



Extract, Transform, and Load (ETL)

Alex PA. Nguyen
Taipei Medical University
Taiwan OHDSI
alexnthhp@tmu.edu.tw



Instructors





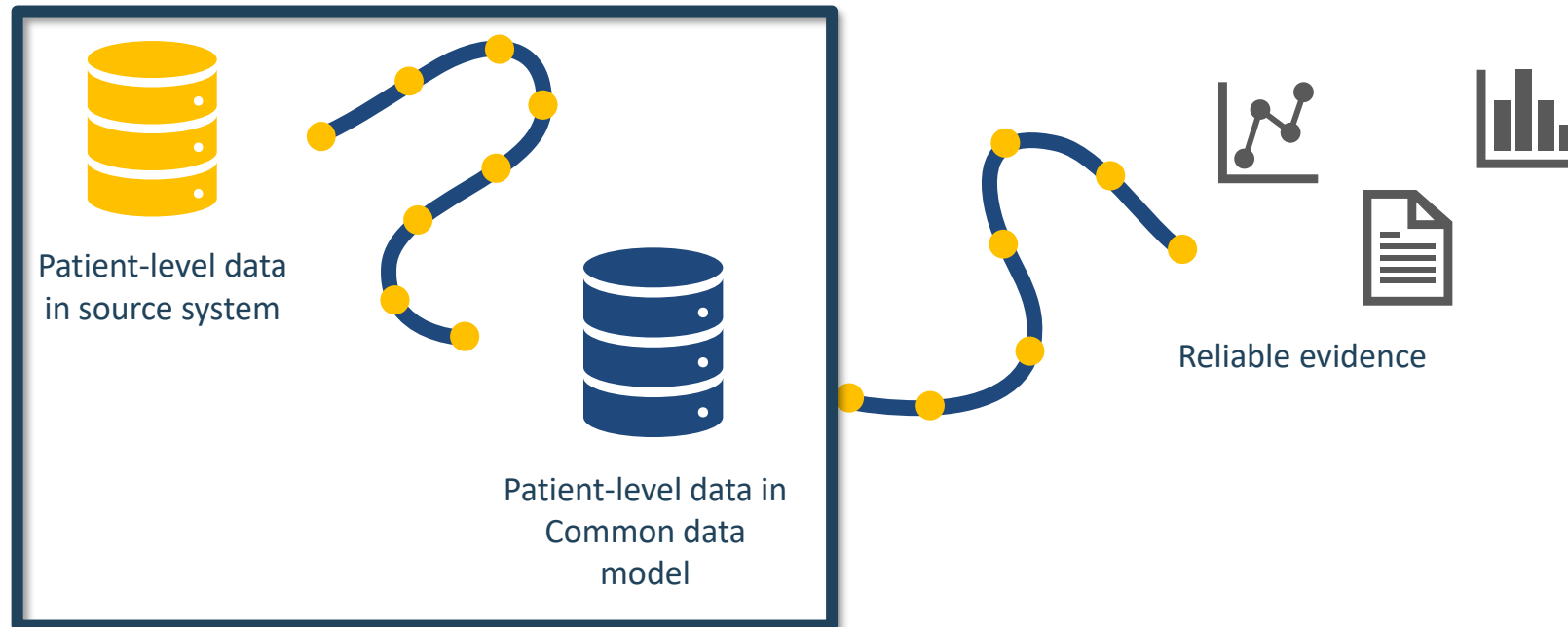
ETL Process and Tools

- Best Practices
- ETL Process
- ETL Tools
 - White Rabbit tool - review the output
 - Rabbit in a Hat tool - document the conceptual logic
 - Usagi - mapping custom source values



ETL

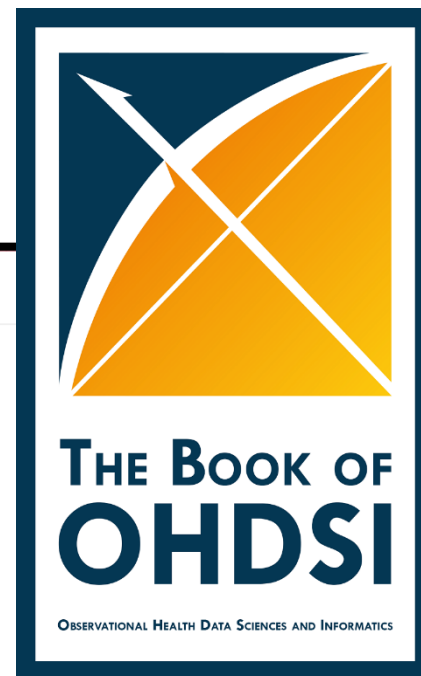
- Extract Transform Load
- In order to get from our native/raw data into the OMOP CDM we need to design and develop an ETL process



- Goal in ETLing is to standardize the format and terminology



ETL Process



The Book of OHDSI

Preface

I The OHDSI Community

1 The OHDSI Community

2 Where to Begin

3 Open Science

II Uniform Data Representation

4 The Common Data Model

5 Standardized Vocabularies

6 Extract Transform Load

6.1 Introduction

6.2 Step 1: Design the ETL

6.3 Step 2: Create the Code Map...

6.4 Step 3: Implement the ETL

6.5 Step 4: Quality Control

6.6 ETL Conventions and THEMIS

6.7 CDM and ETL Maintenance

Chapter 6 Extract Transform Load

Chapter leads: Clair Blacketer & Erica Voss

6.1 Introduction

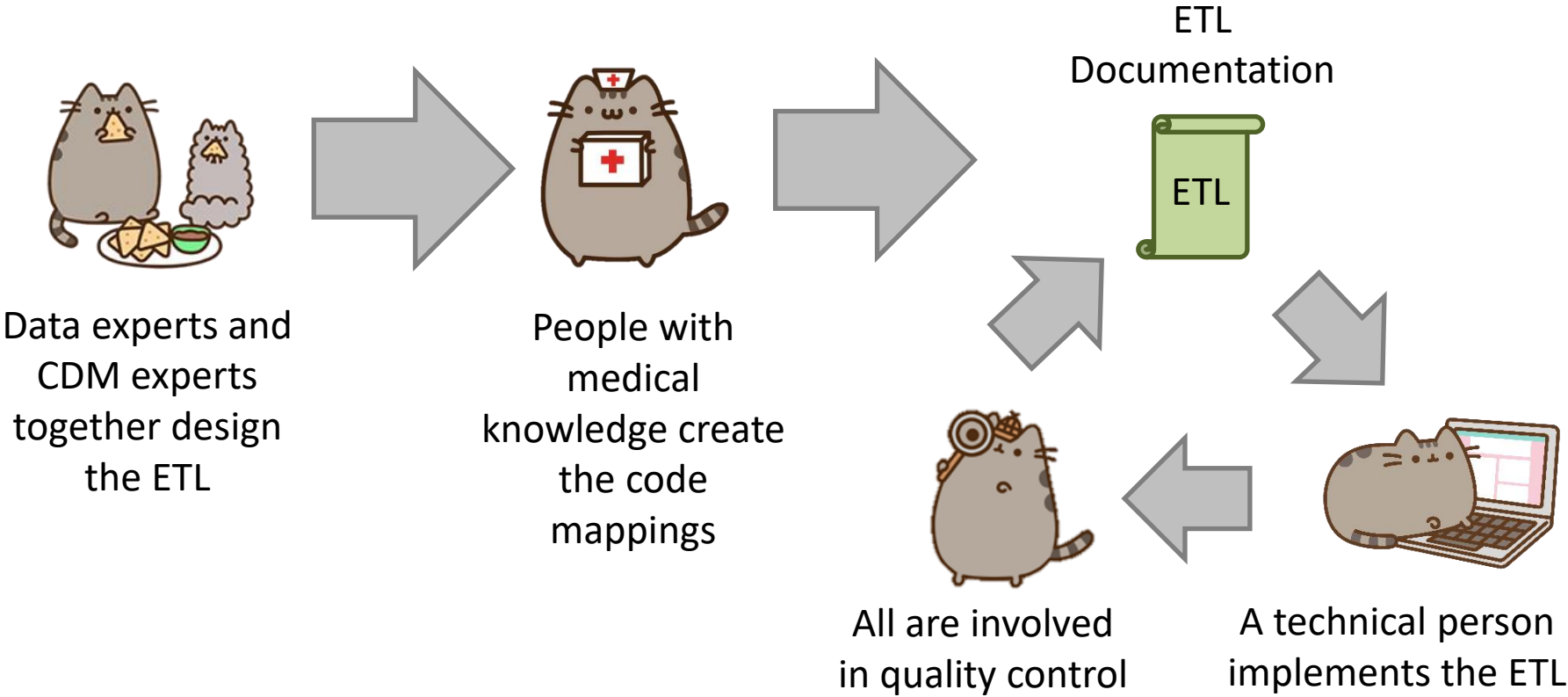
In order to get from the native/raw data to the OMOP Common Data Model (CDM) we have to create an extract, transform, and load (ETL) process. This process should restructure the data to the CDM, and add mappings to the Standardized Vocabularies, and is typically implemented as a set of automated scripts, for example SQL scripts. It is important that this ETL process is repeatable, so that it can be rerun whenever the source data is refreshed.

Creating an ETL is usually a large undertaking. Over the years, we have developed best practices, consisting of four major steps:


1. Data experts and CDM experts together design the ETL.
2. People with medical knowledge create the code mappings.
3. A technical person implements the ETL.



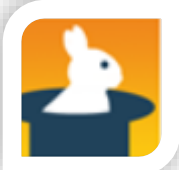
ETL Process




OHDSI Tools




White Rabbit




Rabbit In a Hat




Usagi



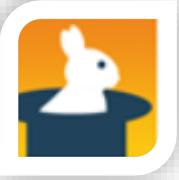
White Rabbit



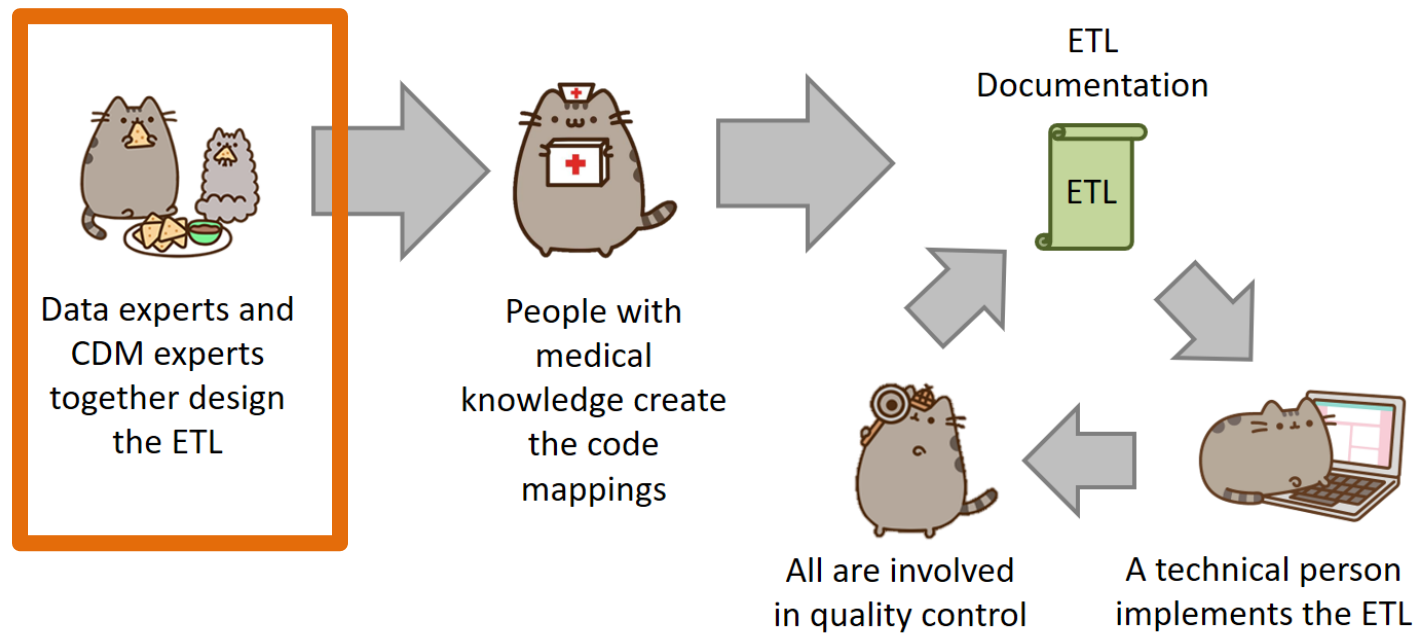
ACHILLES



DQD

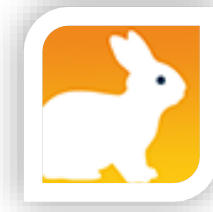


Rabbit In a Hat

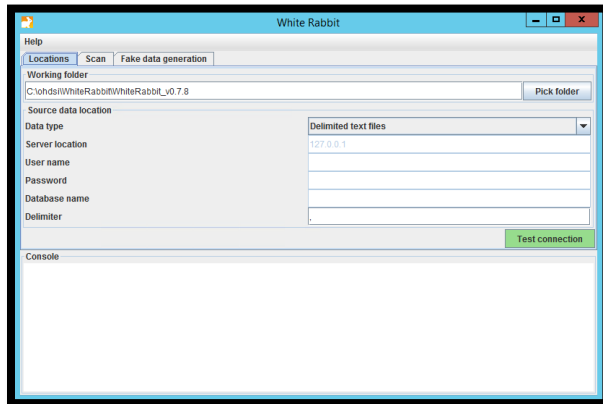




White Rabbit

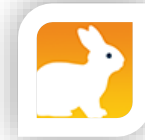


- White Rabbit scans source data & creates a csv report on the source data
- The scan can be used to:
 - Learn about your source data
 - Help design the ETL
 - Used by Rabbit In a Hat





WR Output – ScanReport.xlsx



Table/Field Overview

Table	Field	Description	Type	Max length	N rows
pop	der_sex		character	1	16374539
pop	der_yob		double pre	6	16374539
pop	pat_id		character	64	16374539
pop	pat_hash_id		character	16	16374539
pop	pmtx_flag		numeric	1	16374539
pop	anon_ims_pat_id		character	11	16374539
pop	pat_region		character	2	16374539
pop	pat_state		character	2	16374539
pop	pat_zip3		character	3	16374539
pop	grp_indv_cd		character	1	16374539
pop	mh_cd		character	1	16374539
pop	enr_rel		character	2	16374539
pop	temp_col1		character	0	16374539
pop	temp_col2		character	0	16374539
pop	load_row_id		bigint	9	16374539
claims_diag_lk	person_source_valu		character	64	2992046684
claims_diag_lk	event_start_date		date	10	2992046684
claims_diag_lk	event_end_date		date	10	2992046684

Value counts

	A	B	C	D	
1	der_sex	Frequency	der_yob	Frequency	pa
2	F	50479	1991.0	2030	Li
3	M	49514	1992.0	1970	
4	U	7	1990.0	1947	
5			1989.0	1908	
6			1988.0	1873	
7			1994.0	1872	
8			1995.0	1806	
9			1993.0	1805	
10			1996.0	1716	
11			1986.0	1676	
12			1987.0	1643	
13			1985.0	1633	
14			1983.0	1588	
15			1981.0	1581	
16			1984.0	1576	
17			1970.0	1555	
18			1980.0	1553	

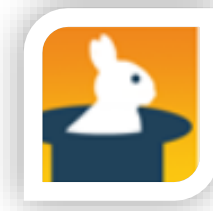
pop

claims_diag_lk

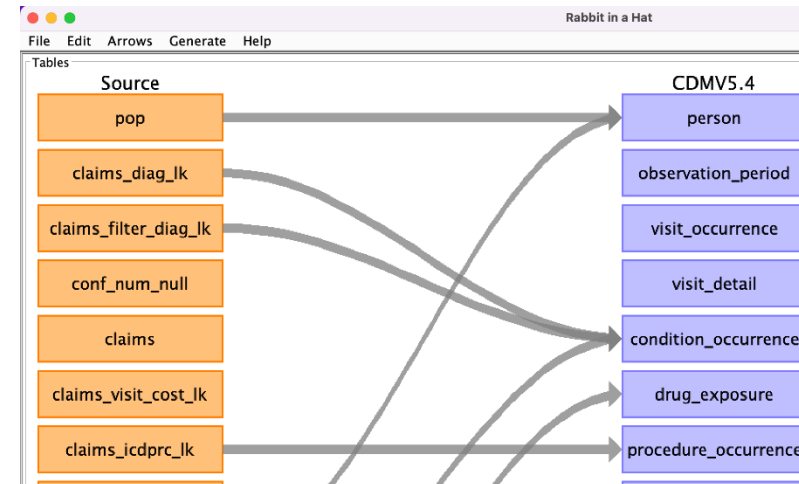
claims



Rabbit in a Hat

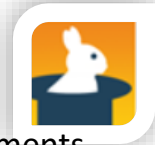


- Read and display a White Rabbit scan document
- Provides a graphical interface to allow a user to connect source data to CDM tables





RiaH - Output



Word document

Mapping Document MM RMG - Compatibility Mode

measurement_time		
measurement_type_concept_id		
operator_concept_id		
unit_concept_id	serum_protein	Standard unit: mg/dL. Create conversion
	urine_protein_mg	New unit concept
range_low		
range_high		
provider_id		
visit_occurrence_id		
visit_detail_id		
measurement_source_concept_id		
unit_source_value		

Table name: observation
Reading from diagnostics
'History of'

Source

*subject_id → *person_id (CDMV5.3.1_adj)

date_diag_875_i1 → *observation_concept_id

history_solitary_plasmacyt_i1 → *observation_date

Destination Field	Source Field	Logic	Comment
observation_id			Auto-increment
person_id	subject_id		
observation_concept_id	history_solitary	Map to a custom concept	
		'History of solitary plasmacytoma'	
observation_date	date_diagnosis		
observation_datetime	date_diagnosis		
observation_type_concept_id		380015486	Registered from EHR
value_as_number			
value_as_string			
value_as_concept_id			

Page 10 of 51 2442 words English (United States)

Html

Person - Tutorial-ETL

Search Tutorial-ETL

CDM Synthea v1 / Person

Person

Reading from Synthea table patients.csv

Destination Field	Source field	Logic	Comment field
person_id		Autogenerate	
gender_concept_id	gender	When gender = 'M' then set gender_concept_id to 8507, when gender = 'F' then set to 8532	Drop any rows with missing/unknown gender.
year_of_birth	birthdate	Take year from birthdate	
month_of_birth	birthdate	Take month from birthdate	
day_of_birth	birthdate	Take day from birthdate	
birth_datetime	birthdate	With midnight as time 00:00:00	
		When race = 'WHITE' then set as 8527, when	

This site uses [Just the Docs](#), a documentation theme for Jekyll.

Markdown documents

```
Measurement.md
Person.md

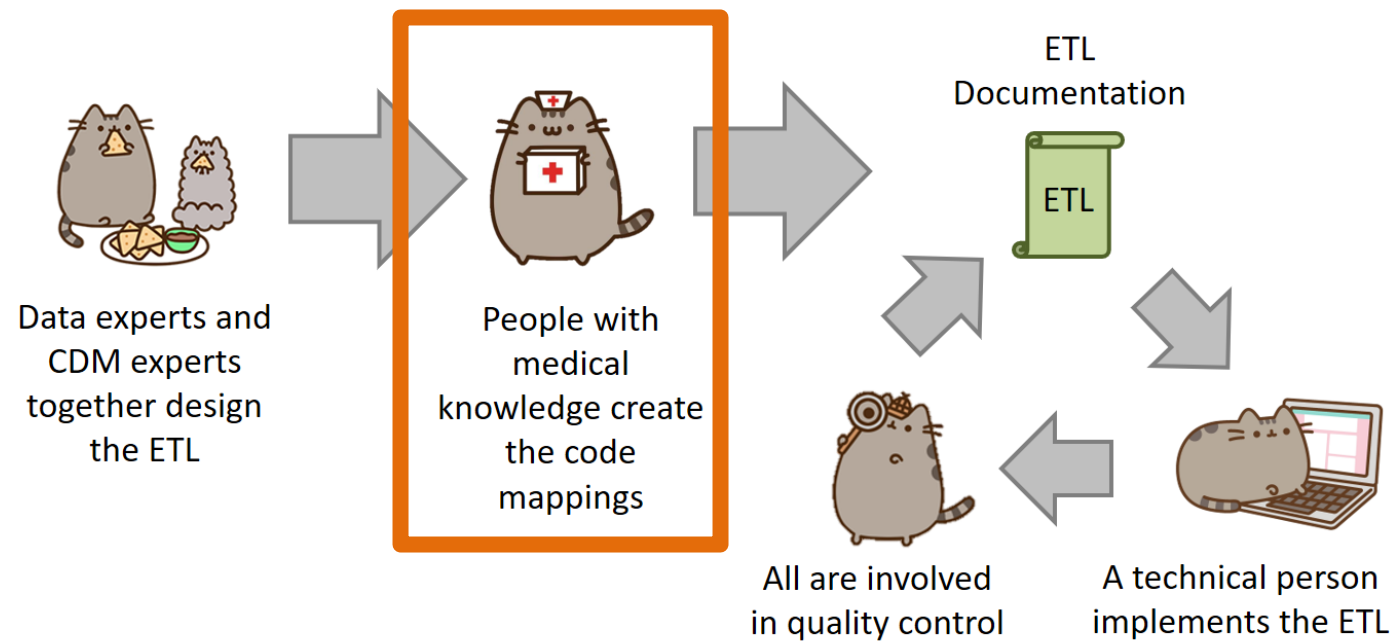
layout: default
title: Person
nav_order: 1
parent: CDM Synthea v1
description: "Person mapping from patients.csv"

# Person

## Reading from Synthea table patients.csv

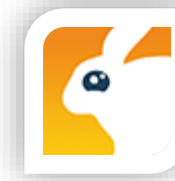


| Destination Field | Source field | Logic | Comment field | |
|---|---|---|---|---|
| person_id | | Autogenerate | |
| gender_concept_id | gender | When gender = 'M' then set gender_concept_id to 8507, when gender = 'F' then set to 8532 | Drop any rows with missing/unknown gender. |
| year_of_birth | birthdate | Take year from birthdate | |
| month_of_birth | birthdate | Take month from birthdate | |
| day_of_birth | birthdate | Take day from birthdate | |
| birth_datetime | birthdate | With midnight as time 00:00:00 | |
| race_concept_id | race | When race = 'WHITE' then set as 8527, when race = 'BLACK' then set as 8516, when race = 'ASIAN' then set as 8515, otherwise set as | |
| ethnicity_concept_id | race | ethnicity | When race = 'HISPANIC', or when ethnicity in ('CENTRAL_AMERICAN', 'DOMINICAN', 'MEXICAN', 'PUERTO_RICAN', 'SOUTH_AMERICAN') then set as 38803563, otherwise set as 0 | |
| location_id | | | |
| provider_id | | | |
| care_site_id | | | |
| person_source_value | id | | |
| gender_source_value | gender | | |
| gender_source_concept_id | | | |
| race_source_value | race | | |
| race_source_concept_id | | | |
| ethnicity_source_value | ethnicity | | |
| ethnicity_source_concept_id | | | |
```





Usagi



- When the Vocabulary does not contain your source terms you will need to create a map to OMOP Vocabulary Concepts
- Usagi helps you to:
 - Find best matches, automatically and/or manually
 - Automatic matching based on text similarities (itf/df)
 - Create 'source to concept map'

The screenshot shows the Usagi application window. At the top is a menu bar with 'File', 'Edit', 'View', and 'Help'. Below it is a table with columns: Status, Source code, Source term, Frequency, ICPC_DES, Match score, Concept ID, Concept name, Domain, Concept class, Vocabulary, Concept code, Standard concept, Parents, Children, and Comment. The table lists several source terms like 'No illness', 'Dermatoph...', 'Acute phary...', 'Cystitis / ur...', 'Acute bronc...', 'Pregnancy...', 'overgewicht', 'Acute infect.', 'episode on...', 'Immunisat.', and 'Cough', each with its corresponding match in the OMOP Vocabulary. Below the table are sections for 'Source code', 'Target concepts', 'Search', and 'Results'. The 'Search' section has a 'Query' field and 'Filters' for 'Filter by user selected concepts', 'Filter by concept class', 'Filter by standard concepts', 'Filter by vocabulary', 'Include source terms', and 'Filter by domain'. The 'Results' section shows a table of search results with columns: Score, Term, Concept ID, Concept name, Domain, Concept class, Vocabulary, Concept code, Standard concept, Parents, and Children. The bottom of the window has a 'Comment' field and an 'Approve' button. The status bar at the bottom indicates 'Approved / total: 0 / 12 0.0% of total frequency' and 'Vocabulary version: v5.0.19-NOV-18'.

Status	Source code	Source term	Frequency	ICPC_DES	Match score	Concept ID	Concept name	Domain	Concept class	Vocabulary	Concept code	Standard concept	Parents	Children	Comment
Unchecked	A97	No illness	500000	Geen ziekte	0.82	4192174	Illness	Condition	Clinical Fin...	SNOMED	39104002	S	1	3	
Unchecked	S74	Dermatoph...	100000	Dermatoph...	0.81	135473	Dermatoph...	Condition	Clinical Fin...	SNOMED	47382004	S	4	25	
Unchecked	L99	Other disa...	100000	Andere ziek...	0.77	4244662	Disorder of...	Condition	Clinical Fin...	SNOMED	928000	S	3	84	
Unchecked	R74.02	Acute phary...	800000	Acute phary...	1.00	25207	Acute phary...	Condition	Clinical Fin...	SNOMED	363746003	S	6	10	
Unchecked	U71	Cystitis / ur...	500000	Cystitis/urin...	0.71	81902	Urinary trac...	Condition	Clinical Fin...	SNOMED	6856005	S	5	17	
Unchecked	R78.00	Acute bronc...	300000	Acute bronc...	0.84	260125	Acute bronc...	Condition	Clinical Fin...	SNOMED	5505005	S	5	4	
Unchecked	W78.00	Pregnancy...	100000	Zwangersc...	0.84	4295535	Pregnant	Condition	Clinical Fin...	SNOMED	77385006	S	2	17	
Unchecked	T83.0	overgewicht	100000	overgewicht	1.00	437525	Overweight	Observation	Clinical Fin...	SNOMED	238131007	S	2	5	
Unchecked	R74	Acute uppe...	800000	Acute infect...	1.00	257011	Acute uppe...	Condition	Clinical Fin...	SNOMED	54398005	S	6	22	
Unchecked	R65.00	episode on...	1	episode op...	0.35	444406	Acute sube...	Condition	Clinical Fin...	SNOMED	70422006	S	4	0	
Unchecked	R44	Immunisat...	100000	Immunisat...	0.70	4144375	Active imm...	Procedure	Clinical Fin...	SNOMED	33679002	S	2	19	
Unchecked	R05	Cough	880000	Hoessten	1.00	254761	Cough	Condition	Clinical Fin...	SNOMED	49727002	S	2	38	

Source code

Source code	Source term	Frequency	ICPC_DESCRIPTION_DUTCH
A97	No illness	500000	Geen ziekte

Target concepts

Concept ID	Concept name	Domain	Concept class	Vocabulary	Concept code	Standard concept	Parents	Children
4192174	Illness	Condition	Clinical Finding	SNOMED	39104002	S	1	3

Remove concept

Search

Query

Use source term as query

Query:

Filters

Filter by user selected concepts

Filter by concept class:

Filter by standard concepts

Filter by vocabulary:

Include source terms

Filter by domain:

Results

Score	Term	Concept ID	Concept name	Domain	Concept class	Vocabulary	Concept code	Standard concept	Parents	Children
0.82	Illness	4192174	Illness	Condition	Clinical Finding	SNOMED	39104002	S	1	3
0.80	Mental illness	4214703	Mental illness	Observation	Qualifier Value	SNOMED	394916006	S	1	0
0.80	Mental illness	432586	Mental disorder	Condition	Clinical Finding	SNOMED	74732009	S	2	41
0.78	Viral illness	440029	Viral disease	Condition	Clinical Finding	SNOMED	34014006	S	3	31
0.77	Mass illness	45883959	Mass illness	Mass Value	Answer	LOINC	LA18096-0	S	0	0
0.75	Sickness	4092256	Sickness	Condition	Clinical Finding	SNOMED	247902006	S	3	1

Replace concept

Add concept

Comment:

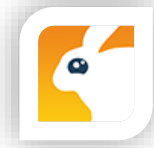
Approved / total: 0 / 12 0.0% of total frequency

Approve

Vocabulary version: v5.0.19-NOV-18

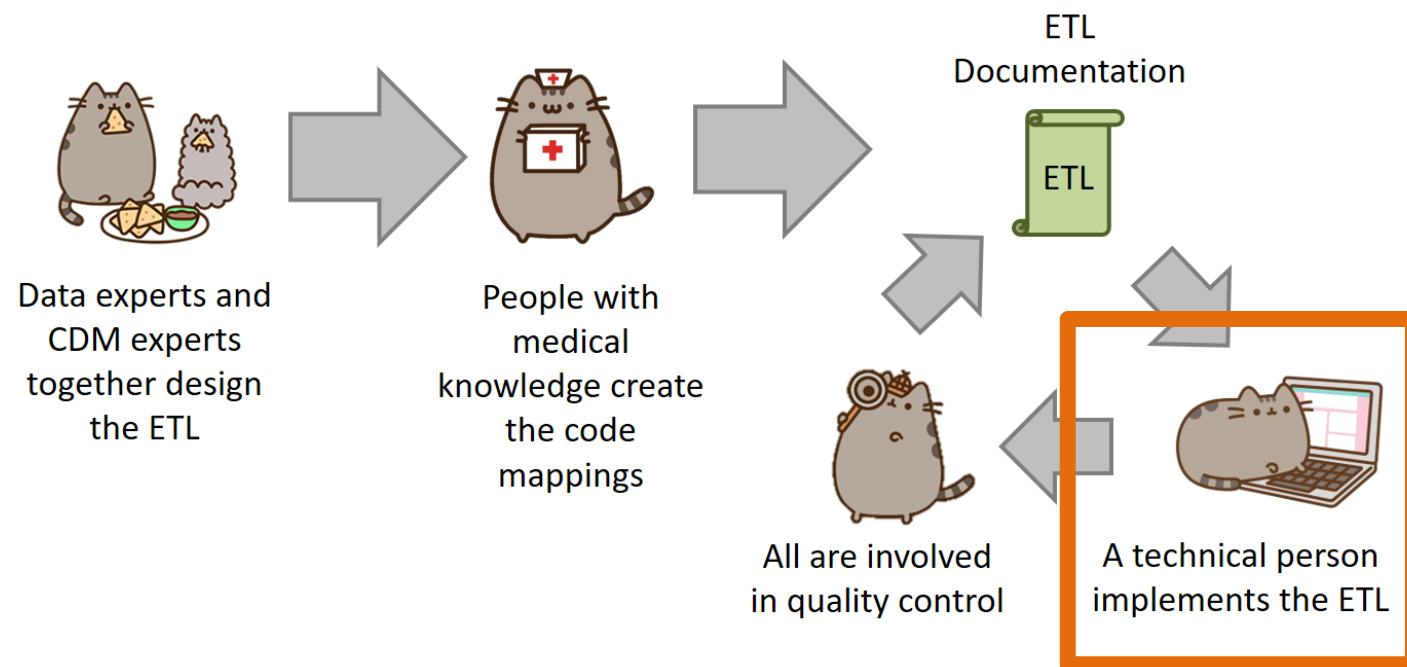


Overview - Steps



1. Get a copy of the Vocabulary from ATHENA
2. Download Usagi
- 3. Have Usagi build an index on the Vocabulary**
4. Load your source codes and let Usagi process them
5. Review and update suggested mappings with someone who has medical knowledge
6. Export codes into the SOURCE_TO_CONCEPT_MAP

} One-time
setup





ETL Implementation



There are multiple tools available to implement your ETL

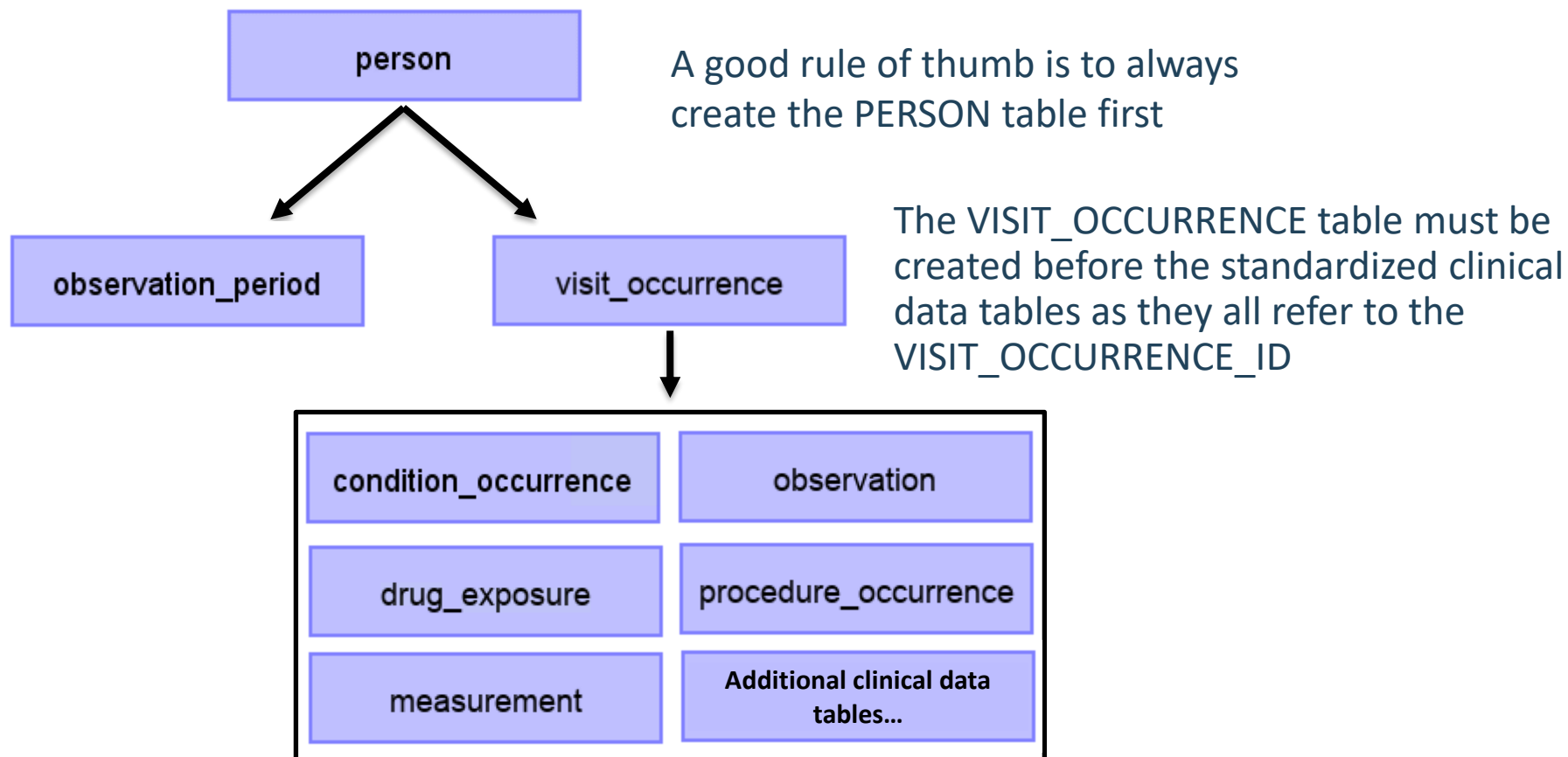


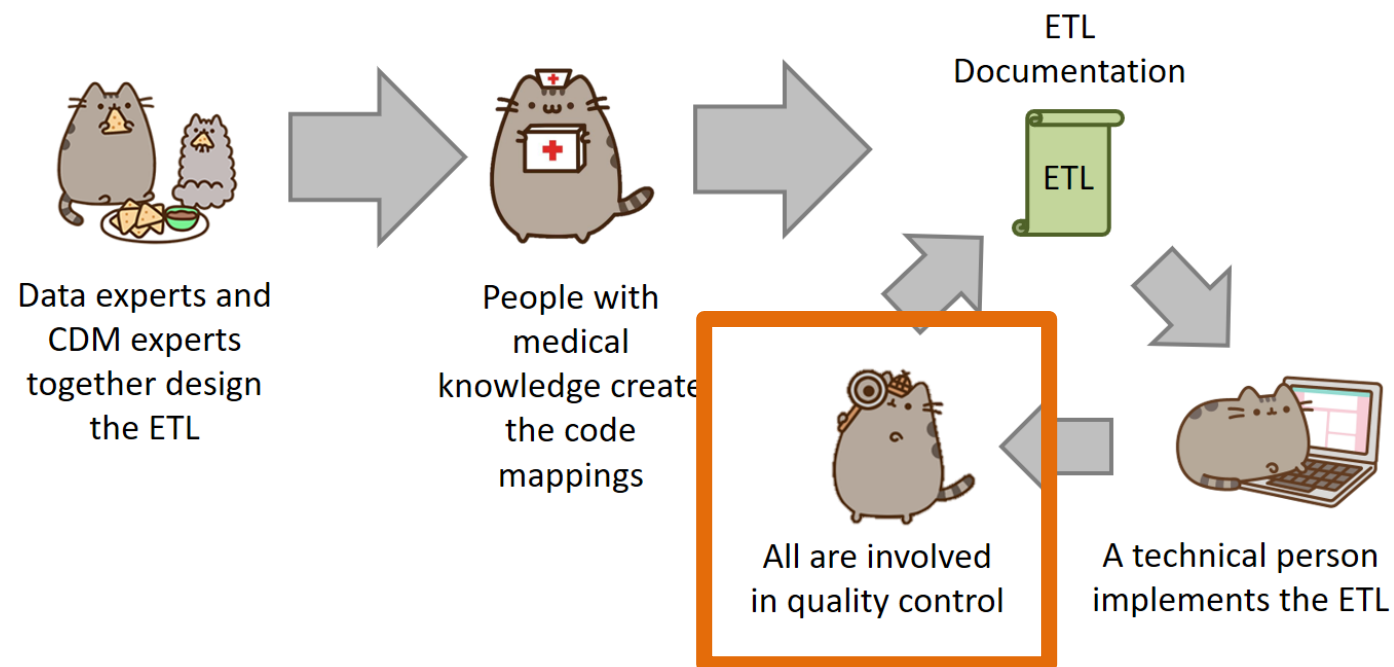
Your choice will largely depend on the size and complexity of the ETL design. And the tools available to you.



ETL Implementation

General Flow of Implementation



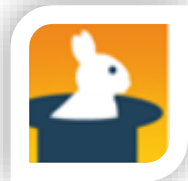




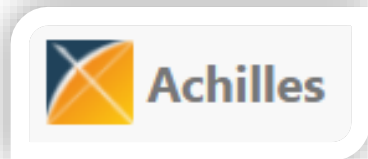
Quality



What tools are available to check that the CDM logic was implemented correctly?



Rabbit-in-a-Hat Test Case Framework



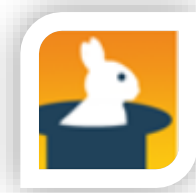
Achilles



DataQualityDashboard (DQD)



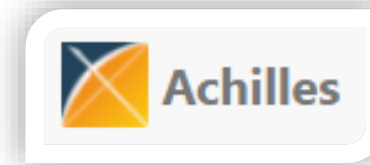
Unit Test Cases



- Testing your CDM builder is important:
 - ETL is often complex, increasing the danger of making mistakes that go unnoticed
 - CDM can update
 - Source data structure/contents can change over time
- Rabbit-In-a-Hat can construct unit tests, or small pieces of code that can automatically check single aspects of the ETL design



Achilles



Achilles is a data characterization and quality tool available for download here:

<https://github.com/OHDSI/Achilles>

For an example of how it was run for some sample data, that R script is located here:

<https://github.com/OHDSI/Tutorial-ETL/blob/master/materials/Achilles/achillesRun.R>



DataQualityDashboard (DQD)



- Runs a prespecified set of data quality checks and thresholds on the CDM



DATA QUALITY ASSESSMENT

SYNTHETIC HEALTH DATABASE

Results generated at 2019-08-22 14:15:06 in 29 mins

	Verification				Validation				Total			
	Pass	Fail	Total	% Pass	Pass	Fail	Total	% Pass	Pass	Fail	Total	% Pass
Plausibility	159	21	180	88%	283	0	283	100%	442	21	463	95%
Conformance	637	34	671	95%	104	0	104	100%	741	34	775	96%
Completeness	369	17	386	96%	5	10	15	33%	374	27	401	93%
Total	1165	72	1237	94%	392	10	402	98%	1557	82	1639	95%



10 Minute Energy Break

Upon your return, we will work
through it in the next session
Q/A?



Extract, Transform, and Load (ETL)

Alex PA. Nguyen
Taipei Medical University
Taiwan OHDSI



ETL



OHDSI

OBSERVATIONAL HEALTH DATA SCIENCES AND INFORMATICS

Who We Are ▾

Updates & News ▾

Standards

Software Tools

Network Studies ▾

Community Forums ▾

Education ▾

New To OHDSI? ▾

Community Calls ▾

Events ▾

Workgroups ▾

Our Journey: Where We Have Been & Where We Are Going (PDF)

Community Dashboards ▾

This Week In OHDSI

Support & Sponsorship ▾

2022 Global Symposium ▾

Github

YouTube

Twitter

LinkedIn

Newsletters ▾


[Home](#) > [Software](#) > WhiteRabbit for ETL design















WhiteRabbit for ETL design


WhiteRabbit is a software tool to help prepare for ETLs (Extraction, Transformation, Loading) of longitudinal healthcare databases into the OMOP Common Data Model (CDM). The source data can be in comma-separated text files, or in a database (MySQL, SQL Server, ORACLE, PostgreSQL); the CDM will be in a database (MySQL, SQL Server, PostgreSQL).


The main function of WhiteRabbit is to perform a scan of the source data, providing detailed information on the tables, fields, and values that appear in a field. This scan will generate a report that can be used as a reference when designing the ETL, for instance when using the Rabbit-In-a-Hat tool. Rabbit-In-a-Hat uses the scan document and displays source data information through a graphical user interface to allow a user to connect source data structure to the CDM data structure. The function of Rabbit-In-a-Hat is to generate documentation for the ETL process, not generate code to create an ETL.

Download WhiteRabbit: <https://github.com/OHDSI/WhiteRabbit>


 MaximMoinat Update documentation to align wording with application ✓ f26a94f on Aug 8 🕒 705 commits

 .github/ISSUE_TEMPLATE	Add issue templates	17 months ago
 .settings	Converting Eclipse to Maven project	4 years ago
 docs	Update documentation to align wording with application	3 months ago
 iniFileExamples	Merge branch 'develop' into fake-data-fix	3 years ago
 lib/com	Restored lib/com/simba/googlebigquery/jdbc jars for local repo	3 years ago
 rabbit-core	Merge branch 'develop' into release-0.10.7	9 months ago
 rabbitinahat	Merge branch 'develop' into release-0.10.7	9 months ago
 whiterabbit	bump version	9 months ago
 .gitignore	Updated dependencies	2 years ago
 .project	Converting Eclipse to Maven project	4 years ago
 LICENSE	Added copyright	9 years ago
 README.md	update documentation urls	3 years ago
 examples.zip	update examples	3 years ago
 pom.xml	bump version	9 months ago

 README.md



White Rabbit



Rabbit in a Hat



White Rabbit

White Rabbit - Location



White Rabbit

Help

Locations Scan Fake data generation

Working folder

C:\ohdsi\WhiteRabbit\WhiteRabbit_v0.8.1\bin Pick folder

Source data location

Data type Delimited text files

Server location 127.0.0.1

User name

Password

Database name

Delimiter ,

Test connection

Console



White Rabbit - Scan



White Rabbit

Help

Locations Scan Fake data generation

Tables to scan

Add all in DB

Add

Remove

☒ Scan field values Min cell count 5 Max distinct values 1,000 Rows per table 100,000

Scan tables

Console

White Rabbit - Scan



White Rabbit

Help

Locations Scan Fake data generation

Tables to scan

Add all in DB

Add

Remove

☒ Scan field values

Min cell count 5

Max distinct values 1,000

Rows per table 100,000

Scan tables

Console



White Rabbit – Open the Scan file

ScanReport.xlsx - Excel																			Tell me what you want to do...										Sign in		Share																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																											
File Home Insert Page Layout Formulas Data Review View																			Tell me what you want to do...										Sign in		Share																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																											
Clipboard Font Paragraph Alignment Number Conditional Formatting Styles Cells Editing																			Tell me what you want to do...										Sign in		Share																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																											
G5																			Tell me what you want to do...										Sign in		Share																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																											
A B C D E F G H I J K L M N O P Q R S T U V W X																			Tell me what you want to do...										Sign in		Share																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																											
1	Table	Field	Description	Type	Max length	N rows	N rows checked	Fraction empty	N unique values	Fraction unique																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																



Rabbit in a Hat



RiaH – Load the Scan file

Rabbit in a Hat
File Edit Arrows Generate Help

Tables

Source

CDMV5.3.1

- condition_occurrence
- death
- device_exposure
- drug_exposure
- fact_relationship
- measurement
- note
- note_nlp
- observation
- observation_period
- person**
- procedure_occurrence
- specimen
- visit_detail
- visit_occurrence
- cohort
- cohort_attribute
- condition_era

Details

General information

Table name: person

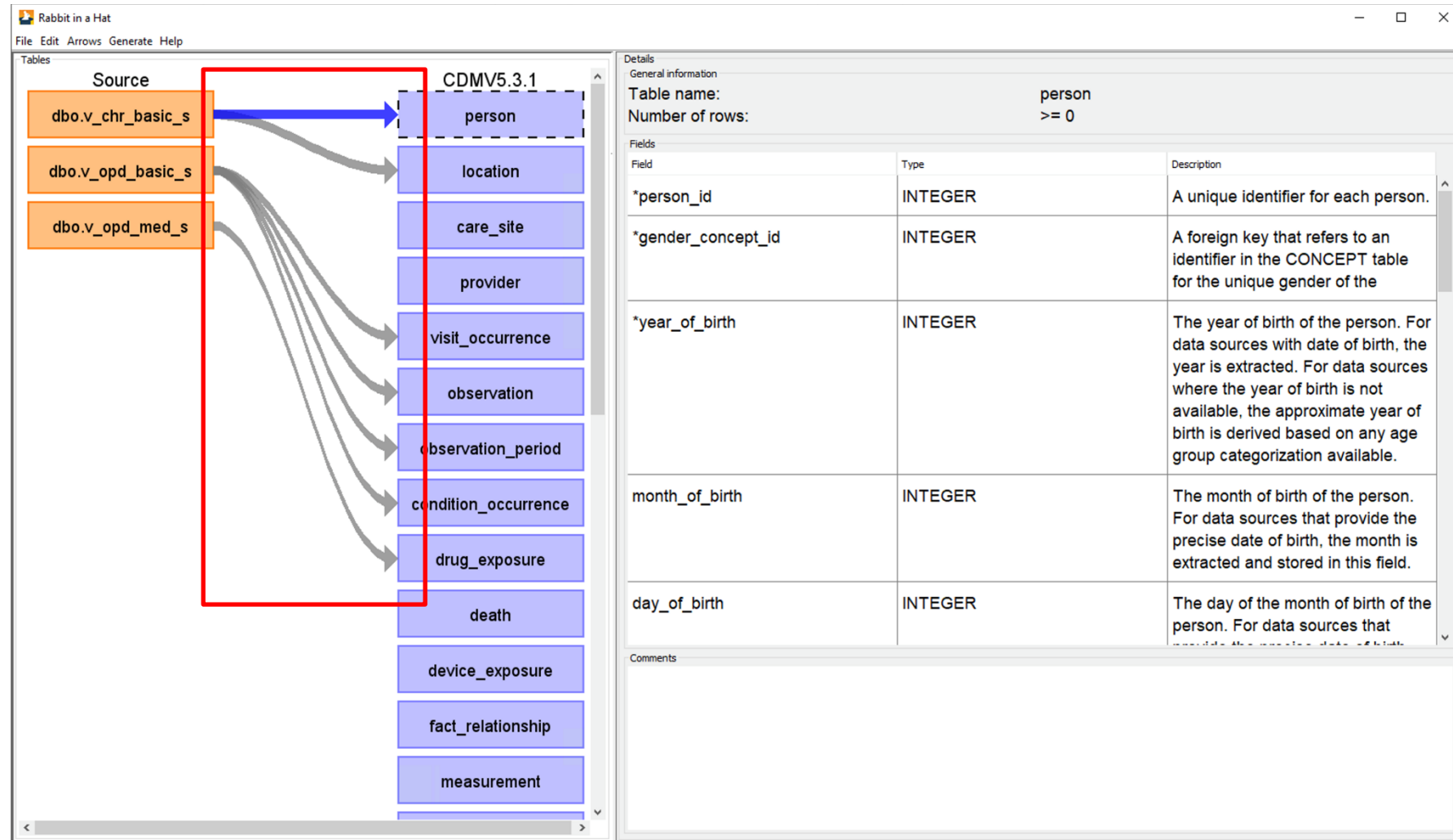
Number of rows: >= 0

Field	Type	Description
*person_id	INTEGER	A unique identifier for each person.
*gender_concept_id	INTEGER	A foreign key that refers to an identifier in the CONCEPT table for the unique gender of the person.
*year_of_birth	INTEGER	The year of birth of the person. For data sources with date of birth, the year is extracted. For data sources where the year of birth is not available, the approximate year of birth is derived based on any age group categorization available.
month_of_birth	INTEGER	The month of birth of the person. For data sources that provide the precise date of birth, the month is extracted and stored in this field.
day_of_birth	INTEGER	The day of the month of birth of the person. For data sources that provide the precise date of birth, the day is extracted and stored in this field.
birth_datetime	DATETIME	The date and time of birth of the person.
*race_concept_id	INTEGER	A foreign key that refers to an identifier in the CONCEPT table for the unique race of the person.
*ethnicity_concept_id	INTEGER	A foreign key that refers to the standard concept identifier in the Standardized Vocabularies for the ethnicity of the person.

Comments



RiaH – Mapping the Source tables to CDM tables





RiaH – Mapping the Source variables to CDM

Rabbit in a Hat

File Edit Arrows Generate Help

Tables

Source CDMV5.3.1

dbo.v_chr_basic_s person

Fields

Source CDMV5.3.1

id_no *person_id

*sex_type person_source_value

*birth_date *gender_concept_id

*hosp_type gender_source_value

pat_name gender_source_concept_id

*pat_id *year_of_birth

birth_place month_of_birth

*ac_type day_of_birth

blood_type location_id

rh_type care_site_id

edu_code provider_id

Details

General information

Field name: race_concept_id

Field type: INTEGER

Description: A foreign key that refers to an identifier in the CONCEPT table for the unique race of the person.

Concept ID Hints

Concept ID	Concept Name	Class	Standard?
38003600	African	Race	S
38003599	African American	Race	S
38003573	Alaska Native	Race	S
38003572	American Indian	Race	S
8657	American Indian or Alaska Native	Race	S
38003616	Arab	Race	S
8515	Asian	Race	S
38003574	Asian Indian	Race	S
38003601	Bahamian	Race	S
38003575	Bangladeshi	Race	S
38003602	Barbadian	Race	S
38003576	Bhutanese	Race	S
38003598	Black	Race	S
8516	Black or African American	Race	S
38003577	Burmese	Race	S
38003578	Cambodian	Race	S
38003579	Chinese	Race	S
38003604	Dominica Islander	Race	S
38003603	Dominican	Race	S

Comments



RiaH – Save and generate the document

File

Edit

Arrows

Generate

Help

ETL Document

Overview Table

ETL Test Framework (R)

SQL Skeleton Files

Word Document

HTML Documents

Markdown Documents

Tables

dbo.v_...

dbo.v_opd_basic_s

dbo.v_opd_med_s

CDMV5.3.1

person

location

care_site

provider

visit_occurrence

observation

observation_period

condition_occurrence

drug_exposure

death

device_exposure

fact_relationship

measurement

Details

General information

Table name: person

Number of rows: >= 0

Fields

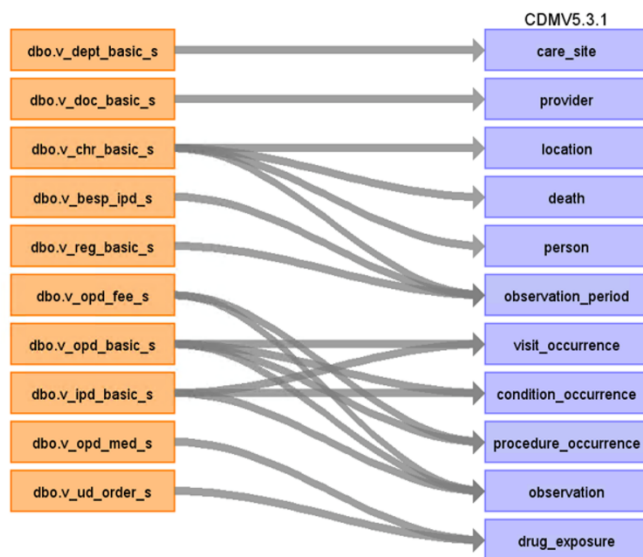
Field	Type	Description
*person_id	INTEGER	A unique identifier for each person.
person_source_value	VARCHAR(50)	An (encrypted) key derived from the person identifier in the source data. This is necessary when a use case requires a link back to the person data at the source dataset.
*gender_concept_id	INTEGER	A foreign key that refers to an identifier in the CONCEPT table for the unique gender of the
gender_source_value	VARCHAR(50)	The source code for the gender of the person as it appears in the source data. The person's gender is mapped to a standard gender concept in the Standardized Vocabularies; the original value is stored here for reference.

Comments



Example of generate document

Data Mapping Approach to CDMV5.3.1 for Clinical Research Database of TMU



Contents

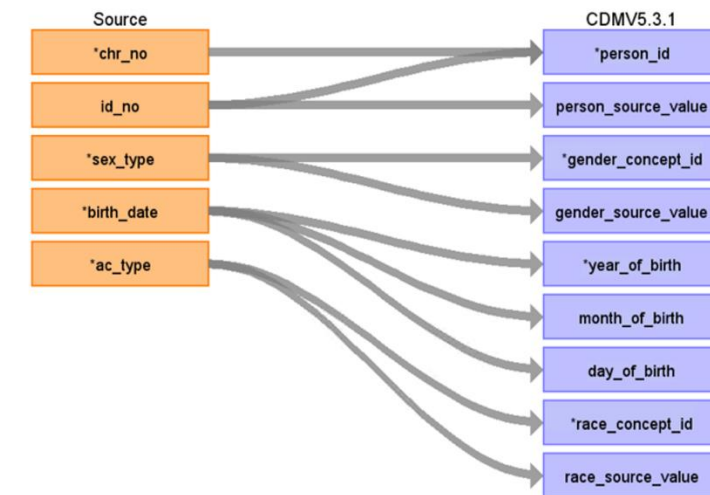
Table name: care_site	4
Reading from dbo.v_dept_basic_s	4
Table name: fact_relationship	5
Table name: provider	6
Reading from dbo.v_doc_basic_s	6
Table name: location	8
Reading from dbo.v_chr_basic_s	8
Table name: death	9
Reading from dbo.v_chr_basic_s	9
Table name: person	10
Reading from dbo.v_chr_basic_s	10
Reading from dbo.v_reg_basic_s	12
Reading from dbo.v_besp_ipd_s	13
Reading from dbo.v_chr_basic_s	14
Table name: visit_occurrence	15
Reading from dbo.v_opd_basic_s	15
Reading from dbo.v_ipd_basic_s	17
Table name: condition_occurrence	19
Reading from dbo.v_opd_basic_s	19
Reading from dbo.v_ipd_basic_s	22
Table name: procedure_occurrence	27
Reading from dbo.v_opd_fee_s	27
Reading from dbo.v_opd_basic_s	28
Table name: observation	29
Reading from dbo.v_opd_basic_s	29
Reading from dbo.v_ipd_basic_s	33
Reading from dbo.v_opd_fee_s	37
Table name: drug_exposure	39
Reading from dbo.v_opd_med_s	39
Reading from dbo.v_ud_order_s	41
Table name: measurement	43
Table name: device_exposure	44

Table name: person

This table serves as the central identity management for all Persons in the database. It contains records that uniquely identify each person or patient, and some demographic information.

https://ohdsi.github.io/CommonDataModel/cdm531.html#clinical_data_tables

Reading from dbo.v_chr_basic_s



Destination Field	Source Field	Logic	Comment
person_id	id_no		
	chr_no		
person_source_value	id_no		
gender_concept_id	sex_type	Male: 8507 Female: 8532 Unknown: 8551	



Usagi



Open Usagi

Usagi

File Edit View Help

Status	Source code	Source term	Frequency	Match score	Concept ID	Concept na...	Domain	Concept cla...	Vocabulary	Concept code	Standard co...	Parents	Children	Comment
--------	-------------	-------------	-----------	-------------	------------	---------------	--------	----------------	------------	--------------	----------------	---------	----------	---------

Source code

Source code	Source term	Frequency
-------------	-------------	-----------

Target concepts

Concept ID	Concept name	Domain	Concept class	Vocabulary	Concept code	Standard concept	Parents	Children
------------	--------------	--------	---------------	------------	--------------	------------------	---------	----------

Remove concept

Search

Query

☒ Use source term as query

☐ Query:

Filters

☐ Filter by user selected concepts

☒ Filter standard concepts

☒ Include source terms

☐ Filter by concept class: 2-dig nonbill code

☐ Filter by vocabulary: ABMS

☐ Filter by domain: Condition

Results

Score	Term	Concept ID	Concept name	Domain	Concept class	Vocabulary	Concept code	Standard concept	Parents	Children
-------	------	------------	--------------	--------	---------------	------------	--------------	------------------	---------	----------

Replace concept Add concept

Comment:

Approved / total: 0/0 0% of total frequency

Vocabulary version: v5.0 26-AUG-17

Approve



Import csv

Usagi

File Edit View Help

Open Ctrl-O
Import codes Ctrl-I
Apply previous mapping
Export source_to_concept_map
Save Ctrl-S
Save As...
Exit

Frequency	Match score	Concept ID	Concept na...	Domain	Concept cla...	Vocabulary	Concept code	Standard co...	Parents	Children	Comment
-----------	-------------	------------	---------------	--------	----------------	------------	--------------	----------------	---------	----------	---------

Source code

Source code	Source term	Frequency
-------------	-------------	-----------

Target concepts

Concept ID	Concept name	Domain	Concept class	Vocabulary	Concept code	Standard concept	Parents	Children
------------	--------------	--------	---------------	------------	--------------	------------------	---------	----------

Remove concept

Search

Query

☒ Use source term as query
☐ Query:

Filters

☐ Filter by user selected concepts
☒ Filter standard concepts
☒ Include source terms

☐ Filter by concept class: 2-dig nonbill code
☐ Filter by vocabulary: ABMS
☐ Filter by domain: Condition

Results

Score	Term	Concept ID	Concept name	Domain	Concept class	Vocabulary	Concept code	Standard concept	Parents	Children
-------	------	------------	--------------	--------	---------------	------------	--------------	------------------	---------	----------

Replace concept Add concept

Comment:

Approved / total: 0/0 0% of total frequency

Vocabulary version: v5.0 26-AUG-17

Approve



Import & map columns

Import codes from Melanie_Demo.xlsx

Count	Source value
49160	Family Medicine
32692	Internal Medicine
4471	Gastroenterology
3586	Pulmonary Medicine
3024	Neurological Surgery
1255	Allergy and Immunology
61	Hematopathology
60	Pediatric Hospitalist
14	ENDOVASCULAR MEDICINE
12	Epidemiology
2	Cosmetic Surgery

Column mapping

Source code column	Source value
Source name column	Source value
Source frequency column	Count
Auto concept ID column	
Additional info column	

Filters

☐ Filter by user selected concepts

☒ Filter standard concepts

☒ Include source terms

☐ Filter by concept class: 2-dig nonbill code

☐ Filter by vocabulary: ABMS

☒ Filter by domain: Provider Specialty

Cancel Import



Select mapping & approve

Usagi

File Edit View Help

Status	Source code	Source term	Frequency	Match score	Concept ID	Concept name	Domain	Concept class	Vocabulary	Concept code	Standard conc...	Parents	Children	Comment
Unchecked	Family Medicine	Family Medicine	49160	0.59	8004453	Family Practice	Provider Speci...	Specialty	Specialty	8	S	0	0	
Unchecked	Internal Medici...	Internal Medici...	32692	1.00	8004456	Internal Medici...	Provider Speci...	Specialty	Specialty	11	S	0	0	
Unchecked	Gastroenterol...	Gastroenterol...	4471	1.00	8004455	Gastroenterol...	Provider Speci...	Specialty	Specialty	10	S	0	0	
Unchecked	Pulmonary Me...	Pulmonary Me...	3586	0.63	8004472	Pulmonary Di...	Provider Speci...	Specialty	Specialty	29	S	0	0	
Unchecked	Neurological ...	Neurological ...	3024	0.69	8004459	Neurosurgery	Provider Speci...	Specialty	Specialty	14	S	0	0	
Unchecked	Allergy and Im...	Allergy and Im...	1255	0.60	8004448	Allergy/immun...	Provider Speci...	Specialty	Specialty	3	S	0	0	
Unchecked	Hematopathol...	Hematopathol...	61	0.81	5756767	Dermatopatho...	Provider Speci...	Specialty	ABMS	OMOP genera...	S	0	0	
Unchecked	Pediatric Hos...	Pediatric Hos...	60	0.65	8004284	Hospitals, Psy...	Provider Speci...	NUCC	NUCC	283Q00000X	S	0	0	
Unchecked	ENDOVASCU...	ENDOVASCU...	14	0.52	5756789	Neuromuscul...	Provider Speci...	Specialty	ABMS	OMOP genera...	S	0	0	
Unchecked	Epidemiology	Epidemiology	12	0.45	5756825	Radiology	Provider Speci...	Specialty	ABMS	OMOP genera...	S	0	0	
Unchecked	Cosmetic Sur...	Cosmetic Sur...	2	0.54	8004464	Oral Surgery	Provider Speci...	Specialty	Specialty	19	S	0	0	

Source code

Source code	Source term	Frequency
Internal Medicine	Internal Medicine	32692

Target concepts

Concept ID	Concept name	Domain	Concept class	Vocabulary	Concept code	Standard concept	Parents	Children
38004456	Internal Medicine	Provider Specialty	Specialty	Specialty	11	S	0	0

Remove concept

Search

Query

☒ Use source term as query

☐ Query:

Filters

☐ Filter by user selected concepts

☒ Filter standard concepts

☒ Include source terms

Filter by concept class: 2-dig nonbill code

Filter by vocabulary: ABMS

Filter by domain: Provider Specialty

Results

Score	Term	Concept ID	Concept name	Domain	Concept class	Vocabulary	Concept co...	Standard concept	Parents	Children
1.00	Internal Medicine	38004456	Internal Medicine	Provider Specialty	Specialty	Specialty	11	S	0	0
0.74	Internal Medicine ...	45756778	Internal Medicine - Critical Care Medicine	Provider Specialty	Specialty	ABMS	OMOP gen...	S	0	0
0.69	Maternal and Feta...	45756780	Maternal and Fetal Medicine	Provider Specialty	Specialty	ABMS	OMOP gen...	S	0	0
0.48	Allopathic & Oste...	38003863	Allopathic & Osteopathic Physicians, Internal Medicine, Addiction Medici...	Provider Specialty	NUCC	NUCC	207RA0401X	S	0	0
0.48	Allopathic & Oste...	38003867	Allopathic & Osteopathic Physicians, Internal Medicine, Critical Care Me...	Provider Specialty	NUCC	NUCC	207RC0200X	S	0	0
0.47	Allopathic & Oste...	38003883	Allopathic & Osteopathic Physicians, Internal Medicine, Sports Medicine	Provider Specialty	NUCC	NUCC	207RS0010X	S	0	0

Replace concept

Add concept

Comment:

Approved / total: 0 / 11 0.0% of total frequency

Vocabulary version: v5.0 26-AUG-17

Approve

Searching for concept_ids

Usagi

File Edit View Help

Status	Source code	Source term	Frequency	Match score	Concept ID	Concept name	Domain	Concept class	Vocabulary	Concept code	Standard co...	Parents	Children	Comment
Unchecked	Family Medicine	Family Medicine	49160	0.59	8004453	Family Practi...	Provider Spe...	Specialty	Specialty	8	S	0	0	
Approved	Internal Medi...	Internal Medicine	32692	1.00	8004456	Internal Medi...	Provider Spe...	Specialty	Specialty	11	S	0	0	
Unchecked	Gastroenterol...	Gastroenterology	4471	1.00	8004455	Gastroentero...	Provider Spe...	Specialty	Specialty	10	S	0	0	
Unchecked	Pulmonary Me...	Pulmonary Medicine	3586	0.63	8004472	Pulmonary D...	Provider Spe...	Specialty	Specialty	29	S	0	0	
Unchecked	Neurological ...	Neurological Surgery	3024	0.69	8004459	Neurosurgery	Provider Spe...	Specialty	Specialty	14	S	0	0	
Unchecked	Allergy and Im...	Allergy and Immunology	1255	0.60	8004448	Allergy/Immu...	Provider Spe...	Specialty	Specialty	3	S	0	0	
Unchecked	Hematopathol...	Hematopathology	61	0.81	5756767	Dermatopath...	Provider Spe...	Specialty	ABMS	OMOP gener...	S	0	0	
Unchecked	Pediatric Hos...	Pediatric Hospitalist	60	0.65	8004284	Hospitals, P...	Provider Spe...	NUCC	NUCC	283Q0000X	S	0	0	
Unchecked	ENDOVASCU...	ENDOVASCULAR MEDICINE	14	0.52	5756789	Neuromuscu...	Provider Spe...	Specialty	ABMS	OMOP gener...	S	0	0	
Unchecked	Epidemiology	Epidemiology	12	0.45	5756825	Radiology	Provider Spe...	Specialty	ABMS	OMOP gener...	S	0	0	
Unchecked	Cosmetic Sur...	Cosmetic Surgery	2	0.54	8004464	Oral Surgery	Provider Spe...	Specialty	Specialty	19	S	0	0	

Source code

Source code	Source term	Frequency
Cosmetic Surgery	Cosmetic Surgery	2

Target concepts

Concept ID	Concept name	Domain	Concept class	Vocabulary	Concept code	Standard concept	Parents	Children
38004464	Oral Surgery	Provider Specialty	Specialty	Specialty	19	S	0	0

Remove concept

Search

Query

☒ Use source term as query

☐ Query:

Filters

☐ Filter by user selected concepts

☒ Filter standard concepts

☒ Include source terms

☐ Filter by concept class: 2-dig nonbill code

☐ Filter by vocabulary: ABMS

☒ Filter by domain: Provider Specialty

Results

Score	Term	Concept ID	Concept name	Domain	Concept class	Vocabulary	Concept ...	Standard co...	Parents	Children
0.38	Congenital C...	45756765	Congenital Cardiac Surgery	Provider Spe...	Specialty	ABMS	OMOP ge...	S	0	0
0.38	Thoracic and ...	45756830	Thoracic and Cardiac Surgery	Provider Spe...	Specialty	ABMS	OMOP ge...	S	0	0
0.37	Vascular Surg...	38004496	Vascular Surgery	Provider Spe...	Specialty	Specialty	77	S	0	0
0.35	Allopathic & O...	38003971	Allopathic & Osteopathic Physicians, Plastic Surgery, Surgery of the Hand	Provider Spe...	NUCC	NUCC	2082S01...	S	0	0
0.34	Allopathic & O...	38004014	Allopathic & Osteopathic Physicians, Surgery, Pediatric Surgery	Provider Spe...	NUCC	NUCC	2086S01...	S	0	0
0.34	Plastic And R...	38004467	Plastic And Reconstructive Surgery	Provider Spe...	Specialty	Specialty	24	S	0	0

Replace concept Add concept

Comment:

Approved / total: 1 / 11 34.7% of total frequency

Vocabulary version: v5.0 26-AUG-17

Approve

Save & export

Usagi

File Edit View Help

Open Ctrl-O
Import codes Ctrl-I
Apply previous mapping
Export source_to_concept_map
Save Ctrl-S
Save As...
Exit

Source term	Frequency	Match score	Concept ID	Concept name	Domain	Concept class	Vocabulary	Concept code	Standard co...	Parents	Children	Comment
Family Practi...	49160	0.59	38004453	Family Practi...	Provider Spe...	Specialty	Specialty	8	S	0	0	
Internal Medi...	32692	1.00	38004456	Internal Medi...	Provider Spe...	Specialty	Specialty	11	S	0	0	
Gastroentero...	4471	1.00	38004455	Gastroentero...	Provider Spe...	Specialty	Specialty	10	S	0	0	
Pulmonary D...	3586	0.63	38004472	Pulmonary D...	Provider Spe...	Specialty	Specialty	29	S	0	0	
Neurosurgery	3024	0.69	38004459	Neurosurgery	Provider Spe...	Specialty	Specialty	14	S	0	0	
Allergy/Immu...	1255	0.60	38004448	Allergy/Immu...	Provider Spe...	Specialty	Specialty	3	S	0	0	
Dermatopath...	61	0.81	45756767	Dermatopath...	Provider Spe...	Specialty	ABMS	OMOP gener...	S	0	0	
Hospitals, P...	60	0.65	38004284	Hospitals, P...	Provider Spe...	NUCC	NUCC	283Q00000X	S	0	0	
Neuromuscu...	14	0.52	45756789	Neuromuscu...	Provider Spe...	Specialty	ABMS	OMOP gener...	S	0	0	
Radiology	12	0.45	45756825	Radiology	Provider Spe...	Specialty	ABMS	OMOP gener...	S	0	0	
Oral Surgery	2	0.54	38004464	Oral Surgery	Provider Spe...	Specialty	Specialty	19	S	0	0	

Source code

Source code	Source term	Frequency
Cosmetic Surgery	Cosmetic Surgery	2

Target concepts

Concept ID	Concept name	Domain	Concept class	Vocabulary	Concept code	Standard concept	Parents	Children
38004464	Oral Surgery	Provider Specialty	Specialty	Specialty	19	S	0	0

Remove concept

Search

Query

☒ Use source term as query
☐ Query:

Filters

☐ Filter by user selected concepts
☒ Filter standard concepts
☒ Include source terms

☐ Filter by concept class: 2-dig nonbill code
☐ Filter by vocabulary: ABMS
☒ Filter by domain: Provider Specialty

Results

Score	Term	Concept ID	Concept name	Domain	Concept class	Vocabulary	Concept ...	Standard co...	Parents	Children
0.54	Oral Surgery	38004464	Oral Surgery	Provider Spe...	Specialty	Specialty	19	S	0	0
0.51	Hand Surgery	38004480	Hand Surgery	Provider Spe...	Specialty	Specialty	40	S	0	0
0.51	Thoracic Surg...	38004473	Thoracic Surgery	Provider Spe...	Specialty	Specialty	33	S	0	0
0.47	Pediatric Surg...	45756819	Pediatric Surgery	Provider Spe...	Specialty	ABMS	OMOP ge...	S	0	0
0.46	General Surg...	38004447	General Surgery	Provider Spe...	Specialty	Specialty	2	S	0	0
0.46	Cardiac Surg...	38004497	Cardiac Surgery	Provider Spe...	Specialty	Specialty	78	S	0	0

Replace concept Add concept

Comment:

Unapprove

Approved / total: 11 / 11 100.0% of total frequency

Vocabulary version: v5.0 26-AUG-17

Finished product

SCOPOLIA EXTRACT																										
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
1	sourceCode	sourceName	sourceFre	source ADD	If match	mappingS	equivalen	statusSet	statusSetOr	conceptId	conceptN	domainId	mappingT	comment	createdBy	createdOn	atc_code	drcon_nar	obs	drug_id	drug_multi					
2	A02AD01	SULCAINE(ETHYL-P-PIPERIDYLACETYLAMINO BEN	142909		1	0.54	APPROVE	EQUAL	Alex	1.6528E+12	35198124	ethyl pipe Drug	MAPS_TO		Alex	1.65277E+12	A02AD01	SULCAINE	142909	A0464011(Y						
3	A02AD01	SULCAINE(ETHYL-P-PIPERIDYLACETYLAMINO BEN	143136		1	0.54	APPROVE	EQUAL	Alex	1.6528E+12	35198124	ethyl pipe Drug	MAPS_TO		Alex	1.65277E+12	A02AD01	SULCAINE	143136	AC464011(Y						
4	A02AF02	ACT-AL EQUIVALENT TO ALUMINUM HYDROXIDE	141863		1	0.55	APPROVE	EQUAL	Alex	1.6528E+12	985247	aluminum Drug	MAPS_TO	<auto>		1.65277E+12	A02AF02	ACT-AL EC	141863	A0057731(Y						
5	A02AG	PIPERILATE(=PIPETHANATE) HCL (=PENSANTE)	141816		1	0.23	APPROVE	EQUAL	Alex	1.6528E+12	35198151	pipethane Drug	MAPS_TO	<auto>		1.65277E+12	A02AG	PIPERILAT	141816	A0035511(Y						
6	A02AG	PIPERILATE(=PIPETHANATE) HCL (=PENSANTE)	141876		1	0.23	APPROVE	EQUAL	Alex	1.6528E+12	35198151	pipethane Drug	MAPS_TO	<auto>		1.65277E+12	A02AG	PIPERILAT	141876	A0060871(Y						
7	A02AG	DRIED ALUMINUM HYDROXIDE GEL	141918		1	0.77	APPROVE	EQUAL	Alex	1.6528E+12	985247	aluminum Drug	MAPS_TO	<auto>		1.65277E+12	A02AG	DRIED ALL	141918	A0093521(Y						
8	A02AG	SCOPOLIA EXTRACT	141938		1	0.76	APPROVE	EQUAL	Alex	1.6528E+12	45776031	Scopolia c Drug	MAPS_TO		Alex	1.65277E+12	A02AG	SCOPOLIA	141938	A0103471(Y						
9	A02AG	VITAMIN U (=METHYLMETHIONINE SULFONIUM I	142033		1	0.58	APPROVE	EQUAL	Alex	1.6528E+12	35198112	methylme Drug	MAPS_TO	<auto>		1.65277E+12	A02AG	VITAMIN I	142033	A0138001(Y						
10	A02AG	BUTINOLIN PHOSPHATE	142051		1	0.7	APPROVE	EQUAL	Alex	1.6528E+12	19037401	calcium pl Drug	MAPS_TO		Alex	1.65277E+12	A02AG	BUTINOLI	142051	A0141081(Y						
11	A02AG	VITAMIN U (=METHYLMETHIONINE SULFONIUM I	142223		1	0.58	APPROVE	EQUAL	Alex	1.6528E+12	35198112	methylme Drug	MAPS_TO	<auto>		1.65277E+12	A02AG	VITAMIN I	142223	A0196631(Y						
12	A02AG	PIPERILATE(=PIPETHANATE)	142305		1	0.32	APPROVE	EQUAL	Alex	1.6528E+12	35198151	pipethane Drug	MAPS_TO	<auto>		1.65277E+12	A02AG	PIPERILAT	142305	A0221261(Y						
13	A02AG	PIPERILATE(=PIPETHANATE) HCL (=PENSANTE)	142374		1	0.23	APPROVE	EQUAL	Alex	1.6528E+12	35198151	pipethane Drug	MAPS_TO	<auto>		1.65277E+12	A02AG	PIPERILAT	142374	A0240791(Y						
14	A02AG	BUTINOLIN PHOSPHATE	142443		1	0.7	APPROVE	EQUAL	Alex	1.6528E+12	19037401	calcium pl Drug	MAPS_TO		Alex	1.65277E+12	A02AG	BUTINOLI	142443	A0265881(Y						
15	A02AG	SCOPOLIA EXTRACT	142686		1	0.76	APPROVE	EQUAL	Alex	1.6528E+12	45776031	Scopolia c Drug	MAPS_TO		Alex	1.65277E+12	A02AG	SCOPOLIA	142686	A0332721(Y						
16	A02AG	BUTINOLIN PHOSPHATE	142897		1	0.7	APPROVE	EQUAL	Alex	1.6528E+12	19037401	calcium pl Drug	MAPS_TO		Alex	1.65277E+12	A02AG	BUTINOLI	142897	A0462541(Y						
17	A02AG	BUTINOLIN PHOSPHATE	142928		1	0.7	APPROVE	EQUAL	Alex	1.6528E+12	19037401	calcium pl Drug	MAPS_TO		Alex	1.65277E+12	A02AG	BUTINOLI	142928	A0471551(Y						
18	A02AG	PIPERILATE(=PIPETHANATE) HCL (=PENSANTE)	142948		1	0.23	APPROVE	EQUAL	Alex	1.6528E+12	35198151	pipethane Drug	MAPS_TO	<auto>		1.65277E+12	A02AG	PIPERILAT	142948	A0486851(Y						
19	A02AG	BUTINOLIN PHOSPHATE	142958		1	0.7	APPROVE	EQUAL	Alex	1.6528E+12	19037401	calcium pl Drug	MAPS_TO		Alex	1.65277E+12	A02AG	BUTINOLI	142958	A0489311(Y						
20	A02AG	BUTINOLIN PHOSPHATE	142963		1	0.7	APPROVE	EQUAL	Alex	1.6528E+12	19037401	calcium pl Drug	MAPS_TO		Alex	1.65277E+12	A02AG	BUTINOLI	142963	A0490461(Y						
21	A02AG	BUTINOLIN PHOSPHATE	142968		1	0.7	APPROVE	EQUAL	Alex	1.6528E+12	19037401	calcium pl Drug	MAPS_TO		Alex	1.65277E+12	A02AG	BUTINOLI	142968	A0497631(Y						
22	A02AG	BUTINOLIN PHOSPHATE	142978		1	0.7	APPROVE	EQUAL	Alex	1.6528E+12	19037401	calcium pl Drug	MAPS_TO		Alex	1.65277E+12	A02AG	BUTINOLI	142978	A0497661(Y						
23	A02AG	DRIED ALUMINUM HYDROXIDE GEL	143006		0	0.77	APPROVE	EQUAL	Alex	1.6528E+12	985247	aluminum Drug	MAPS_TO	<auto>		1.65277E+12	A02AG	DRIED ALL	143006	AC093521(Y						
24	A02AG	BUTINOLIN PHOSPHATE	143020		1	0.7	APPROVE	EQUAL	Alex	1.6528E+12	19037401	calcium pl Drug	MAPS_TO		Alex	1.65277E+12	A02AG	BUTINOLI	143020	AC141081(Y						
25	A02AG	BUTINOLIN PHOSPHATE	143061		1	0.7	APPROVE	EQUAL	Alex	1.6528E+12	19037401	calcium pl Drug	MAPS_TO		Alex	1.65277E+12	A02AG	BUTINOLI	143061	AC265881(Y						
26	A02AG	BUTINOLIN PHOSPHATE	143122		1	0.7	APPROVE	EQUAL	Alex	1.6528E+12	19037401	calcium pl Drug	MAPS_TO		Alex	1.65277E+12	A02AG	BUTINOLI	143122	AC462541(Y						
27	A02AG	BUTINOLIN PHOSPHATE	143157		1	0.7	APPROVE	EQUAL	Alex	1.6528E+12	19037401	calcium pl Drug	MAPS_TO		Alex	1.65277E+12	A02AG	BUTINOLI	143157	AC497631(Y						
28	A02AG	BUTINOLIN PHOSPHATE	143170		1	0.7	APPROVE	EQUAL	Alex	1.6528E+12	19037401	calcium pl Drug	MAPS_TO		Alex	1.65277E+12	A02AG	BUTINOLI	143170	AC497661(Y						
29	A02AG	BUTINOLIN PHOSPHATE	143171		1	0.7	APPROVE	EQUAL	Alex	1.6528E+12	19037401	calcium pl Drug	MAPS_TO		Alex	1.65277E+12	A02AG	BUTINOLI	143171	AC497661(Y						
30	A02AG	PIPERILATE(=PIPETHANATE) HCL (=PENSANTE)	143227		1	0.23	APPROVE	EQUAL	Alex	1.6528E+12	35198151	pipethane Drug	MAPS_TO	<auto>		1.65277E+12	A02AG	PIPERILAT	143227	B0104831(Y						
31	A02AG	PIPERILATE(=PIPETHANATE) HCL (=PENSANTE)	143267		1	0.23	APPROVE	EQUAL	Alex	1.6528E+12	35198151	pipethane Drug	MAPS_TO		Alex	1.65277E+12	A02AG	PIPERILAT	143267	B0170711(Y						
32	A02AG	PIPERILATE(=PIPETHANATE) HCL (=PENSANTE)	143290		1	0.23	APPROVE	EQUAL	Alex	1.6528E+12	35198151	pipethane Drug	MAPS_TO	<auto>		1.65277E+12	A02AG	PIPERILAT	143290	B0177131(Y						
33	A02AG	PIPERILATE(=PIPETHANATE) HCL (=PENSANTE)	143339		1	0.23	APPROVE	EQUAL	Alex	1.6528E+12	35198151	pipethane Drug	MAPS_TO	<auto>		1.65277E+12	A02AG	PIPERILAT	143339	B0222671(Y						
34	A02AG	PIPERILATE(=PIPETHANATE) HCL (=PENSANTE)	143344		1	0.23	APPROVE	EQUAL	Alex	1.6528E+12	35198151	pipethane Drug	MAPS_TO	<auto>		1.65277E+12	A02AG	PIPERILAT	143344	B0222671(Y						
35	A02AG	METHYLMETHIONINE SULFONIUM BROMIDE	143381		1	0.82	APPROVE	EQUAL	Alex	1.6528E+12	19004724	methioninr Drug	MAPS_TO		Alex	1.65277E+12	A02AG	METHYLM	143381	N0027511(Y						
36	A02AG	GINGER OIL	143383		1	1	APPROVE	EQUAL	Alex	1.6528E+12	1314928	ginger roc Drug	MAPS_TO		Alex	1.65277E+12	A02AG	GINGER O	143383	N0031861(Y						
37	A02AG	SCOPOLIA EXTRACT	143417		1	0.76	APPROVE	EQUAL	Alex	1.6528E+12	45776031	Scopolia c Drug	MAPS_TO		Alex	1.65277E+12	A02AG	SCOPOLIA	143417	N0047661(Y						
38	A02AG	BUTINOLIN PHOSPHATE	199042		1	0.7	APPROVE	EQUAL	Alex	1.6528E+12	19037401	calcium pl Drug	MAPS_TO		Alex	1.65277E+12	A02AG	BUTINOLI	199042	A0193411(Y						
39	A02AG	BUTINOLIN PHOSPHATE	199955		1	0.7	APPROVE	EQUAL	Alex	1.6528E+12	19037401	calcium pl Drug	MAPS_TO		Alex	1.65277E+12	A02AG	BUTINOLI	199955	A0467261(Y						
40	A02AH	PEPPER	142121		1	1	APPROVE	EQUAL	Alex	1.6528E+12	19086712	pepperperi Drug	MAPS_TO		Alex	1.65277E+12	A02AH	PEPPER	142121	A0164091(Y						
41	A02AH	FENNEL (=FOENICULUM)	143327		1	0.32	APPROVE	EQUAL	Alex	1.6528E+12	43013629	fennel ext Drug	MAPS_TO		Alex	1.65277E+12	A02AH	FENNEL (=	143327	B0199541(Y						
42	A02AX	METHYLMETHIONINE SULFONIUM CHLORIDE DL-	137266		1	0.87	APPROVE	EQUAL	Alex	1.6528E+12	35198112	methylme Drug	MAPS_TO	<auto>		1.65277E+12	A02AX	METHYLM	137266	A0299271(Y						
43	A02AX	OXETHAZINE	141869		1	0.59	APPROVE	EQUAL	Alex	1.6528E+12	36879092	Oxetacain Drug	MAPS_TO		Alex	1.65277E+12	A02AX	OXETHAZ	141869	A0058111(Y						
44	A02AX	POLYMAGAL	141993		1	0.61	APPROVE	EQUAL	Alex	1.6528E+12	908360	magaldrat Drug	MAPS_TO		Alex	1.65277E+12	A02AX	POLYMAG	141993	A0128281(Y						
45	A02AX	POLYMAGAL	141995		1	0.61	APPROVE	EQUAL	Alex	1.6528E+12	908360	magaldrat Drug	MAPS_TO		Alex	1.65277E+12	A02AX	POLYMAG	141995	A0128281(Y						



Create code mappings

- Need to map codes to one of the OMOP CDM standards:
 - Drugs: RxNorm
 - Conditions: SNOMED
 - Lab values: LOINC
 - Specialties: CMS
 - ...



Implement the ETL

- No standard. Depends on the expertise of the person doing the work:
 - SQL
 - SAS
 - C++ (CDMBuilder)
 - Java (JCDMBuilder)
 - Kettle



Q/A?