

OMOP Standardized Vocabularies



Christian Reich

Timur Vakhitov, Dima Dimschits, Team (Odysseus)

Erica Voss

Patrick Ryan

Martijn Schuemie

Rimma Belenkaya

Chris Knoll

Vojtech Huser

Christophe Lambert

Evanette Burrows

Jen Duryea

14-Jun-2016



Overview

- Recent Development
- Condition mappings
- International Drugs
- Plans



Recent Development

- Athena
 - 1 year old (11-Mar-2015)
 - 268 Downloaders (by domain name in email)
 - 70% US
 - 4% India
 - 2.5% Korea
 - 2% UK, Holland, Belgium
 - 1% Japan, Sweden, Australia, Switzerland, France, Germany,
...
 - License Handling
 - New Production Server
 - More regular releases



Goals

- ✓ **Domains:** Every Standard Concept belongs to the right Domain
- **No duplicates:** For every entity exists one Standard Concept
- **Comprehensive:** For every Domain exists a complete finite set of Concepts covering all possible entities in this domain
- **Hierarchy:** All Concepts are connected through a comprehensive hierarchy
- **Mapping:** For every existing code in a vocabulary there is a map to a Standard Concept or a map to 0



Achieving Goal #1: Domains

For every Standard Concept exists one **Domain** Non-standard ones can be multi-Domain

HCPSC G8879	Clinically node negative (t1n0m0) or t2n0m0) invasive breast cancer	SNOMED 254837009	Malignant tumor of breast	Condition
----------------	---	------------------	---------------------------	-----------

ICD9CM V67.01	Following surgery, follow-up vaginal pap smear	SNOMED 440615002	Postoperative care	Procedure
		SNOMED 133899007	Microscopic examination of vaginal Papanicolaou smear	Measurement

CPT4 90655	Influenza virus vaccine, split virus, preservative free, for children 6-35 months of age, for intramuscular use	CPT4 90655	Influenza virus vaccine, split virus, preservative free, for children 6-35 months of age, for intramuscular use	Procedure
		RxNorm 5806	Influenza virus vaccine	Drug



Distribution of Domains in Vocabularies

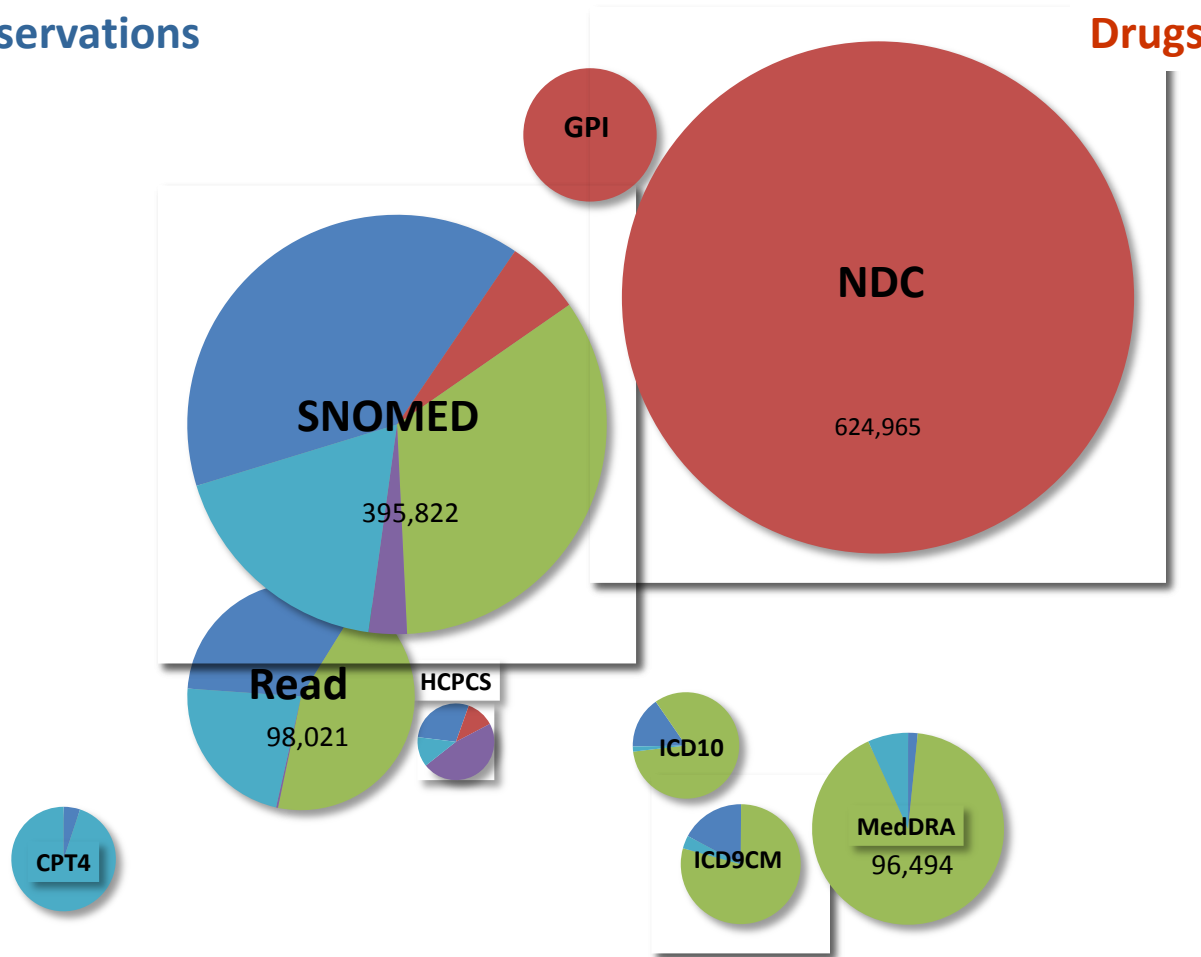
Observations

Drugs

Procedures

Devices

Conditions





Achieve Goal #2: No Duplicates

For every medical entity (condition, drug, procedure etc), there is **only one Standard Concept**

- **Drug:** easy unique combination of ingredient, strength, form, and we got RxNorm, but
 - Forms are not unambiguous
 - Ingredients are easy for patented drugs, but hard for herbal, traditional, excipients, etc
 - Strength is not uniform (% , vol-% , g% , mg/dL)
 - RxNorm is US-only
- **Conditions, lab tests:** harder
 - SNOMED is trying, but
 - Duplications (4 times "Leukemic infiltration of skin")
 - Constant churn of introduction and deprecation
 - Local SNOMEDs don't help
 - LOINC good for clinical labs, too detailed for clinicians and researchers
- **Procedures, observations:** hardest
 - Procedure code systems not comprehensive, cross-links between procedures sporadic and unreliable
 - Observations: Wild West
- **Specialties, place of service:** Messy
- **Devices, disposables:** Impossible



Achieve Goal #3: Comprehensive

- Comprehensive is defined at hierarchical level if possible:
 - Comprehensive set of eg drug classes, ingredients, clinical drugs, branded drugs
 - When hierarchies are extended down to more granularity (eg packaged drugs), the comprehensiveness rule will extend with it
 - Only stratified hierarchies support that approach

Good

- Condition, Drug, Procedure, Measurement
- Type Concepts (on demand)

Bad

- Device, Specialty,
- Observation outside the rule



Achieve Goal #4: Hierarchy

For every medical domain (condition, drug, procedure etc), there is a **comprehensive hierarchy**

- **Drug:** Well established and clinically used drug classes, but
 - No authority or agreement what falls under
 - Many parallel classification systems
 - Many drugs not covered
 - RxNorm has no classes
- **Conditions, Procedures, Tests:**
 - SNOMED is trying, but sometimes contorted lattice
 - Between "Neoplasm and/or hamartoma" and "Suprasellar germ cell tumor" are 3 to 11 levels of separation
 - MedDRA easy to use, but duplications and overlaps
 - "Non-site specific gastrointestinal haemorrhages", "Gastrointestinal haemorrhages"
 - CPT4: 252 codes have no hierarchical connections
- **Observations, Devices**
 - No meaningful hierarchies



Achieve Goal #5: Mapping

For every code that exists there is a **map** to a Standard Concept (including 0 if no useful mapping is possible)


- Existing maps
 - NDC to RxNorm
 - ICD-9-CM to SNOMED
 - SNOMED to MedDRA
 - CPT-4 to SNOMED
 - Read to SNOMED
 - ICD-9-Proc to SNOMED
 - ICD-9-Proc, CPT-4 and HCPCS to RxNorm (procedure drugs)
 - ICD-10-CM to SNOMED
 - DPD to RxNorm/Extension
- Working on
 - ICD10PCS to SNOMED
 - DM+D to RxNorm/Extension
 - Gemscript to RxNorm/Extension
 - AMIS to RxNorm/Extension
 - JDBC to RxNorm/Extension
 - Other national drug schemes to RxNorm/E
 - Other national ICD-10 dialects to SNOMED
 - HCPCS to all sorts of things
 - Units to UCUM
- Need
 - OCPS-4 to SNOMED
 - Comprehensive CPT-4, LOINC, OCPS-4 and HCPCS to SNOMED



Mapping

- <http://www.ohdsi.org/web/wiki/doku.php?id=documentation:vocabulary:mapping>

Logged in as: Christian Reich (cgreich) [Update Profile](#) [Log Out](#)

 Observational Health Data Sciences and Informatics

Search [Recent Changes](#) [Media Manager](#) [Sitemap](#)

Trace: [sidebar](#) • [mapping](#)

documentation:vocabulary:mapping

Standardized Vocabularies

- Introduction
 - Background
 - Principles
 - Definitions of Terms
 - Changes from Version 4
- General Structure and Use
 - [Availability, License, Download](#)
 - Domains and Vocabularies
 - Standard, Classification and Source Concepts
 - Mapping
 - Data ETL
- Domains
 - Condition
 - Currency
 - Device
 - Drug
 - Ethnicity
 - **Gender**
 - Measurement

Mapping of Concepts

Mapping is the process to transform one Concept to another. The clinical data tables of the CDM allow only Standard Concepts. All other codes used in the source databases have to be translated to Standard Concepts.

Mapping is done through records in the CONCEPT_RELATIONSHIP table. They connect each Concept to a Standard Concept through a number of special relationship_id:

Relationship ID	Purpose
Maps to	Mapping Concept to Standard Concept to be placed in the *_concept_id
Maps to value	Special mapping between a Concept and a Standard Concept to be placed into the value_as_concept field of the MEASUREMENT and OBSERVATION tables

[Edit](#)

"Maps to" Relationships

Concepts participating in the "Maps to" mapping are Source Concepts and Standard Concepts. The mapping attempts to map to the equivalent target Concept. Equivalent means it carries the same meaning, and



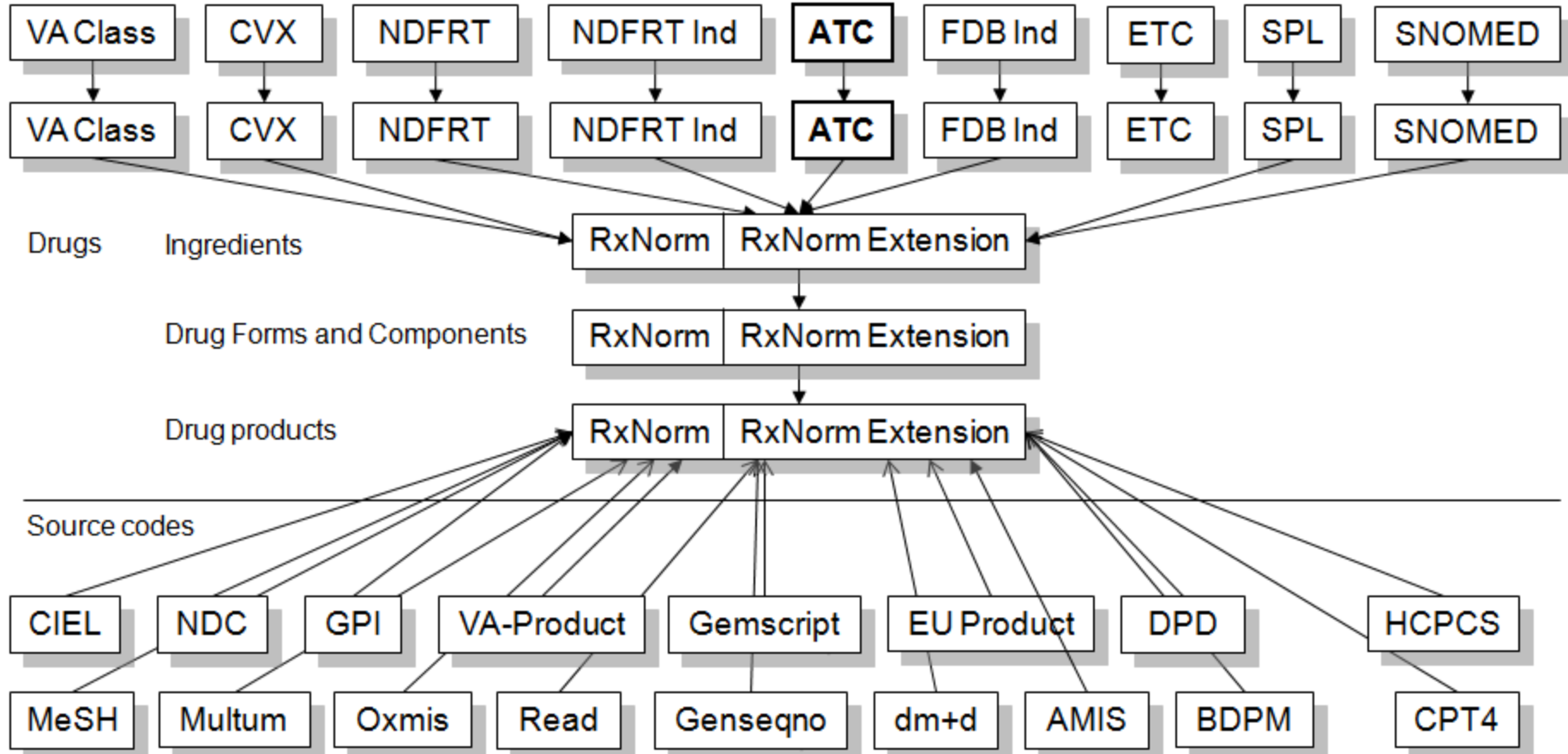
Special Condition Mappings

- **Aftercare after procedures**
 - **Late effects or sequelae of other conditions**
 - **Family history of a condition**
 - **History of medical**
 - **Underdosing of medication**
 - **Pre-coordinated Concepts listing two or more semantic components through AND or OR**
 - **Maternal care**
 - **Need for immunization**
 - **Conditions indicating abnormal levels of a test**
 - **Conditions resulting from the use of Devices**
-



International Drugs

Classifications





Plans

- Fixing Goals
 - Procedures
 - Devices
 - Vocabularies with re-used codes (DRG, NDC, ATC flavors)
 - Automating generation
 - Outreach
-



Links

Sign up for Vocabulary release notifications:

<https://github.com/OHDSI/OMOP-Standardized-Vocabularies/releases>

Click on "Watch" and select "Watching".

Athena – Vocabulary Download:

<http://www.ohdsi.org/web/athena/>

Let us know if you have a license to a proprietary vocabulary.

Vocabulary Documentation (in progress):

<http://www.ohdsi.org/web/wiki/doku.php?id=documentation:vocabulary:sidebar>

Please help making it correct and meaningful.