



CDM-ETL WG

OHDSI Europe Symposium

June 2nd 2024



Program

- 15' Introduction to the OMOP CDM (Anne)
 - 15' Introduction to the Themis convention library (Melanie)
 - 15' Introduction to Data Quality Tooling (Maxim)
 - Convention Quiz (All)
-



Leads

Anne Van Winzum



Melanie Philofsky



Maxim Moinat





Your Journey

- Who already has an OMOP CDM instance completed?
 - Who is in the process of converting their data to the OMOP CDM?
 - Who has done the EHDEN Academy course for OMOP CDM?
-

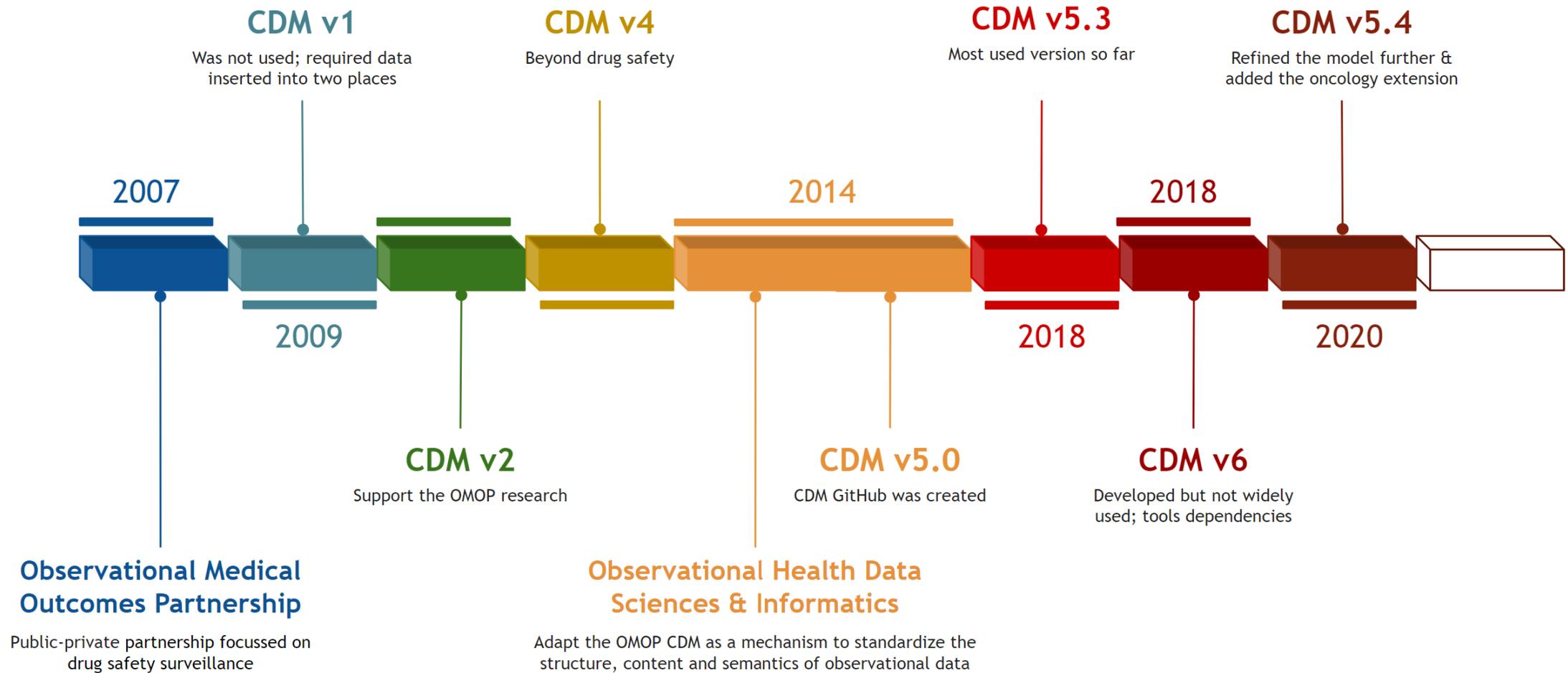


OHDSI

OBSERVATIONAL HEALTH DATA SCIENCES AND INFORMATICS

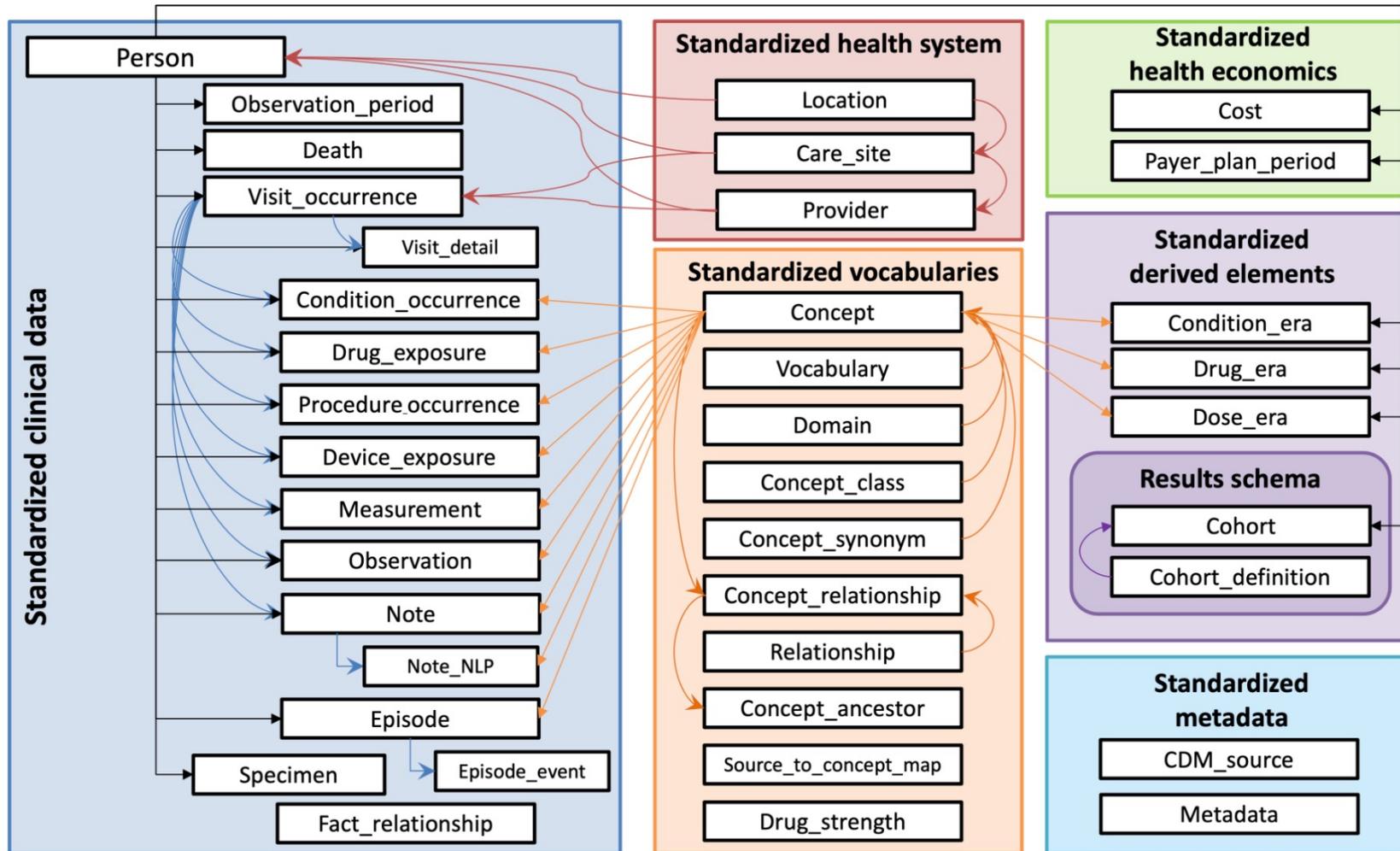
Introduction to the OMOP CDM and Vocabularies

Anne van Winzum



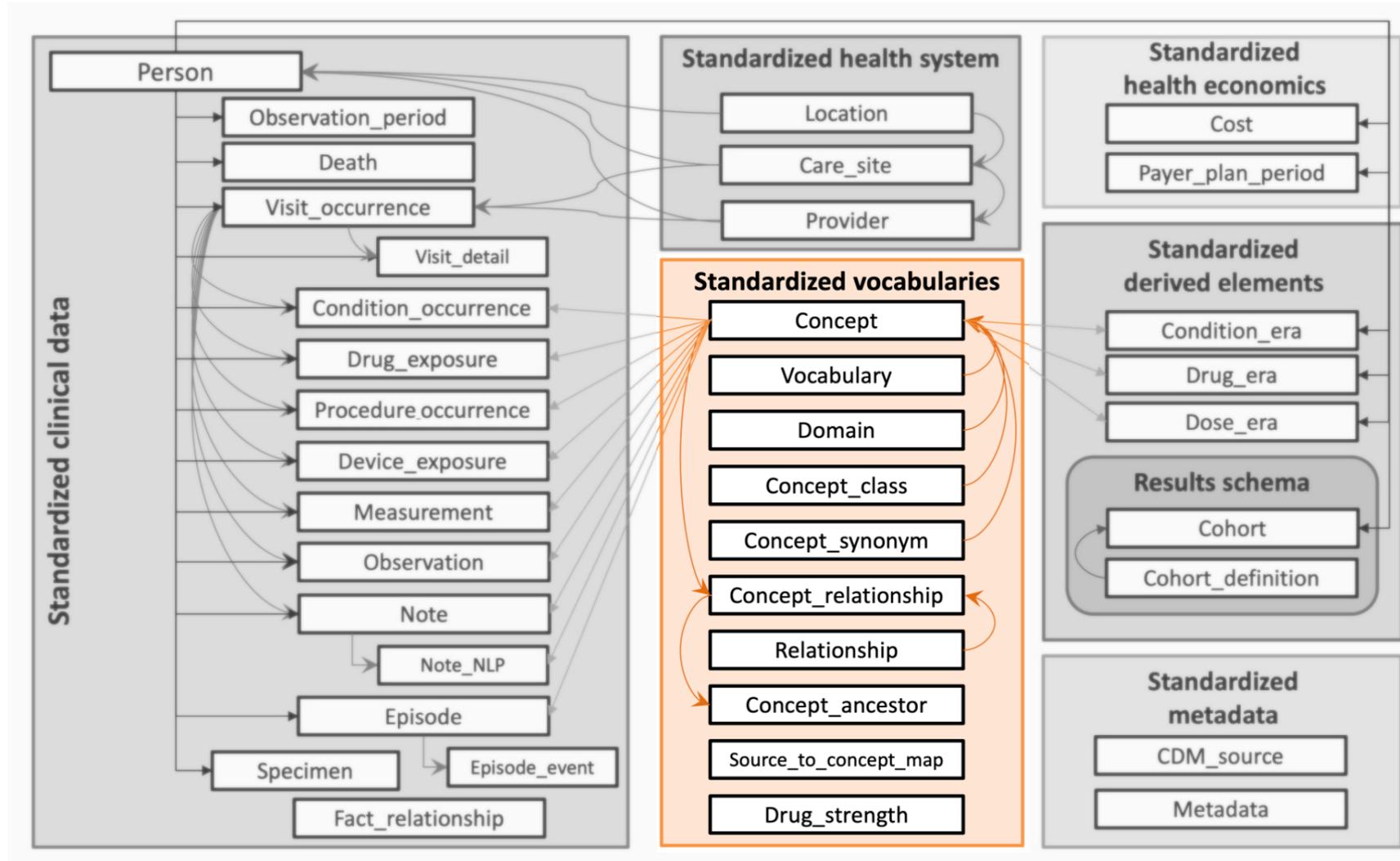


OMOP CDM 5.4





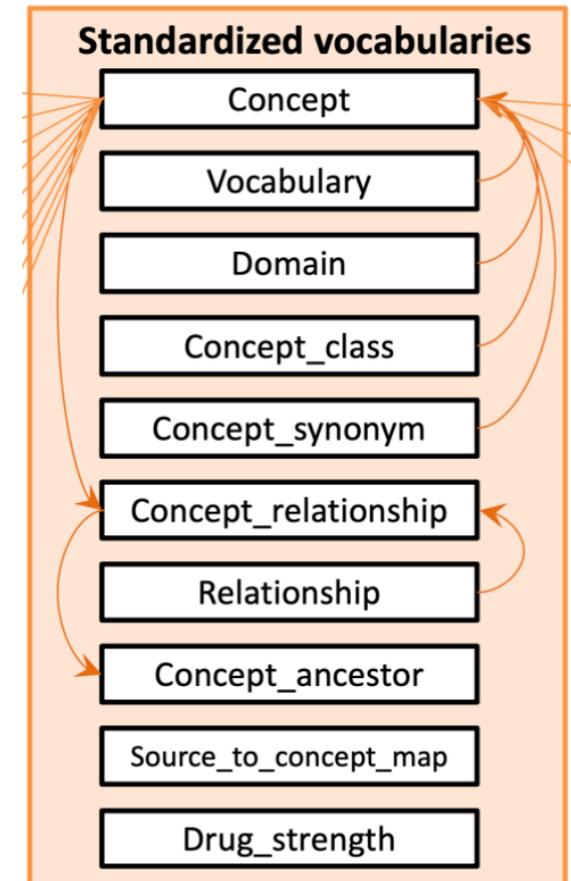
OMOP CDM 5.4





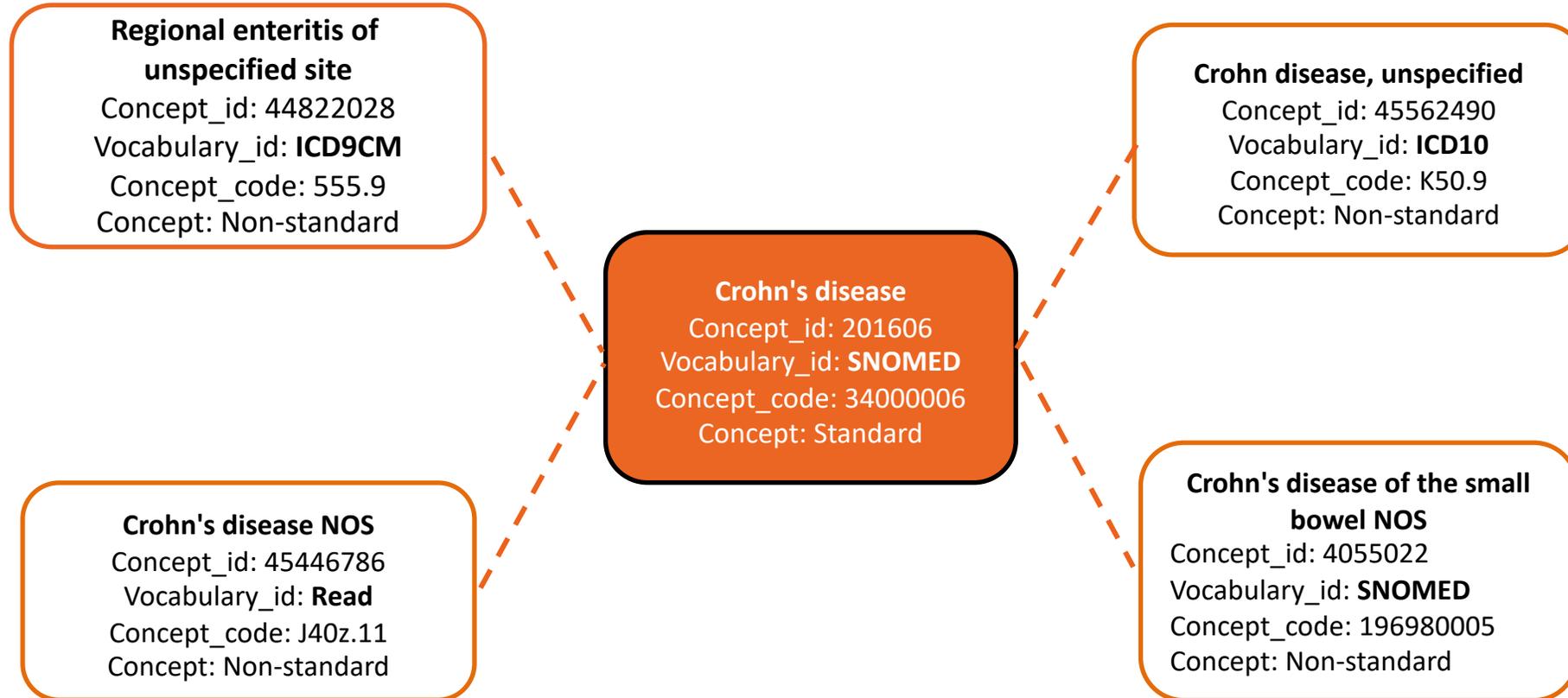
OMOP Standardized vocabularies

- **Standard concept:** One concept representing the meaning of each clinical event
- **Non standard concepts:** rest of the concepts of the standardized vocabularies which represent the same clinical event
- **CONCEPT_RELATIONSHIP table:** Mapping between non standard and standard concepts





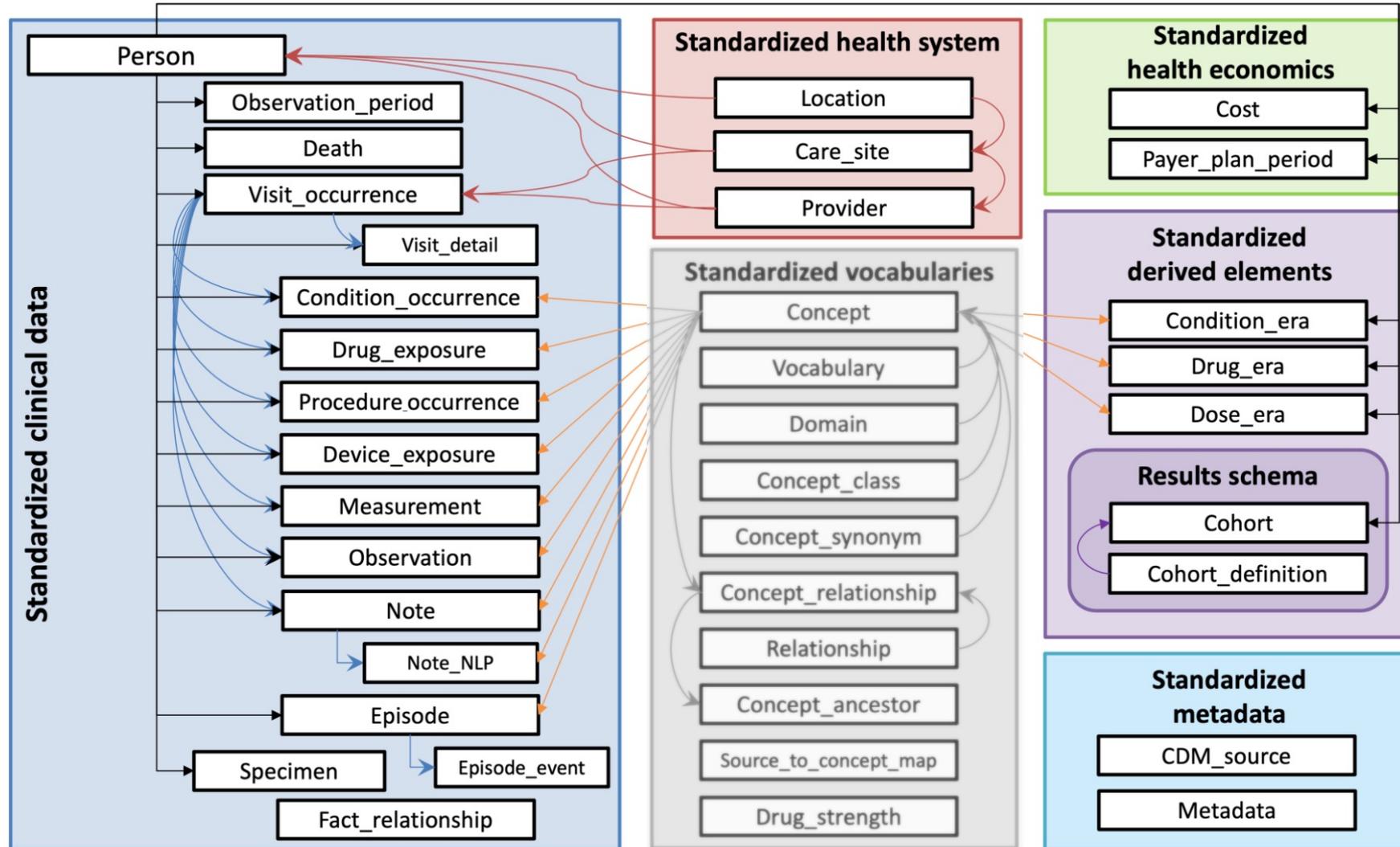
Standardized vocabulary example





Concept attributes

	Medication Example	Lab Example	Diagnosis Example
<i>concept_id</i>	1503297	3012888	45591453
<i>concept_name</i>	Metformin	Diastolic blood pressure	Essential (primary) hypertension
<i>domain_id</i>	Drug	Measurement	Condition
<i>vocabulary_id</i>	RxNorm	LOINC	ICD10
<i>concept_class_id</i>	Ingredient	Clinical Observation	ICD10 Hierarchy
<i>concept_code</i>	6809	8462-4	I10
<i>standard_concept</i>	S	S	NULL





Core tables

- Person (mandatory)
- Observation Period (mandatory)
- CDM source (mandatory)
- Medical events
 - Visit Occurrence
 - Death
 - Clinical event domains



Example Person and Observation Period record

Column	Value
person_id	123
gender_concept_id	8532
year_of_birth	1992
race_concept_id	8527
gender_source_value	F
race_source_value	white



Column	Value
observation_period_id	1
person_id	123
observation_period_start_date	2010-07-01
observation_period_end_date	2014-12-05



Eight Clinical Event Domain tables

Domain	Type of data
Condition Occurrence	Diagnoses
Drug Exposure	Medications
Procedure Occurrence	Diagnostics or Surgical operations
Measurement	Lab results
Observation	Other clinical facts
Specimen	Sample, biopt
Device Exposure	Implantation, med. equipment, supplies
Note	Free text



Clinical Event Structure

- Core elements
 - `<>_id`
 - `person_id`
 - `<>_concept_id`
 - `<>_source_value`
 - `<>_date/datetime`
 - `<>_source_concept_id`
 - `<>_type_concept_id`
 - `visit_occurrence_id`



Example record

Column	Value
conditon_occurrence_id	12345
person_id	123
condition_start_date	2010-07-01
condition_concept_id	261236 (SNOMED 372135002: Primary malignant neoplasm of upper lobe, bronchus or lung)
condition_source_value	C34.1
condition_source_concept_id	45595646 (ICD10 C34.1: Malignant neoplasm: Upper lobe, bronchus or lung)
condition_type_concept_id	32817 (OMOP4976890: EHR)
visit_occurrence_id	51234



Themis

Work Group and Convention Library

Melanie Philofsky



Themis: What is it?

The goal of Themis is to provide *conventions* on how source data should be standardized to the OMOP CDM to support the OHDSI community to generate the evidence that promotes better health decisions and better care. When there is *ambiguity* on how data should be inserted into the CDM, Themis will examine the issue, create a convention and document it.

We follow the FAIR principles: Findable, Accessible, Interoperable and Reusable.

Themis makes decisions for the good of the whole community. We must compromise. We can always revisit and modify the convention. Don't let perfect be the enemy of great. And interoperability between different OMOP CDMs is great!



State of the Conventions at the start of 2024

- CDM requirements and THEMIS conventions were housed on the CDM website, CDM github, THEMIS github and OHDSI forum.
- The documentation was not clear or concise
- Some conventions had conflicting advice and were ambiguous



Goals of April Olympians

OHDSI Data Standards Collab-a-thon

1. Identified all *currently ratified* CDM and THEMIS conventions for every CDM table and field
2. Wrote clear documentation for each THEMIS convention
3. Created a repository for THEMIS conventions
4. Updated the CDM documentation to link to relevant THEMIS repository entries
5. Created CDM documentation related to expansion module efforts around the community



Themis Convention Library

THEMIS | THEMIS Convention x +

ohdsi.github.io/Themis/index.html#

Gmail YouTube Maps

THEMIS Repository

Nav THEMIS GitHub CDM Docs DQD Docs Feedback

search...

THEMIS Conventions

- General Conventions
- CDM Tables
- Tag Browser**
- Address
- Birthdate
- Dates
- Gender
- Gender Identity
- Labs
- Lab Values
- Location
- Measurement
- Negative
- Observation Period
- Patient Reported
- Person
- Place of Service
- Provider
- Quantity

THEMIS

Summary: This page provides an overview of the THEMIS project.

Table of Contents

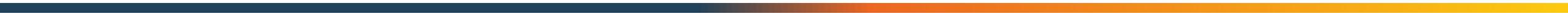
- Introduction
 - What falls under Themis and what does not
- Topic Processing
- Getting Involved

Introduction

OHDSI Community: Themis makes decisions for the good of the whole community. We must compromise. We can always revisit and modify the convention. Don't let perfect be the enemy of great. And interoperability between different OMOP CDMs



Let's explore the the Themis website





Themis details

- **Documentation:** MS Teams
- **GitHub:** <https://github.com/OHDSI/Themis>
- **Meetings:** 1st and 3rd Thursday of the month at 10:30am Eastern Time/ 4:30pm CEST/UTC-04:00



High Quality OMOP CDM transformation

Maxim Moinat



OHDSI

OBSERVATIONAL HEALTH DATA SCIENCES AND INFORMATICS



Extract Load Transform (ETL) journey

Exploration & ETL Design ★ A

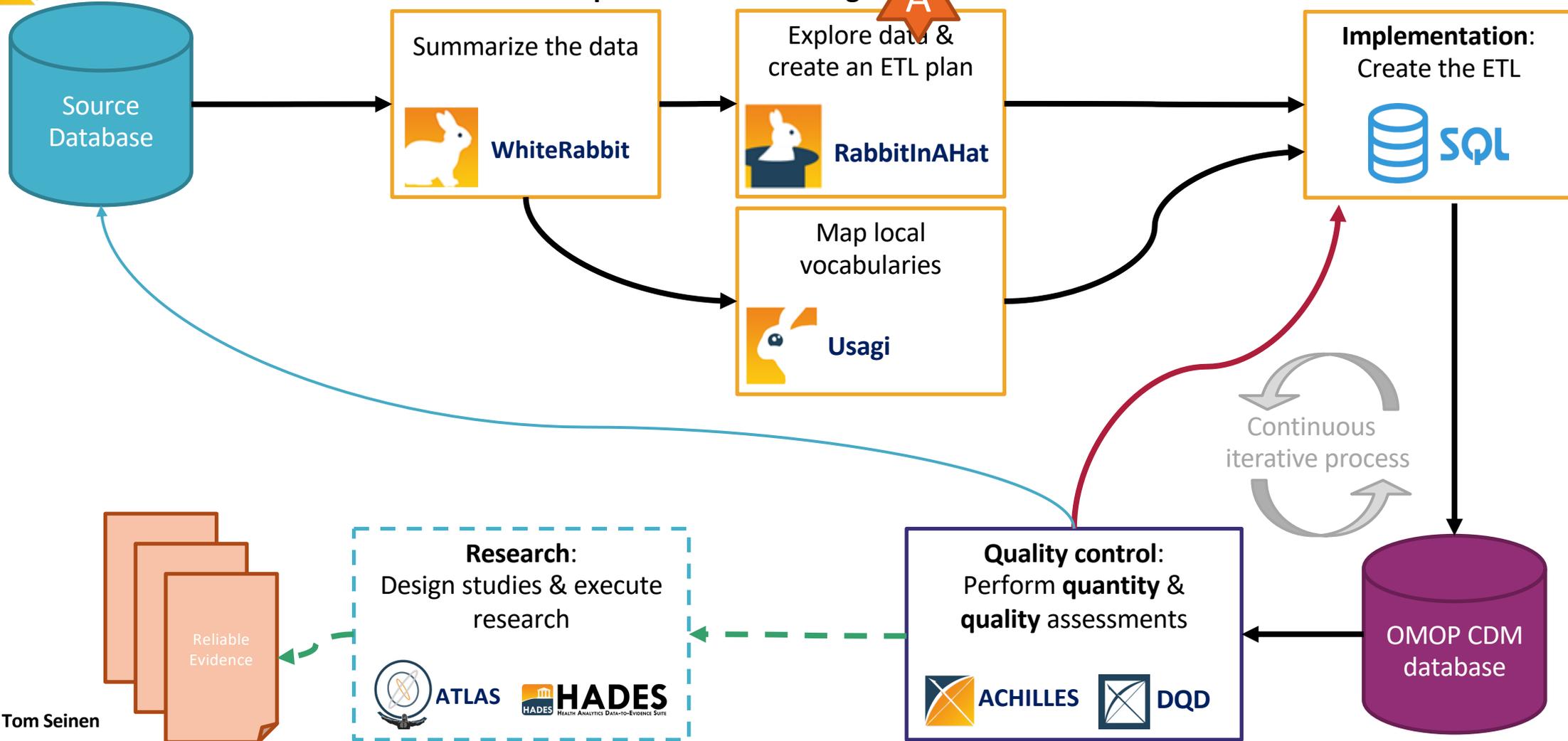


Image credit: Tom Seinen (ErasmusMC)



A - ETL design

- Explore Source data

Table/Field Overview

Table	Field	Type	Max length	N rows	N rows ch	Fraction emp
allergies	start	date	10	619	619	0
allergies	stop	date	10	619	619	0.904685
allergies	patient	character	36	619	619	0
allergies	encounte	character	36	619	619	0
allergies	code	character	9	619	619	0
allergies	descriptio	character	24	619	619	0

Value frequencies

W	X	Y	Z	AA	AB	AC	AD	AE	AF
marital	Frequency	race	Frequency	ethnicity	Frequency	gender	Frequency	birthplace	Frequency
M	622	white	846	irish	233	M	572	Boston	142
	344	hispanic	112	italian	145	F	558	Springfiel	30
S	166	black	82	english	102		2	Worceste	28
		asian	70	puerto_ri	72			Lowell	22
		native	20	french	72			Brockton	21
		other	1	german	64			Cambridg	18
		Unknown	1	chinese	51			Methuen	18
				polish	49			Newton	17
				american	39			Quincy	16
				portugues	37			Framingh	16
				french_ca	35			Lynn	12
				african	33			Arlington	12
				west_indi	28			Weymout	12
				dominicar	21			New Bedf	12
				american_	20			Lawrence	11
				russian	20			Haverhill	11
				scottish	19			Fitchburg	11
				asian_indi	19			Marshfiel	10
				mexican	18			Somervill	10
				swedish	17			Barnstabl	10
				central_ar	13			Fall River	9
				greek	12			Waltham	9

Two records with an empty gender value

Document ETL logic

Tutorial-ETL

Search Tutorial-ETL Synthea OMOP ETL

CDM Synthea v1 / 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	



Semantic Harmonisation



• Lookup table (Source

Overview table

The screenshot displays the Usagi application interface for semantic harmonisation. It includes a main table of mappings, a 'Selected mapping' view, and a 'Search facility'.

Overview table (Main Table):

Status	Source code	Source term	Frequency	Dutch term	Match score	Concept ID	Concept name	Domain	Concept class	Vocabulary	Concept code	Standard con.	Parents	Children	Comment
Unchecked	A99.00	General dise.	5774012	Andere gege.	0.78	435569	Disorder of jaw	Condition	Clinical Findi.	SNOMED	37156001	S	1	39	
Unchecked	K85.00	Hypertension	3987206	Essentiële	0.72	42709887	Hypertensie	Condition	Clinical Findi.	SNOMED	449759005	S	1	8	
Unchecked	R44.00	Preventive im.	3702922	Immunisatie	0.41	4102757	Response to	Condition	Clinical Findi.	SNOMED	26529003	S	1	0	
Unchecked	I90.02	Diabetes mell.	2275799	Diabetes me.	0.91	201826	Type 2 diabet.	Condition	Clinical Findi.	SNOMED	44054006	S	1	14	
Auto mapped	R05.00	Cough	1268829	Hoesten	0.58	4158493	C/O - cough	Condition	Clinical Findi.	SNOMED	272039006	S	2	0	
Unchecked	R74.00	Upper respir.	1061504	Acute infectie	1.00	260427	Common cold	Condition	Clinical Findi.	SNOMED	82272006	S	1	0	
Unchecked	A20.00	General sym.	1035167	Andere alege.	0.67	4937322	C/O - a gener.	Condition	Clinical Findi.	SNOMED	162409008	S	1	1	
Unchecked	L03.00	Low back sy.	998249	Lage rugpin	0.82	4155085	C/O - a back	Condition	Clinical Findi.	SNOMED	272009001	S	2	0	
Unchecked	U71.00	Cystitis/urina	970719	Cystitis/urine	0.64	81902	Urinary tract i.	Condition	Clinical Findi.	SNOMED	68566005	S	2	17	

Selected mapping:

Source code	Source term	Frequency	Dutch term
A99.00	General disease	5774012	

Search facility:

Query:

Filters:

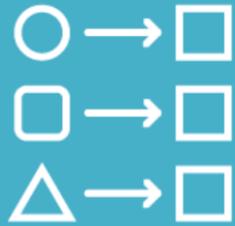
- Filter by user selected concepts
- Filter standard concepts
- Include source terms
- Filter by concept class: 2-dig nonbill code
- Filter by vocabulary: ABMS
- Filter by domain: Condition

Results:

Score	Term	Concept ID	Concept name	Domain	Concept class	Vocabulary	Concept code	Standard concept	Parents	Children
0.78	Oral disease	435569	Disorder of jaw	Condition	Clinical Finding	SNOMED	37156001	S	1	39
0.75	Veneral disease	440647	Sexually transmitted	Condition	Clinical Finding	SNOMED	80980009	S	1	18
0.75	Veneral disease	433957	Syphilitic infection	Condition	Clinical Finding	SNOMED	41115009	S	1	15
0.67	Viral disease	440029	Viral disease	Condition	Clinical Finding	SNOMED	34014006	S	1	36
0.64	Nil disease	4260398	Minimal change dis.	Condition	Clinical Finding	SNOMED	44785005	S	1	3
0.64	Nil disease	199071	Nephrotic syndrom.	Condition	Clinical Finding	SNOMED	266549004	S	1	0
0.62	Viral diseases	140020	Viral exanthem	Condition	Clinical Finding	SNOMED	49882001	S	2	7
0.61	Genetic disease	443916	Hereditary disease	Condition	Clinical Finding	SNOMED	32895009	S	1	34
0.59	Degenerative disc d.	80816	Degeneration of int.	Condition	Clinical Finding	SNOMED	77547008	S	3	5

Local 'custom' concepts

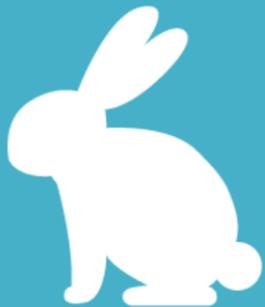
1	concept_id_1	concept_id_2	relationship_id	valid_start_date	valid_end_date	omain_id	concept_class_id	standard_concept
2	2103000001	36203185	Maps to	2023-01-01	1999-12-31	'isit	Visit	
3	36203185	2103000001	Mapped from	2023-01-01	1999-12-31	'isit	Visit	
4	2103000010	3027018	Maps to	2023-01-01	1999-12-31	'isit	Visit	
1	vocabulary_id	vocabulary_name	vocabulary_reference		vocabulary_version			
2	CS Meting Parameter	ChipSoft Meting Parameter	https://metadata.erasmusmc.nl/ids/MeasurementParameter/Hix		2020-12-02 15:44:26.130			
7	2103000001	2103000001	mapped from	2023-01-01	1999-12-31	'isit	Visit	
8	2103000012	3024171	Maps to	2023-01-01	1999-12-31	'isit	Visit	
9	3024171	2103000012	Mapped from	2023-01-01	1999-12-31	'isit	Visit	
10	2103000013	40762499	Maps to	2023-01-01	1999-12-31	'isit	Visit	
11	40762499	2103000013	Mapped from	2023-01-01	1999-12-31	'isit	Visit	
12	2103000014	3020891	Maps to	2023-01-01	1999-12-31	'isit	Visit	
13	3020891	2103000014	Mapped from	2023-01-01	1999-12-31	'isit	Visit	
14	2103000020	42529374	Maps to	2023-01-01	1999-12-31	'isit	Visit	
15	42529374	2103000020	Mapped from	2023-01-01	1999-12-31	'isit	Visit	
16	2103000021	3036277	Maps to	2023-01-01	1999-12-31	'isit	Visit	
17	3036277	2103000021	Mapped from	2023-01-01	1999-12-31	'isit	Visit	
18	2103000023	3025315	Maps to	2023-01-01	1999-12-31	'isit	Visit	
19	3025315	2103000023	Mapped from	2023-01-01	1999-12-31	'isit	Visit	



Extract, Transform and Load

Tool

Map raw observational data to the OMOP CDM.



Introduction to Usagi & Code Mappings for an ETL

Tool

Introduction to the Usagi tool, importing codes, review and output the code mapp...





Extract Load Transform (ETL) journey

Exploration & ETL Design

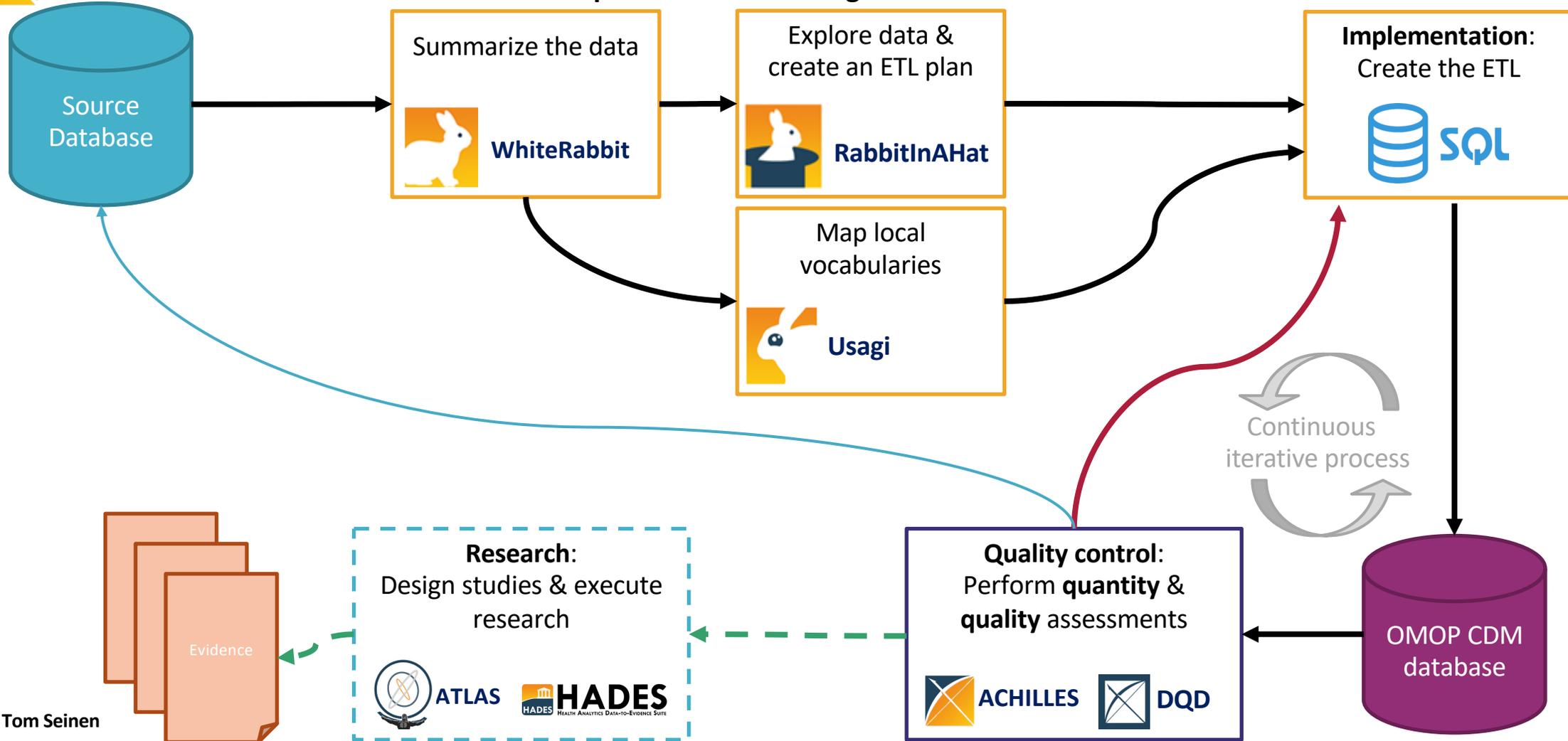
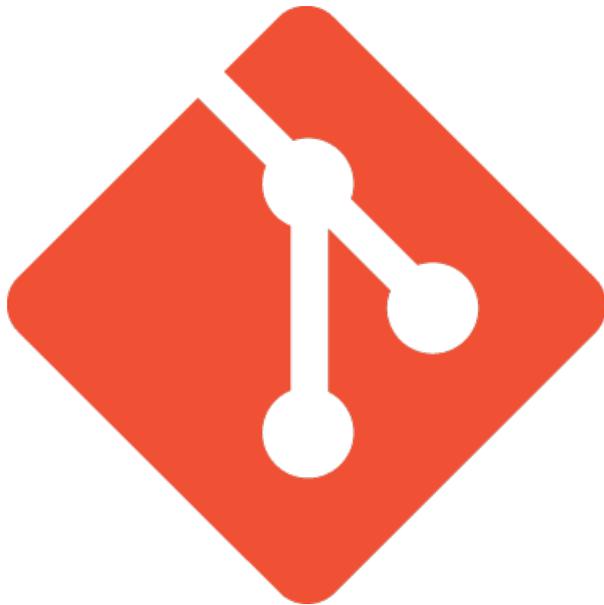


Image credit: Tom Seinen (ErasmusMC)



B – ETL implementation

- **Version Control**



Software validity

- Unit tests
- End-to-end tests

```
48 declareTest(801, 'Conditon to Condition Occurrence - record')
49 add_patient(id = '801', custom_id = 801)
50 add_condition(subject_patient = '801')
51 expect_condition_occurrence(person_id = lookup_person('person_id', person_source_value = '801'))
52
53 declareTest(802, 'Conditon to Condition Occurrence - code')
54 add_patient(id = '802', custom_id = 802)
55 add_condition(subject_patient = '802', code_code_original = 'CL10I12.0')
56 expect_condition_occurrence(person_id = lookup_person('person_id', person_source_value = '802'),
57                             condition_source_value = 'CL10I12.0')
58
59 declareTest(803, 'Conditon to Condition Occurrence - date')
60 add_patient(id = '803', custom_id = 803)
61 add_condition(subject_patient = '803', onsetdatetime = '2022-05-01')
62 expect_condition_occurrence(person_id = lookup_person('person_id', person_source_value = '803'),
63                             condition_start_date = '2022-05-01')
64
```



ETL Implementation

There are multiple tools available to implement your ETL

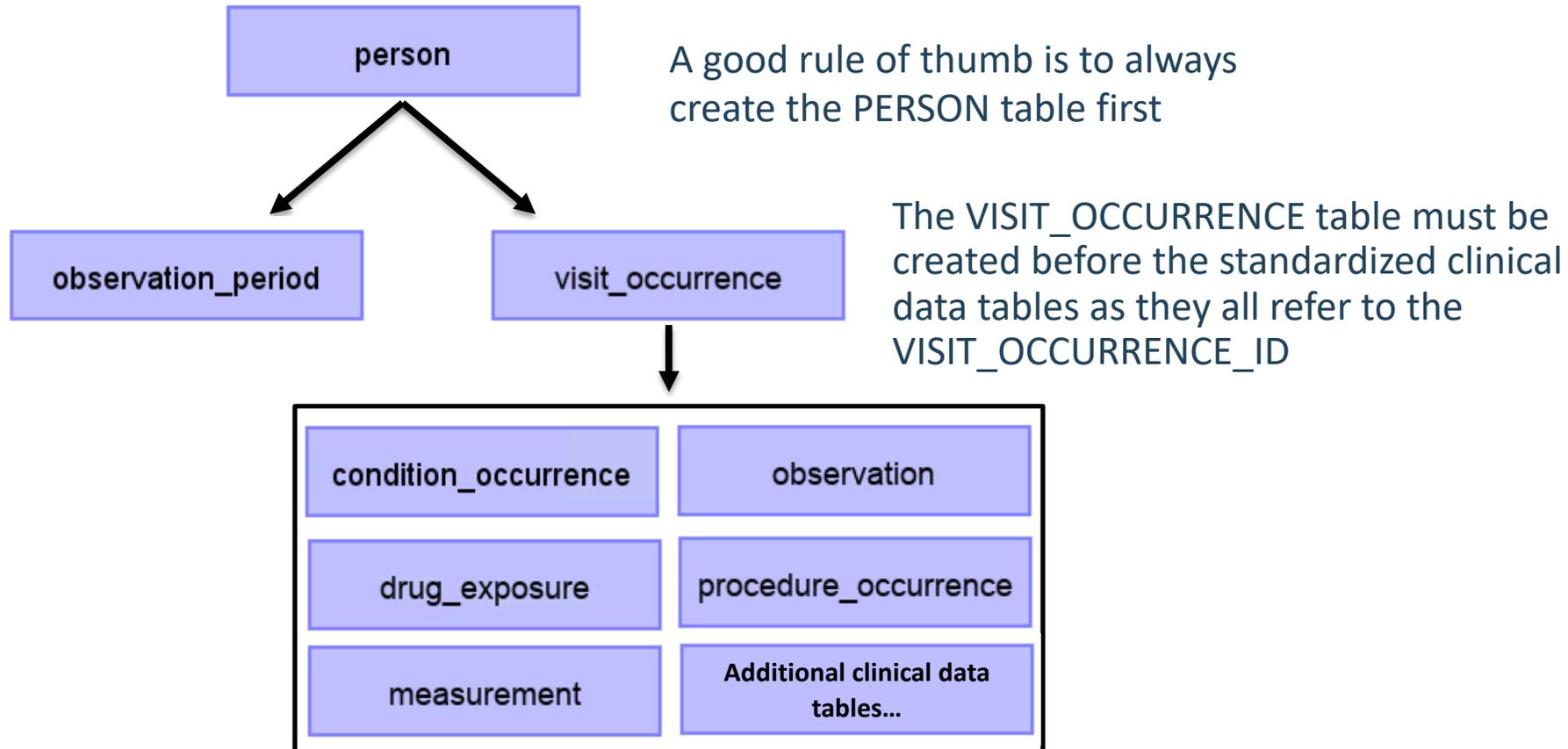


In this example we created a builder using SQL and R, though your choice will largely depend on the size and complexity of the ETL design



ETL Implementation

General Flow of Implementation





OMOP ETL Initiatives

- SQL - RiaH exports simple SQL skeleton
- Python - delphyne
- R – Synthea ETL
- Java - JCdmBuilder
- C# - ETL-CDMBuilder
- Frameworks: dbt, Pentaho, bonobo
- *Commercial offerings (e.g. from EHDEN trained SMEs)*

<https://www.ehden.eu/business-directory/>



Repositories

Q etl Type Language Sort

19 results for all repositories matching etl sorted by last updated

Clear filter

dbt-synthea Public [Under development] A dbt ETL project to convert a Synthea synthetic data set into the OMOP CDM ☆ 6 🍴 1 🔄 21 📄 2 Updated 9 hours ago	
ETL-Synthea Public A package supporting the conversion from Synthea CSV to OMOP CDM ● R ☆ 93 🍴 69 🔄 2 📄 0 Updated yesterday	
ETL-LambdaBuilder Public CDM Builder leveraging AWS Lambda ● C# ☆ 10 🍴 Apache-2.0 🍴 14 🔄 3 📄 0 Updated 3 days ago	
WhiteRabbit Public WhiteRabbit is a small application that can be used to analyse the structure and contents of a database as preparation for designing an ETL. It comes with RabbitnAHat, an application for interactive design of an ETL to the OMOP Common Data Model with the help of the the scan report generated by White Rabbit. ● Java ☆ 171 🍴 Apache-2.0 🍴 83 🔄 103 (3 issues need help) 📄 2 Updated last week	
ETL-German-FHIR-Core Public ETL process from FHIR (defined by the German Medical Informatics Initiative) to OMOP ● Java ☆ 6 🍴 3 🔄 1 📄 0 Updated on Mar 11	
Perseus Public [under development] Tools for ETL into OMOP CDM and deployment of OHDSI toolstack ● TypeScript ☆ 31 🍴 Apache-2.0 🍴 8 🔄 13 📄 0 Updated on Nov 10, 2023	
ETL--LupusRegistries Public Conventions for Lupusnet registries ☆ 0 🍴 0 🔄 0 📄 0 Updated on May 5, 2023	
ETL-CDMBuilder Public ETL-CDMBuilder is a repo containing a .NET Core application to perform ETL to OMOP CDM for multiple databases ● C# ☆ 49 🍴 Apache-2.0 🍴 30 🔄 1 📄 3 Updated on Apr 16, 2023	
MIMIC Public MIMIC (Medical Information Mart for Intensive Care) is a large, single-center database comprising information relating to patients admitted to critical care units at a large tertiary care hospital. This repository contains the ETL to the OMOP CDM. ● Python ☆ 65 🍴 Apache-2.0 🍴 21 🔄 20 📄 2 Updated on Apr 4, 2023	
ETL-German-Claims Public ETL process for German claims data to OMOP ● Shell ☆ 3 🍴 1 🔄 0 📄 0 Updated on Nov 23, 2022	

<https://github.com/OHDSI?q=etl>



YouTube NL Search

OMOP ELT using dbt

<https://omop-lsc.surge.sh/>

dbt
Data Build Tool
SQL + Jinja + Git
www.getdbt.com

Data Lineage, Self-Documenting, Metadata Generation
Diverse source and target architectures
Shallow Learning Curve

- stash On vc/main: HGithub_Desktop-i
- index on vc/main: ac50207 Addition of r
- qa/main -remotes/origin/qa/main
- Change of workflow
- Created of vocab layer to condition occ
- Added in cons code to flex provider id
- Change of colours in lineage graph
- remotes/origin/qa/main Correcting de
- Creation of new drug era tables
- removal of duplicate field name
- removal of duplicate field name
- Addition of new seed file
- Merge branch 'qa/main' of https://github
- Update README.md
- Addition of drug route mappings
- Adding drugs seed to yml file
- formatting
- New workflow for DRUG_EXPOSUR
- Workarounds for drug exposure
- PROCEDURE_OCCURRENCE
- PROCEDURE_OCCURRENCE
- Removal of old proc_occurrence que
- Addition of full seed files

Git + GitHub
Version Control
CI/CD

George Hripcak
Craig Sachson
Chandrabalan Vishnu (LTH...
Martijn Schuemie
Paul Nagy Frank Def...
Erica Voss Katy Sado...
Patrick R... +130

DevCon 2024 Review



Subscribed

0 0 Share Download Clip

All From OHDSI Computers and information t

[May 7th Community Call – DevCon 2024 Review](https://www.youtube.com/watch?v=Ki5fZjCboS0)
<https://www.youtube.com/watch?v=Ki5fZjCboS0>



Perseus ETL Tool

April Olympians #1/Perseus ETL Tool Demo

 OHDSI
2.15K subscribers

 Subscribed

 0



 Share

 Download

 Clip



All

From OHDSI

Presentations

Related

[April 3rd Community Call – April Olympians #1/Perseus ETL Tool Demo](https://www.youtube.com/watch?v=D70KYxuL5p0)

<https://www.youtube.com/watch?v=D70KYxuL5p0>



Extract Load Transform (ETL) journey

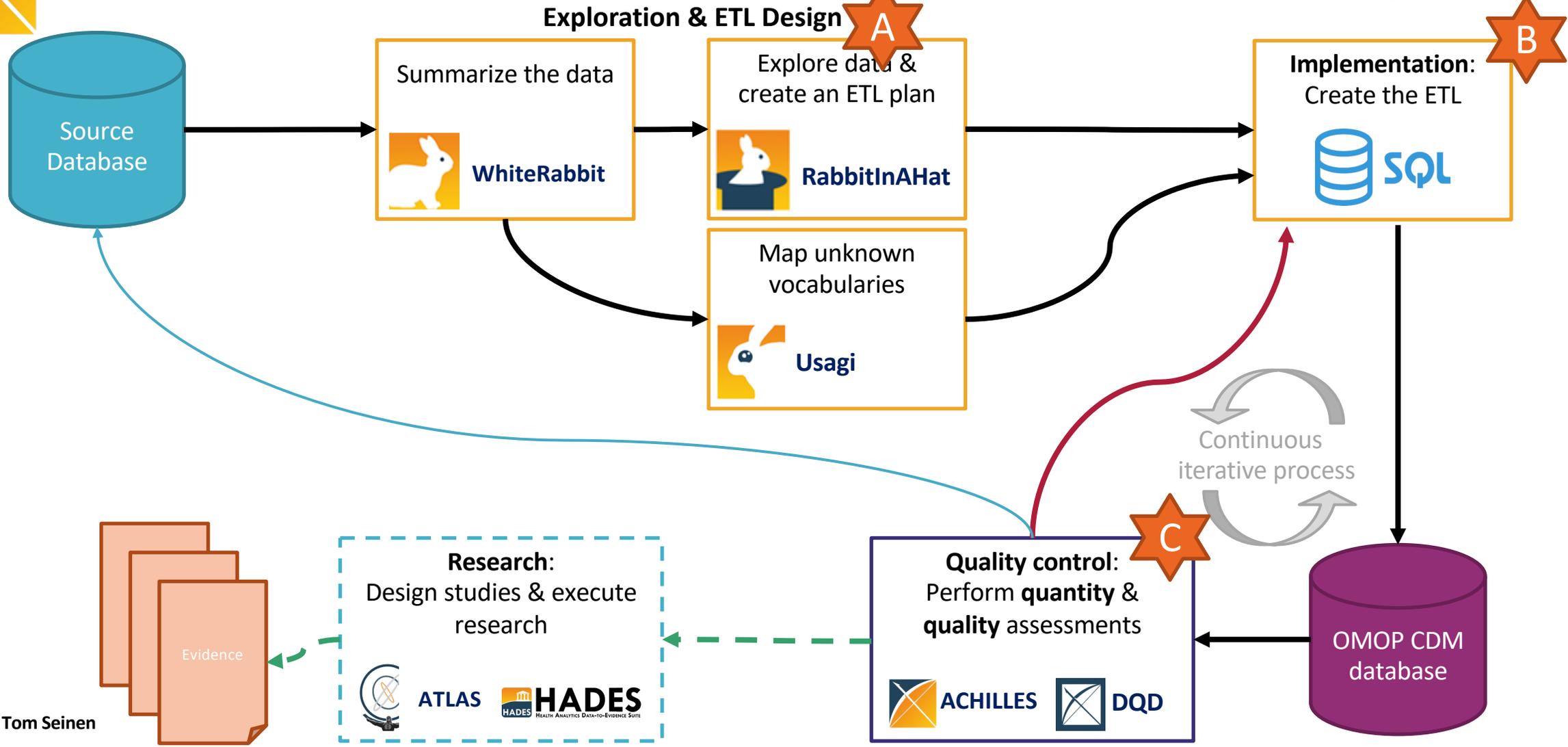


Image credit: Tom Seinen (ErasmusMC)



C – OHDSI Quality Control tools

Standardised data quality checks



DataQualityDashboard Version: 1.0.0
Results generated at 2021-06-29 09:58:20 in 1 hours

	Verification				Validation				Total			
	Pass	Fail	Total	% Pass	Pass	Fail	Total	% Pass	Pass	Fail	Total	% Pass
Plausibility	1999	14	2013	99%	286	1	287	100%	2285	15	2300	99%
Conformance	676	5	681	99%	104	0	104	100%	780	5	785	99%
Completeness	379	7	386	98%	8	7	15	53%	387	14	401	97%
Total	3054	26	3080	99%	398	8	406	98%	3452	34	3486	99%

Standardised database characterisation



OHDSI_Sample_Database
Dashboard





Data Quality Dashboard (DQD)



	Verification				Validation				Total			
	Pass	Fail	Total	% Pass	Pass	Fail	Total	% Pass	Pass	Fail	Total	% Pass
Plausibility	1999	14	2013	99%	286	1	287	100%	2285	15	2300	99%
Conformance	676	5	681	99%	104	0	104	100%	780	5	785	99%
Completeness	379	7	386	98%	8	7	15	53%	387	14	401	97%
Total	3054	26	3080	99%	398	8	406	98%	3452	34	3486	99%

- 3000+ checks
- Plausibility
 - temporal: e.g. dates not in the future, atemporal: e.g. lab value within expected extremes
- Conformance
 - adhering to OMOP CDM standards, e.g. person id exists, standard concepts
- Completeness
 - missing data, e.g. missing gender values, empty measurement table



DataQualityDashboard



ST	Name:	plausibleGender	IDS
Fr	Description:	For a CONCEPT_ID 193437 (NEOPLASM OF UNCERTAIN BEHAVIOR OF FEMALE GENITAL ORGAN), the number and percent of records associated with patients with an implausible gender (correct gender = FEMALE). (Threshold=5%).	
	Notes:		1.80%
+	Level:	CONCEPT	
	Table:	CONDITION_OCCURRENCE	1.31%
+	Field:	CONDITION_CONCEPT_ID	
	Concept Id:	193437	1.31%
+	Unit Concept Id:		
	Check Id:	concept_plausiblegender_condition_occurrence_condition_concept_id_193437	
+	# Rows Violated:	2	1.33%
	% Rows Violated:	33.33%	
+	# Denominator Rows:	6	1.56%
	Execution Time:	0.038004 secs	



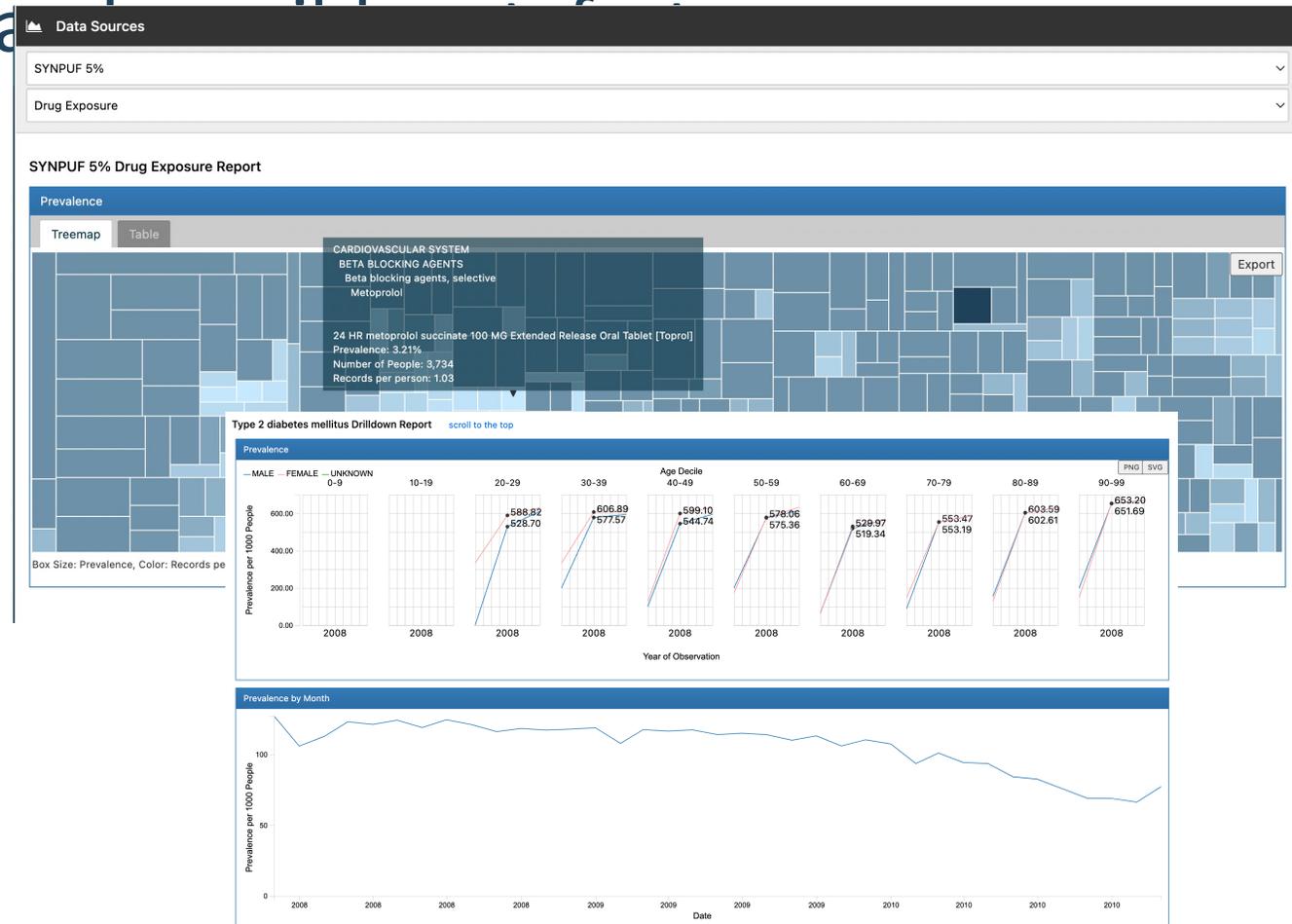
Future developments

- Linking conventions to checks
- New visualisation
- New checks
 - StartBeforeEnd
 - AfterBirth
 - BeforeDeath
 - GenderPlausibility



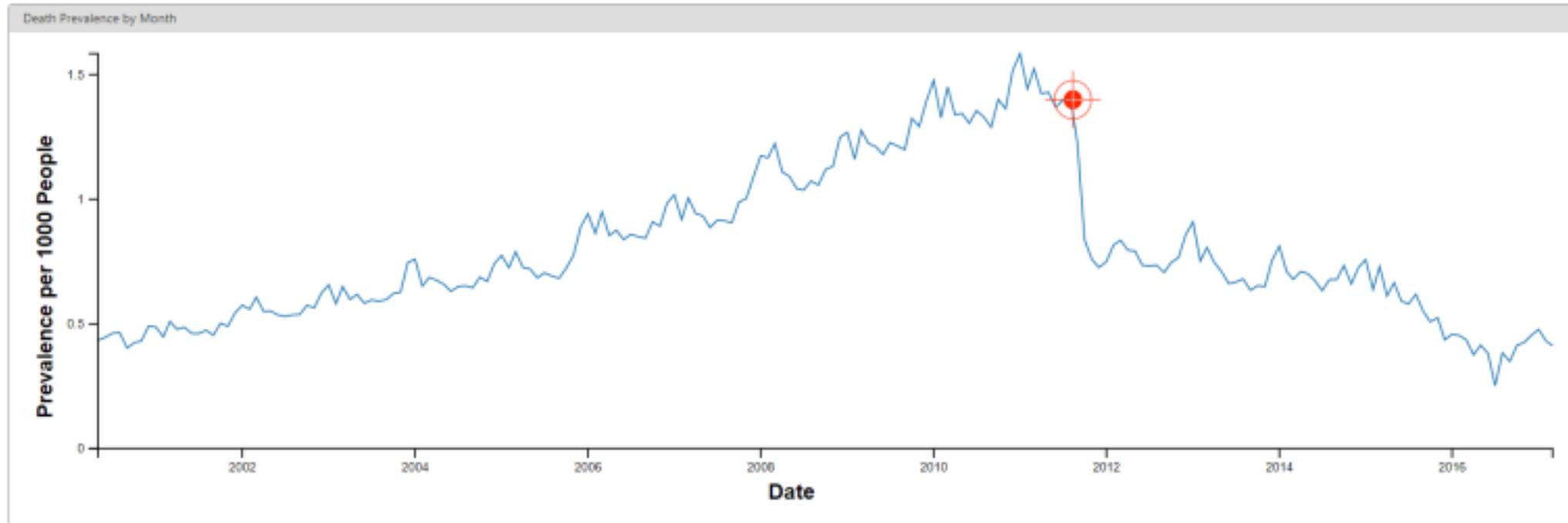
Achilles

- Broad database aggregation and characterization
- Visualize interesting trends and patterns





Example 1 – Death record incidence



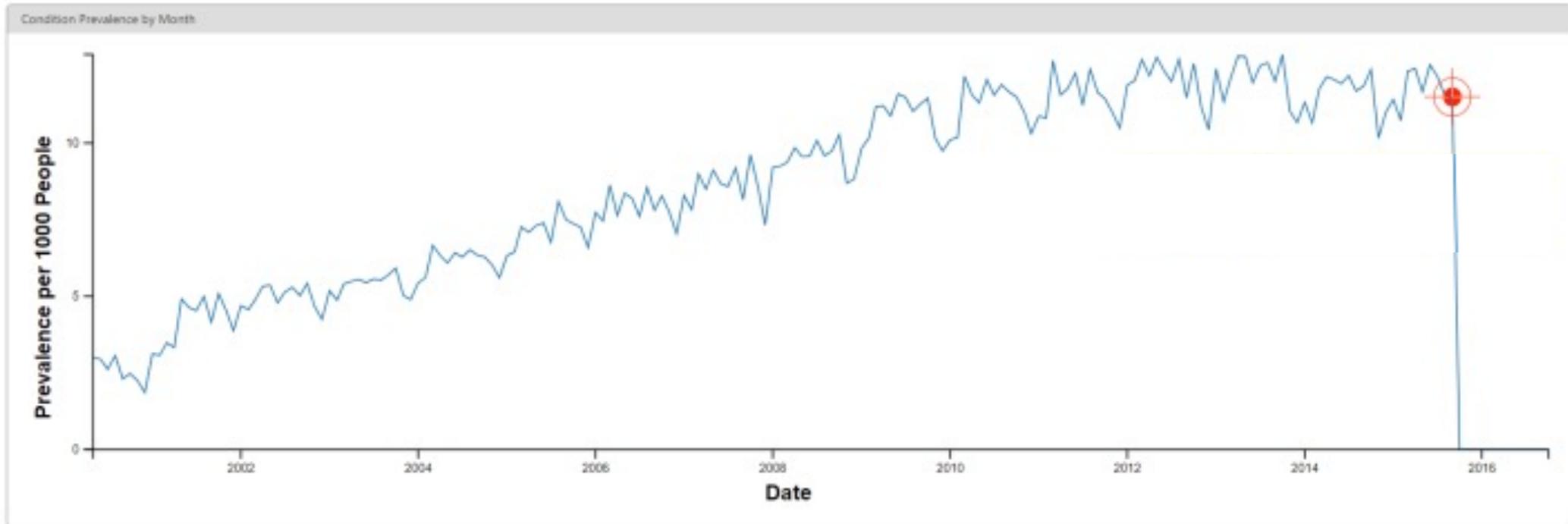
The Impact of Data Quality Annotations on Observational Data Research, Ajit A. Londhe, MPH, Vojtech Huser, MD, PhD, Erica A. Voss, MPH, OHDSI Symposium 2017

https://www.ohdsi.org/web/wiki/lib/exe/fetch.php?media=symposium_2017:ajitlondhe_annotationsposter_final.pdf



Example 2 – Prevalence of Condition

SNOMED:271795006 (Malaise and fatigue)



The Impact of Data Quality Annotations on Observational Data Research, Ajit A. Londhe, MPH, Vojtech Huser, MD, PhD, Erica A.

Voss, MPH, OHDSI Symposium 2017

https://www.ohdsi.org/web/wiki/lib/exe/fetch.php?media=symposium_2017:ajitlondhe_annotationsposter_final.pdf



Extract Load Transform (ETL) journey

Exploration & ETL Design

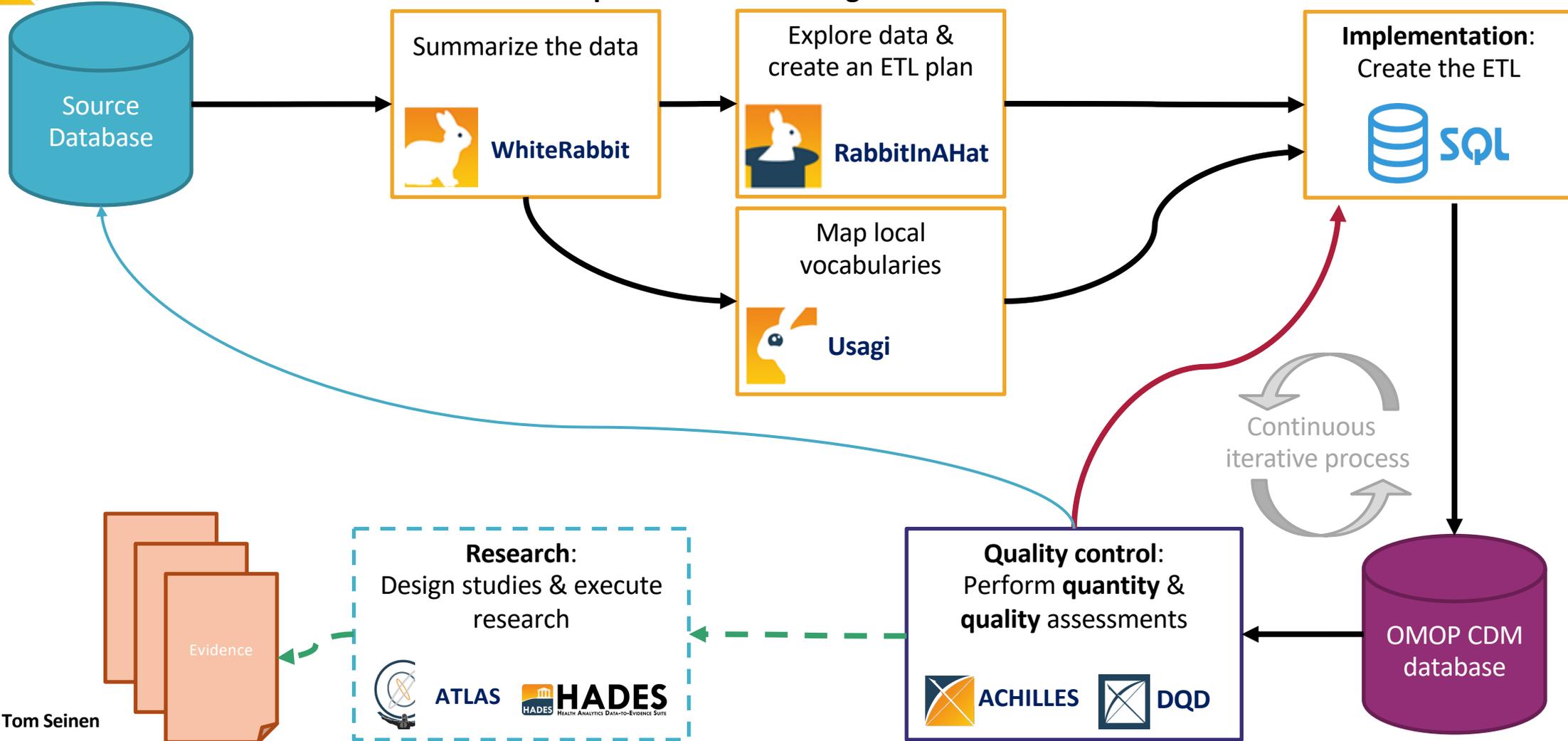


Image credit: Tom Seinen (ErasmusMC)



OHDSI
OBSERVATIONAL HEALTH DATA SCIENCES AND INFORMATICS

Conventions Quiz



The latest release of the OMOP CDM is v5.4

True or False?

Where can I find this convention?

- CDM Documentation
- Themis convention library
- Forums / Github

Checked by which DQ tool?

- Achilles
- DQD
- Other
- Not checked



The latest release of the OMOP CDM is v5.4

True or False?

Where can I find this convention?

- **CDM Documentation**
- Themis convention library
- Forums / Github

Checked by which DQ tool?

- Achilles
- **DQD**
- Other
- Not checked



If no *year_of_birth* is available and it cannot be approximated, the person should be dropped from the CDM.

True or False?

Where can I find this convention?

- CDM Documentation
- Themis convention library
- Forums / Github

Checked by which DQ tool?

- Achilles
- DQD
- Other
- Not checked



If no *year_of_birth* is available and it cannot be approximated, the person should be dropped from the CDM.

True or False?

Where can I find this convention?

- CDM Documentation
- **Themis convention library**
- Forums / Github

Checked by which DQ tool?

- Achilles
- **DQD**
- Other
- Not checked



Technically only two tables must be populated in the OMOP CDM

True or False?

Where can I find this convention?

- CDM Documentation
- Themis convention library
- Forums / Github

Checked by which DQ tool?

- Achilles
- DQD
- Other
- Not checked



Technically only two tables must be populated in the OMOP CDM

True or **False**?

Where can I find this convention?

- **CDM Documentation**
- Themis convention library
- Forums / Github

Checked by which DQ tool?

- Achilles
- DQD
- Other
- Not checked



The Observation Period for an individual person can overlap in time

True or False?

Where can I find this convention?

- CDM Documentation
- Themis convention library
- Forums / Github

Checked by which DQ tool?

- Achilles
- DQD
- Other
- Not checked



The Observation Period for an individual person can overlap in time

True or **False**?

Where can I find this convention?

- **CDM Documentation**
- Themis convention library
- Forums / Github

Checked by which DQ tool?

- Achilles
- DQD
- Other
- Not checked



The *value_as_number* in the Measurement table cannot be a negative value

True or False?

Where can I find this convention?

- CDM Documentation
- Themis convention library
- Forums / Github

Checked by which DQ tool?

- Achilles
- DQD
- Other
- Not checked



The *value_as_number* in the Measurement table cannot be a negative value

True or **False**?

Where can I find this convention?

- CDM Documentation
- **Themis convention library**
- Forums / Github

Checked by which DQ tool?

- Achilles
- DQD
- Other
- Not checked



Condition_Occurrence.condition_start_date is a mandatory field

True or False?

Where can I find this convention?

- CDM Documentation
- Themis convention library
- Forums / Github

Checked by which DQ tool?

- Achilles
- DQD
- Other
- Not checked



Condition_Occurrence.condition_start_date is a mandatory field

True or False?

Where can I find this convention?

- **CDM Documentation**
- Themis convention library
- Forums / Github

Checked by which DQ tool?

- Achilles
- **DQD**
- Other
- Not checked



Condition_Occurrence.condition_end_date is a mandatory field

True or False?

Where can I find this convention?

- CDM Documentation
- Themis convention library
- Forums / Github

Checked by which DQ tool?

- Achilles
- DQD
- Other
- Not checked



Condition_Occurrence.condition_end_date is a mandatory field

True or **False**?

Where can I find this convention?

- **CDM Documentation**
- Themis convention library
- Forums / Github

Checked by which DQ tool?

- Achilles
- DQD
- Other
- Not checked



A person can have more than one date of death

True or False?

Where can I find this convention?

- CDM Documentation
- Themis convention library
- Forums / Github

Checked by which DQ tool?

- Achilles
- DQD
- Other
- Not checked



A person can have more than one date of death

True or **False?**

Where can I find this convention?

- CDM Documentation
- **Themis convention library**
- Forums / Github

Checked by which DQ tool?

- Achilles
- DQD
- Other
- **Not checked**



It is ok to leave `visit_end_datetime` null when `visit_start_datetime` is populated

True or False?

Where can I find this convention?

- CDM Documentation
- Themis convention library
- Forums / Github

Checked by which DQ tool?

- Achilles
- DQD
- Other
- Not checked



It is ok to leave `visit_end_datetime` null when `visit_start_datetime` is populated

True or False?

Where can I find this convention?

- CDM Documentation
- **Themis convention library**
- Forums / Github

Checked by which DQ tool?

- Achilles
- DQD
- Other
- **Not checked**



When populating your CDM from more than one source, you create one person record from each source

True or False?

Where can I find this convention?

- CDM Documentation
- Themis convention library
- Forums / Github

Checked by which DQ tool?

- Achilles
- DQD
- Other
- Not checked



When populating your CDM from more than one source, you create one person record from each source

True or **False?**

Where can I find this convention?

- **CDM Documentation**
- Themis convention library
- Forums / Github

Checked by which DQ tool?

- Achilles
- DQD
- Other
- **Not checked**



Concepts that are not invalid are assigned valid_end_date of '31-Dec-2099'

True or False?

Where can I find this convention?

- CDM Documentation
- Themis convention library
- Forums / Github

Checked by which DQ tool?

- Achilles
- DQD
- Other
- Not checked



Concepts that are not invalid are assigned valid_end_date of '31-Dec-2099'

True or False?

Where can I find this convention?

- **CDM Documentation**
- Themis convention library
- Forums / Github

Checked by which DQ tool?

- Achilles
- DQD
- Other
- **Not checked**



Drug Exposures should always contain strength and dose form information

True or False?

Where can I find this convention?

- CDM Documentation
- Themis convention library
- Forums / Github

Checked by which DQ tool?

- Achilles
- DQD
- Other
- Not checked



Drug Exposures should always contain strength and dose form information

True or **False**?

Where can I find this convention?

- **CDM Documentation**
- Themis convention library
- Forums / Github

Checked by which DQ tool?

- Achilles
- DQD
- Other
- Not checked



I can use the same type concept id across all tables in the CDM

True or False?

Where can I find this convention?

- CDM Documentation
- Themis convention library
- Forums / Github

Checked by which DQ tool?

- Achilles
- DQD
- Other
- Not checked



I can use the same type concept id across all tables in the CDM

True or False?

Where can I find this convention?

- CDM Documentation
- Themis convention library
- Forums / Github

Checked by which DQ tool?

- Achilles
- **DQD**
- Other
- Not checked



You have to use WhiteRabbit and Usagi in designing your ETL

True or False?

Where can I find this convention?

- CDM Documentation
- Themis convention library
- Forums / Github

Checked by which DQ tool?

- Achilles
- DQD
- Other
- Not checked



You have to use WhiteRabbit and Usagi in designing your ETL

True or **False**?

Where can I find this convention?

- CDM Documentation
- Themis convention library
- Forums / Github

Checked by which DQ tool?

- Achilles
- DQD
- Other
- Not checked



OHDSI
OBSERVATIONAL HEALTH DATA SCIENCES AND INFORMATICS

Thank You!