OMOP-CDM

The Observational Medical Outcomes Partnership Common Data Model

Ronald Cornet

Amsterdam UMC

HemaFAIR

doi:10.5281/zenodo.14846882

Funded by the European Union

About me

- Full Professor of Medical Informatics at Amsterdam UMC
- Principal Educator on FAIR Data
- Principal Investigator on Reusable Health Data

ERDERA
European Rare Diseases Research Alliance

HELIOS

HemaFAIR

LEAPfROG

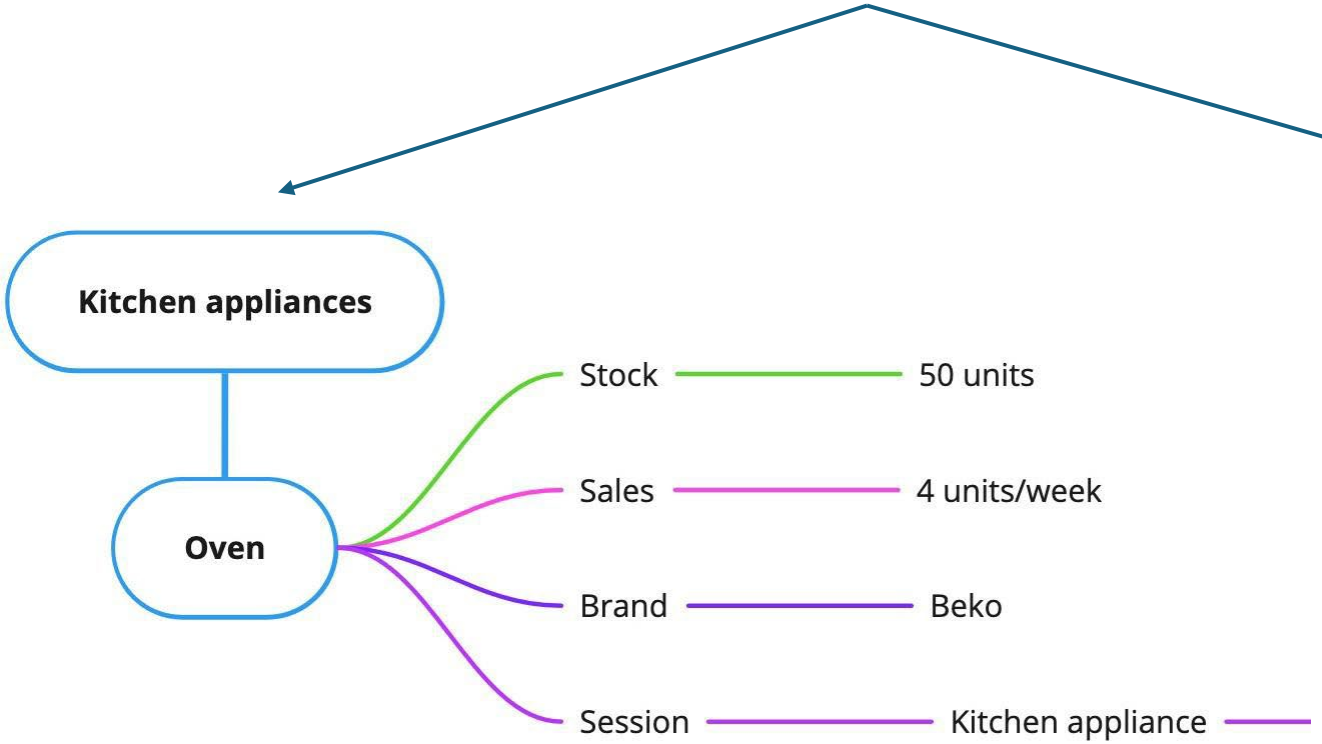
PaLaDIn
Patient Lifestyle & Disease Data Interactium

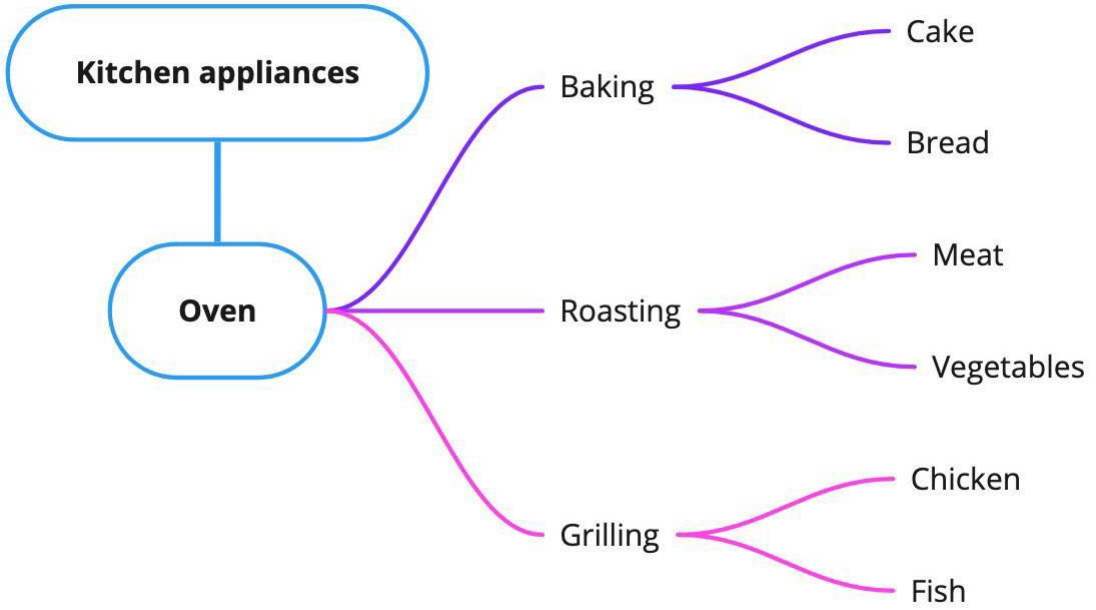
LEARNFAIR

HemaFAIR

Funded by the European Union

Models are context sensitive





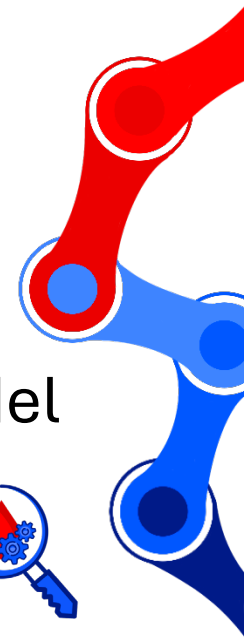
Credits: César Bernabé

HemaFAIR

Funded by the European Union

Information models – examples

Model	Context
• Set of common data elements for Rare Diseases Registration	• Rare Diseases Registration
• OMOP-CDM	• Real-world data
• OpenEHR	• Clinical data storage
• Phenopackets	• Bioinformation
• CDISC	• Clinical trials
• HL7 FHIR ← March 11	• Clinical data exchange
• CARE-SM	• Semantic information model



Funded by the European Union

HemaFAIR

The good thing about standards

- … *is that there are so many to choose from*

  Andrew Tanenbaum

Source: https://commons.wikimedia.org/wiki/File:Toothbrushes_Sizes.jpg
CC-BY-SA – Author: JolieADavis

HemaFAIR

Funded by the European Union

“Any standard is better than no standard”

HemaFAIR

Funded by the European Union

Mapping …

From OMOP to CDISC SDTM: Successes, Challenges, and Future Opportunities of using EHR Data for Drug Repurposing in COVID-19

Wesley Anderson[1], Ruth Kurtycz, Tahsin Farid[2], Shermarke Hassan[3], Kalynn Kennon[3], Pam Dasher[1], Danielle Boyce[4], Will Roddy[1], Smith F. Heavner[1,5]
[1]CURE Drug Repurposing Collaboratory, Critical Path Institute, [2]U.S. Food and Drug Administration, [3]Infectious Disease Data Observatory, [4]Johns Hopkins University, [5]Department of Public Health Sciences, Clemson University

Cropped from: https://www.ohdsi.org/wp-content/uploads/2023/10/3-AndersonBriefreport-Wes-Anderson.pdf

HL7® FHIR® DevDays

Mapping FHIR to OMOP Using Open Source Tools

Vivian Neilley, Sebastiaan van Sandijk

Cropped from: https://www.youtube.com/watch?v=-SBs8fK-Iv4

openFHIR bridging openEHR and HL7® FHIR®

Cropped from: https://www.youtube.com/watch?v=lx_SrbCdg_o

Home › HL7 International and OHDSI Announce Collaboration to Provide Single Common Data Model for Sharing Information in Clinical Care and Observational Research

HL7 International and OHDSI Announce Collaboration to Provide Single Common Data Model for Sharing Information in Clinical Care and Observational Research

Cropped from: https://www.ohdsi.org/ohdsi-hl7-collaboration/

CDISC and HL7 Jointly Release Mapping Guide to Facilitate the Use of Electronic Health Record Data in Clinical Research

Cropped from: https://www.cdisc.org/news/cdisc-and-hl7-jointly-release-mapping-guide-facilitate-use-electronic-health-record-data

Mapping …

Mapping OHDSI OMOP Common Data Model and GA4GH Phenopackets for COVID-19 disease epidemics and analytics

Núria Queralt-Rosinach[1], Pablo Alarcón[2], Tiffany Callahan[3], Giovanni Delussu[4], Charlotte Fraboulet[5], Romain Goussault[5], Jules Jacobsen[6], Leyla Jael Castro[7], Rajaram Kaliyaperumal[1], Maxat Kulmanov[8], Peter Robinson[9], Venkata Satagopam[10], Anastasios Siapos[11], Vasundra Touré[12], and Danielle Welter[10]

Cropped from: https://osf.io/preprints/biohackrxiv/ep3xh_v1

Funded by the European Union

HemaFAIR

Today’s pick: OMOP-CDM

- Context: OHDSI
- Aim: Real-world evidence
- Approach
  - Model: OMOP-CDM
  - Tools
- Experiences

HemaFAIR

Funded by
the European Union

Context

OHDSI
OBSERVATIONAL HEALTH DATA SCIENCES AND INFORMATICS

- 4,200 collaborators
- 83 countries
- health records for about 810 million unique patients from around the world

HemaFAIR

Welcome to OHDSI!

The Observational Health Data Sciences and Informatics (or OHDSI, pronounced "Odyssey") program is a multi-stakeholder, interdisciplinary collaborative to bring out the value of health data through large-scale analytics. All our solutions are open-source.

OHDSI has established an international network of researchers and observational health databases with a central coordinating center housed at Columbia University.

Read more about us, about our goals, and how you can help support the OHDSI community.

Join the Journey

Cropped from: https://ohdsi.org/

Aim: from real-world data to evidence

Different types of observational data:

Populations
- Pediatric vs. elderly
- Socioeconomic disparities

Care setting
- Inpatient vs. outpatient
- Primary vs. secondary care

Data capture process
- Administrative claims
- Electronic health records
- Clinical registries

Health system
- Insured vs. uninsured
- Country policies

Types of evidence desired:

Clinical characterization
- Clinical trial feasibility
- Treatment utilization
- Disease natural history
- Quality improvement

Population-level effect estimation
- Safety surveillance
- Comparative effectiveness

Patient-level prediction
- Precision medicine
- Disease interception

HemaFAIR

Source: https://ohdsi.github.io/TheBookOfOhdsi (public domain)

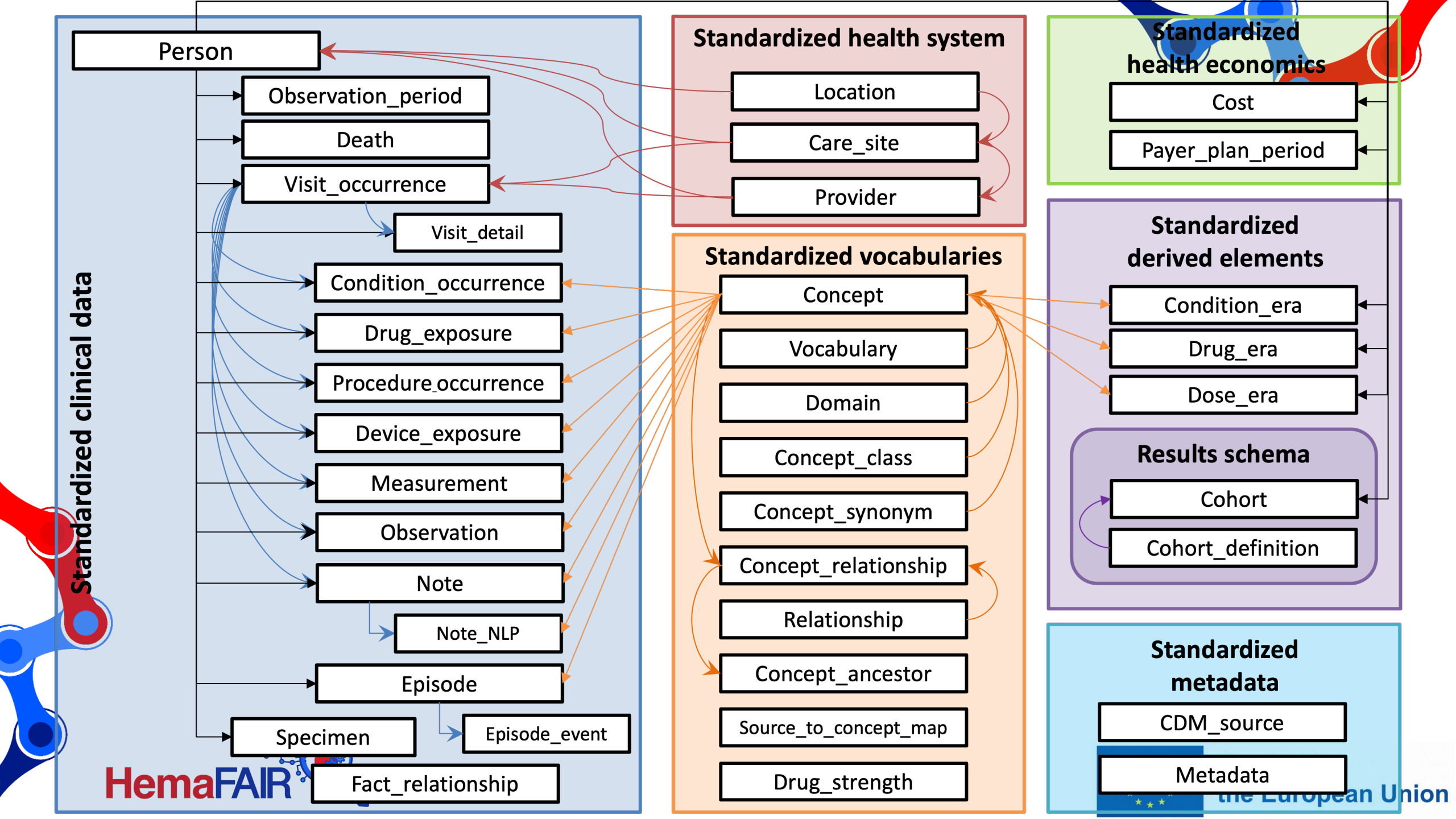
Funded by the European Union

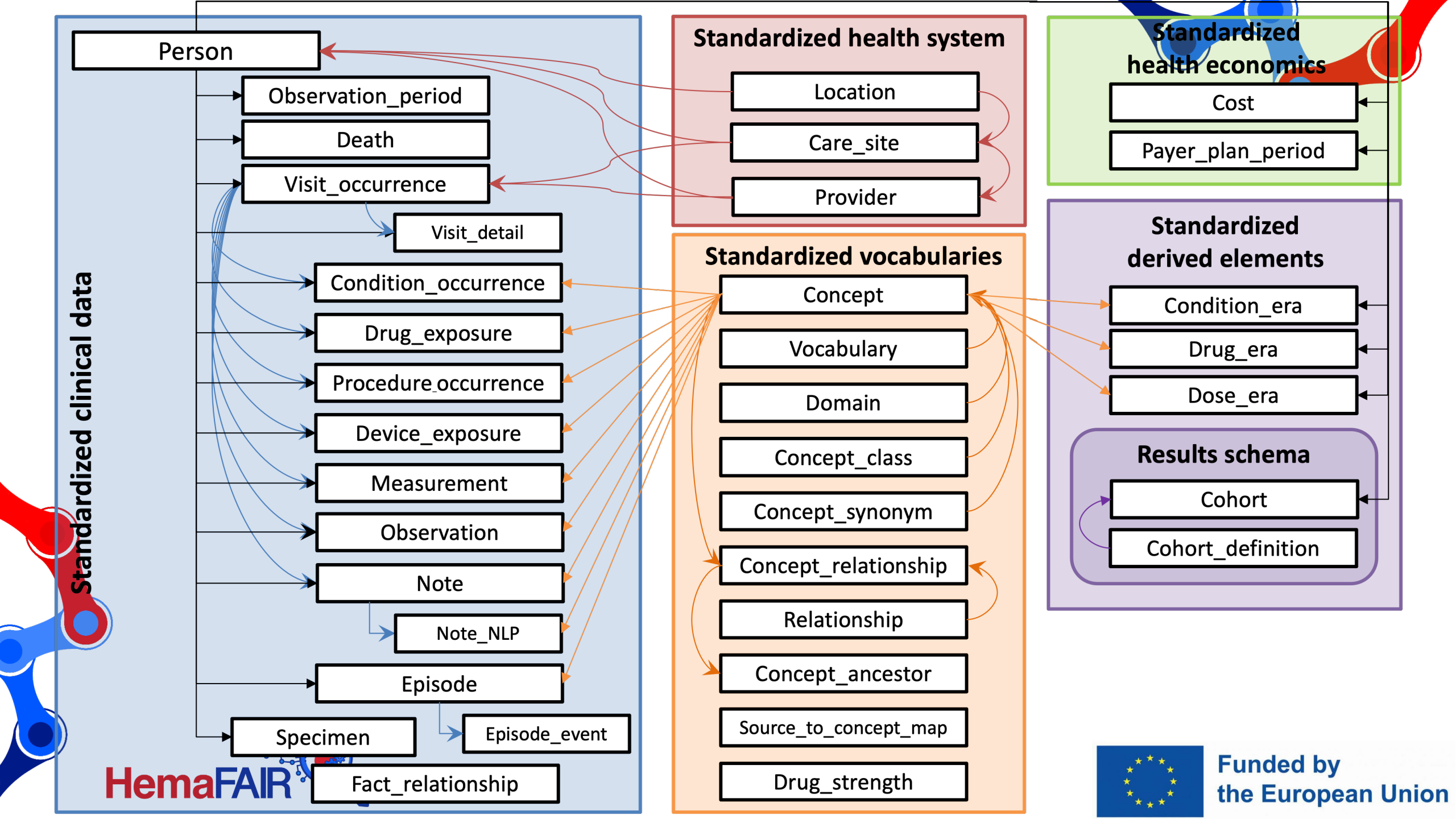
Approach: OMOP-CDM

- OHDSI was previously called OMOP: Observational Medical Outcomes Partnership
- CDM: Common Data Model

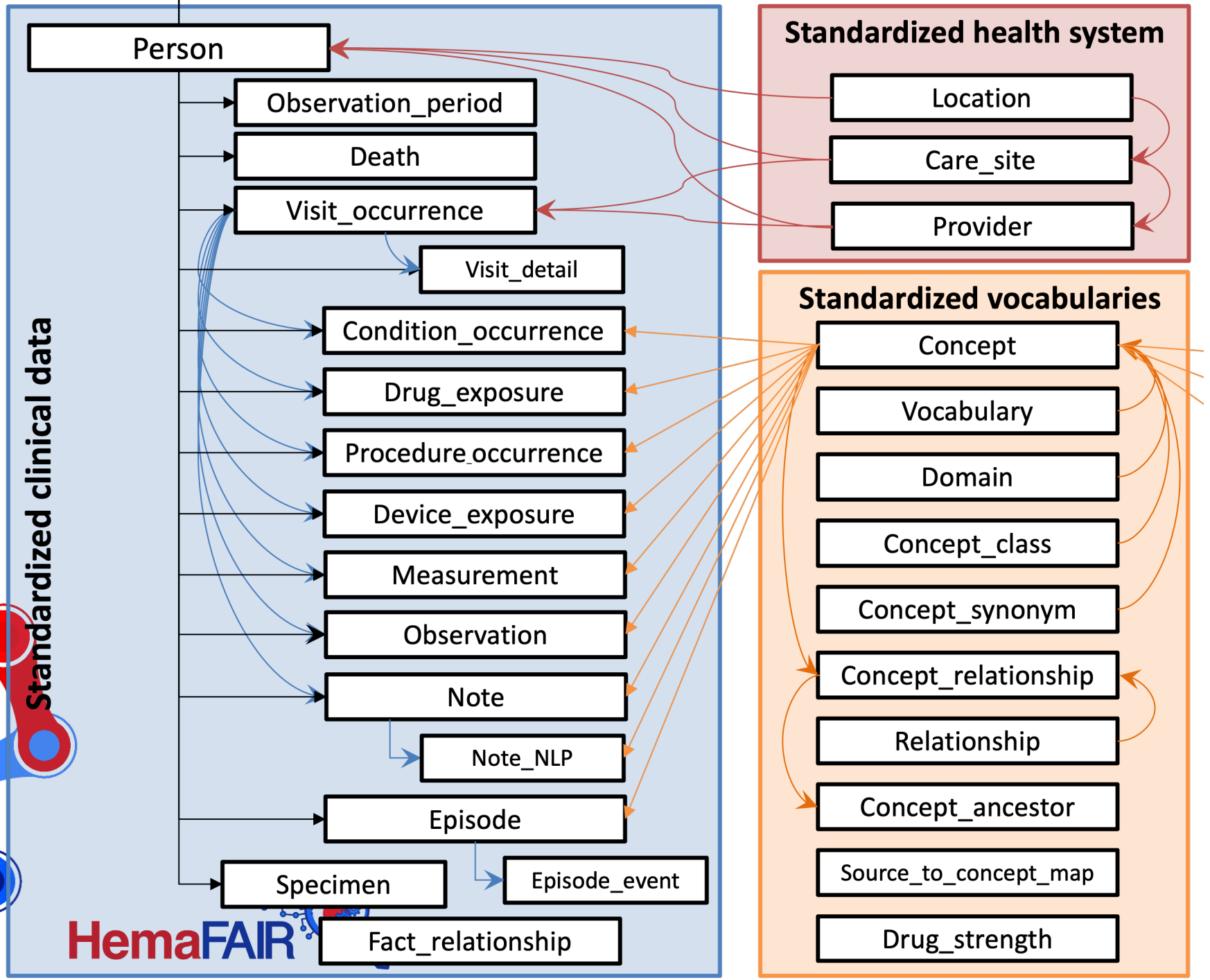
HemaFAIR

Funded by the European Union


Standardized clinical data
Person
Observation_period
Death
Visit_occurrence
Visit_detail
Condition_occurrence
Drug_exposure
Procedure_occurrence
Device_exposure
Measurement
Observation
Note
Note_NLP
Episode
Episode_event
Specimen
Fact_relationship
Standardized health system
Location
Care_site
Provider
Standardized vocabularies
Concept
Vocabulary
Domain
Concept_class
Concept_synonym
Concept_relationship
Relationship
Concept_ancestor
Source_to_concept_map
Drug_strength
Standardized health economics
Cost
Payer_plan_period
Standardized derived elements
Condition_era
Drug_era
Dose_era
Results schema
Cohort
Cohort_definition
Standardized metadata
CDM_source
Metadata
HemaFAIR
the European Union

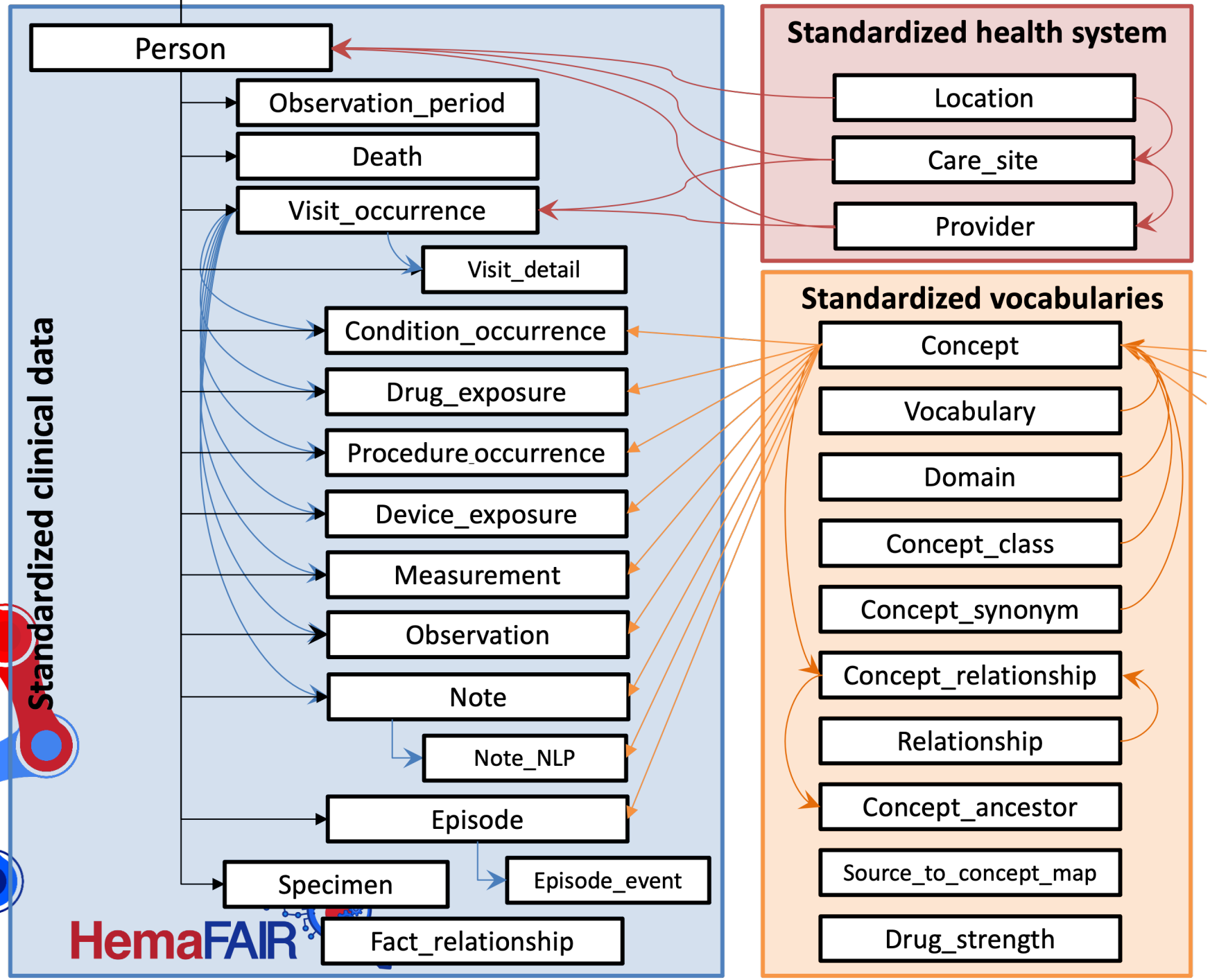


Funded by the European Union



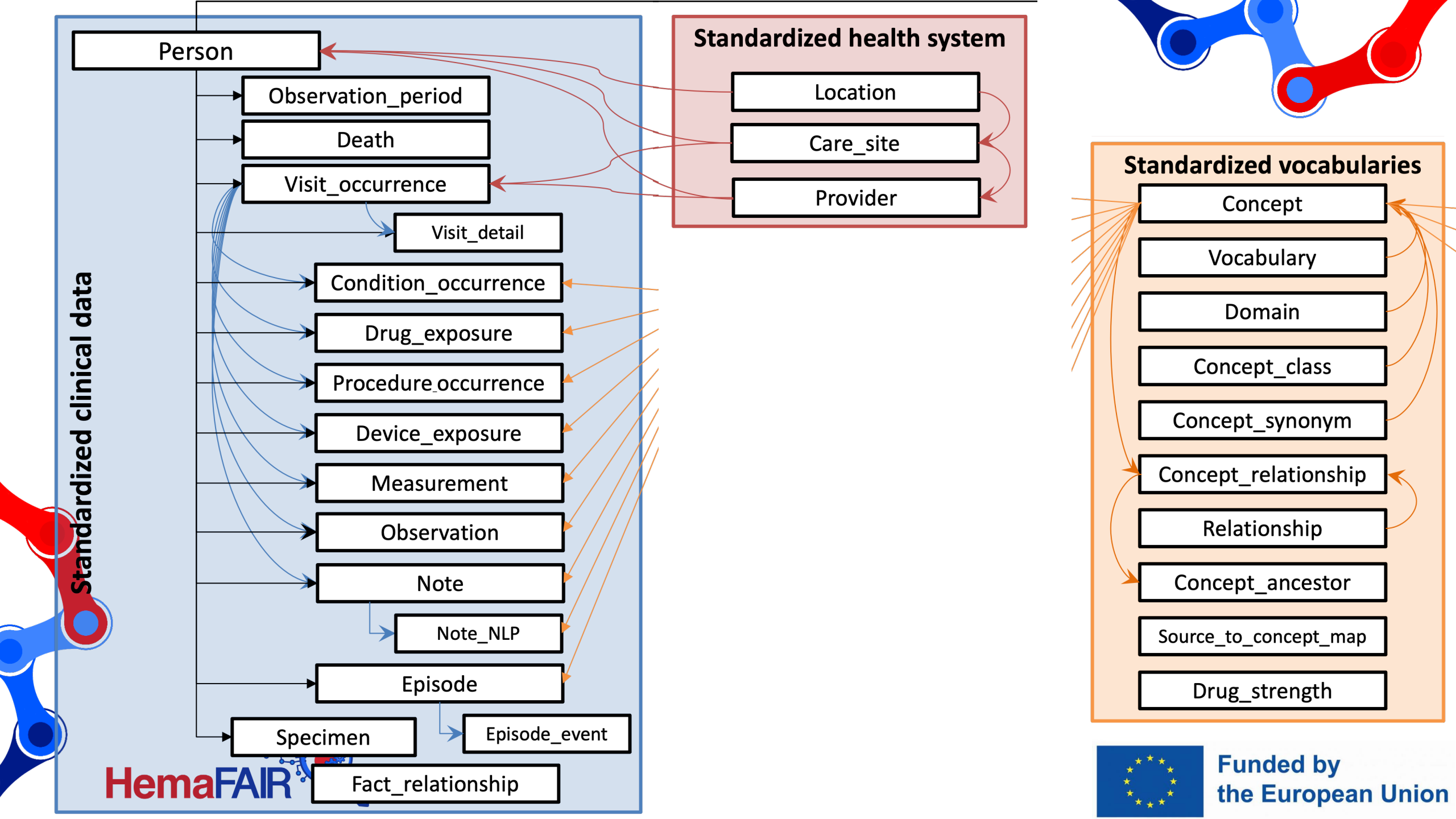
HemaFAIR

Funded by the European Union

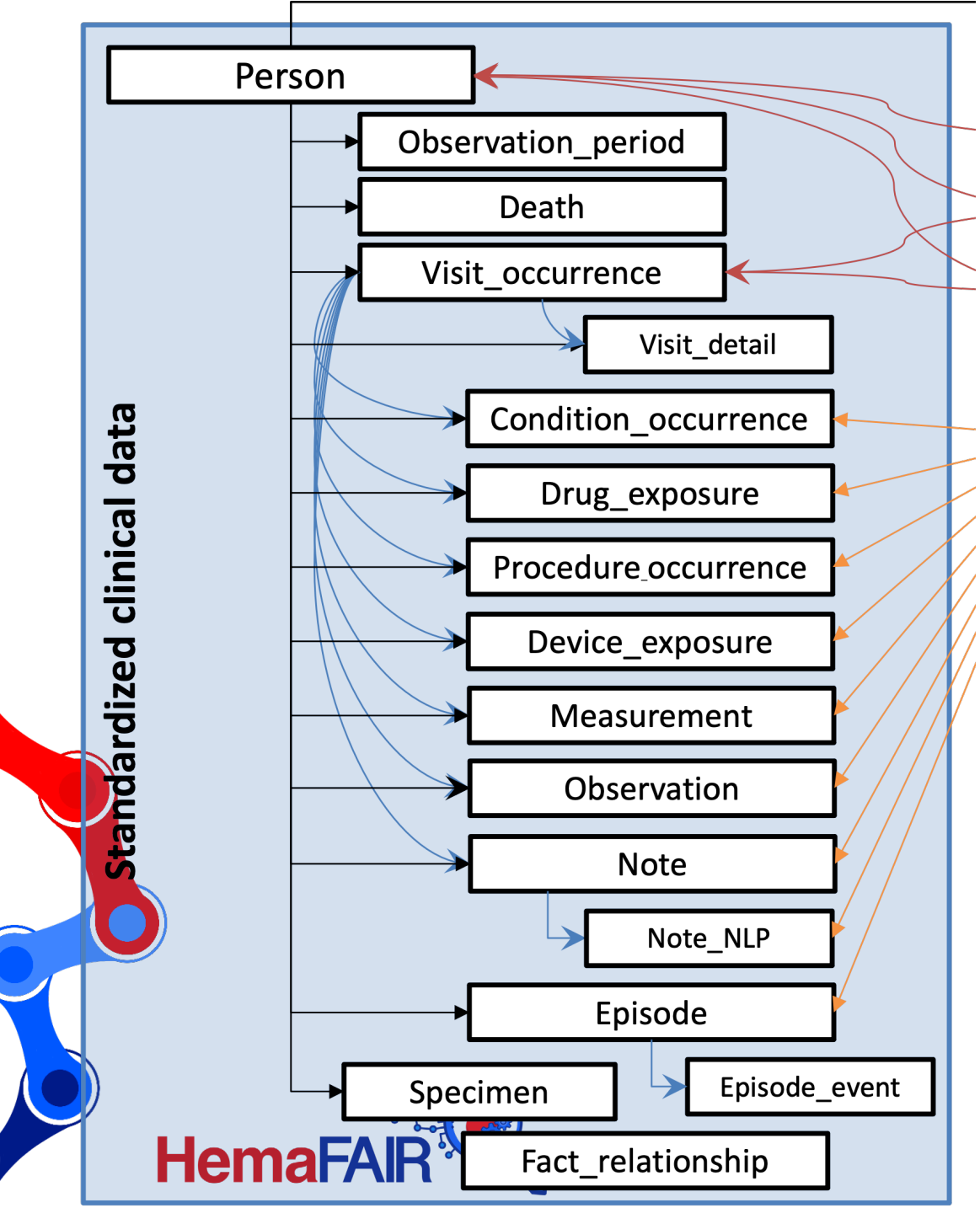


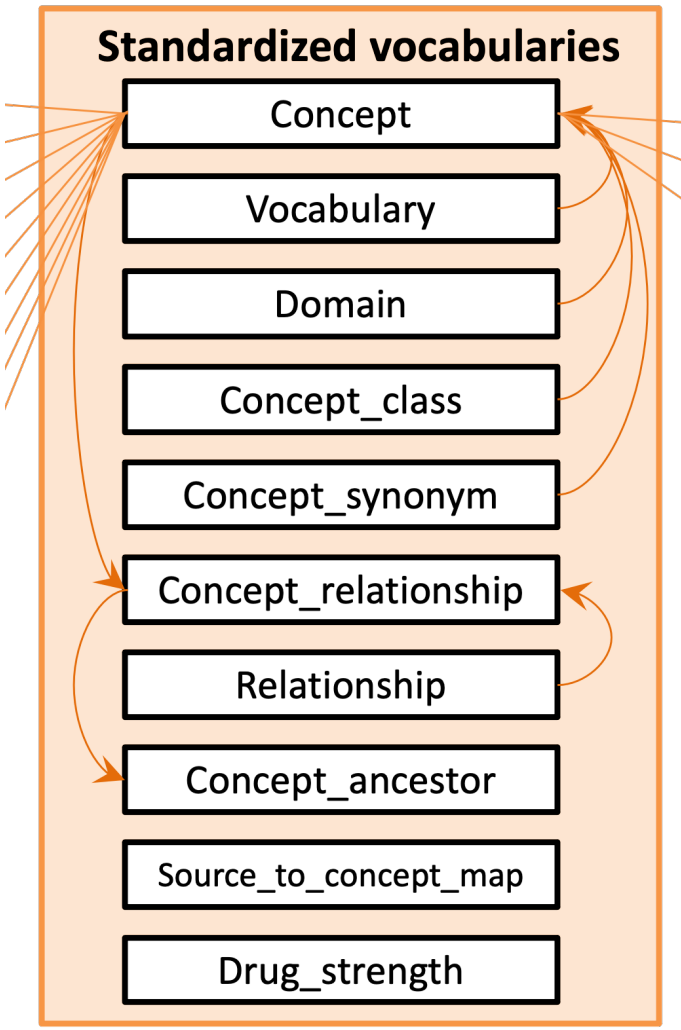
HemaFAIR

Funded by the European Union



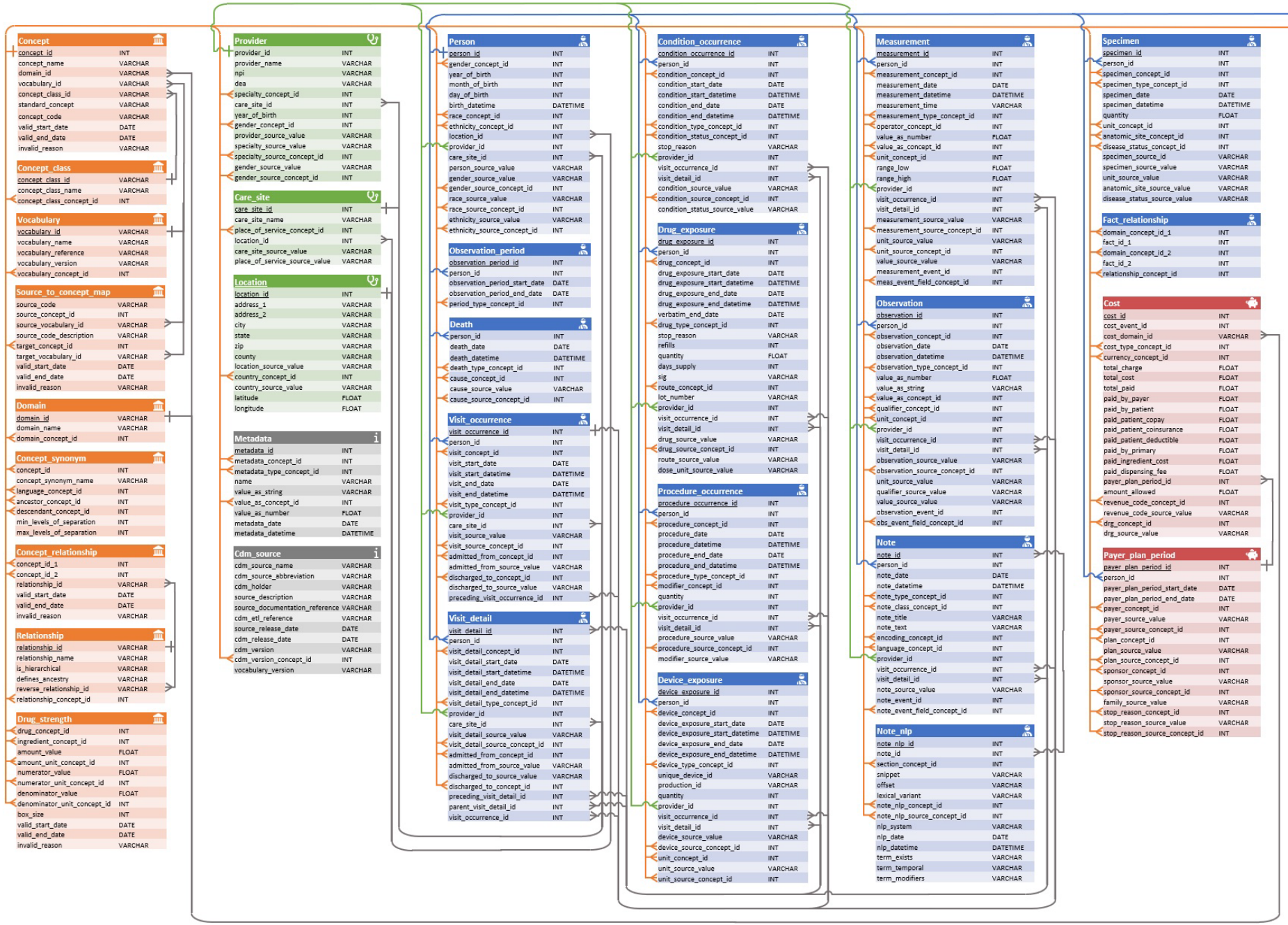
Funded by the European Union

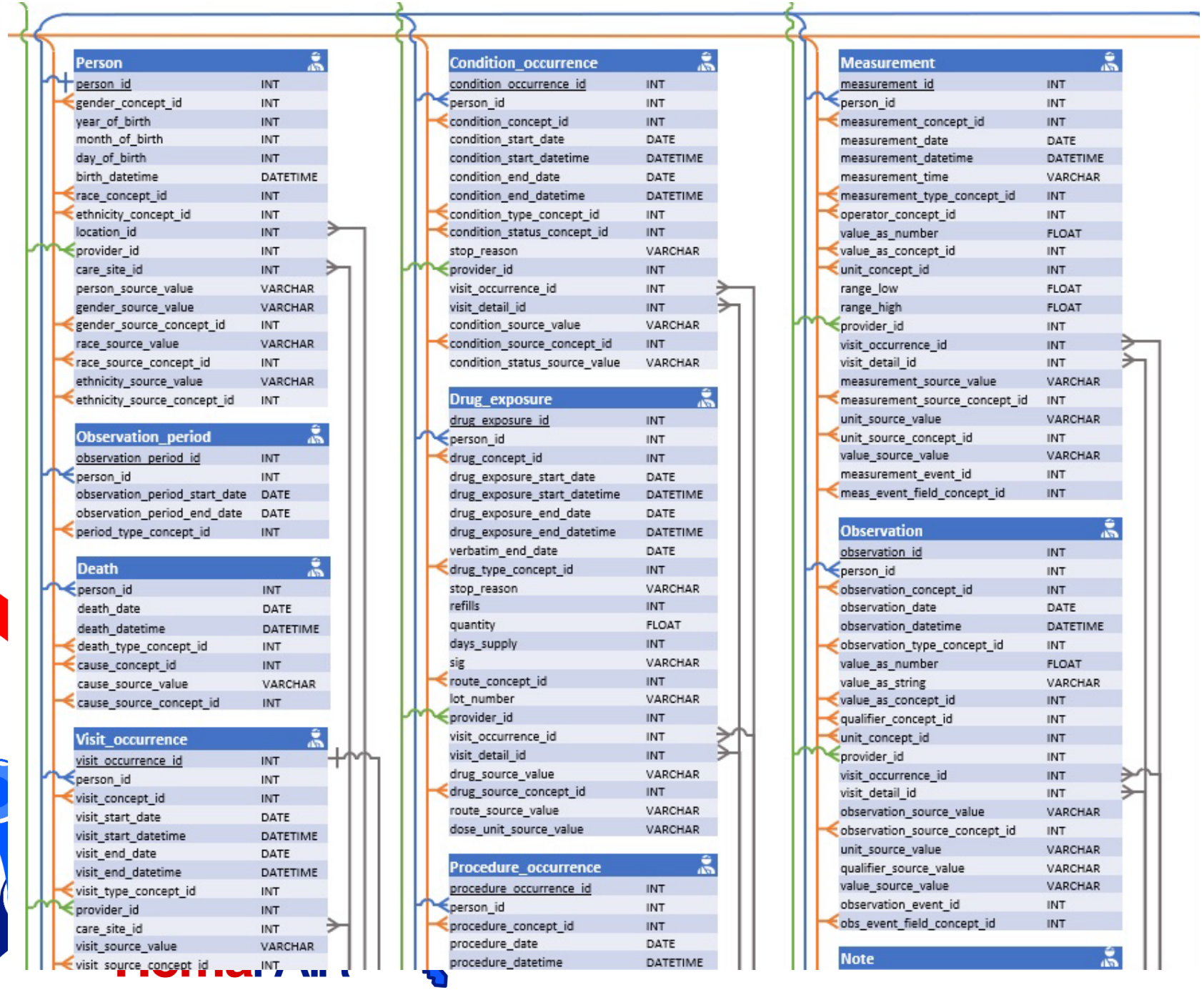




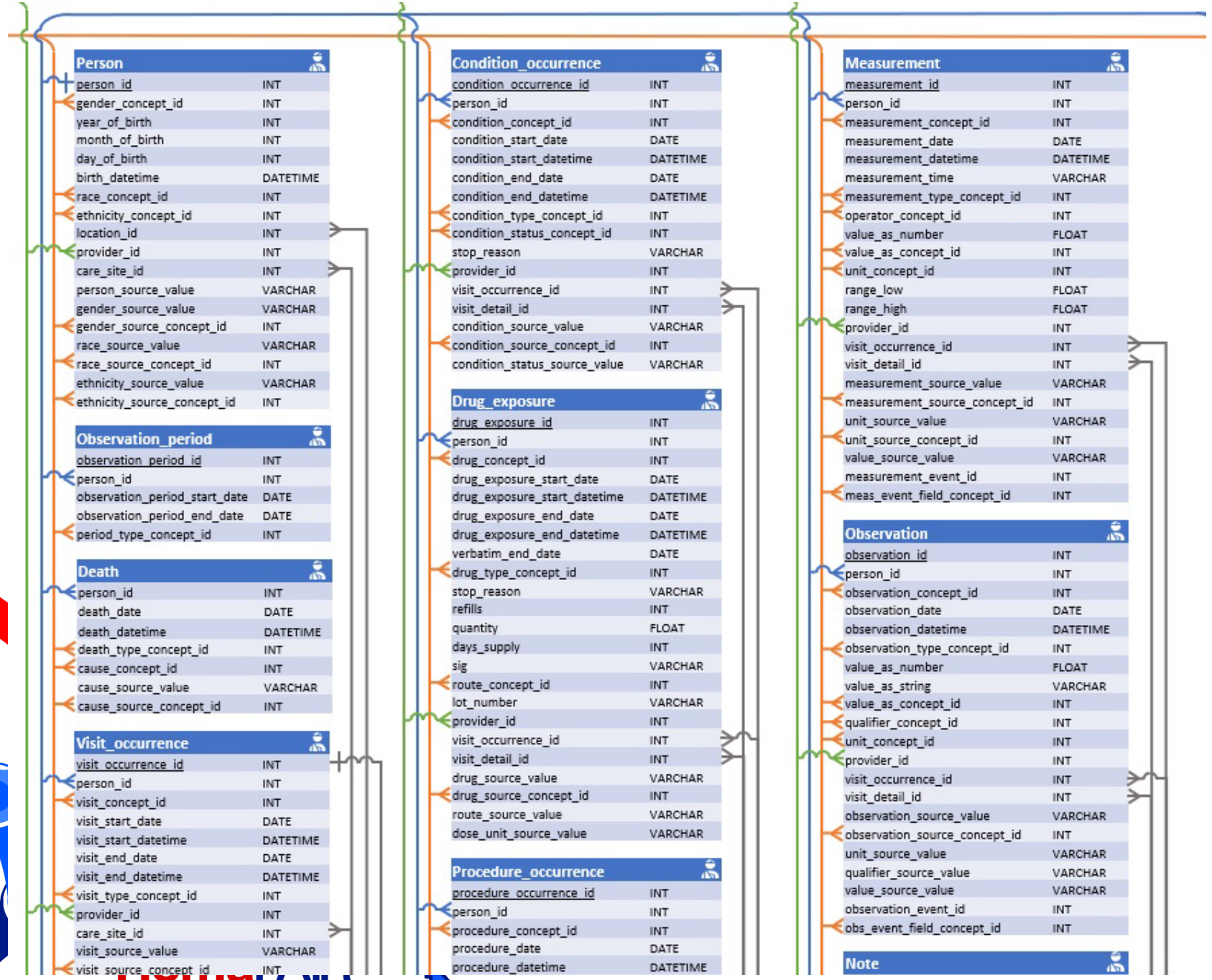
Funded by
the European Union

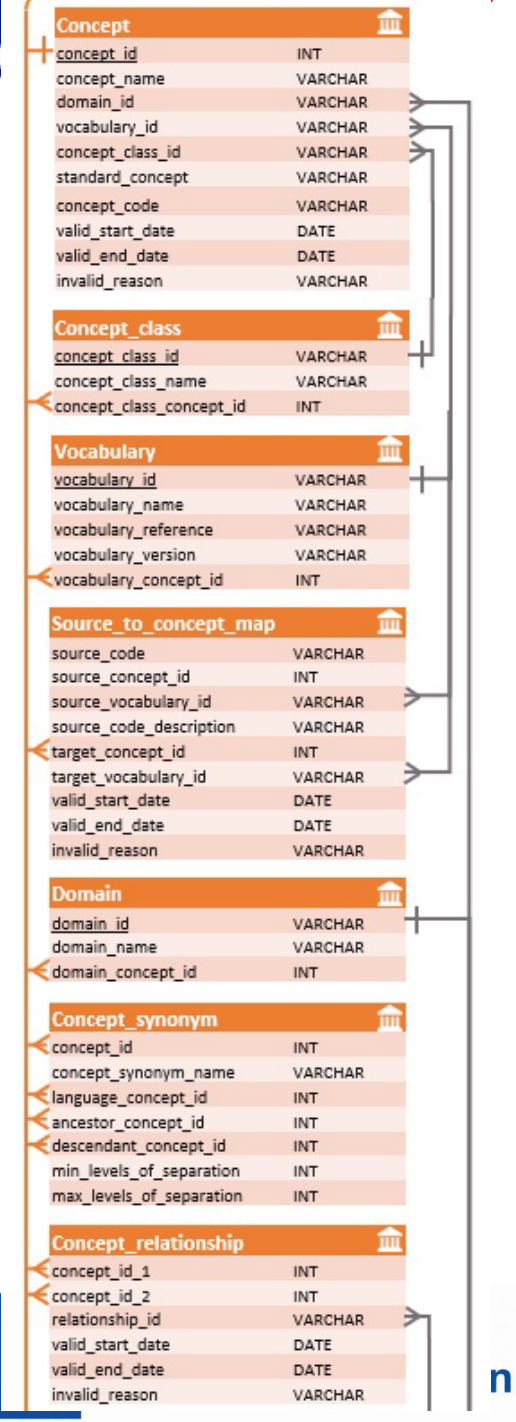
OMOP Common Data Model 5.4





Funded by
the European Union


Person
person_id INT
gender_concept_id INT
year_of_birth INT
month_of_birth INT
day_of_birth INT
birth_datetime DATETIME
race_concept_id INT
ethnicity_concept_id INT
location_id INT
provider_id INT
care_site_id INT
person_source_value VARCHAR
gender_source_value VARCHAR
gender_source_concept_id INT
race_source_value VARCHAR
race_source_concept_id INT
ethnicity_source_value VARCHAR
ethnicity_source_concept_id INT
Observation_period
observation_period_id INT
person_id INT
observation_period_start_date DATE
observation_period_end_date DATE
period_type_concept_id INT
Death
person_id INT
death_date DATE
death_datetime DATETIME
death_type_concept_id INT
cause_concept_id INT
cause_source_value VARCHAR
cause_source_concept_id INT
Visit_occurrence
visit_occurrence_id INT
person_id INT
visit_concept_id INT
visit_start_date DATE
visit_start_datetime DATETIME
visit_end_date DATE
visit_end_datetime DATETIME
visit_type_concept_id INT
provider_id INT
care_site_id INT
visit_source_value VARCHAR
visit_source_concept_id INT
Condition_occurrence
condition_occurrence_id INT
person_id INT
condition_concept_id INT
condition_start_date DATE
condition_start_datetime DATETIME
condition_end_date DATE
condition_end_datetime DATETIME
condition_type_concept_id INT
condition_status_concept_id INT
stop_reason VARCHAR
provider_id INT
visit_occurrence_id INT
visit_detail_id INT
condition_source_value VARCHAR
condition_source_concept_id INT
condition_status_source_value VARCHAR
Drug_exposure
drug_exposure_id INT
person_id INT
drug_concept_id INT
drug_exposure_start_date DATE
drug_exposure_start_datetime DATETIME
drug_exposure_end_date DATE
drug_exposure_end_datetime DATETIME
verbatim_end_date DATE
drug_type_concept_id INT
stop_reason VARCHAR
refills INT
quantity FLOAT
days_supply INT
sig VARCHAR
route_concept_id INT
lot_number VARCHAR
provider_id INT
visit_occurrence_id INT
visit_detail_id INT
drug_source_value VARCHAR
drug_source_concept_id INT
route_source_value VARCHAR
dose_unit_source_value VARCHAR
Procedure_occurrence
procedure_occurrence_id INT
person_id INT
procedure_concept_id INT
procedure_date DATE
procedure_datetime DATETIME
Measurement
measurement_id INT
person_id INT
measurement_concept_id INT
measurement_date DATE
measurement_datetime DATETIME
measurement_time VARCHAR
measurement_type_concept_id INT
operator_concept_id INT
value_as_number FLOAT
value_as_concept_id INT
unit_concept_id INT
range_low FLOAT
range_high FLOAT
provider_id INT
visit_occurrence_id INT
visit_detail_id INT
measurement_source_value VARCHAR
measurement_source_concept_id INT
unit_source_value VARCHAR
unit_source_concept_id INT
value_source_value VARCHAR
measurement_event_id INT
meas_event_field_concept_id INT
Observation
observation_id INT
person_id INT
observation_concept_id INT
observation_date DATE
observation_datetime DATETIME
observation_type_concept_id INT
value_as_number FLOAT
value_as_string VARCHAR
value_as_concept_id INT
qualifier_concept_id INT
unit_concept_id INT
provider_id INT
visit_occurrence_id INT
visit_detail_id INT
observation_source_value VARCHAR
observation_source_concept_id INT
unit_source_value VARCHAR
qualifier_source_value VARCHAR
value_source_value VARCHAR
observation_event_id INT
obs_event_field_concept_id INT
Note


Concept
concept_id INT
concept_name VARCHAR
domain_id VARCHAR
vocabulary_id VARCHAR
concept_class_id VARCHAR
standard_concept VARCHAR
concept_code VARCHAR
valid_start_date DATE
valid_end_date DATE
invalid_reason VARCHAR
Concept_class
concept_class_id VARCHAR
concept_class_name VARCHAR
concept_class_concept_id INT
Vocabulary
vocabulary_id VARCHAR
vocabulary_name VARCHAR
vocabulary_reference VARCHAR
vocabulary_version VARCHAR
vocabulary_concept_id INT
Source_to_concept_map
source_code VARCHAR
source_concept_id INT
source_vocabulary_id VARCHAR
source_code_description VARCHAR
target_concept_id INT
target_vocabulary_id VARCHAR
valid_start_date DATE
valid_end_date DATE
invalid_reason VARCHAR
Domain
domain_id VARCHAR
domain_name VARCHAR
domain_concept_id INT
Concept_synonym
concept_id INT
concept_synonym_name VARCHAR
language_concept_id INT
ancestor_concept_id INT
descendant_concept_id INT
min_levels_of_separation INT
max_levels_of_separation INT
Concept_relationship
concept_id_1 INT
concept_id_2 INT
relationship_id VARCHAR
valid_start_date DATE
valid_end_date DATE
invalid_reason VARCHAR

Tools

- Preparing: USAGI / Rabbit-in-a-hat / White Rabbit
- Data loading: PostgreSQL / MSSQL scripts
- Athena – vocabulary repository
- Atlas – cohort builder

- The Book of OHDSI: https://ohdsi.github.io/TheBookOfOhdsi/

HemaFAIR

Funded by the European Union

Some experiences

- Great potential
- Great tools and technology
- Modelling differences may (still) occur

HemaFAIR

Funded by
the European Union

Sources

- The Book of OHDSI: https://ohdsi.github.io/TheBookOfOhdsi/
- EHDEN Academy: https://academy.ehden.eu/

HemaFAIR

Funded by
the European Union

Prof. Ronald Cornet, PhD
Full Professor ; Principal Investigator ; Principal Educator ;
Education director Medical informatics; Program Leader APH Digital Health
Department of Medical Informatics - Reusable Health Data
Amsterdam Public Health Research Institute
Amsterdam UMC
Amsterdam Public Health Research Institute
Location AMC | Meibergdreef 15, 1105 AZ Amsterdam
+31 20-566 52 05 | r.cornet@amsterdamumc.nl
kik.amsterdamumc.org/home/rcornet | www.amsterdamumc.nl

HemaFAIR

Amsterdam UMC

Funded by the European Union