



# Addressing biomedical research needs with the Data2Evidence platform

DATA4LIFE.CARE

An initiative by the  
**HASSO PLATTNER**  
FOUNDATION  
Public 1

# About Data4Life

## Mission-driven and Collaborative.

In collaboration with leading domain experts, we digitalize health research to advance public health & personalized medicine.

## Non-profit.

We are 100% nonprofit, funded by the Hasso-Plattner Foundation

## Open source.

We open source our software to make it accessible to everyone.



# Our co-created solutions address health researchers' specific needs

**D2E** DATA2EVIDENCE.

**D4L** COLLECT.

Sensor Data Collection  
Platform

Easy and efficient  
collection of data from  
sensors, wearables and  
questionnaires

**D2E** INTEGRATE.

Health Data Management  
Solution

Solutions to extract and  
combine data into one  
ready-to-use interoperable  
dataset

**D2E** ANALYZE.

Health Data Analytics  
Suite

Direct and collaborative  
analysis of data to generate  
insights and drive evidence-  
based medicine

**D4L** FAIR.

Health Research  
Metadata Catalog

Easy identification of  
datasets suitable for  
secondary research

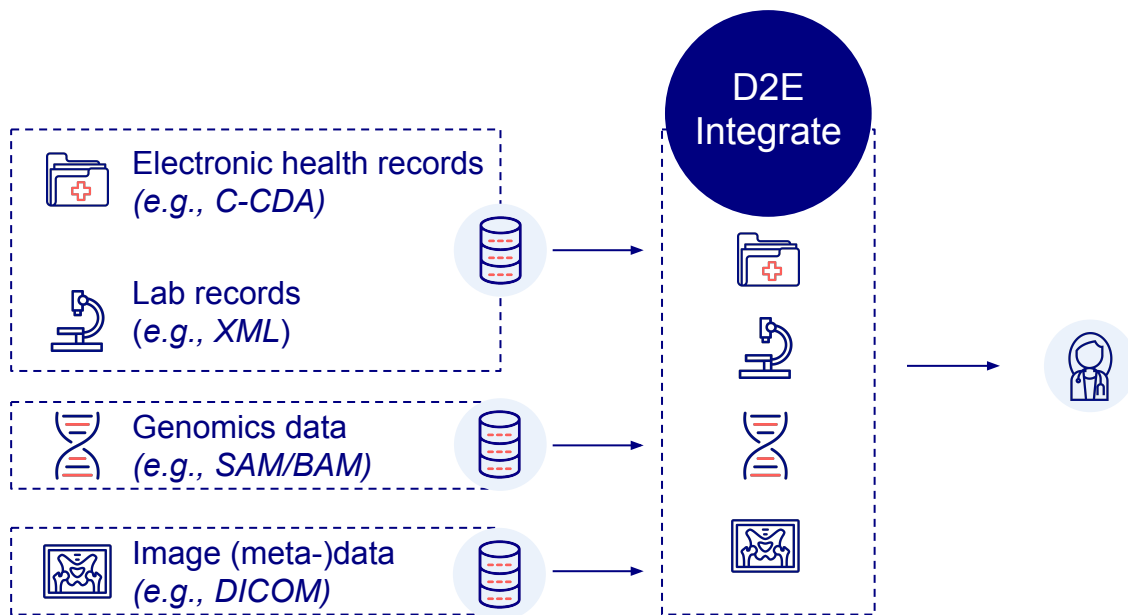
# What is Data2Evidence?

An end to end solution for healthcare data management and analysis to make siloed data readily available for researchers



# D2E Integrate

Data is integrated in a harmonized dataset

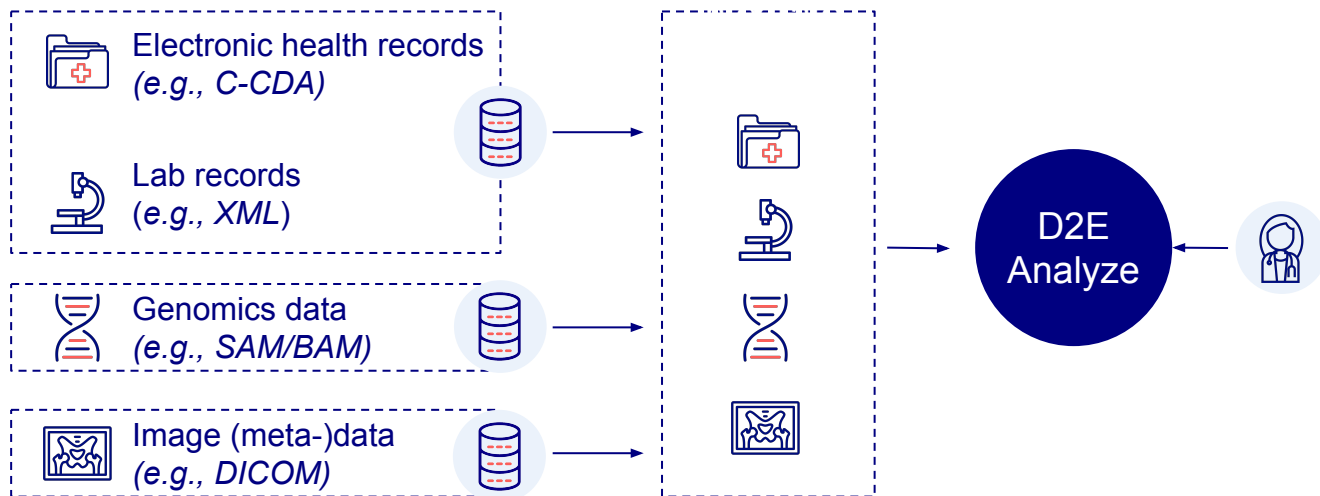


## Features

- Tooling to extract, transform & load data
- Data storage in interoperable format with continuous updates
- Tooling for quality checks included
- Standardized and streamlined access requests
- Integrated user identification working with existing sign-ons

# D2E Analyze

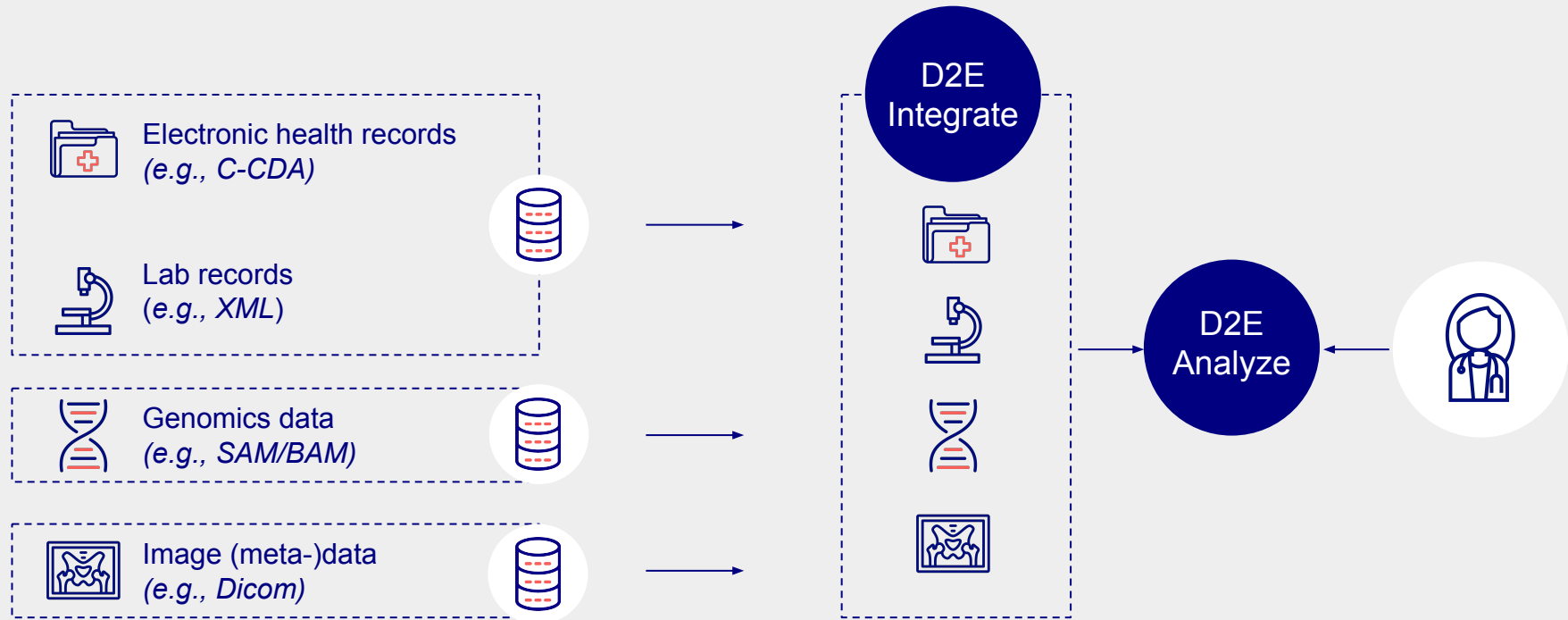
Access and analyze datasets to generate insights



## Features

- Tooling for data exploration
- Interactive, no code cohort building tooling
- Tooling for cohort quality checks
- Interactive notebooks and integration of Azure ML / existing HPCC
- Easy collaboration on analysis and data subsets

# With Data2Evidence real world data is readily accessible as one harmonized dataset



# D2E in the data lifecycle - Data Owner

DESIGN

IMPLEMENTATION

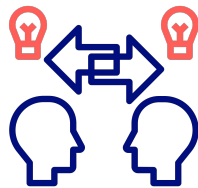
QUALITY ASSESSMENT



Medical experts

D2E Concept Mapping

Help to ease the code mapping from source data to OMOP vocabularies.



Data experts and CDM experts

D2E Data Mapping

To provide a better insight about the structure of the data, the values and the frequencies and help to identify how information is recorded.

Create field mappings and generate ETL documentation specification.



Technical experts

D2E Data Integration

ETL tool with built-in data transformation capabilities and connectors to different sources.



Quality Controller

D2E DQD & DC

Quality assessment platform with pre-programmed rules and visualization of observational health databases.



# D2E in the data lifecycle - Data Owner



**Schema evolution** Manages changes to the data model, with migration scripts



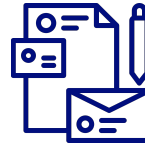
**Access management** Fine granular access control to datasets



**Metadata** Store additional information and metadata about each dataset



**Dataset version** manage multiple versions of the dataset leveraging history tables



**Datamart creation** create subset of the data based on selecting cohorts and specific attributes of the data

# D2E in the data lifecycle - Researcher

DATA SELECTION



Clinical researcher

Portal

Quick look at datasets available, their statistical characteristics, their metadata, download sample dataset, and request for access.

CONCEPTS AND COHORTS



Clinical researcher and data scientists

Cohort Manager/Explorer

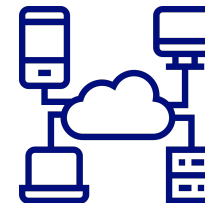
Interactively explore dataset in depth

Build a reusable set of concepts to be used in research

Search and add concepts to set, get recommendations for similar concepts to enrich concept set

Create a cohort using concepts and concept sets using inclusion, exclusion criteria, boolean logic, and temporal relationships

ANALYSIS



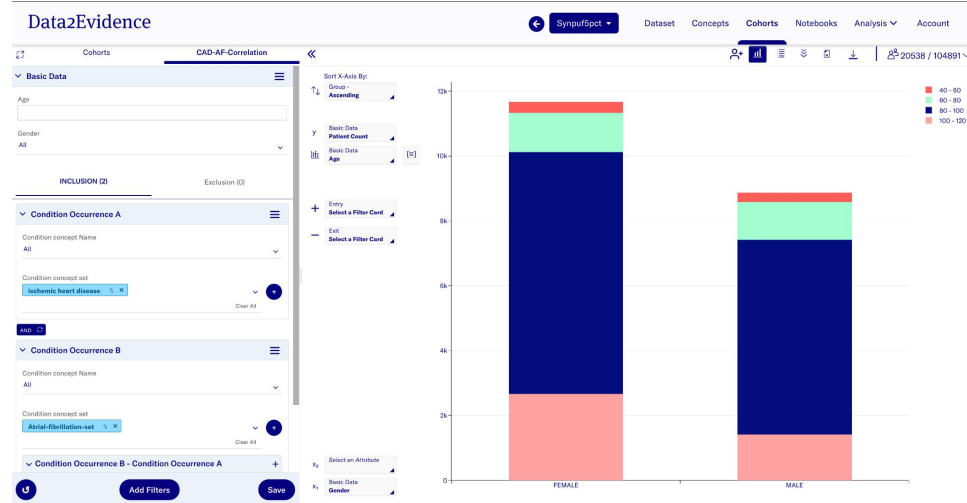
Data Scientist

D2E Plugins and Notebooks

Perform analysis on cohorts using pre-defined analysis or creating new analysis

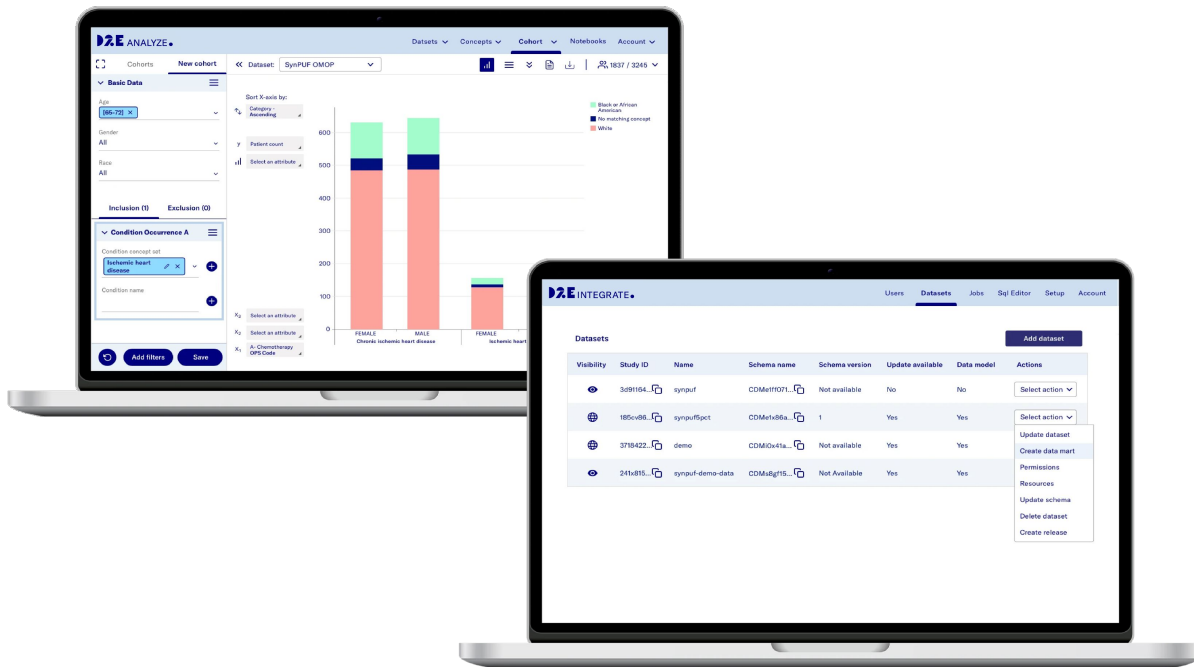
# Cohort Manager/Explorer (Patient Analytics)

- Easy to use query tool for researchers to explore the data and perform phenotyping
- Developed specifically for the clinical research domain
  - Supports any data model, incl. OMOP CDM
- Visualize aggregated / non-aggregated data
- Export data for further analysis



# Your research. Our platform.

## **D2E** DATA2EVIDENCE. is open source!



Get started: [www.d2e.sg](http://www.d2e.sg)

Github repo:  
<https://github.com/OHDSI/D2E>

# Thank you.

Dr. Karthik Seetharaman  
Lead PM (D2E and FAIR)

[karthik.seetharaman@data4life-asia.care](mailto:karthik.seetharaman@data4life-asia.care)

