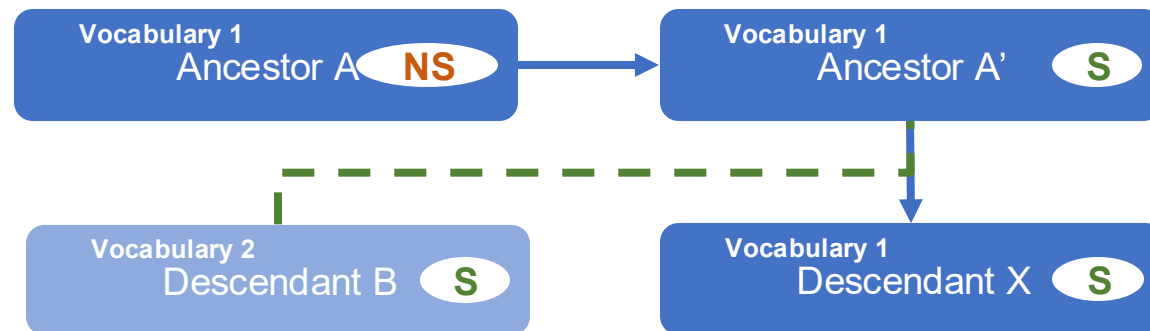


Hierarchy changes between releases



Release 1

Release 2



w/ Hier. Propagation



Propagation of hierarchy



MaximMoinat opened on Jun 18

TERM CONNECTIONS (2)



HIERARCHY

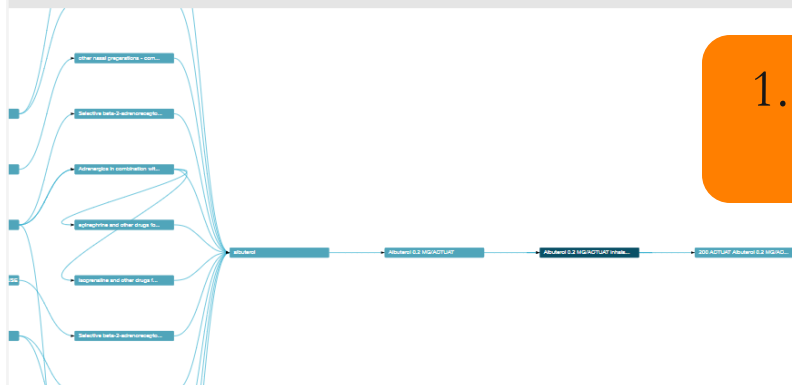
RELATED CONCEPTS

NUMBER OF PARENT LEVELS

10

LEVEL OF DETAILS

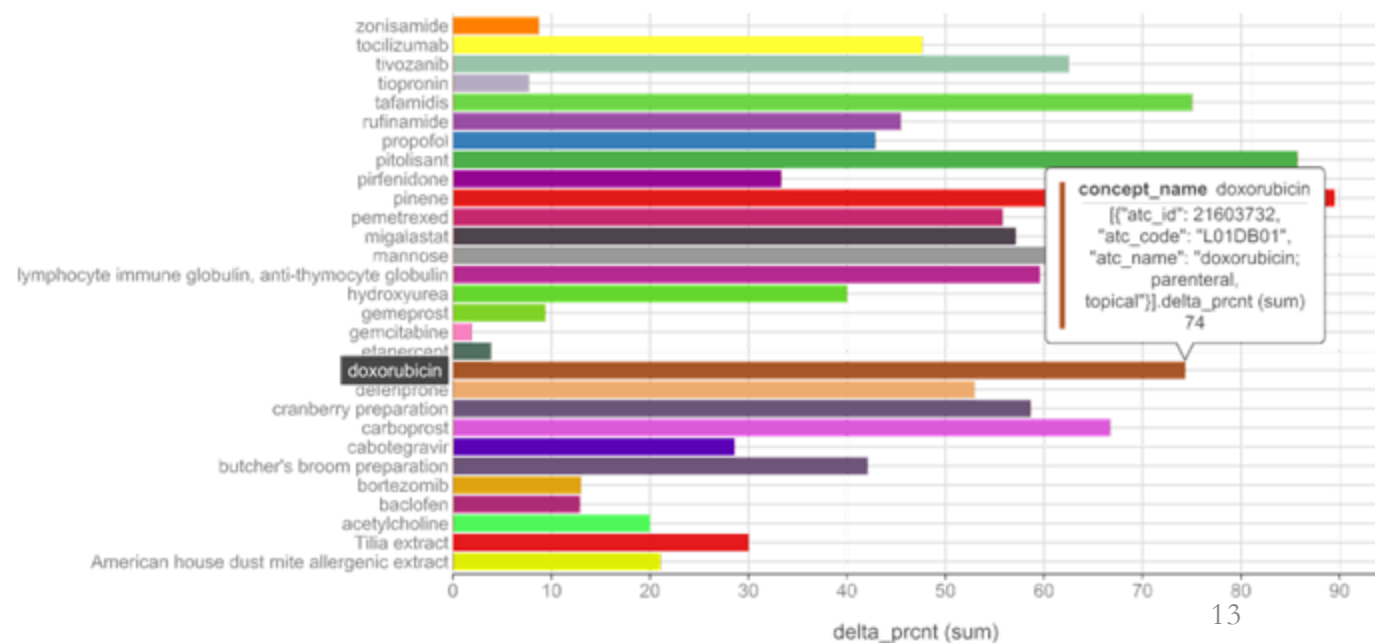
MAXIMUM



1. Specific issue fixed!

2. 162 Rx/RxE Ingredients (0.5%) regained their descendant drugs!
3. >50% of them gained $\geq 20\%$ descendants (+ 116K new concept ancestor records OR 0.5% of full concept_ancestor table)!

- First machinery step to systematically fulfill hierarchical gaps
- Mapping-dependent – not a panacea
- Will need your help (watch out for a forum post)





Vocabularies maintenance



Vocabulary refresh



- SNOMED package
- LOINC
- CPT4
- HCPCS
- ICD10PCS
- ICD10CM
- MedDRA
- CVX
- RxNorm/RxNorm Extension
- NUCC



SNOMED refresh



Int - Feb 2025

US - Mar 2025

UK - Apr 2025



+ 4,059 new concepts

+ 17.5k new mappings



**Legacy replacement
relationships
deprecation**



SNOMED refresh and impact on phenotyping

- *Substance use disorder* branch of the hierarchy was revised by SNOMED and subsequently remapped to standard concepts by us

4175635, Non-dependent drug abuse	Upgraded, replaced by SNOMED	4175635, Nondependent harmful pattern of substance use
436954, Drug abuse	Deprecated	
4279309, Substance abuse	Deprecated	



SNOMED refresh and impact on phenotyping

- *Substance use disorder* branch of the hierarchy was revised by SNOMED and subsequently remapped to standard concepts by us

4175635, Non-dependent drug abuse	Upgraded, replaced by SNOMED	4175635, Nondependent harmful pattern of substance use
436954, Drug abuse	Deprecated, maps to by us	1448779, Harmful pattern of substance use
4279309, Substance abuse	Deprecated, maps to by us	1448779, Harmful pattern of substance use



SNOMED refresh and impact on phenotyping

- *Angioplasty* branch of the hierarchy was revised by SNOMED and subsequently partially remapped to standard concepts by us

4169110, Fluoroscopy guided angioplasty of carotid artery with contrast	Upgraded, replaced by SNOMED	4020941, Fluoroscopy guided percutaneous transluminal angioplasty of carotid artery with contrast
4169110, Angioplasty of carotid artery	Deprecated	



SNOMED refresh and impact on phenotyping

- *Angioplasty* branch of the hierarchy was revised by SNOMED and subsequently partially remapped to standard concepts by us

4169110, Fluoroscopy guided angioplasty of carotid artery with contrast	Upgraded, replaced by SNOMED	4020941, Fluoroscopy guided percutaneous transluminal angioplasty of carotid artery with contrast
4169110, Angioplasty of carotid artery	Deprecated, maps to by us	4020941, Fluoroscopy guided percutaneous transluminal angioplasty of carotid artery with contrast



SNOMED refresh and impact on phenotyping

- *Transplantation* branch of the hierarchy was deprecated by SNOMED and subsequently partially remapped to standard concepts by us

4001893, Transplantation of tissue of nasopharynx	Upgraded, replaced by SNOMED	1447937, Grafting of tissue to nasopharynx
4001892, Transplantation of tissue of nose	Deprecated	



SNOMED refresh and impact on phenotyping

- ***Transplantation*** branch of the hierarchy was deprecated by SNOMED and subsequently partially remapped to standard concepts by us

4001893, Transplantation of tissue of nasopharynx	Upgraded, replaced by SNOMED	1447937, Grafting of tissue to nasopharynx
4001892, Transplantation of tissue of nose	Deprecated, maps to by us	4122430, Graft to nose



SNOMED refresh and impact on phenotyping

- *Radiographic imaging* branch of the hierarchy was revised by SNOMED

4192850, Radiography of gastrointestinal tract	Upgraded, alternative to	1447660, CT of gastrointestinal tract
4056855, Selective bronchography	Upgraded, replaced by SNOMED	1447384, Fluoroscopic selective bronchography with contrast
4340237, Ambigraide operative pancreatography	Deprecated	?



SNOMED refresh and impact on phenotyping

- *SNOMED* replaces concepts with loss of information as a part of their work to separate complications

	Mapped to in previous version	Mapped to in the current version
ICD9CM 675.13 Abscess of breast associated with childbirth, antepartum condition or complication	SNOMED 200376001 Obstetric breast abscess with antenatal complication (Disorder)	SNOMED 270502007 Obstetric breast abscess (Disorder)



SNOMED refresh and impact on phenotyping

We use the tool by **Dima Dymshyts** that evaluates the impact of vocabulary changes on phenotypes as part of our QA routine.

Example of deprecated concept detected by the tool:

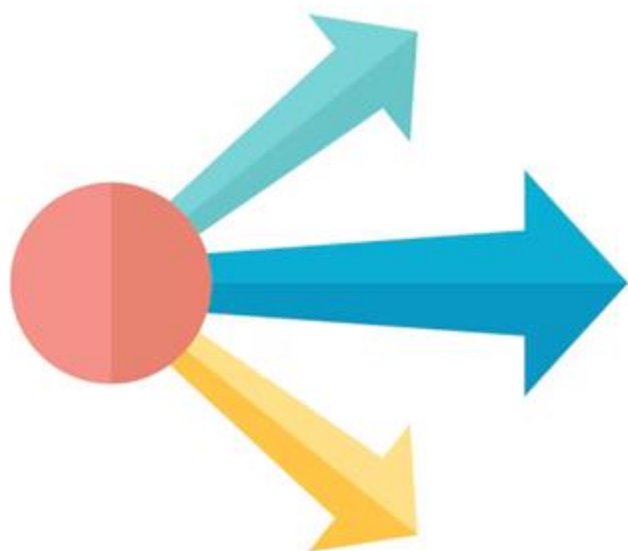
432271000124109, Tonic seizures, refractory (Clinical Finding)	Deprecated, possibly equivalent to	192993002, Tonic epileptic seizure (Clinical Finding) + 1263935000, Drug resistance to anti-seizure medication (Disorder)
---	--	--

PhenotypeChangesInVocab-Update package





We updated and curated mapping



- ICD10
- ICD9CM
- KCD7*
- ICD10GM*
- ICD10CN*
- CIM10*
- OMOP Invest Drug
- CO-CONNECT TWINS
- MeSH
- OPCS4
- UK Biobank
- GRR
- CDISC
- CIEL
- Read

* Mapping are automatically propagated to new standards using machinery



Vocabulary refresh: conditions



MedDRA version 28.0:

+791 new concepts

ICD10CM updated to the 2026 version and mappings curated

+ 614 new concepts



MedDRA-SNOMED hierarchy:

~ 2k relationships revised and improved

Wrong hierarchical connections SNOMED concept 'Is a' MedDRA 'Is a' SNOMED concept #1114

dimshitz opened on Mar 24

ⓧ Closed

⬅️ Bronchitis haemophilus

DETAILS		TERM CONNECTIONS (22)			
Domain ID	Condition	RELATIONSHIP	RELATES TO	CONCEPT ID	VOCABULARY
Concept Class ID	PT	Concept replaces	MedDRA LLT duplicate of PT Concept, do not use, use PT Concept indicated by the CONCEPT_RELATIONSHIP table instead	36146657	MedDRA
Vocabulary ID	MedDRA	Is a	Bacterial lower respiratory tract infections	37203766	MedDRA
Concept ID	36110136		Bronchitis	256451	SNOMED
Concept code	10005460		Haemophilus infection	4211821	SNOMED
Validity	Valid		Haemophilus infections	36102914	MedDRA
Concept	Classification	Non-standard to Standard map (CMOP)	Bronchitis	256451	SNOMED
Valid start	01-Jan-1970		Haemophilus infection	4211821	SNOMED
Valid end	31-Dec-2099				

select * from concept c
join concept_ancestor ca on ca.ancestor_concept_id = c.concept_id
join concept_ancestor ca2 on ca2.descendant_concept_id = c.concept_id
where ca2.ancestor_concept_id = 434277 -- Haem
and c.vocabulary_id = 'MedDRA'
and ca2.ancestor_concept_id = 36110584 --bronch
:
In this case bronchitis becomes a parent of haemo
And I believe there are more cases like this.

Create sub-issue



Vocabulary refresh: drugs



RxNorm/RxE version June 2025:

+1,759 new concepts

+ more than 4k replacement mappings

- 7,6k orphan concepts

NDC/SPL version 20250817 +47k (21k+26k) new concepts

CVX version 20250502 + 5 new concepts



Vocabulary refresh: procedures and measurements



LOINC version 2.80:
+ 2,860 new concepts

CPT4 version 2025AA UMLS:
+ 319 new concepts

HCPCS version July 2025:
+ 122 new concepts + no reused concepts detected

ICD10PCS version 2026:
+212 new concepts



Expanded LOINC - SNOMED hierarchy:
+ 49 new 'Is a' relationships

CPT4 - SNOMED hierarchy:
+ 585 new 'Is a' relationships

HCPCS-SNOMED hierarchy:
+314 'Is a' relationships



Other vocabularies



- **NUCC** was updated to the 25.0 version:
- + 25 new concepts
- + internal hierarchy



Community in action: races and ethnicities



New Standard Races and Ethnicities

Melanie Philofsky and Piper Ranallo in collaboration with Vocabulary WG



Objective 3: Enable vocabulary-related collaboration in the community

KR1: Run a race/ethnicity vocabulary-a-thon

Dealing with multiple races and other exceptions

Vocabulary Users



Pulver Gerald Pulver

Sep 2023

Sep 2023

We are trying to adhere closely to OHDSI standards for the REDS study.

1 / 90



aostroplets Anna Ostroplets

Dec 2023

Sep 2023

Absolutely! You can also watch the recording [here](#).

In this release:

- + 1,356 new standard Races
- + 148 new standard Ethnicities
- + 732 synonyms

Reached out to the OHDSI community via our usual communication channels and personally to all who had contributed to the Race/Ethnicity discussions and to the international chapter leads: APAC, Africa, EU, etc.

Special thank you to Piper for wrangling the disparate source data contributions!

Health Systems, AMCs, and CROs contributed their value sets and we incorporated publicly available past and present sources



Community contributions



Community Contribution

- **Complex contributions: dm+d update**

by Dani Prieto Alhambra, University of Oxford and Erasmus MC, and the SciForce team

- **Complex converted to simple: SINGAPORE COHORTS introduction**

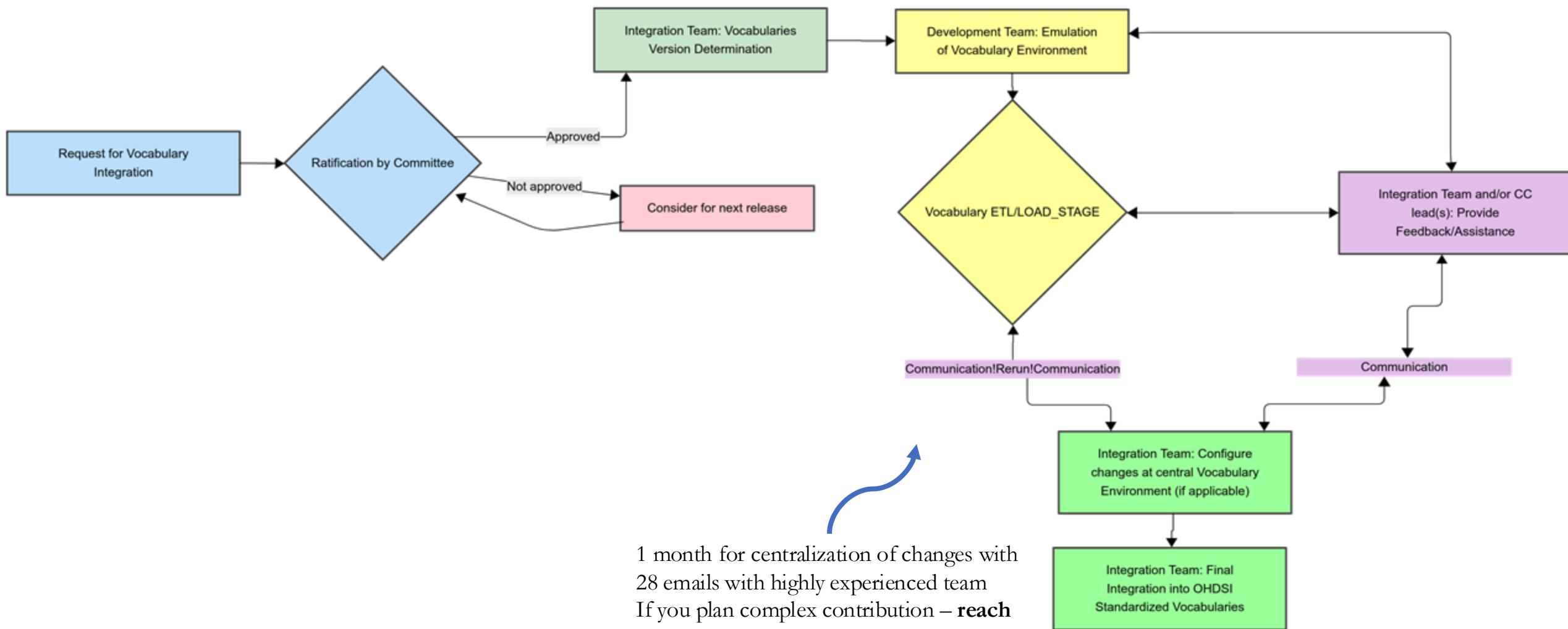
by the team of Nanyang Technological University in association with Esmond Urwin and Phil Quinlan, University of Nottingham

- **Simple contributions**

3 changes in mappings



Complex community contribution flow: dm+d example



1 month for centralization of changes with
28 emails with highly experienced team
If you plan complex contribution – **reach
out early!**



dm+d update

by Dani Prieto Alhambra, University of Oxford and Erasmus MC, and the SciForce team



Version 6.0.0 (June 2025)
First update since 2023!



+ **36k new concepts**
+ **26k new mappings**



Iterative scripts refining



SINGAPORE COHORTS

by the team of Nanyang Technological University in association with Esmond Urwin and Phil Quinlan, University of Nottingham



- New source vocabulary!
- + 579 new concepts with mappings

DHLQ285_E6_1, Are you colour blind?: 1 (SG_COHORTS)	Maps to	193683001, Color blindness (SNOMED)
--	---------	--

If you have a survey you want to add please talk to the Survey WG Lead **Nicole Gerlanc!**

<https://www.ohdsi.org/workgroups/>



Simple contributions: easy filling with Excel spreadsheets

# GitHub	What
1 1120	Change mapping for shoulder fracture ICD family (Anna Ostropolets)
2 1154	Incorrect mapping of Multiple Myeloma in Relapse (Karthik Natarajan)
3 1159	Incorrect mapping of vaccination Read codes (Edward Burn)



Many thanks to all the contributors!





Encourage the Community Contribution



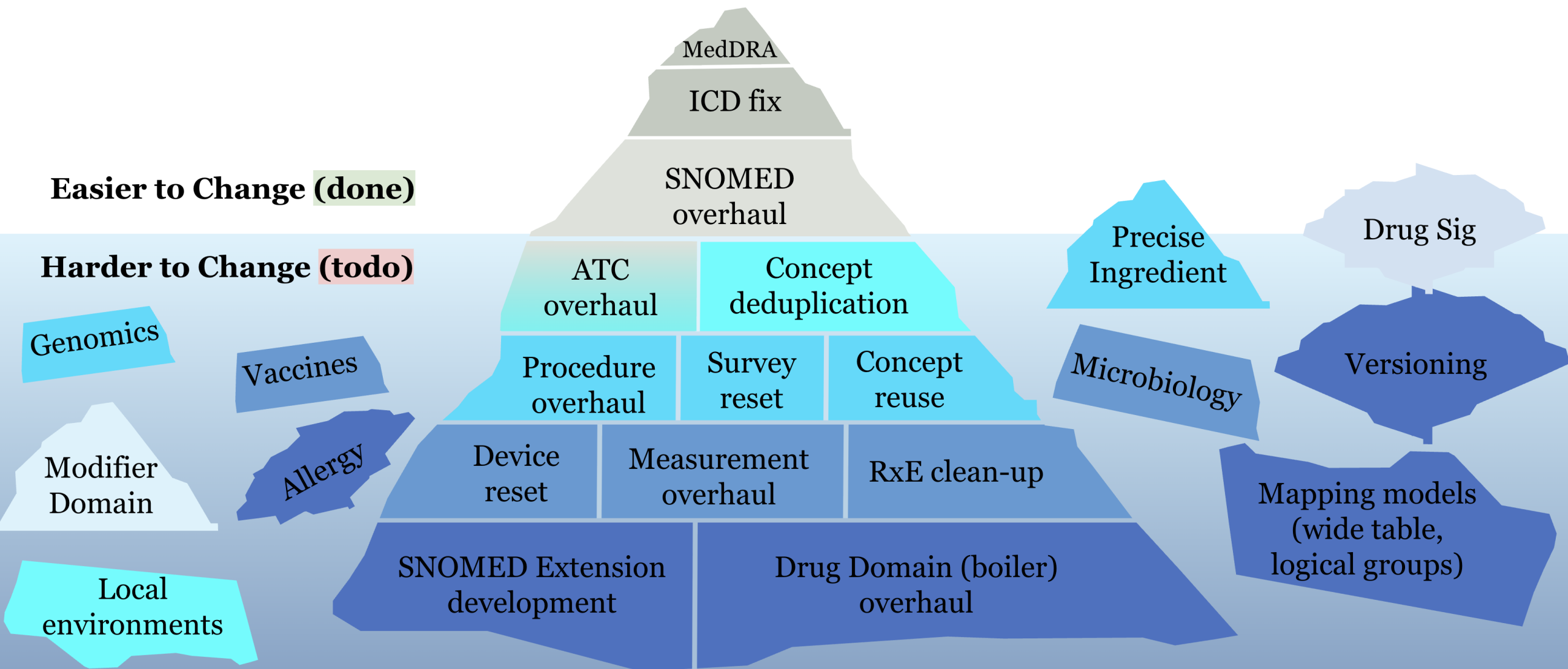
Alexander Davydov:
Community Contribution Lead
davydov@ohdsi.org



Vocabulathon



Vocabulathon: why





Vocabulathon: format

- In-person event at the **Global Symposium**: Tue, Oct 7 | 8 a.m. – 12 p.m.
- Participants join teams by **topic of interest**
- Each team **guided** by vocabulary expert
- Teams prepare the **problem space** in advance
- In-person: focus on **brainstorming** and **design**
- **Outcome**: shared understanding, plans for long-term work
- Not about fixing local codes - focus is on **modeling** and **design**



Vocabulathon: topics confirmed

Surveys & EAV-type data – Nicole Gerlanc

Vaccine vocabulary – Oliver He

Races/Ethnicities – Melanie Philofsky, Piper Ranallo

Visit definitions – Melanie Philofsky, Piper Ranallo

Non-same-as standard concept mapping – Dmitry Dymshyts

Procedure modifiers – Oleg Zhuk



Vocabulathon: topics under discussion

- Microbiology & drug susceptibility
- Rare diseases (Orphanet, Monarch)
- Oncology-related subjects
- Missing & reused NDCs
- Family history – outlined, looking for a lead
- Others? → Let us know!



Vocabulathon: how to sign up

- Sign up when registering for the Symposium
- Already registered, but not signed up
→ Log into Eventbrite and edit your registration, or
→ Email symposium@ohdsi.org to be added
- Come with ideas, use cases, or just curiosity
- Everyone is welcome – we're all vocabulary users

What is your relationship to the OHDSI Community? (please check all that apply) *

- ☐ I actively participate in OHDSI meetings and workgroups
- ☐ I am in the process of converting my data into the OMOP CDM
- ☐ I have an OMOP CDM instance
- ☐ I use OHDSI tools and methods to support my research
- ☐ I contribute code to the OHDSI GitHub
- ☐ I actively participate in discussions on the OHDSI forum
- ☐ I am participating in an OHDSI network research study
- ☐ I am new to OHDSI and curious to learn more
- ☐ This is my first OHDSI Symposium

Do you plan to attend the Vocabulathon on Tuesday, Oct. 7 from 8am-12pm (if you are taking the Intro Tutorial, please answer No)

- ☐ Yes, I plan to attend the Vocabulathon
- ☐ No, I do not plan to attend the Vocabulathon


Which workgroup session do you plan to attend on Thursday, October 9, at 8am ET? (you may select only one below) *

☐ 1. OHDSI 101: Getting started with OHDSI

Your organizer needs more details

To issue your ticket, your organizer needs some more information. Any personal details will remain private.

 Order successful

 Answer questions

 Get your tickets

