

HiRID to OMOP CDM report

By: Julio Portella

Index

- HiRID dataset
- OMOP-CDM
- ETL Process
- Vocabulary translation
- Technical Implementation
- Quality Assurance
- Results
- Observations
- Next Steps

HiRID Database

- Dataset from the Department of Intensive Care Medicine of Bern University Hospital
 - Continuous measurements every 2 minutes
 - Contains Lab values, administered drugs, fluids and nutrition
 - Measurements from bedside monitoring
 - Measurements and settings of medical devices
- Currently available at Physionet (Requires Credentials)
- It has 776'921,131 observations, close to a billion. MIMIC has 312 million of observations
- <https://physionet.org/content/hirid/1.0/>

HiRID Database

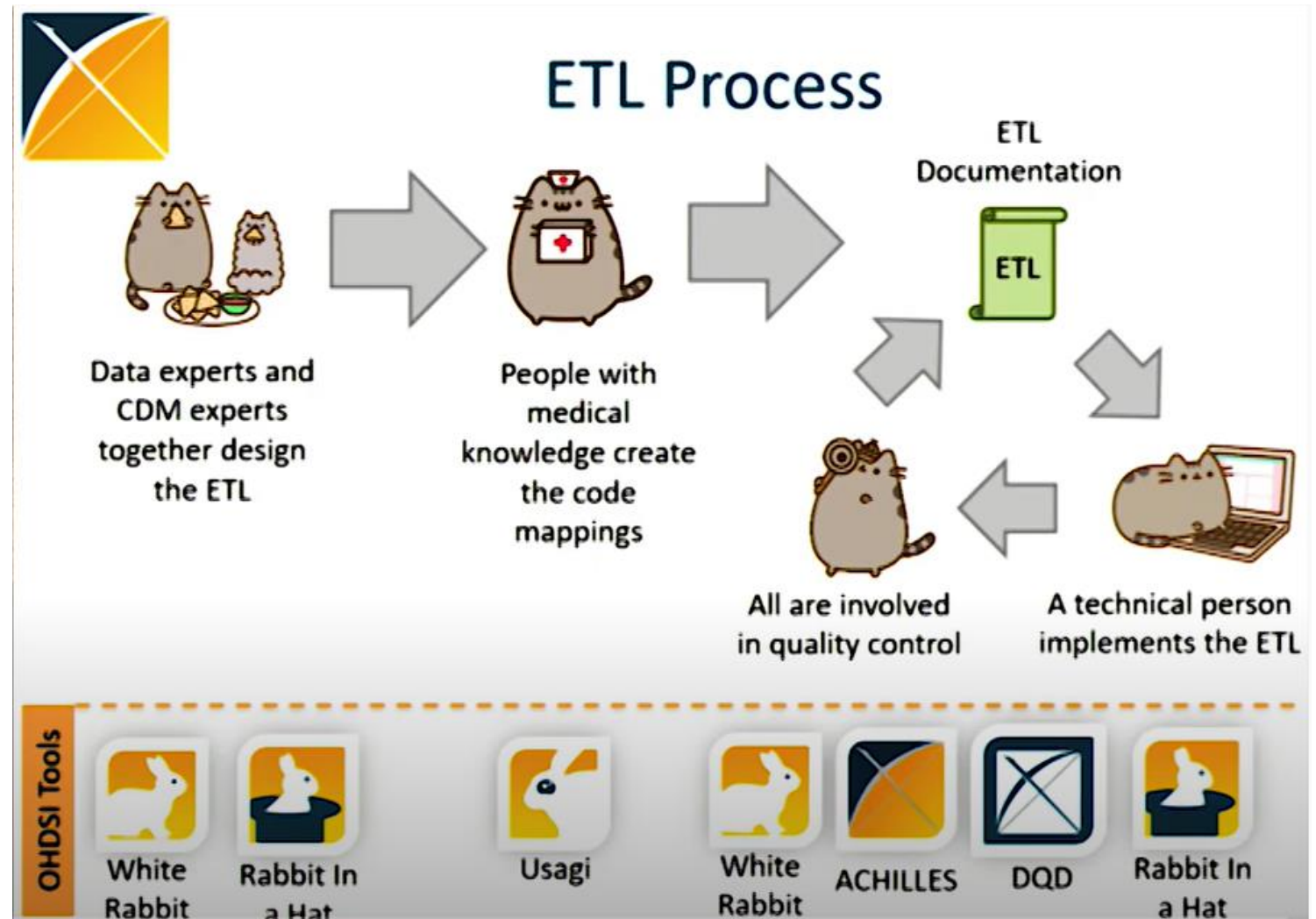
- The dataset contains data of 34 thousand people
- From this group, 64.2% are male and 35.8% are female
- All the patients are over 40 and the median is 65 years old
- The observations and pharmaceutical records are divided in 250 groups, each group has a different amount of random patients.
- There is a summarized version of the HiRID database with less variables and downsampled data used in their published paper. See: <https://www.nature.com/articles/s41591-020-0789-4>
- The downsampled data has measurements every 5 minutes

OMOP CDM

- Managed by the Observational Health Data Sciences and Informatics (OHDSI)
- Most popular common data model
- Created in 2008 and funded by the FDA
- In 2014 moved to OHDSI at Columbia University
- Queries and tools freely available and Open Source
- Patient centered data model
- Currently in version 6.0.0
- Designed for Relational databases but it can be translated to column databases and Apache Hive

ETL Process

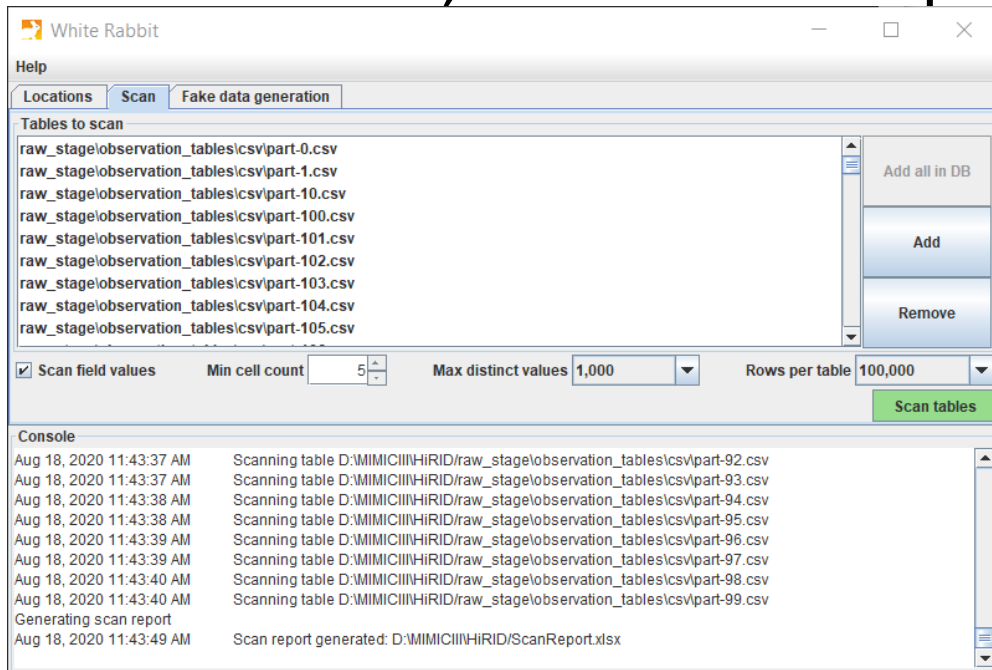
- Cyclical process with 5 parts
 - Design
 - Code mapping
 - Documentation
 - Implementation
 - Validation
- This process was used to translate the database



<https://www.youtube.com/watch?v=4vytGDrhHpk>

ETL Process

- WhiteRabbit: Tool used to map all the values from a csv file or a database connection
- It scan details can be configured, the more detailed the count the more it will take to scan
- In this case, each csv file represents a table



ETL Process

- Once the scan is done it generates an excel spreadsheet named ScanReport.xlsx
- The Scan Report has an overview and an analysis of the most popular values for each table.
- Very useful to know which observations and drugs are the most used

	A	B	C	D	E	F	G	H
1	Table	Field	Type	Max length	N rows	N rows chi	Fraction empty	
2	part-0.csv	datetime	varchar	23	-1	99999	0	
3	part-0.csv	entertime	varchar	23	-1	99999	0	
4	part-0.csv	patientid	int	4	-1	99999	0	
5	part-0.csv	status	int	2	-1	99999	0	
6	part-0.csv	stringvalue	empty	0	-1	99999	1	
7	part-0.csv	type	varchar	1	-1	99999	0.99062	
8	part-0.csv	value	real	11	-1	99999	0	
9	part-0.csv	variableid	int	8	-1	99999	0	
10								
11	part-1.csv	datetime	varchar	23	-1	99999	0	
12	part-1.csv	entertime	varchar	23	-1	99999	0	
13	part-1.csv	patientid	int	4	-1	99999	0	
14	part-1.csv	status	int	3	-1	99999	0	
15	part-1.csv	stringvalue	varchar	4	-1	99999	0.99983	
16	part-1.csv	type	varchar	1	-1	99999	0.98312	
17	part-1.csv	value	real	11	-1	99999	0.00017	
18	part-1.csv	variableid	int	8	-1	99999	0	
19								
20	part-10.csv	datetime	varchar	23	-1	99999	0	
21	part-10.csv	entertime	varchar	23	-1	99999	0	
22	part-10.csv	patientid	int	3	-1	99999	0	
23	part-10.csv	status	int	2	-1	99999	0	
24	part-10.csv	stringvalue	varchar	4	-1	99999	0.99997	
25	part-10.csv	type	varchar	1	-1	99999	0.98231	
26	part-10.csv	value	real	11	-1	99999	3E-05	
27	part-10.csv	variableid	int	8	-1	99999	0	
28								
29	part-100.c	datetime	varchar	23	-1	99999	0	

Overview

part-0.csv

part-1.csv

part-10.csv

part-100.csv

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	datetime	Frequency	entertime	Frequency	patientid	Frequency	status	Frequency	stringvalue	Frequency	type	Frequency	value	Frequency	variableid	Frequency
2	2146-07-0	117	2128-02-1	30	307	52126	16	70086		99999		99061	0.0	5259	200	10780
3	2146-07-0	51	2128-02-1	29	766	10865	8	29710		F		938	100.0	2198	110	10488
4	2183-05-1	45	2128-02-1	28	1088	10490	4	126				96.0	2023	120		10476
5	2130-07-0	35	2128-02-1	26	229	8057	12	64				97.0	1701	100		10473
6	2128-02-2	33	2128-02-2	26	148	6155	0	13				99.0	1506	4000		9714
7	2128-02-2	28	2128-02-2	25	496	4981						95.0	1401	700		8110
8	2183-05-1	28	2128-02-1	25	852	4255						5.0	1360	2200		5054
9	2116-12-1	28	2128-02-2	23	833	3070						98.0	1345	211		3220
10	2130-07-0	28	2128-02-2	23								1.0	1235	30005010		3079
11	2169-11-2	27	2128-02-1	21								89.0	1132	30005080		3064
12	2128-02-2	27	2128-02-1	20								94.0	1113	30005075		3064
13	2169-11-2	23	2128-02-1	20								30.0	973	210		2955
14	2130-07-0	20	2169-11-2	19								3.0	970	300		2458
15	2169-11-2	20	2169-11-2	19								4.0	960	2410		1719
16	2128-02-2	19	2128-02-1	18								6.0	920	2010		1628
17	2128-02-1	18	2128-02-2	18								90.0	842	3000		1622
18	2183-05-1	18	2128-02-2	18								93.0	834	3110		1622
19	2169-11-2	18	2169-11-2	18								32.0	826	2610		1622
20	2169-11-2	18	2128-02-2	18								31.0	822	310		1404
21	2169-11-2	17	2130-07-0	17								2.0	804	30010009		1343
22	2169-11-2	17	2128-02-2	17								87.0	787	212		1230
23	2172-07-3	17	2128-02-2	17								83.0	785	3200		841
24	2116-12-1	16	2128-02-1	17								84.0	754	8290		564
25	2116-12-1	16	2128-02-1	17								79.0	750	3845		549
26	2128-02-1	16	2130-07-0	17								8.0	740	2600		309
27	2128-02-1	16	2128-02-2	17								75.0	738	15001565		290
28	2169-11-2	15	2128-02-2	16								74.0	735	30005110		197
29	2169-11-2	15	2128-02-2	16								7.0	735	15001166		184

Overview

part-0.csv

part-1.csv

part-10.csv

part-100.csv

part-101.csv

part-102.csv

part-103.csv

part-104.csv

part-105.csv

part-106.csv

part-107.csv

part-108.csv

part-109.csv

part-110.csv

part-111.csv

part-112.csv

part-113.csv

part-114.csv

part-115.csv

part-116.csv

part-117.csv

part-118.csv

part-119.csv

part-120.csv

part-121.csv

part-122.csv

part-123.csv

part-124.csv

part-125.csv

part-126.csv

part-127.csv

part-128.csv

part-129.csv

part-130.csv

part-131.csv

part-132.csv

part-133.csv

part-134.csv

part-135.csv

part-136.csv

part-137.csv

part-138.csv

part-139.csv

part-140.csv

part-141.csv

part-142.csv

part-143.csv

part-144.csv

part-145.csv

part-146.csv

part-147.csv

part-148.csv

part-149.csv

part-150.csv

part-151.csv

part-152.csv

part-153.csv

part-154.csv

part-155.csv

part-156.csv

part-157.csv

part-158.csv

part-159.csv

part-160.csv

part-161.csv

part-162.csv

part-163.csv

part-164.csv

part-165.csv

part-166.csv

part-167.csv

part-168.csv

part-169.csv

part-170.csv

part-171.csv

part-172.csv

part-173.csv

part-174.csv

part-175.csv

part-176.csv

part-177.csv

part-178.csv

part-179.csv

part-180.csv

part-181.csv

part-182.csv

part-183.csv

part-184.csv

part-185.csv

part-186.csv

part-187.csv

part-188.csv

part-189.csv

part-190.csv

part-191.csv

part-192.csv

part-193.csv

part-194.csv

part-195.csv

part-196.csv

part-197.csv

part-198.csv

part-199.csv

part-200.csv

part-201.csv

part-202.csv

part-203.csv

part-204.csv

part-205.csv

part-206.csv

part-207.csv

part-208.csv

part-209.csv

part-210.csv

part-211.csv

part-212.csv

part-213.csv

part-214.csv

part-215.csv

part-216.csv

part-217.csv

part-218.csv

part-219.csv

part-220.csv

part-221.csv

part-222.csv

part-223.csv

part-224.csv

part-225.csv

part-226.csv

part-227.csv

part-228.csv

part-229.csv

part-230.csv

part-231.csv

part-232.csv

part-233.csv

part-234.csv

part-235.csv

part-236.csv

part-237.csv

part-238.csv

part-239.csv

part-240.csv

part-241.csv

part-242.csv

part-243.csv

part-244.csv

part-245.csv

part-246.csv

part-247.csv

part-248.csv

part-249.csv

part-250.csv

part-251.csv

part-252.csv

part-253.csv

part-254.csv

part-255.csv

part-256.csv

part-257.csv

part-258.csv

part-259.csv

part-260.csv

part-261.csv

part-262.csv

part-263.csv

part-264.csv

part-265.csv

part-266.csv

part-267.csv

part-268.csv

part-269.csv

part-270.csv

part-271.csv

part-272.csv

part-273.csv

part-274.csv

part-275.csv

part-276.csv

part-277.csv

part-278.csv

part-279.csv

part-280.csv

part-281.csv

part-282.csv

part-283.csv

part-284.csv

part-285.csv

part-286.csv

part-287.csv

part-288.csv

part-289.csv

part-290.csv

part-291.csv

part-292.csv

part-293.csv

part-294.csv

part-295.csv

part-296.csv

part-297.csv

part-298.csv

part-299.csv

part-300.csv

part-301.csv

part-302.csv

part-303.csv

part-304.csv

part-305.csv

part-306.csv

part-307.csv

part-308.csv

part-309.csv

part-310.csv

part-311.csv

part-312.csv

part-313.csv

part-314.csv

part-315.csv

part-316.csv

part-317.csv

part-318.csv

part-319.csv

part-320.csv

part-321.csv

part-322.csv

part-323.csv

part-324.csv

part-325.csv

part-326.csv

part-327.csv

part-328.csv

part-329.csv

part-330.csv

part-331.csv

part-332.csv

part-333.csv

part-334.csv

part-335.csv

part-336.csv

part-337.csv

part-338.csv

part-339.csv

part-340.csv

part-341.csv

part-342.csv

part-343.csv

part-344.csv

part-345.csv

part-346.csv

part-347.csv

part-348.csv

part-349.csv

part-350.csv

part-351.csv

part-352.csv

part-353.csv

part-354.csv

part-355.csv

part-356.csv

part-357.csv

part-358.csv

part-359.csv

part-360.csv

part-361.csv

part-362.csv

part-363.csv

part-364.csv

part-365.csv

part-366.csv

part-367.csv

part-368.csv

part-369.csv

part-370.csv

part-371.csv

part-372.csv

part-373.csv

part-374.csv

part-375.csv

part-376.csv

part-377.csv

part-378.csv

part-379.csv

part-380.csv

part-381.csv

part-382.csv

part-383.csv

part-384.csv

part-385.csv

part-386.csv

part-387.csv

part-388.csv

part-389.csv

part-390.csv

part-391.csv

part-392.csv

part-393.csv

part-394.csv

part-395.csv

part-396.csv

part-397.csv

part-398.csv

part-399.csv

part-400.csv

part-401.csv

part-402.csv

part-403.csv

part-404.csv

part-405.csv

part-406.csv

part-407.csv

part-408.csv

part-409.csv

part-410.csv

part-411.csv

part-412.csv

part-413.csv

part-414.csv

part-415.csv

part-416.csv

part-417.csv

part-418.csv

part-419.csv

part-420.csv

part-421.csv

part-422.csv

part-423.csv

part-424.csv

part-425.csv

part-426.csv

part-427.csv

part-428.csv

part-429.csv

part-430.csv

part-431.csv

part-432.csv

part-433.csv

part-434.csv

part-435.csv

part-436.csv

part-437.csv

part-438.csv

part-439.csv

part-440.csv

part-441.csv

part-442.csv

part-443.csv

part-444.csv

part-445.csv

part-446.csv

part-447.csv

part-448.csv

part-449.csv

part-450.csv

part-451.csv

part-452.csv

part-453.csv

part-454.csv

part-455.csv

part-456.csv

part-457.csv

part-458.csv

part-459.csv

part-460.csv

part-461.csv

part-462.csv

part-463.csv

part-464.csv

part-465.csv

part-466.csv

part-467.csv

part-468.csv

part-469.csv

part-470.csv

part-471.csv

part-472.csv

part-473.csv

part-474.csv

part-475.csv

part-476.csv

part-477.csv

part-478.csv

part-479.csv

part-480.csv

part-481.csv

part-482.csv

part-483.csv

part-484.csv

part-485.csv

part-486.csv

part-487.csv

part-488.csv

part-489.csv

part-490.csv

part-491.csv

part-492.csv

part-493.csv

part-494.csv

part-495.csv

part-496.csv

part-497.csv

part-498.csv

part-499.csv

part-500.csv

part-501.csv

part-502.csv

part-503.csv

part-504.csv

part-505.csv

part-506.csv

part-507.csv

part-508.csv

part-509.csv

part-510.csv

part-511.csv

part-512.csv

part-513.csv

part-514.csv

part-515.csv

part-516.csv

part-517.csv

part-518.csv

part-519.csv

part-520.csv

part-521.csv

part-522.csv

part-523.csv

part-524.csv

part-525.csv

part-526.csv

part-527.csv

part-528.csv

part-529.csv

part-530.csv

part-531.csv

part-532.csv

part-533.csv

part-534.csv

part-535.csv

part-536.csv

part-537.csv

part-538.csv

part-539.csv

part-540.csv

part-541.csv

part-542.csv

part-543.csv

part-544.csv

part-545.csv

part-546.csv

part-547.csv

part-548.csv

part-549.csv

part-550.csv

part-551.csv

part-552.csv

part-553.csv

part-554.csv

part-555.csv

part-556.csv

part-557.csv

part-558.csv

part-559.csv

part-560.csv

part-561.csv

part-562.csv

part-563.csv

part-564.csv

part-565.csv

part-566.csv

part-567.csv

part-568.csv

part-569.csv

part-570.csv

part-571.csv

part-572.csv

part-573.csv

part-574.csv

part-575.csv

part-576.csv

part-577.csv

part-578.csv

part-579.csv

part-580.csv

part-581.csv

part-582.csv

part-583.csv

part-584.csv

part-585.csv

part-586.csv

part-587.csv

part-588.csv

part-589.csv

part-590.csv

part-591.csv

part-592.csv

part-593.csv

part-594.csv

part-595.csv

part-596.csv

part-597.csv

part-598.csv

part-599.csv

part-600.csv

part-601.csv

part-602.csv

part-603.csv

part-604.csv

part-605.csv

part-606.csv

part-607.csv

part-608.csv

part-609.csv

part-610.csv

part-611.csv

part-612.csv

part-613.csv

part-614.csv

part-615.csv

part-616.csv

part-617.csv

part-618.csv

part-619.csv

part-620.csv

part-621.csv

part-622.csv

part-623.csv

part-624.csv

part-625.csv

part-626.csv

part-627.csv

part-628.csv

part-629.csv

part-630.csv

part-631.csv

part-632.csv

part-633.csv

part-634.csv

part-635.csv

part-636.csv

part-637.csv

part-638.csv

part-639.csv

part-640.csv

part-641.csv

part-642.csv

part-643.csv

part-644.csv

part-645.csv

part-646.csv

part-647.csv

part-648.csv

part-649.csv

part-650.csv

part-651.csv

part-652.csv

part-653.csv

part-654.csv

part-655.csv

part-656.csv

part-657.csv

part-658.csv

part-659.csv

part-660.csv

part-661.csv

part-662.csv

part-663.csv

part-664.csv

part-665.csv

part-666.csv

part-667.csv

part-668.csv

part-669.csv

part-670.csv

part-671.csv

part-672.csv

part-673.csv

part-674.csv

part-675.csv

part-676.csv

part-677.csv

part-678.csv

part-679.csv

part-680.csv

part-681.csv

part-682.csv

part-683.csv

part-684.csv

part-685.csv

part-686.csv

part-687.csv

part-688.csv

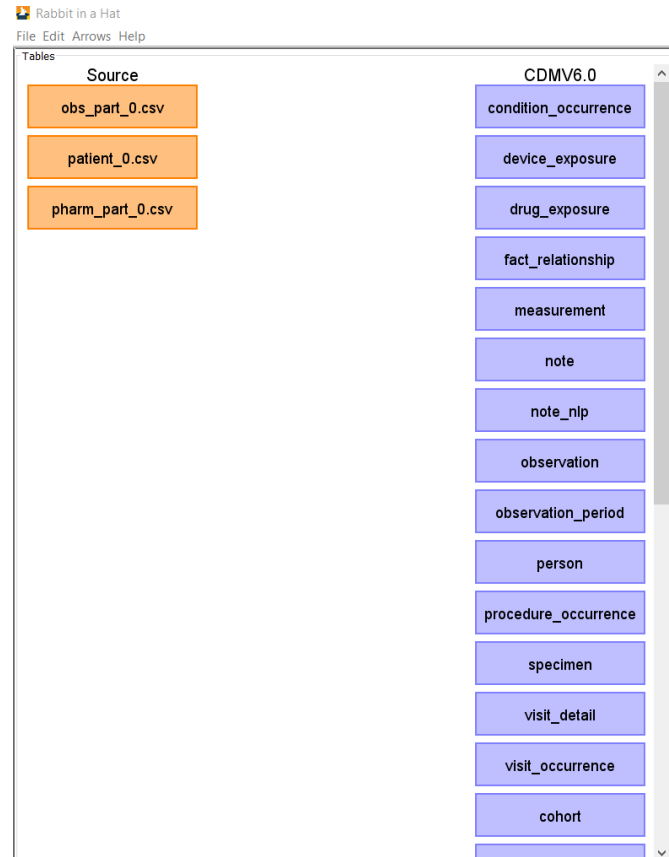
part-689.csv

part-690.csv

part-691.csv</

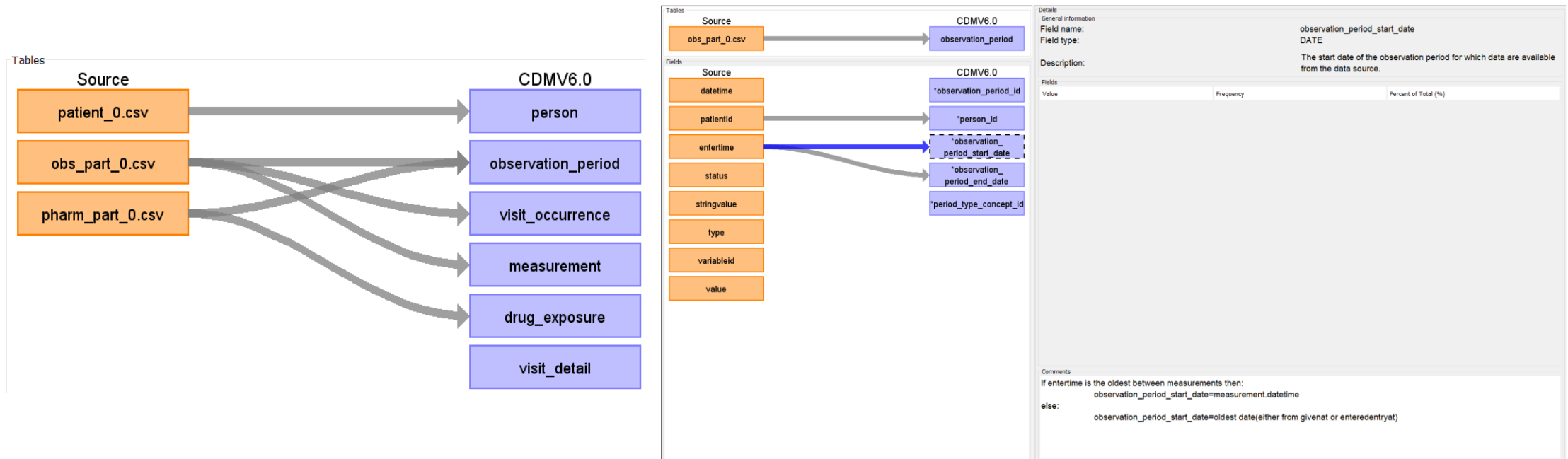
ETL Process

- Rabbit In a Hat: Tool that generates the ETL document.
- Use the ScanReport generated in “White Rabbit” generate the ETL map. Make sure that the ScanReport contains only the required tables



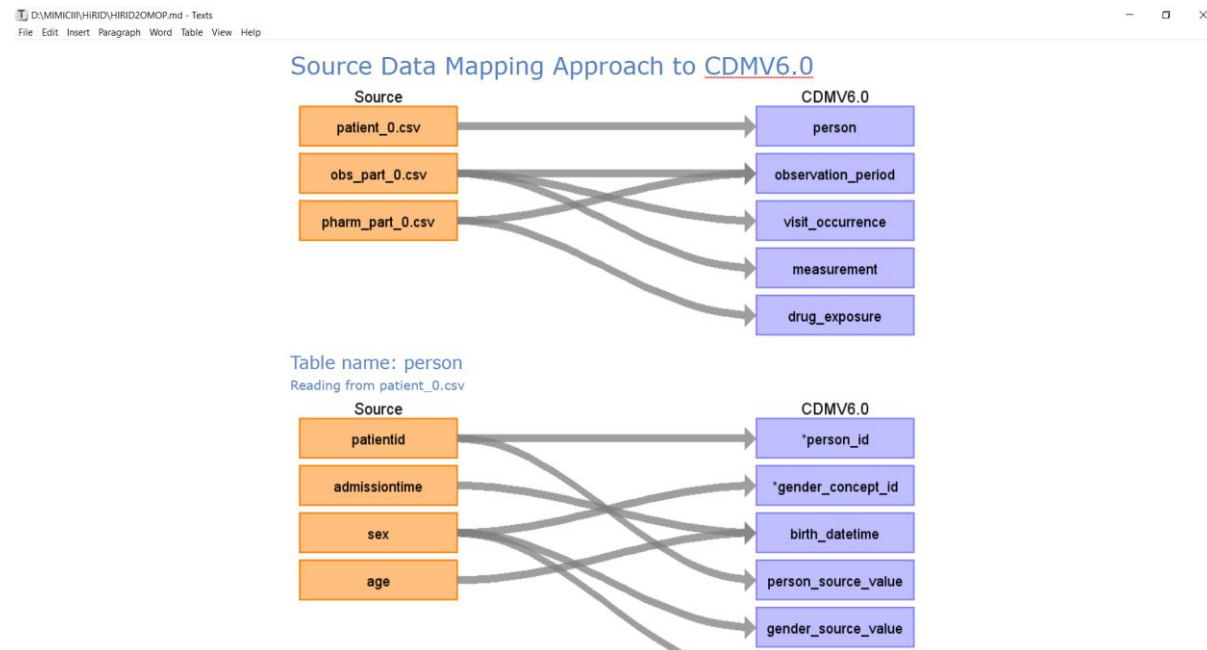
ETL Process

- Given that HiRID has 3 tables, many OMOP tables and fields had to be deduced.
- Tables like measurement, drug_exposure and person can be translated directly from HiRID
- Observation_period and visit_occurrence had to be deduced



ETL Process

- Once the connections and mapping logic was done, the ETL documentation can be generated
- White Rabbit also generates the ETL test framework
- The generated ETL document is a reference, it is expected to change it during the whole process



ETL Process

- Many deductions were done using “Rabbit in a Hat”
- Patient’s birthday were deduced from the admission datetime minus their age at the moment
- The observational period start and end date are from the oldest and newest date from the observations and drugs reports
- The table visit_occurrence has a similar logic but it only uses the dates from the observations
- Drug_exposure’s refills column was deduced by counting the number of times that a drug was taken.
- Many terms had to be translated using OMOP CDM vocabularies or were deduced by the context of the database.

Vocabulary translation

- 70% of the vocabulary was translated, 2 missing observations and 195 drugs
- Missing data by vocabulary items

Vocabulary Type	% of the total data
Observations	0.025% of all observations
Pharmaceuticals	13% of all medications

- Most of the observations were translated into LOINC, the rest used SNOMED

Technical Implementation

- Before translating the tables, the database was created using OMOP CDM queries
- Most relevant tools are DQD and Achilles
- Given that DQD is not available for OMOP v6, Achilles is going to be used
- Given that

Technical Implementation

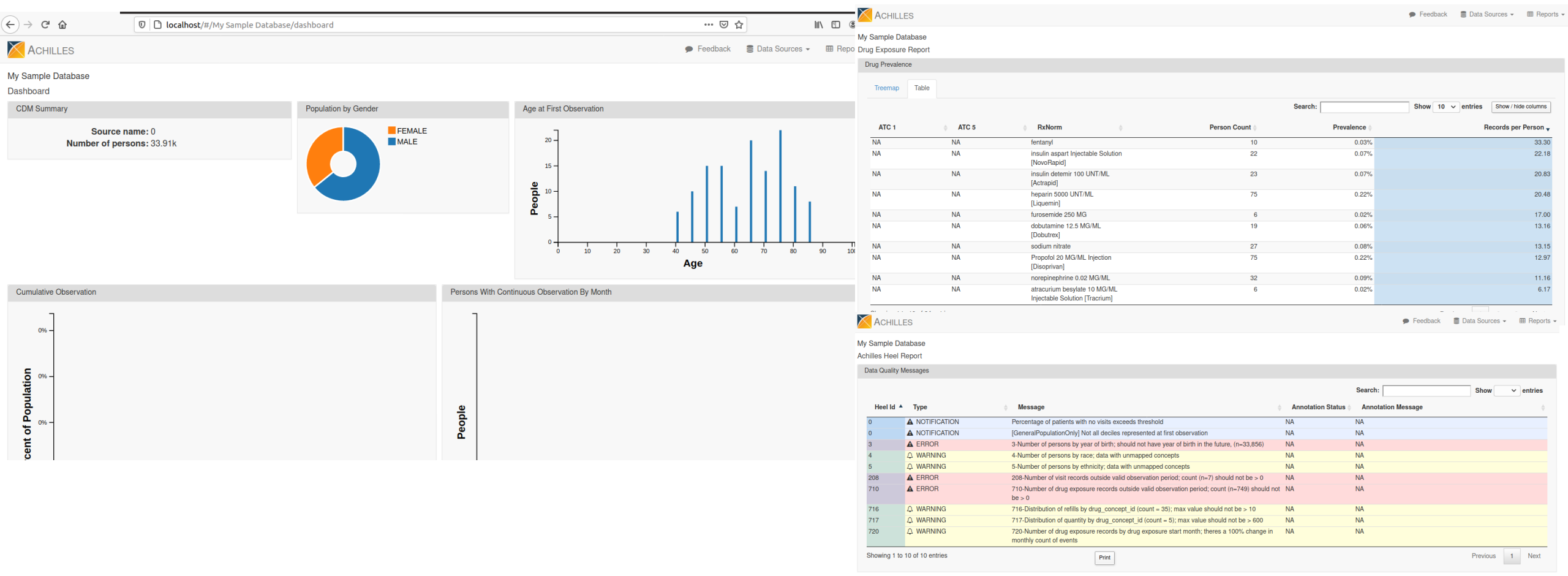
- After translating the dataset and adding the constraints into the database, quality assurance tools were used
- The technical implementation was done using Python
- The data was uploaded and processed using Panda's dataframe
- Rows with unmappable elements and errors were removed
- After creating the tables, they were bulk uploaded into the databse
- Once the data was uploaded the foreign keys and restrictions were added into the databse

Quality Assurance

- Before translating the tables, the database was created using OMOP CDM queries
- Most relevant tools are DQD and Achilles
- Given that DQD is not available for OMOP v6, Achilles is going to be used
- Using the R library for Achilles, the analysis was positive
- For Achilles Web, the data was sampled. The algorithm has a mode that allows the user to upload only 10% of the existing data

Results

- Translation successful with minor errors and warnings



Results

- Observations (HiRID)

datetime	entertime	patientid	status	stringvalue	type	value	variableid
00:00.0	20:27.0	148	8			0	30005080
00:00.0	16:11.6	148	8			0	30005110
00:00.0	20:26.9	148	8			0	30005010
00:00.0	20:27.0	148	8			0	30005075
35:00.0	36:37.7	148	12			73	4000
35:00.0	36:37.6	148	12			98	200
36:00.0	37:37.7	148	8			94	4000
36:00.0	37:37.7	148	8			118	200
37:00.0	38:37.2	148	12			-2.5	211

- Observations (OMOP-CDM)

Data Output		Explain	Messages	Notifications														
	measurement_id [PK] bigint	person_id bigint	measurement_concept_id integer	measureme date	measurement_datetime timestamp without time zone	measurement_time character varying (16)	measurement_type integer	oper integer	value_as_number numeric	value integer	unit integer	range_low numeric	range_high numeric	prov bigint	visit_occurrence_id bigint	visit_date bigint	measurement_source_value character varying (255)	measurement_source_id integer
1	5598	148	36309521	2183-05-13	2183-05-13 08:34:02.35	08:34:02	42530833	[null]	0.0	[null]	[null]	[null]	[null]	[null]	1 [null]	[null]	0.0	210
2	4944	148	3002137	2183-05-13	2183-05-13 05:53:58.96	05:53:58	42530833	[null]	1846.8541	[null]	[null]	[null]	[null]	[null]	1 [null]	[null]	1846.8541	30005010
3	4681	148	3013502	2183-05-13	2183-05-13 04:51:07.09	04:51:07	42530833	[null]	97.0	[null]	[null]	[null]	[null]	[null]	1 [null]	[null]	97.0	4000
4	3887	148	36309521	2183-05-13	2183-05-13 01:46:02.24	01:46:02	42530833	[null]	0.5	[null]	[null]	[null]	[null]	[null]	1 [null]	[null]	0.5	210
5	617	148	36309521	2183-05-12	2183-05-12 13:43:00	13:43:00	42530833	[null]	0.6	[null]	[null]	[null]	[null]	[null]	1 [null]	[null]	0.6	211
6	2378	148	21499295	2183-05-12	2183-05-12 19:54:20.14	19:54:20	42530833	[null]	0.0	[null]	[null]	[null]	[null]	[null]	1 [null]	[null]	0.0	30005080
7	1167	148	3013502	2183-05-12	2183-05-12 14:42:00	14:42:00	42530833	[null]	96.0	[null]	[null]	[null]	[null]	[null]	1 [null]	[null]	96.0	4000
8	793	148	36309521	2183-05-12	2183-05-12 14:02:00	14:02:00	42530833	[null]	0.7	[null]	[null]	[null]	[null]	[null]	1 [null]	[null]	0.7	211
9	1000	148	36309521	2183-05-12	2183-05-12 13:43:00	13:43:00	42530833	[null]	0.5	[null]	[null]	[null]	[null]	[null]	1 [null]	[null]	0.5	210

Results

- Pharma Records (HiRID)

datetime	entertime	patientid	status	stringvalue	type	value	variableid
00:00.0	20:27.0	148	8			0	30005080
00:00.0	16:11.6	148	8			0	30005110
00:00.0	20:26.9	148	8			0	30005010
00:00.0	20:27.0	148	8			0	30005075
35:00.0	36:37.7	148	12			73	4000
35:00.0	36:37.6	148	12			98	200
36:00.0	37:37.7	148	8			94	4000
36:00.0	37:37.7	148	8			118	200
37:00.0	38:37.2	148	12			-2.5	211

- Drug_exposure (OMOP-CDM)

	drug_exposure_id [PK] bigint	person_id bigint	drug_concept_id integer	drug_exposure_start_date	drug_exposure_start_datetime timestamp without time zone	drug_exposure_end_date	drug_exposure_end_datetime timestamp without time zone	verbatim_end_date	drug_type_concept_id integer	stop_reason character varying	refills integer	quantity numeric	days_supply integer	sig text
1	10559	2470	40953611	2155-04-23	2155-04-23 04:11:11.71	2155-04-23	2155-04-23 04:11:11.71	2155-04-23	38000180	0	483	4.166664	0	0
2	24177	11317	1154029	2158-12-10	2158-12-10 07:27:21.62	2158-12-10	2158-12-10 07:27:21.62	2158-12-10	38000180	0	129	2.430555	0	0
3	65263	29217	43712853	2187-02-11	2187-02-11 21:21:32.48	2187-02-11	2187-02-11 21:21:32.48	2187-02-11	38000180	0	16	2.65	0	0
4	15891	6586	35130768	2133-07-31	2133-07-31 07:21:51.1	2133-07-31	2133-07-31 07:21:51.1	2133-07-31	38000180	0	1	0.0	0	0
5	16356	7638	40953611	2150-09-10	2150-09-10 06:13:46.56	2150-09-10	2150-09-10 06:13:46.56	2150-09-10	38000180	0	67	04.62963	0	0
6	53266	24171	19036476	2113-07-16	2113-07-16 08:00:00	2113-07-16	2113-07-16 08:11:25.683	2113-07-16	38000180	0	0	60.0	0	0
7	50899	21805	40953611	2125-09-04	2125-09-04 10:11:35.43	2125-09-04	2125-09-04 10:11:35.43	2125-09-04	38000180	0	19	00000001	0	0
8	363	307	40953611	2128-02-17	2128-02-17 17:46:17.96	2128-02-17	2128-02-17 17:46:17.96	2128-02-17	38000180	0	14	4.166664	0	0

Observations

- The number of rows for observation dataset is in the 100 thousands, uploading 1 group from jupyter dataset to postgres takes around 3 minutes
- With the current methods applied the total upload time will take 21 hours
- The ETL is a cyclical process, there were many iterations before getting the final version of the algorithm
- The table `observational_period` was created because there was an observation from Achilles web
- There are fields that are related with procedures like chemotherapy, they have to be mapped as procedures, thus creating a new table
- The date of birth transformation is right. However, the Achilles json file only shows patients over 40. More research has to be done,

Next Steps

- Publish the algorithm to the OHDSI community and anyone interested
- Upload the data to the cloud and measure the performance.
Currently the work was done locally in a Laptop(Core i7, 32GB of RAM)