"Perseus": Design and run your own ETL to CDM

Anton Ivanov, MS[1], Samus Sergey, MS[1], Alexander Efimov, MS[1], Maxim Draschinsky, MS[1], Dolotova Maria, MS[1], Matvey Chudakov, MS[1],
Karpushina Anna, MS[1], Semenov Artem, MS[1]

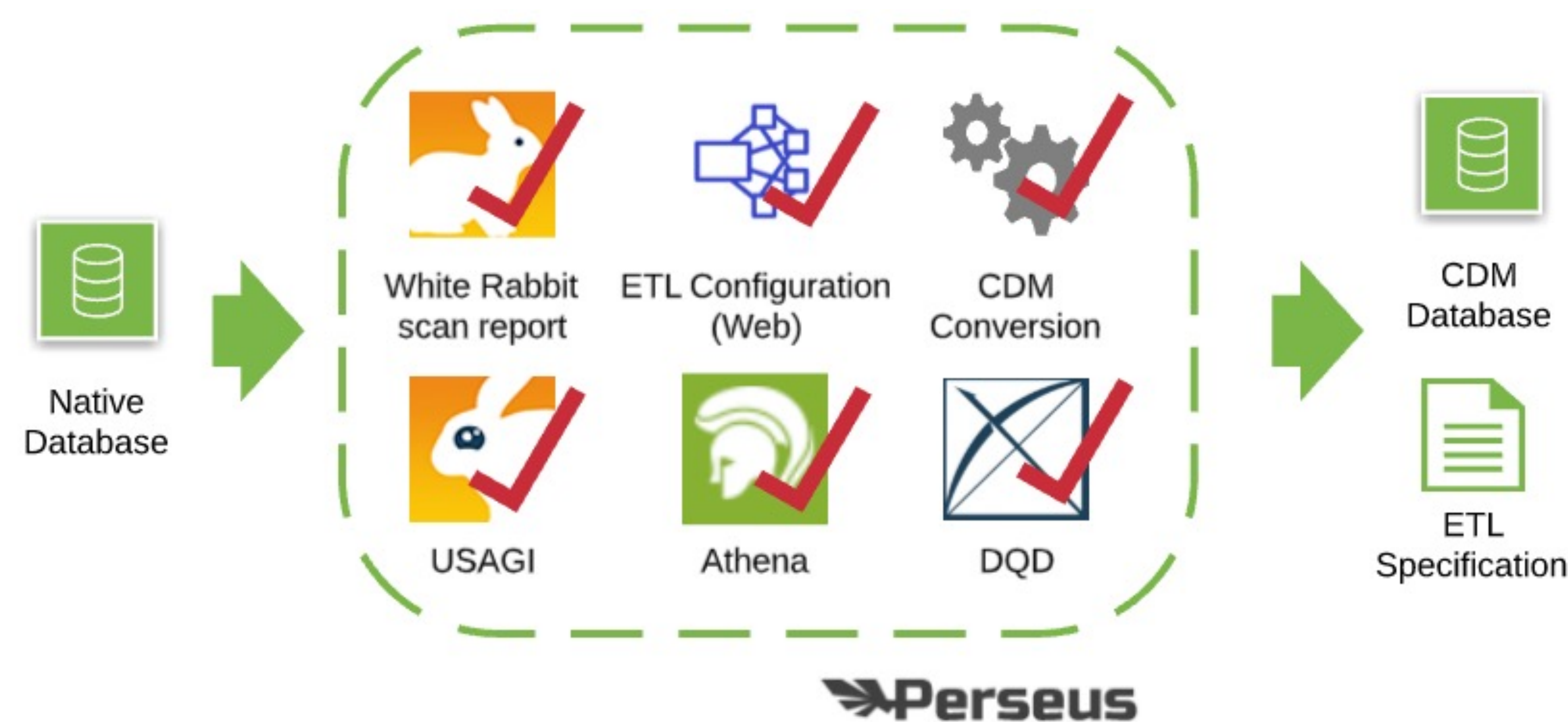[1]Arcadia Inc., St. Petersburg, Russia

Background

There are several opensource ETL tools designed for CDM and available for OHDSI community. Tools for code mapping and quality control provide a good starting point to design ETL process for getting CDM but ETL implementation still requires technical skills in programming languages and platforms. Presented solution addresses this need and provides a visual designer for ETL implementation targeting CDM.
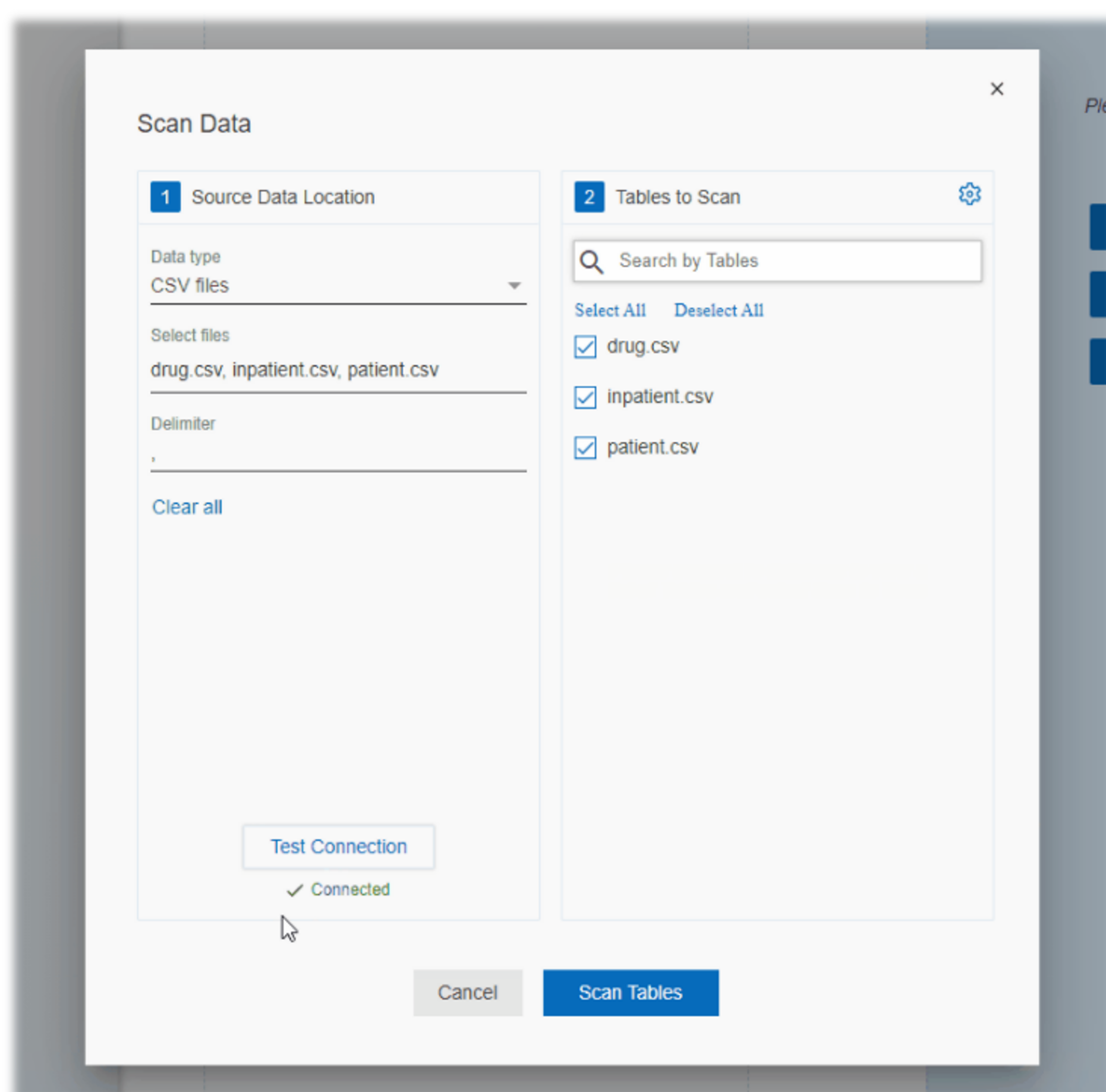
Methods

"Perseus" is combination of Web application for ETL configuration and engine for: conversion of native data into CDM, data quality check, code mapping, generating ETL document and vocabulary search.
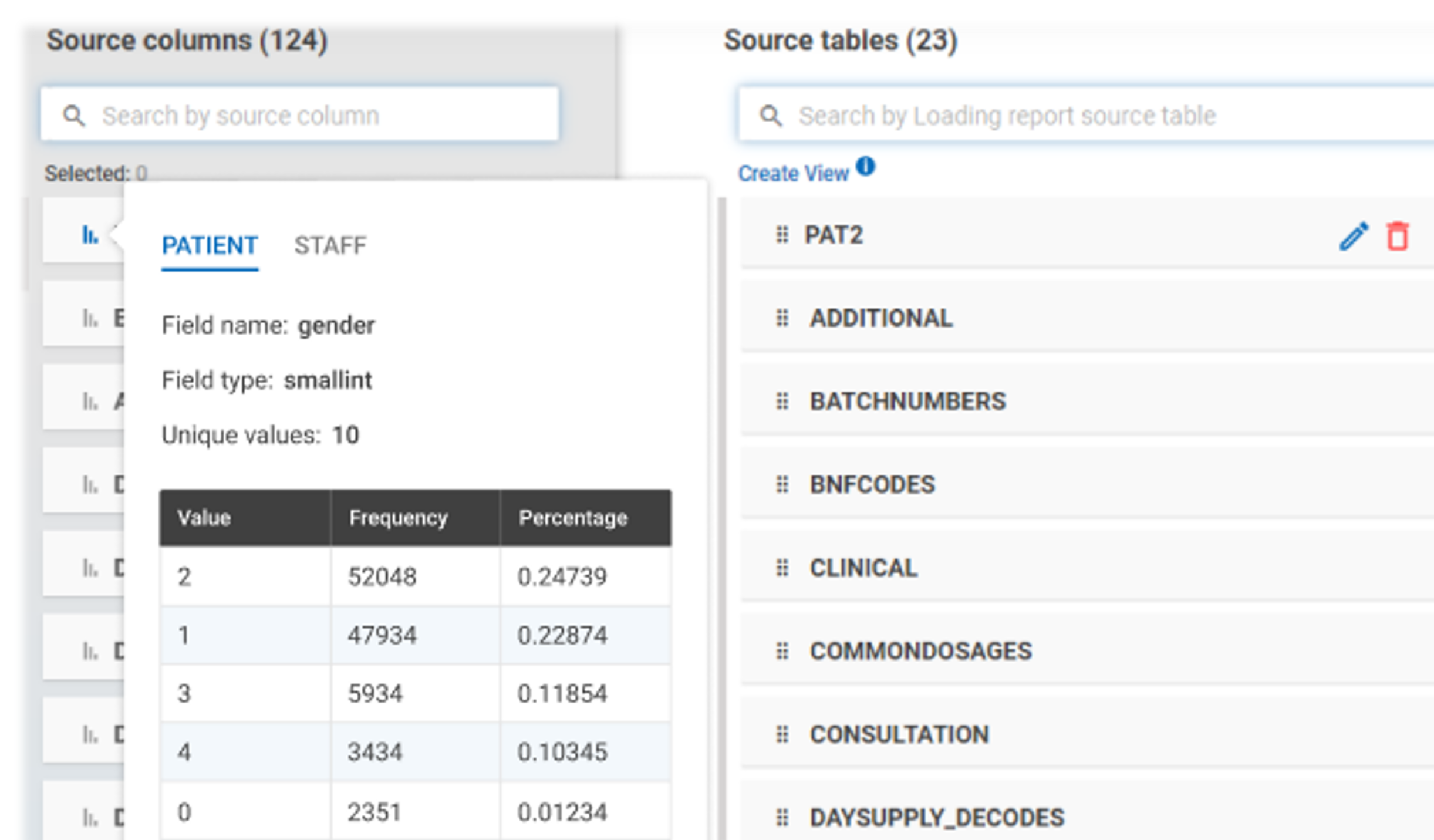

"Perseus" has instruments for all major steps of creating an ETL.
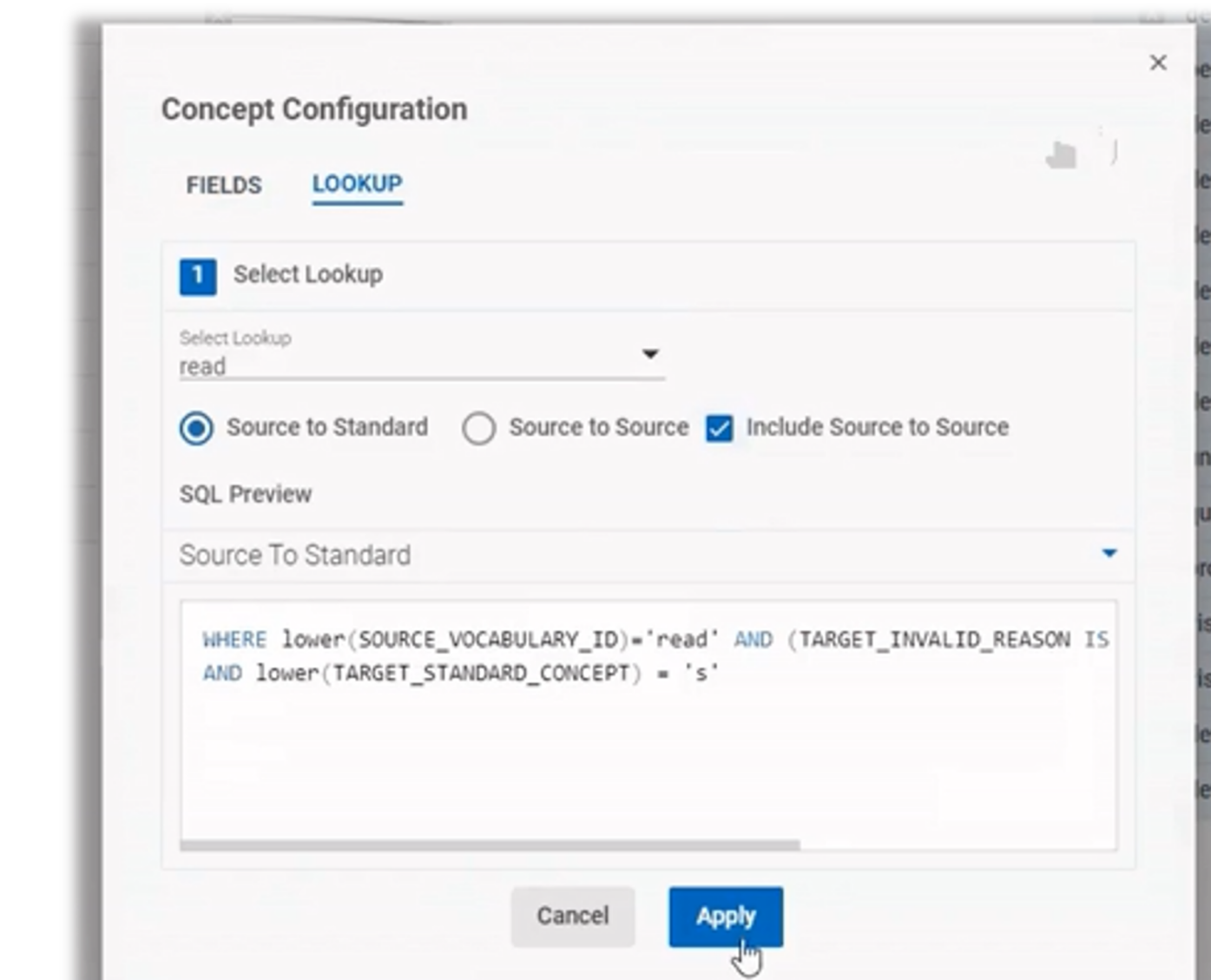
Step 1: Scan source data

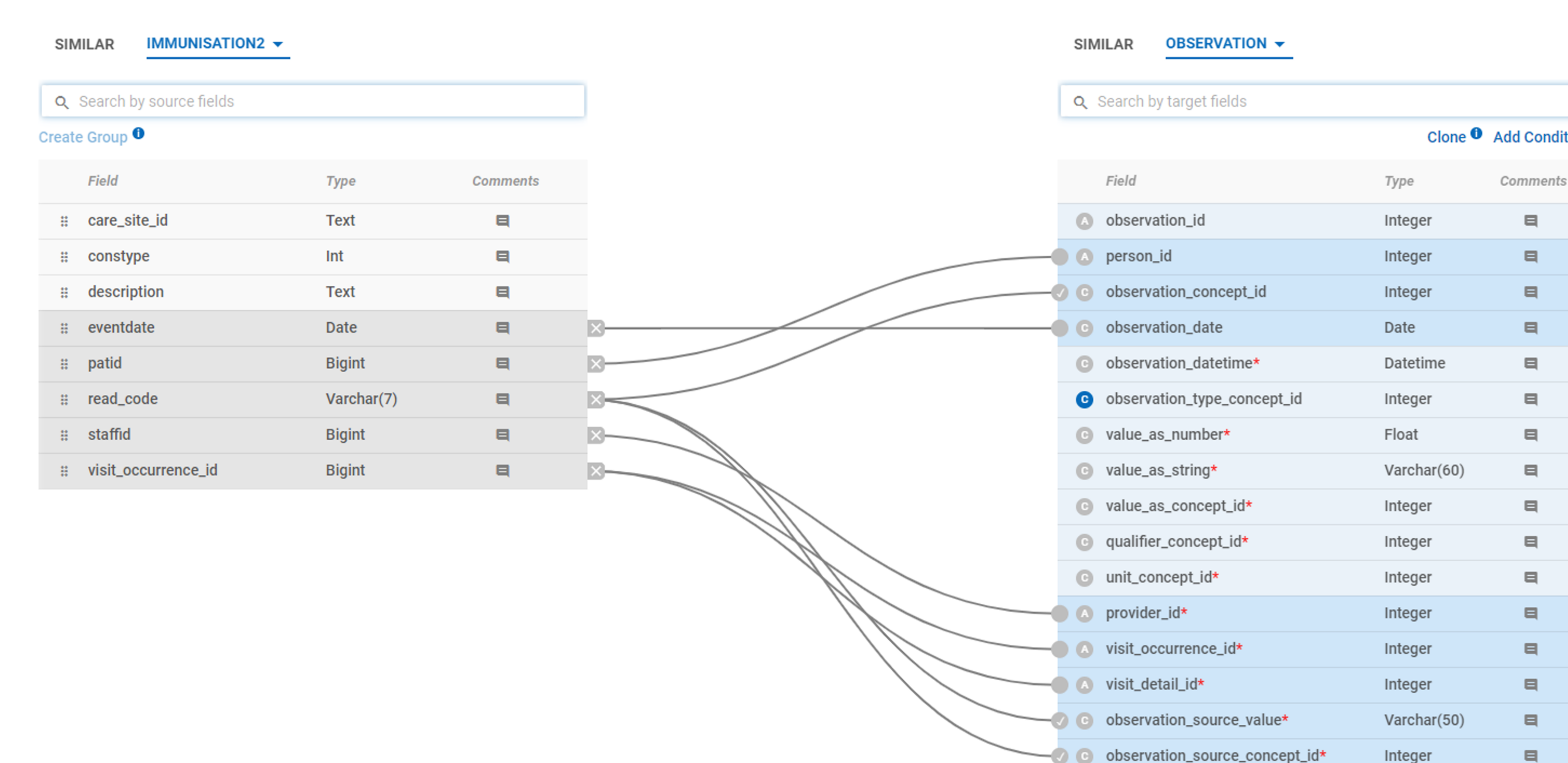✓ Use White Rabbit scan report
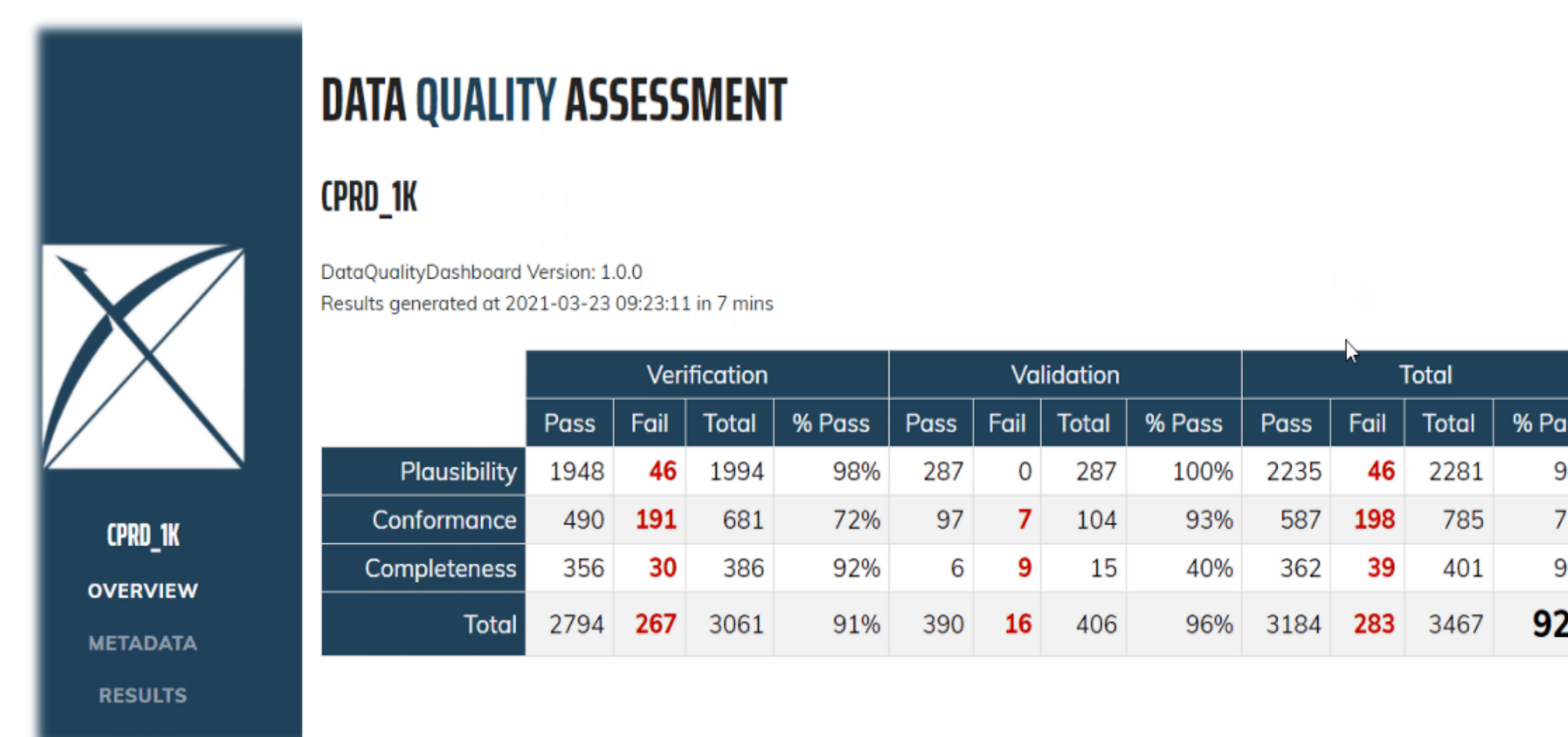✓ or scan the source data directly

Step 2: Create the Code Mappings

Step 3: Design and run ETL

✓ Visually map raw data to CDM
✓ Embedded set of transformations
✓ Embedded set of lookups
✓ Combine source tables
✓ Combine fields into groups
✓ Configure conditional mappings
✓ Automatic era creation
✓ Automatic domain switching

Step 4: Quality Control

✓ Integrated Data Quality Dashboard

Conclusions

"Perseus" reduces the time spent on manual processes such as writing code and mapping source data to target systems. The configuring and running ETLs with "Perseus" makes it repeatable, shareable and saves time.

References:
1. Perseus - https://github.com/SoftwareCountry/Perseus
2. White Rabbit, and Rabbit-in-a-Hat https://github.com/SoftwareCountry/WhiteRabbit
3. Usagi - https://github.com/OHDSI/Usagi
4. Achilles - https://github.com/OHDSI/Achilles
5. DQD - https://github.com/SoftwareCountry/DataQualityDashboard
6. CDM Builder - https://github.com/SoftwareCountry/ETL-CDMBuilder

Contact: info@softwarecountry.com