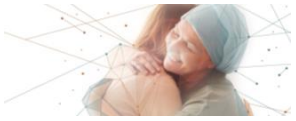


WHITERABBIT

ABOUT WHITERABBIT

Links:	https://github.com/OHDSI/WhiteRabbit
License:	Apache License 2.0
Version:	0.9.0
Last Update:	12/31/2019
OS:	Windows, macOS, Linux
System Requirements:	Requires Java 1.8 or higher and read access to the database to be scanned.

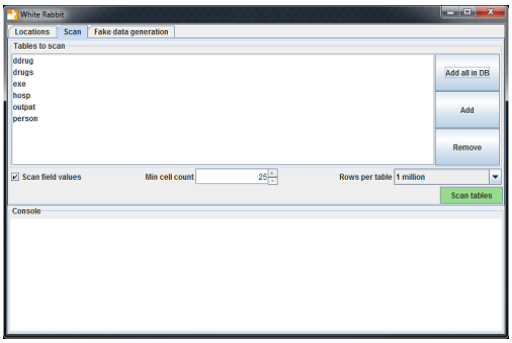
Description:

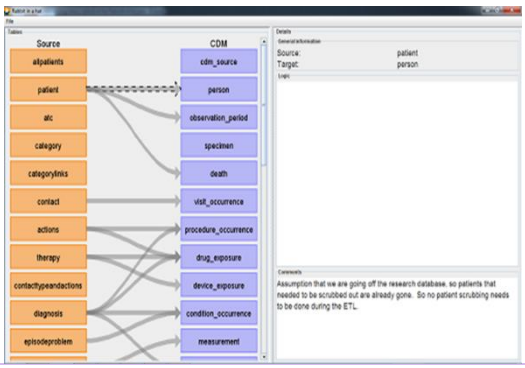
WhiteRabbit is an application that can be used to analyze the structure and contents of a database as preparation for designing an ETL. It comes with RabbitInAHat, an application for interactive design of an ETL to the OMOP Common Data Model with the help of the scan report generated by White Rabbit. WhiteRabbit's main function is to perform a scan of the source data, providing detailed information on the tables, fields, and values that appear in a field. This scan will generate a report that can be used as a reference when designing the ETL, for instance by using the Rabbit-In-a-Hat tool. White Rabbit differs from standard data profiling tools in that it attempts to prevent the display of personally identifiable information (PII) data values in the generated output data file.

Features:

- Can scan databases in SQL Server, Oracle, PostgreSQL, MySQL, MS Access, Amazon RedShift, Google BigQuery, SAS files and CSV files
- The scan report contains information on tables, fields, and frequency distributions of values
- Cutoff on the minimum frequency of values to protect patient privacy
- WhiteRabbit can be run with a graphical user interface or from the command prompt
- Interactive tool (Rabbit in a Hat) for designing the ETL using the scan report as basis
- Rabbit in a Hat generates ETL specification document according to OMOP template

Screenshots:



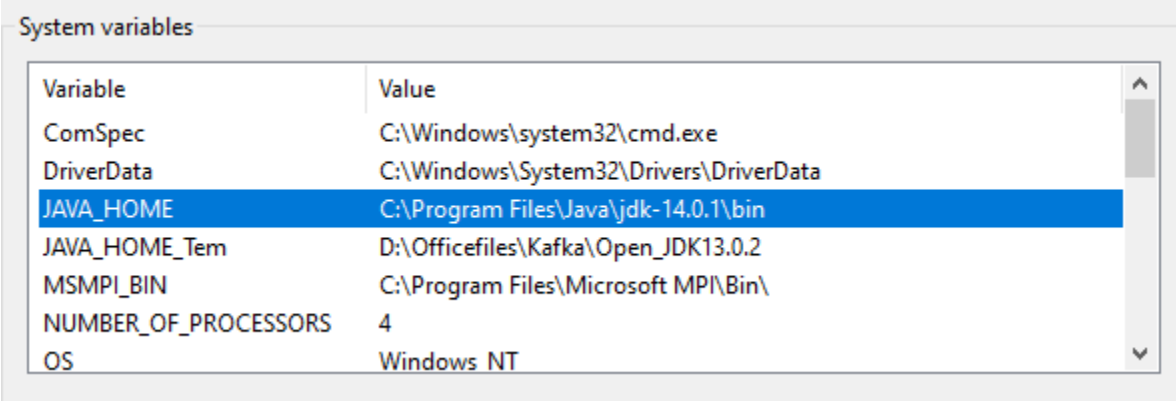


WHITE RABBIT:

1. Download the WhiteRabbit tool from "https://github.com/OHDSI/WhiteRabbit/releases"
2. Browse to the White Rabbit bin folder and start the "WhiteRabbit.bat".
3. Select the "Scan" tab and load the CSV files by clicking the "Add" button.
4. Browse to the Synthea CSV files and select the files which is needed.
5. Press the "Scan Tables" to generate the "ScanReport.xlsx" in the same bin folder.

PREREQUISITES:

Install Oracle JDK latest version from "https://www.oracle.com/java/technologies/javase-downloads.html". After installing, set the "JAVA_HOME" environment variable as shown in the below image,



RABBIT IN HAT:

1. In the same bin folder, start the "RabbitInAHat.bat".
2. From the "File" menu click the "Scan Report" menu item.
3. Select the "ScanReport.xlsx" file which was generated by the White Rabbit.
4. Map the source attributes from the CSV to the OMOP CDM target table attributes (Shown below in the screenshot).
5. In the "File" menu, select "Generate SQL Skeleton" menu item for creating the SQL scripts.

SSMS:

1. In SQL Server Management Studio, load the CSV files by right clicking on the database and selecting "Tasks -> Import Flat File". Load the required CSV files into the SQL Tables.
2. Execute the SQL files generated from the "RabbitInAHat" tool.

Note: There could be some data type conversion errors which needs to adjusted in the generated SQL Script.

Sample Source Data Mapping Approach to CDMV6.0