

Documentatie workflow – week 7

Om de data met elkaar te mappen, is er gebruikt gemaakt van de volgende informatiebron: <https://github.com/thehyve/OHDSI-ETL-Synthea/tree/master/docs>. Deze is als voorbeeld genomen, om zowel figuur 1 en 2 te maken.

PERSON	CONDITION_OCCURENCE
person_id = PGPC_ID int year_of_birth = Birth year int month_of_birth = Birth month int gender_concept_id = gender_concept_id_athena gender_source_value = Sex race_concept_id = Ehtnicity_athena race_source_value = Ethnicity	condition_occurence_id = UUID person_id = PGPC_ID int condition_concept_id = Athena_condition condition_source_value = condition

MEASUREMENT
measurement_id = measurement_id UUID measurement_date = date_file_source_before_filter person_id = PGPC_ID int measurement_concept_id = measurement_athena measurement_source_value = gene value_as_concept_id = value_athena value_source_value = feature

Figuur 1: Algemene beschrijving van welke source data tegen welke OMOP CDM tabel mapt semantische en syntactische mapping

PERSON	CONDITION_OCCURENCE
person_id = PGPC_ID int year_of_birth = Birth year int month_of_birth = Birth month int gender_concept_id = 8507 M, 8532 F gender_source_value = Sex race_concept_id = Ehtnicity (White (45532670)) race_source_value = Ethnicity	condition_occurence_id = UUID person_id = PGPC_ID int condition_concept_id = 4147507 condition_source_value = condition (Bilateral retinal detachments and cataracts)

MEASUREMENT
measurement_id = measurement_id UUID measurement_date = date_file_source_before_filter person_id = PGPC_ID int measurement_concept_id = 4281995 (athena) measurement_source_value = gene value_as_concept_id = 4048365 (Athena) value_source_value = feature

Figuur 2: Gedetailleerde beschrijving van wat voor data tegen de OMOP CDM tabel aan is gemapt.

Om deze tabellen in te vullen zijn de volgende commando's gebruikt:

- INSERT INTO onderwijs.di_groep_6.condition_occurrence(condition_occurrence_id, person_id, condition_concept_id, condition_source_value, condition_start_date, condition_type_concept_id) VALUES (973727460, 26, 4147507, 'Bilateral retinal detachments and cataracts', '2022-05-30', 32020), (5781220, 26, 40320318, 'Osteoarthritis', '2022-05-30', 32020);
- INSERT INTO onderwijs.di_groep_6.person(person_id, person_source_value, year_of_birth, month_of_birth, gender_concept_id, gender_source_value, race_concept_id, race_source_value, ethnicity_concept_id, ethnicity_source_value) VALUES ('3', 'PGPC-3', '1959', NULL, 8507, 'M', 45532670, 'White', 45532670, 'White'), ('26', 'PGPC-26', '1933', '8', 8507, 'M', 45532670, 'White', 45532670, 'White'), ('25', 'PGPC-25', '1944', '12', 8507, 'M', 45532670, 'White', 45532670, 'White');
- INSERT INTO onderwijs.di_groep_6.measurement(measurement_id, person_id, measurement_date, measurement_concept_id, measurement_source_value, value_as_concept_id, value_source_value, measurement_type_concept_id) VALUES (160862948, 3, '2022-04-25', 4281995, 'ENSG00000279864', 4048365, 'ENST00000651618', 5001), (358379748, 3, '2022-04-25', 4281995, 'ENSG00000279864', 4048365, 'ENST00000651618', 5001), (995399908, 3, '2022(580728548, 3, '2022-04-25', 4281995, 'ENSG00000279864', 4048365, 'ENST00000651618', 5001), -04-25', 4281995, 'ENSG00000279864', 4048365, 'ENST00000651618', 5001), (313909988, 3, '2022-04-25', 4281995, 'ENSG00000279864', 4048365, 'ENST00000651618', 5001), (192916708, 3, '2022-04-25', 4281995, 'ENSG00000279864', 4048365, 'ENST00000651618', 5001), (511426788, 3, '2022-04-25', 4281995, 'ENSG00000279864', 4048365, 'ENST00000651618', 5001), (390433508, 3, '2022-04-25', 4281995, 'ENSG00000279864', 4048365, 'ENST00000651618', 5001), (708943588, 3, '2022-04-25', 4281995, 'ENSG00000279864', 4048365, 'ENST00000651618', 5001), (587950308, 3, '2022-04-25', 4281995, 'ENSG00000279864', 4048365, 'ENST00000651618', 5001), (673551588, 26, '2022-04-25', 4281995, 'ENSG00000279864', 4048365, 'ENST00000651618', 5001), (463618788, 26, '2022-04-25', 4281995, 'ENSG00000279864', 4048365, 'ENST00000651618', 5001), (221632228, 26, '2022-04-25', 4281995, 'ENSG00000279864', 4048365, 'ENST00000651618', 5001), (540142308, 26, '2022-04-25', 4281995, 'ENSG00000279864', 4048365, 'ENST00000651618', 5001), (419149028, 26, '2022-04-25', 4281995, 'ENSG00000279864', 4048365, 'ENST00000651618', 5001), (298155748, 26, '2022-04-25', 4281995, 'ENSG00000279864', 4048365, 'ENST00000651618', 5001), (616665828, 26, '2022-04-25', 4281995, 'ENSG00000279864', 4048365, 'ENST00000651618', 5001), (495672548, 26, '2022-04-25', 4281995, 'ENSG00000279864', 4048365, 'ENST00000651618', 5001), (814182628, 26, '2022-04-25', 4281995, 'ENSG00000279864', 4048365, 'ENST00000651618', 5001), (693189348, 26, '2022-04-25', 4281995, 'ENSG00000279864', 4048365, 'ENST00000651618', 5001);

Tabel 1: Beschrijving waar de data vandaan is gekomen, van de Person tabel.

Person_id	Id van de patiënt, deze kan gevonden worden in het PDF-bestand van de patiënt
Year_of_birth Month_of_birth	Beide gegevens zijn uit de PDF-bestand van de patiënt.
Gender_concept_id	Deze concept id is gevonden uit Athena. Waarbij de man staat voor 8507 en de vrouw voor 8532.

Race_concept_id	Deze concept id is gevonden met behulp van Athena.
-----------------	----------------------------------------------------

Tabel 2: Beschrijving waar de data vandaan is gekomen, van de Condition tabel.

Condition_source_value	Condition van de betreffende patiënt. Dat staat in het PDF-bestand van de patiënt.
Condition_concept_id	Deze concept ID is gevonden met behulp van Athena door de condition van het PDF-bestand van de betreffende patiënt in Athena op te zoeken. <ul style="list-style-type: none"> - 4147507: Bilateral retinal detachments and cataracts - 40320318: Osteoarthritis
Condition_occurrence_id	UUID, unieke id dat toegewezen is

Tabel 3: Beschrijving waar de data vandaan is gekomen, van de Measurement tabel.

Measurement_date	Deze datum is uit de VCF-bestanden gehaald van vep
Measurement_concept_id	Deze concept ID is door “gene” in Athena op te zoeken.
Measurement_source_value	Bevat de gennamen die uit de VCF-bestanden komen (kolom “gene” in het VCF-bestand)
Value_source_value	Bevat de features die uit het VCF-bestand komt. (Kolom “feature” in het bestand)
Value_as_concept_id	Deze concept ID is door “feature” in Athena op te zoeken.
Measurement_id	UUID, unieke id dat toegewezen is