



# OMOP Common Data Model

**THL:n OMOP-pilotit**

Gustav Klingstedt, Jukka Jokinen

09/12/2022

**Finnish Institute for Health and Welfare**

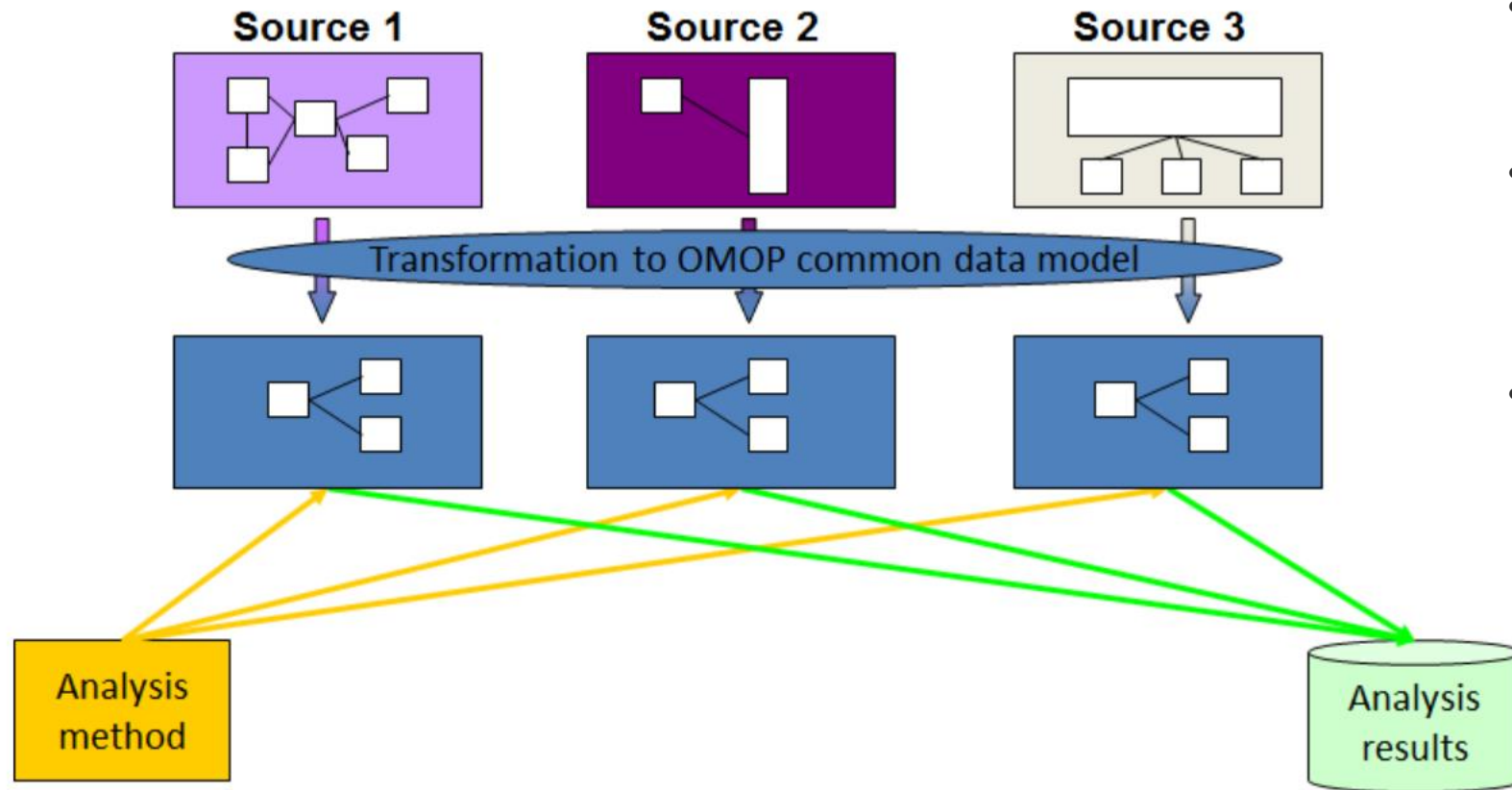
# Agenda

- **Tausta ja terminologia**
  - OHDSI
  - OMOP CDM
  - EHDEN
  - FinOMOP
- **THL:n OMOP-CDM pilotit ja tiekartta**
  - Pilotit:
    - HILMO-CDM (EHDEN-hanke)
    - FinIP rokotetutkimus (HUS-yhteistyö)
  - Kohti yhteistä tietomallia keskeisille rekisteritietovarannoille?

# Tausta ja terminologia

- OHDSI yhteisö ja työkalut
  - OMOP Common Data Model (CDM)

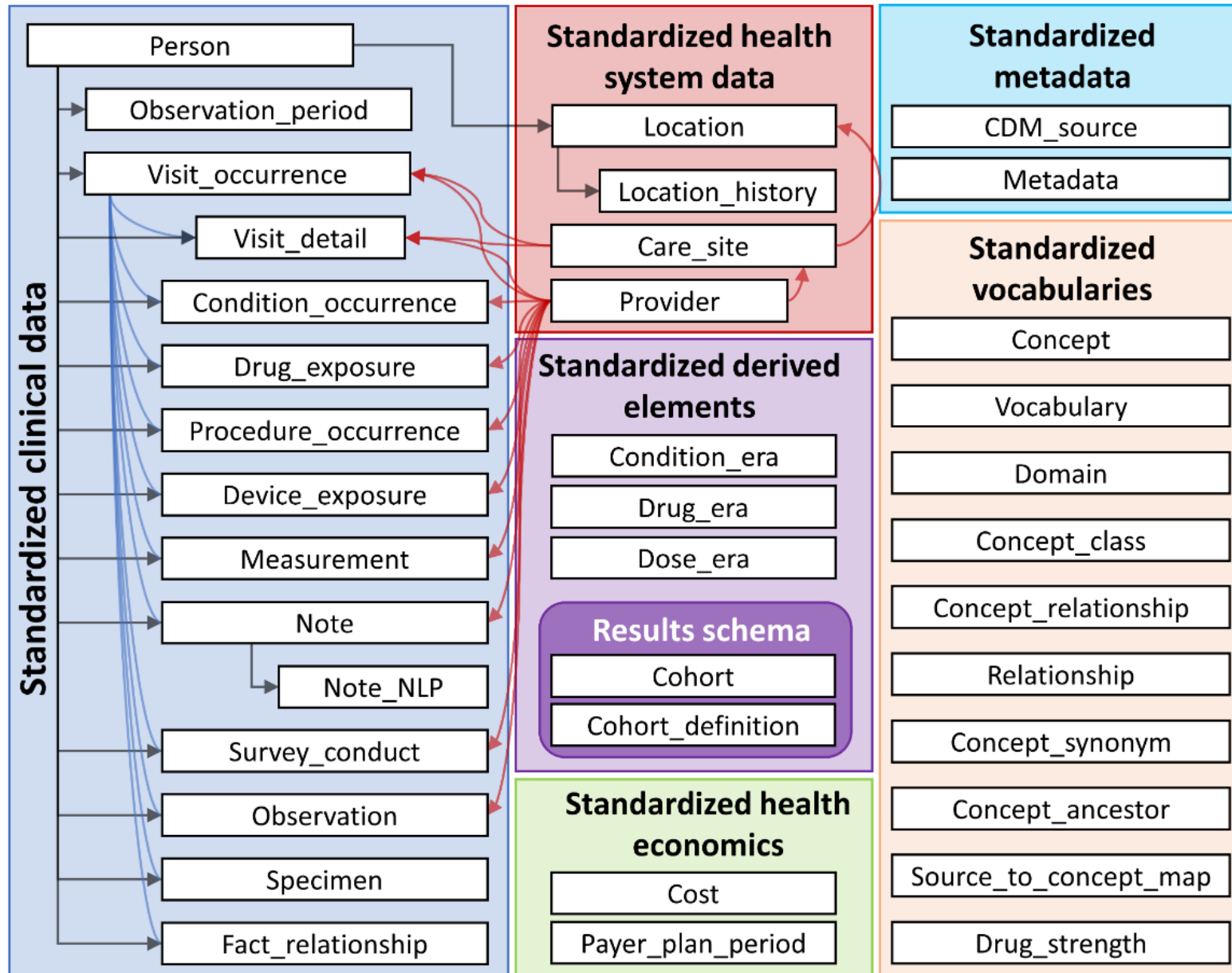
# OMOP-CDM havainnekuva



- ”Analyysidatan harmonisointikerros”
- Ensisijaisesti: Federoitu (hajautettu) käsittely ja analytiikka
- Tuo myös huomattavia etuja yksittäisen aineiston (käyttölupavaraiseseen) käsittelyyn ja analytiikkaan

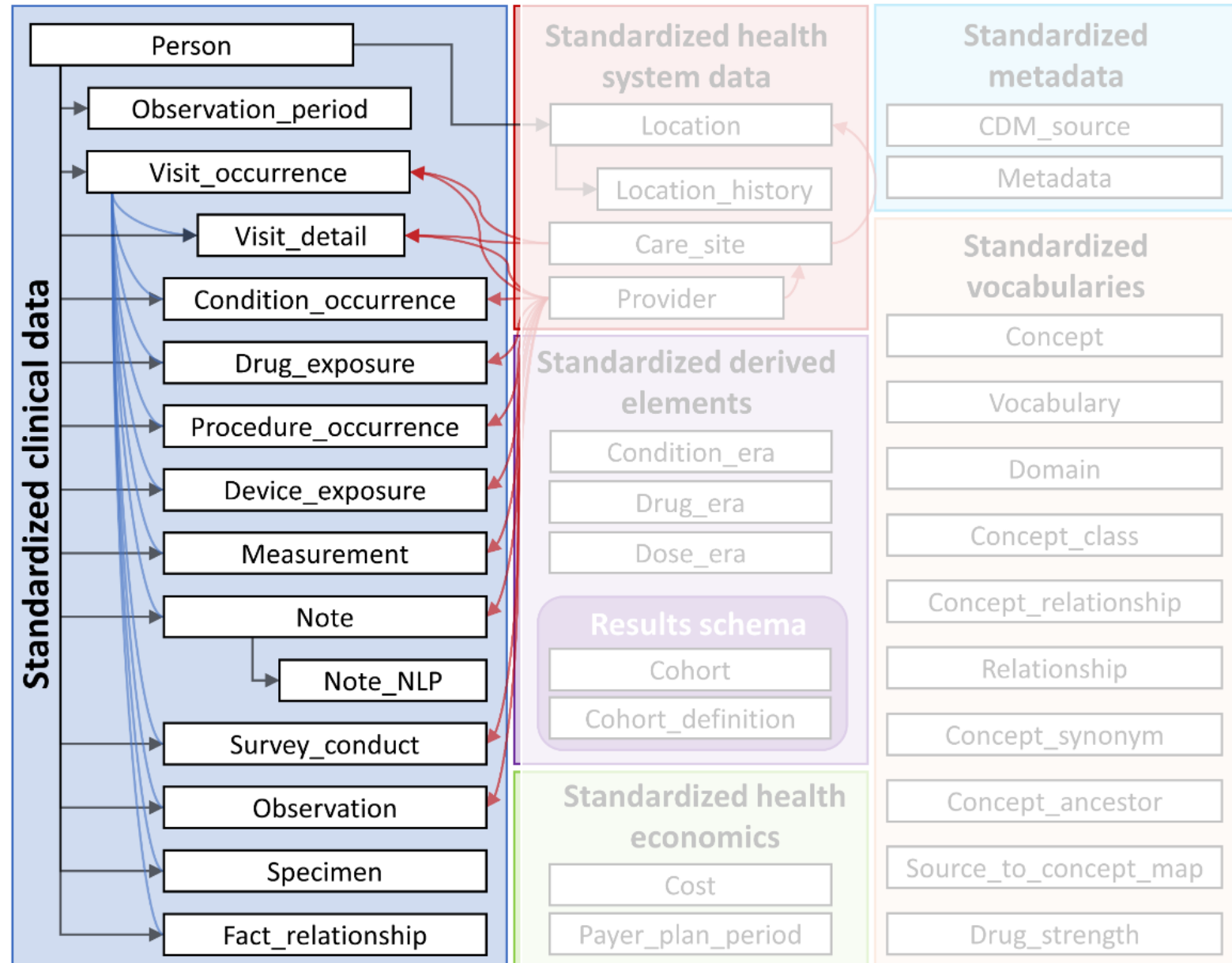
<https://www.ohdsi.org/data-standardization/the-common-data-model/>

# OMOP-CDM Overview



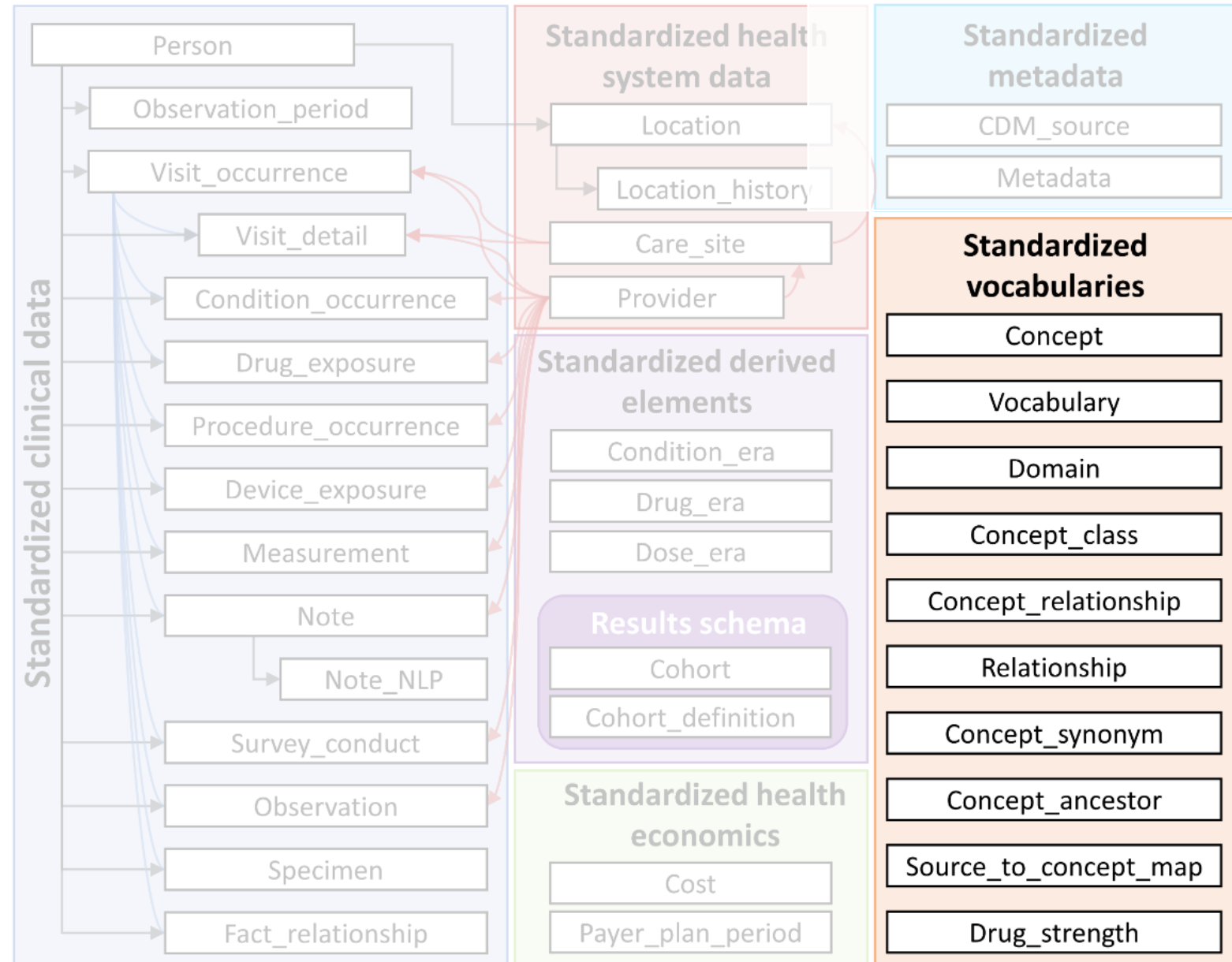
<https://ohdsi.github.io/TheBookOfOhdsi/CommonDataModel.html#CommonDataModel>

# OMOP-CDM Overview 1/2



<https://ohdsi.github.io/TheBookOfOhdsi/CommonDataModel.html#CommonDataModel>

# OMOP-CDM Overview 2/2



<https://ohdsi.github.io/TheBookOfOhdsi/CommonDataModel.html#CommonDataModel>

# Tausta ja terminologia 1/4

- OHDSI yhteisö ja työkalut
  - OMOP Common Data Model (CDM)
  - Standardoidut sanastot ja mäppäykset (vocabularies and mappings)
  - Työkaluja analytiikkaan, tietopoimintamäärittelyyn ja laaduntarkastukseen
  - Menetelmä ja työkalut sanastojen mäppäämiseen ja ETL suunnitteluun



# Tausta ja terminologia 2/4

- OHDSI yhteisö ja työkalut
  - OMOP Common Data Model (CDM)
  - Standardoidut sanastot ja mäppäykset (vocabularies and mappings)
  - Työkaluja analytiikkaan, tietopoimintamäärittelyyn ja laaduntarkastukseen
  - Menetelmä ja työkalut sanastojen mäppäämiseen ja ETL suunnitteluun
- **EHDEN** (European Health Data and Evidence Network)
  - Data Partners and SME:s (certification), Data Catalog (Portal)

# Tausta ja terminologia 3/4

- OHDSI yhteisö ja työkalut
  - OMOP Common Data Model (CDM)
  - Standardoidut sanastot ja mäppäykset (vocabularies and mappings)
  - Työkaluja analytiikkaan, tietopoimintamäärittelyyn ja laaduntarkastukseen
  - Menetelmä ja työkalut sanastojen mäppäämiseen ja ETL suunnitteluun
- EHDEN (European Health Data and Evidence Network)
  - Data Partners and SME:s (certification), Data Catalog (Portal)
- **DARWIN** (Data Analysis and Real World Interrogation Network)

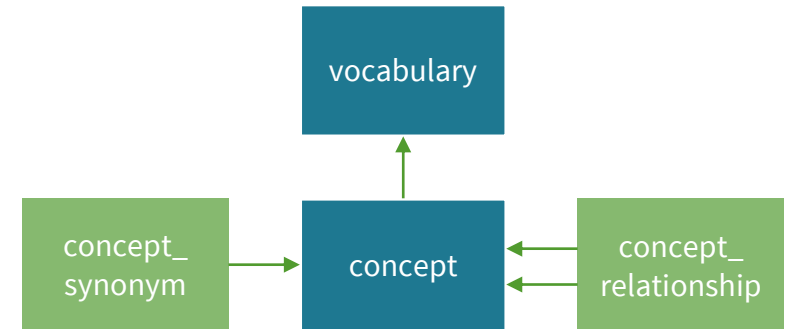
# Tausta ja terminologia 4/4

- OHDSI yhteisö ja työkalut
  - OMOP Common Data Model (CDM)
  - Standardoidut sanastot ja mäppäykset (vocabularies and mappings)
  - Työkaluja analytiikkaan, tietopoimintamäärittelyyn ja laaduntarkastukseen
  - Menetelmä ja työkalut sanastojen mappamiseen ja ETL suunnitteluun
- EHDEN (European Health Data and Evidence Network)
  - Data Partners and SME:s (certification), Data Catalog (Portal)
- DARWIN (Data Analysis and Real World Interrogation Network)
- **FinOMOP**
  - OMOP-perusteisen tutkimustoiminnan tukeminen ja viestintä
  - Kansallinen harmonisointi
    - CDM-soveltamisen käytännöt
    - **Kansallisten sanastojen OMOP-mäppäysten kehittäminen ja ylläpito**
  - Pending: National Node status within OHDSI Europe



# FinOMOP sanastotietovaranto

- Kansalliset sanastot OMOP-muodossa git-jakeluna (csv taulut)
  - Sanastot (vocabulary)
  - Koodit/termit (concept)
  - Kielistykset (concept\_synonym)
  - Mäppäykset (concept\_relationship)  
(= kansainväliset vastaavuudet, esim. Snomed CT)



- Varannon ylläpitoa ja muodostamista tukevat toiminnot
  - Koodistopalvelimen käytettävyys
    - Rajapinta mm. hakutoiminnot, historiatiedot
    - Yhtenäiset metatiedot koodistoille
  - Kansainväliset vastaavuudet tunnistettu kun julkaistaan koodistoja
  - Kielikäännökset
    - Mäppäystä tukee erityisesti laadukkaat englanninkieliset vastaavuudet kansallisesti julkaistuille sanastoille

# Agenda

- Tausta ja terminologia

- OHDSI
- OMOP CDM
- EHDEN
- FinOMOP

- **THL:n OMOP-CDM pilotit ja tiekartta**

- Pilotit:
  - HILMO-CDM (EHDEN-hanke)
  - FinIP rokotetutkimus (HUS-yhteistyö)
- Kohti yhteistä tietomallia keskeisille rekisteritietovarannoille?

# OMOP CDM –pilotit 1/2

- HILMO-CDM konversio
  - EHDEN Data Partner projekti alkanut maaliskuussa 2022
  - Tavoitteet:
    - Kyky tuottaa kuratoitu OMOP-CDM instanssi terveystietokannasta (10/2022)
    - Jalkauttaminen/demot sidosryhmissä (HILMO-ryhmä, Laaturekisterit, Analytiikkatiimi, Tutkijat!)
    - OHDSI-työkalujen ja niiden potentiaalin arviointi tietotuotannossa (12/2022)

# OMOP CDM –pilotit 2/2

- HILMO-CDM konversio
  - EHDEN Data Partner projekti alkanut maaliskuussa 2022
  - Tavoitteet:
    - Kyky tuottaa kuratoitu OMOP-CDM instanssi terveystieteilijöistä (10/2022)
    - Jalkauttaminen/demot sidosryhmissä (HILMO-ryhmä, Laaturekisterit, Analytiikkatiimi, Tutkijat!)
    - OHDSI-työkalujen ja niiden potentiaalin arviointi tietotuotannossa (12/2022)
- FinIP rokotetutkimus HUS tietoaaltaassa
  - FinIP-rokotetutkimuksen (finip.fi) uudelleenanalysointi HUS:n OMOP CDM-ympäristössä
  - Arvioidaan OMOP-mallin ilmaisukykyä uudelleentuottaa rekisteripohjaisen FinIP-tutkimuksen päätulokset
- THL tietotuotannon näkökulma:
  - Suoraan PTJ:stä muodostetun OMOP datan käyttö
  - HUS tietoaaltaalla toimiminen (osaaminen)



# OMOP CDM:n mahdolliset hyödyt

## Analytiikka ja tutkimus

- Parannettu aineistojen käytettävyys ja hyödynnettävyys (FAIR)
- Datat karakterisointi, kohorttiarvioinnit ja -poiminnot
- Vakiintuneiden analytiikkamenetelmien kirjasto

## Data management ja kuratointi

- Systemaattinen ja yhdenmukainen tapa ilmaista luokitellut muuttujat
- Vakiintunut tapa normalisoida TH-data
- Yhtenäiset laadun arvioinnin ja tarkastuksen työkalut

## Tiedonkeruu

- Tiedon yhdenmukaistaminen ja kuratointi lähellä tiedon lähdettä

# Kohti yhteistä tietomallia keskeisille rekisteritietovarannoille?

- Muita mahdollisia sovellusalueita OMOP CDM:lle
  - Muiden rekisterien tuominen yhteiseen tietomalliin
    - Avohilmo
    - Laaturekisterit
    - Kanta Laboratoriotiedot
    - Ensihoidon tietovaranto
  - OMOP CDM periaatteiden hyödyntäminen koodistojen käsittelyssä
  - Tiedonkeruun laaduntarkastuksen työkalut

# Kiitos!