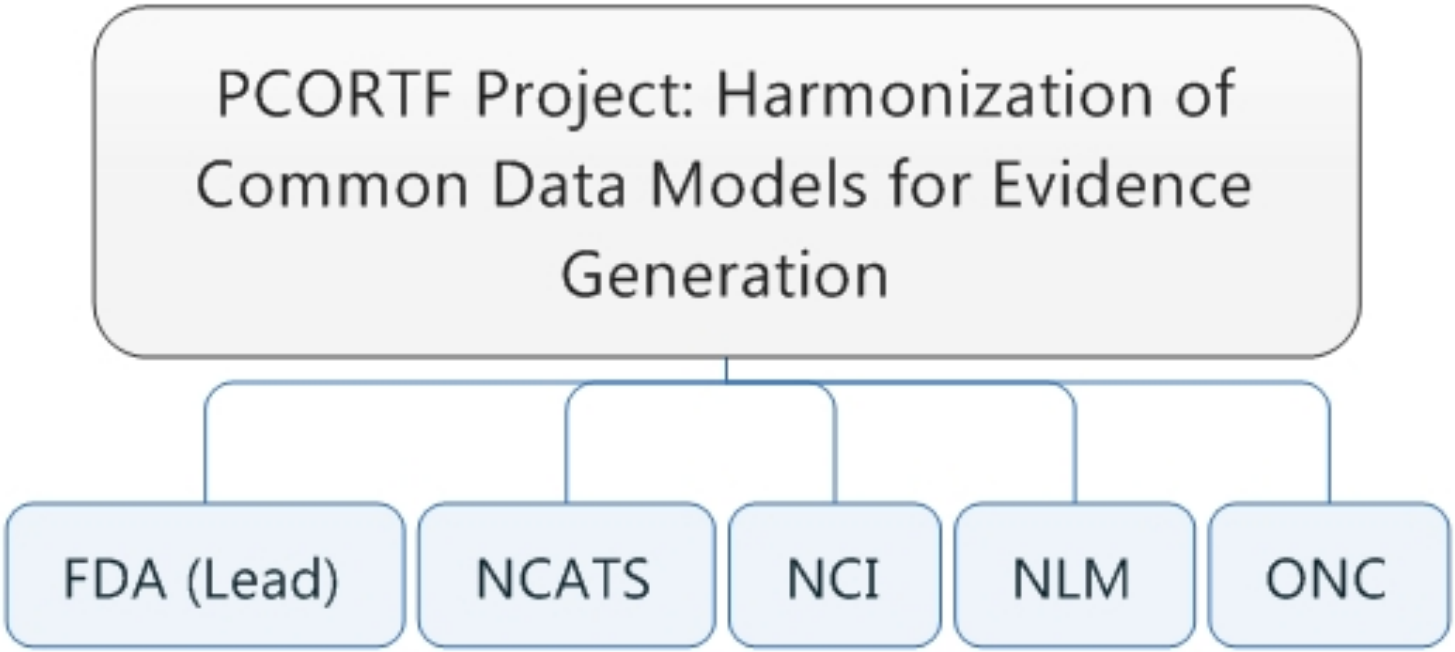
PCORTF: The Common Data Model Harmonization Project



Contents

- Current State and Desired State of RWD for FDA
- Phase I: Common Data Model Harmonization Project
- Phase II: Pluri-Potent Data Model

Real World Data Uses for Clinical Trials

- Rapidly find large numbers of possible participants with all requirements for clinical trial study
  - Health history
  - Treatment history
  - Medicine history
  - Range of demographics
  - Range of geographies and rural/urban
- Rapidly identify site/investigators
- Savings of time and money for manufacturers, clinical investigators, and other stakeholders
- Potential to identify additional “off-label” uses for approved drugs

Current State for Real World Data (RWD)

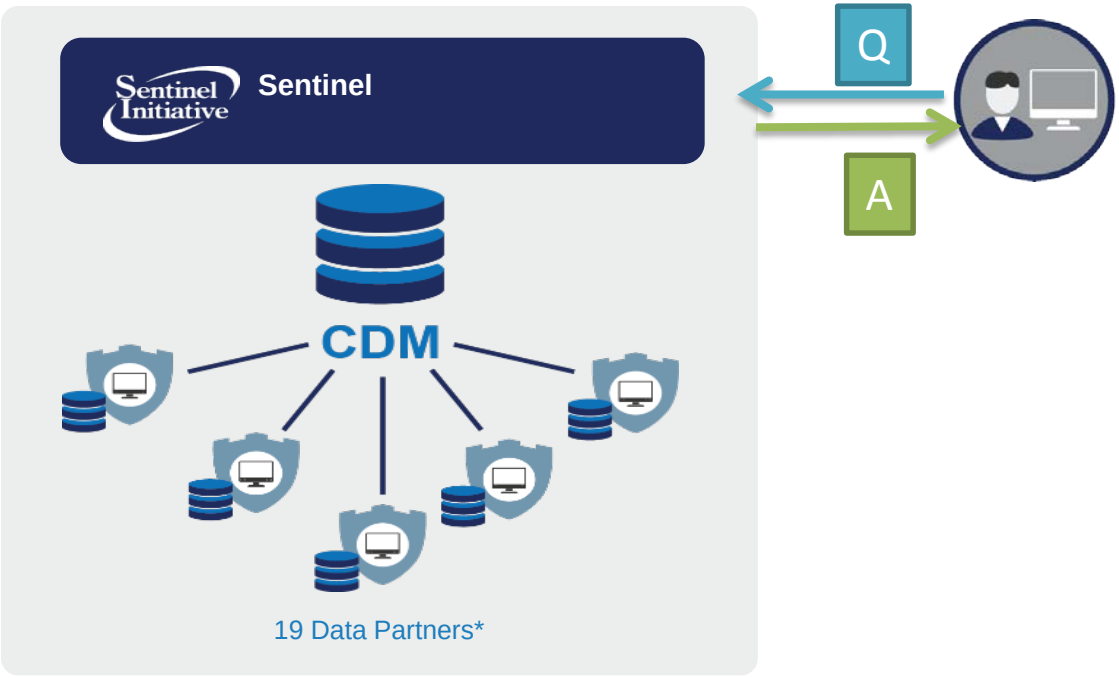
- Thousands of real-world data systems (EHRs, claims data, etc.)
- Limited data connection to FDA research systems
- Cannot effectively conduct large scale RWD research
- Cannot get aggregate different RWD CDMs
- Cannot get timely access to RWD CDMs

Ideal State for Real World Data

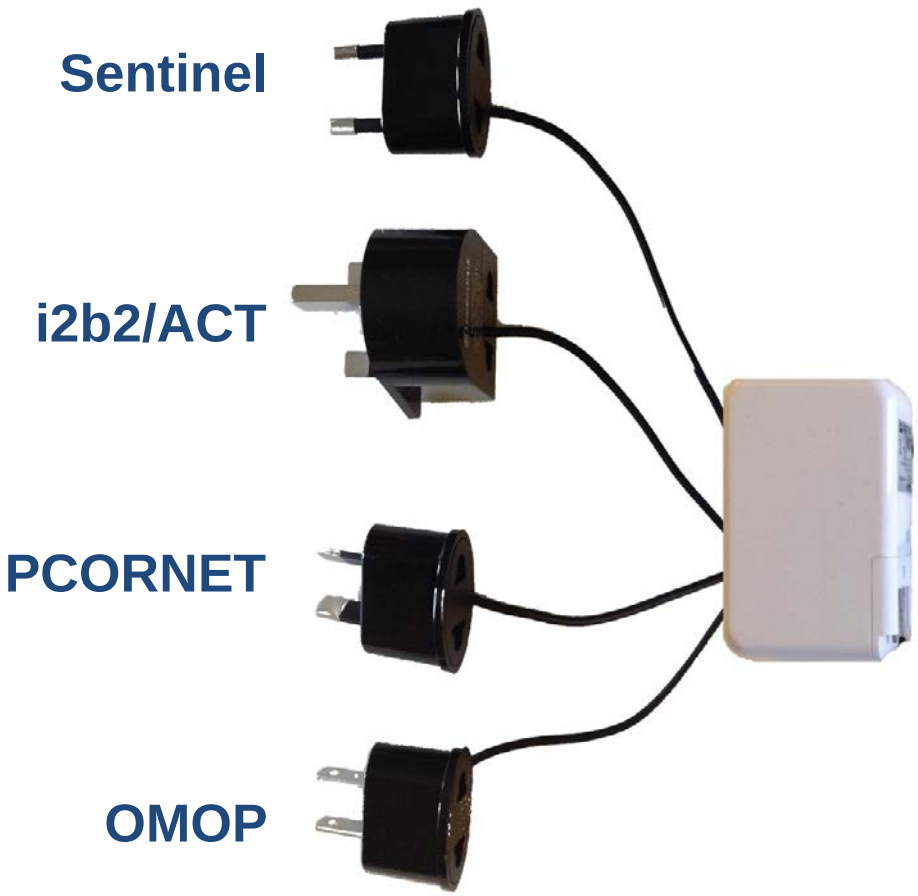
- Real-time access to real world data systems accessible to FDA researchers:
  - Secure
  - Privacy protected / de-identified
  - Sources retain ownership and control of raw data
  - Crosses demographics, geography, etc.
- Mechanisms to get rapid answers to research questions from pools of millions of patients
  - Without risk of any privacy breach
  - Results are de-identified

Current State: Multiple RWD Networks

- Large RWD Networks (PCORNet, Sentinel, OMOP, i2b2/ACT)
  - Use different data models with limited interoperability
  - Network have some unique and some overlapping patients populations
  - No single portal to query all networks
  - Inability to aggerate results data across networks



The solution CDM 'Adapter Model'



- Use a converter between various adapters.
- Allow researchers to ask a question once and receive results from many different sources using a common agreed-upon standard structure, or a Common Data Model.

PCORTF CDM Harmonization Project

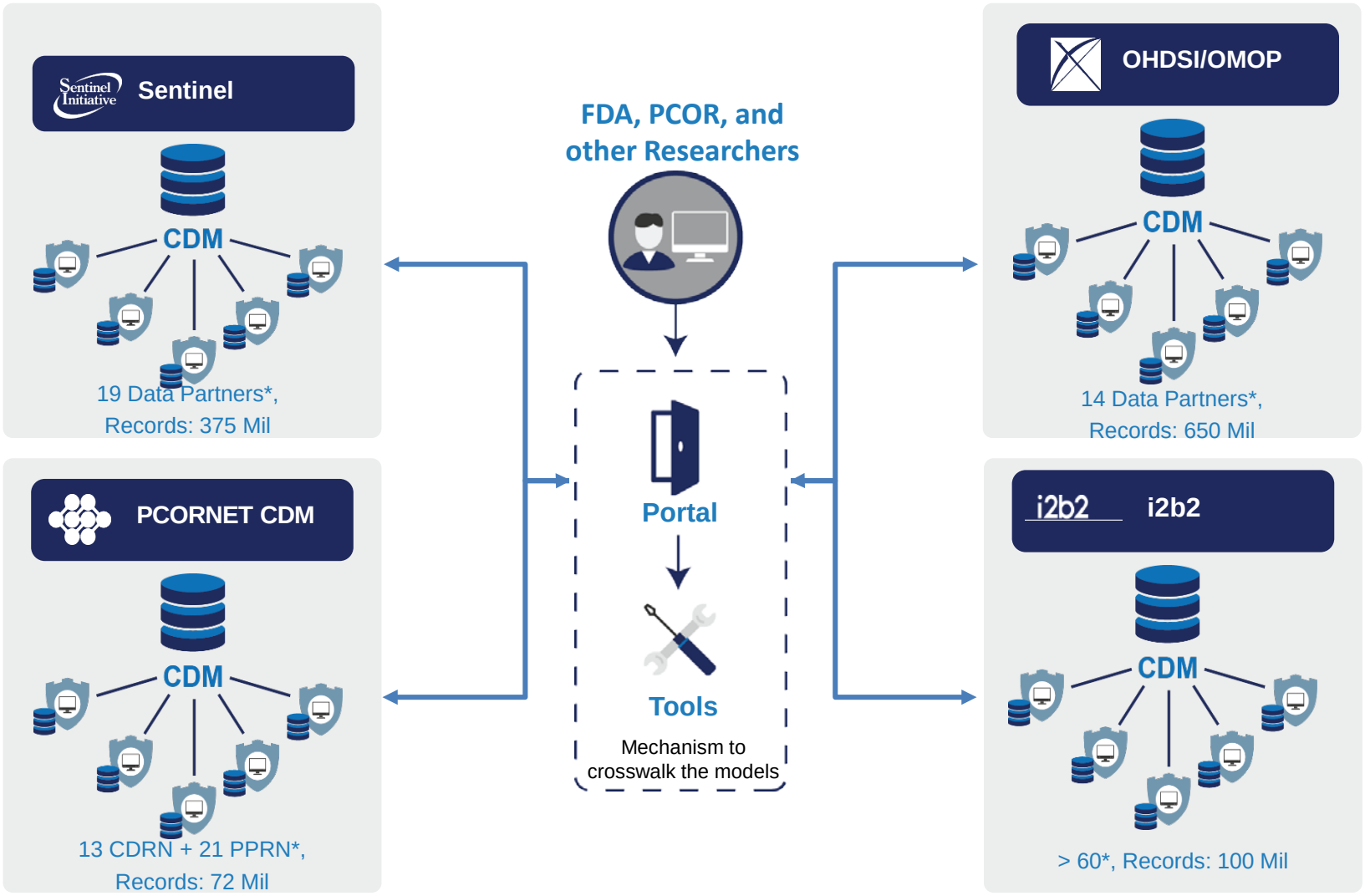
Goal:

Build a data infrastructure for conducting research using Real World Data derived from the delivery of health care in routine clinical settings.

Objective:

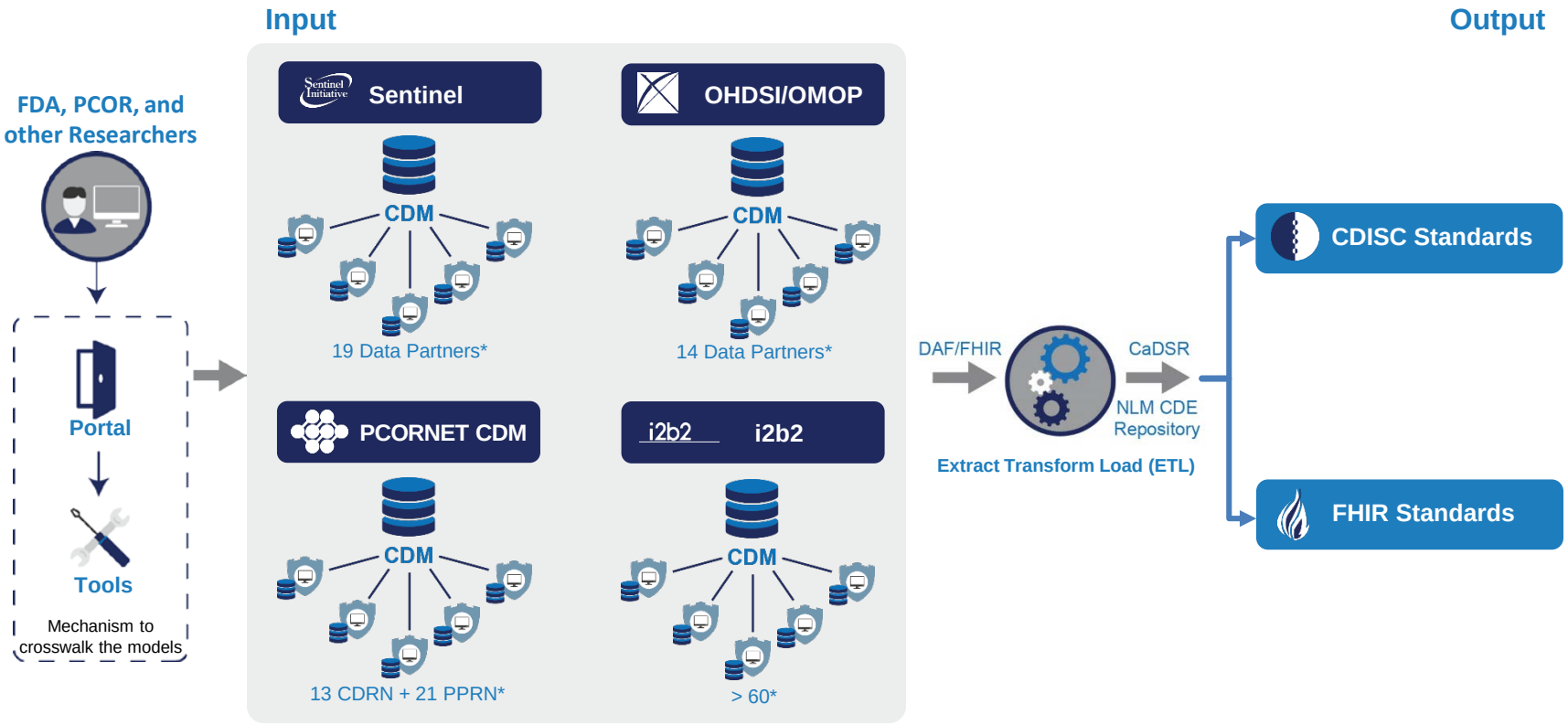
Develop the method to harmonize the Common Data Models of various networks, allowing researchers to simply ask research questions on much larger amounts of Real World Data than currently possible, leveraging open standards and controlled terminologies to advance Patient-Centered Outcomes Research.

The Concept

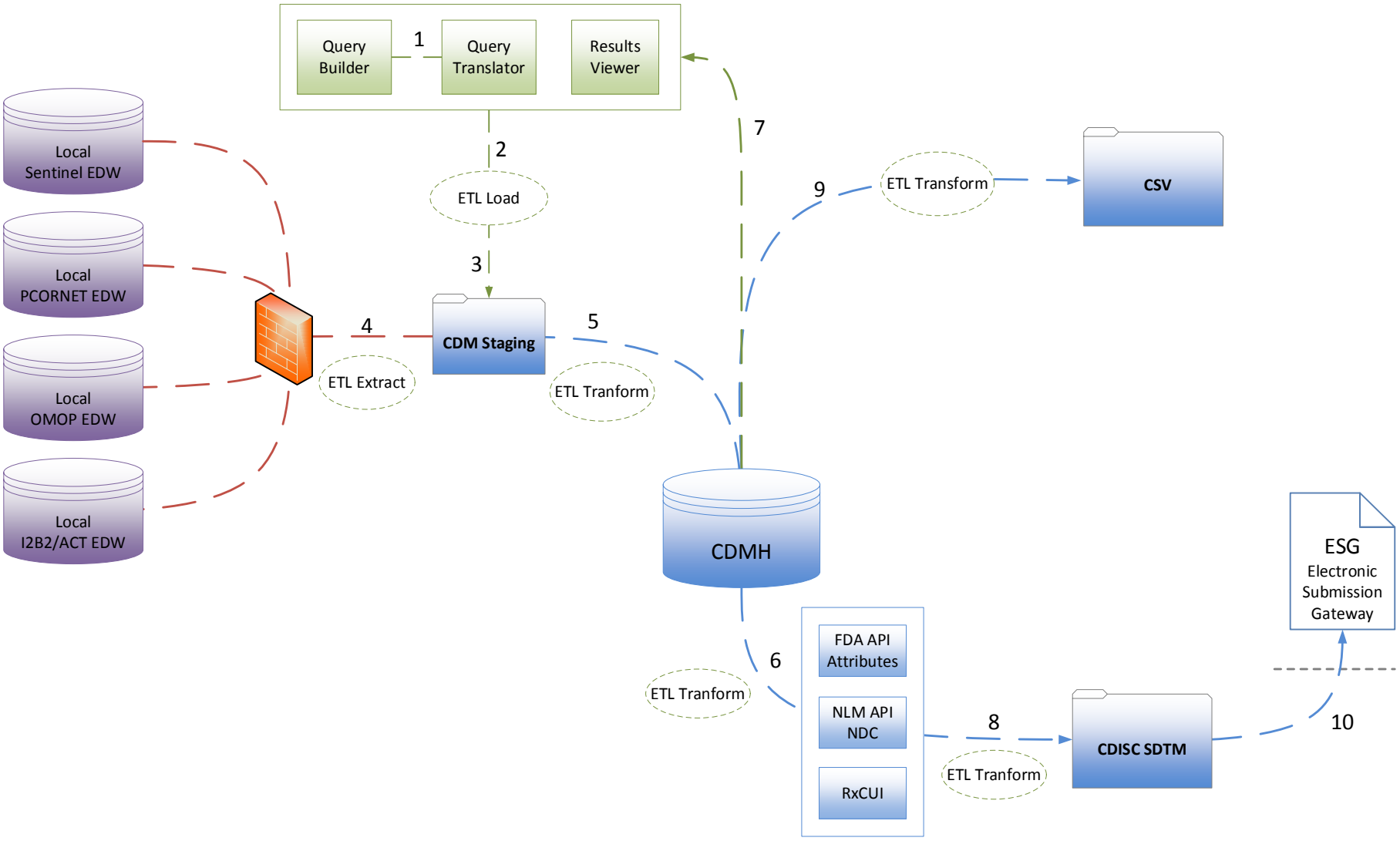

Sentinel
CDM
19 Data Partners*,
Records: 375 Mil
PCORNET CDM
CDM
13 CDRN + 21 PPRN*,
Records: 72 Mil
FDA, PCOR, and
other Researchers
Portal
Tools
Mechanism to
crosswalk the models
OHDSI/OMOP
CDM
14 Data Partners*,
Records: 650 Mil
i2b2
i2b2
CDM
> 60*, Records: 100 Mil

Extended Solution

Develop a general framework (i.e., tools, processes, policies, and standards) for the transformation of various CDMs, curation, maintenance and sustainability.



Phase I: Adaptor Model



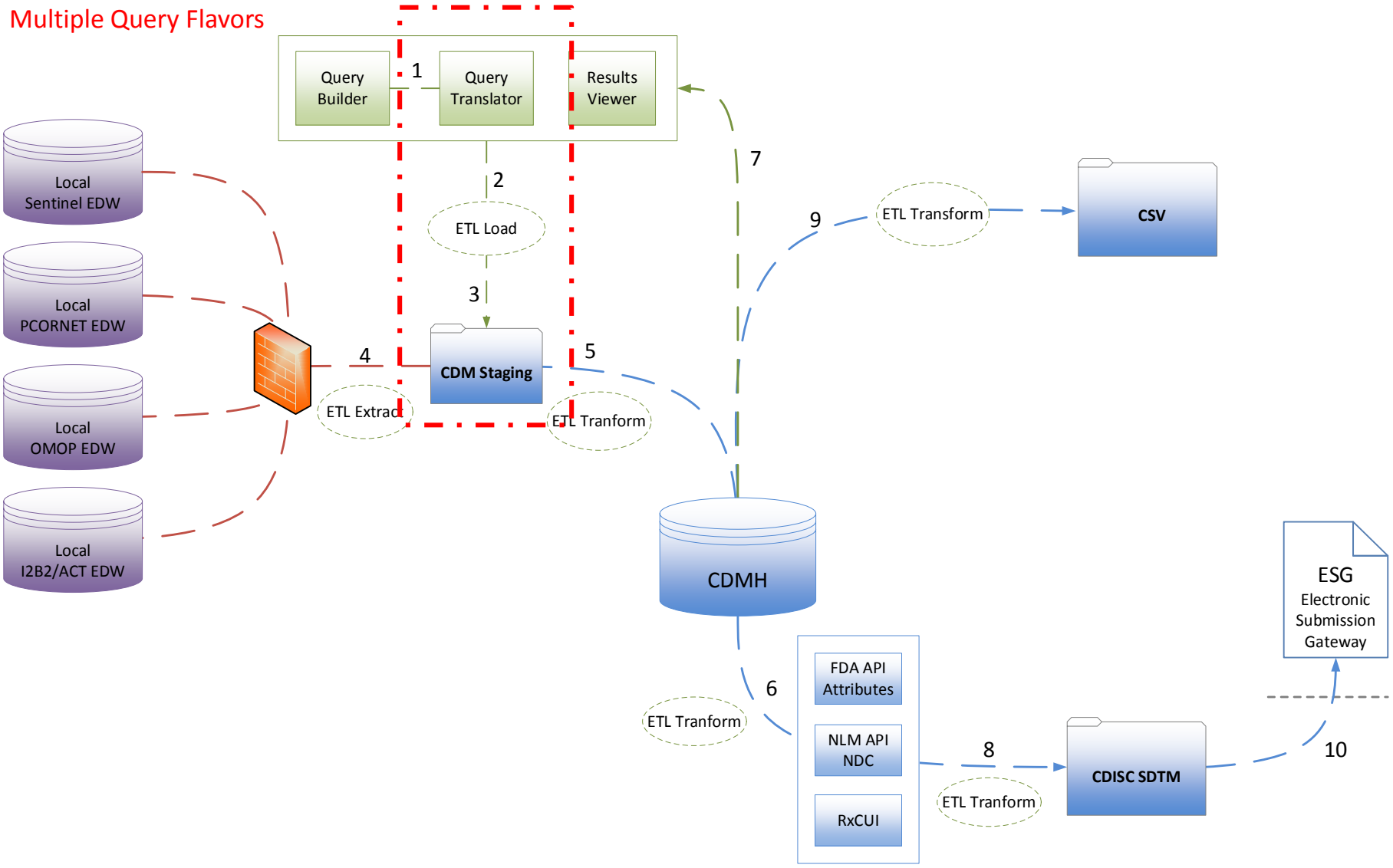
Phase I: Adaptor Model Deliverables

Software development
Query Builder: Gives researcher ability to construct their queries
Query Transformation: Transforms the single query into the multiple CDM versions of the query
CDMH Results Database and Viewer: To receive and analyze results of a query in one or more of the CDM formats
SDTM export: To export record level results into FDA-compliant STDM format.

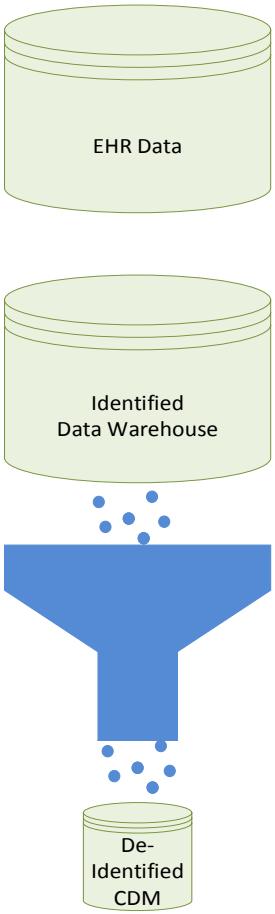
Transformational Mappings
CDMs-to-BIRDG mappings: This provides the rules for transforming from any of the CDMs to a common transitional model (BRIDG)
CDMH/BRIDG-to-SDTM mapping: This provides the rules to export results in support of submissions to FDA
CDMH/BRIDG-to-FHIR mapping: This provides the alignment and gaps of the CDMs to existing FHIR resources.
Initiation of development of FHIR extensions and profiles: Allows implementation with FHIR, built from CDMH-to-BRIDG mapping, to be balloted at HL7.
RxNorm to NDC mapping: Allows mapping of drugs in the most widely used reference database (RxNorm) to the specific formatting and content required for FDA submissions

PCOR Research Community Support
Oncology Use Case Query Development: Development of a query based on the Oncology
Registration of cross-mappings to NCI enterprise vocabulary service EVS: Entry of the Transformational Mappings (above) into NCIs Cancer Data Standards Registry and Repository (CaSDR)
BRIDG Updated with CDMH Data Elements: The BRIDG model was updated to include concepts of the mapped CDMs, promoting implementation strategies for use of BRIDG with the various CDMs allowing use of the BRIDG model in CDM projects as well as linking the CDMs to FHIR and SDTM via BRIDG
Education and Data Governance: Initiation of work to build governance, policy and educational material for use of the CDMH tools. This material will describe the process, for using the CDMH solution, as well as the educational material for using the tools

Phase I: Adaptor Model Problems

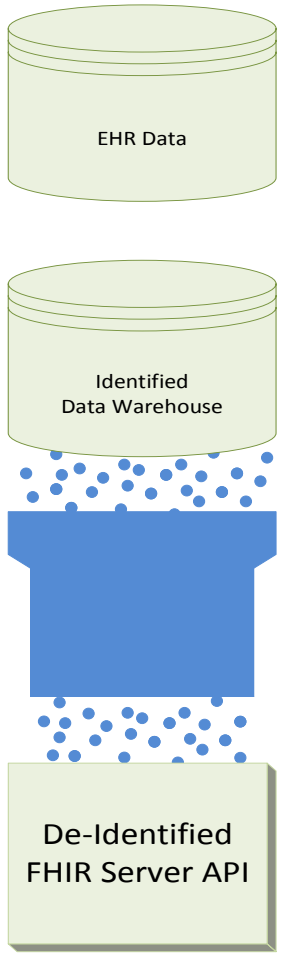


Phase I: Adaptor Model Problems – CDM's



CDM
(PCORnet, OMOP, Sentinel, ACT/i2b2)

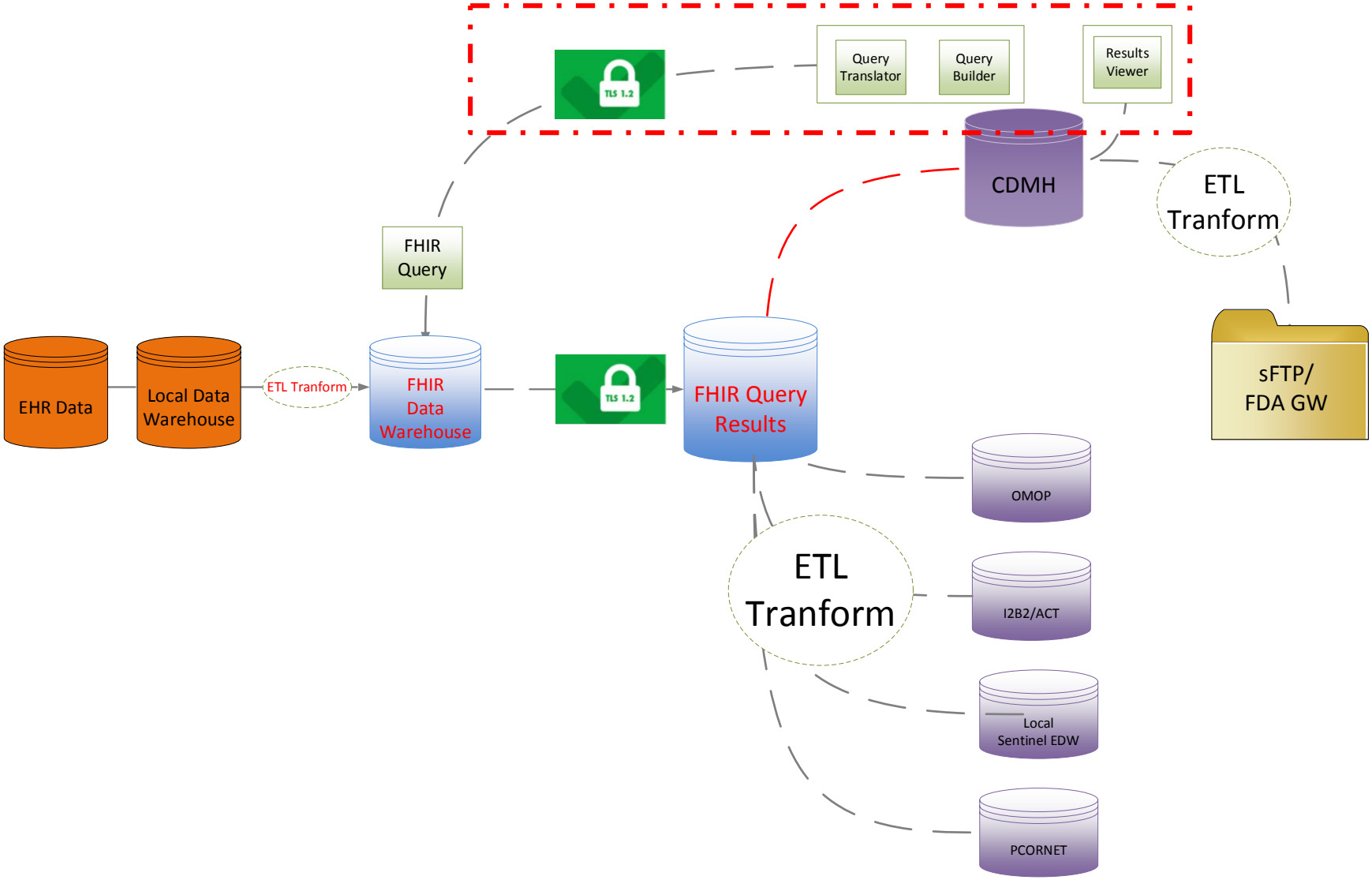
1. Reduce loss of data integrity and granularity
2. Decrease complexity and eliminate need to build separate queries for each data model and proprietary databases
3. Reduce Maintenance by implementing HL7 FHIR standards, eliminates dependence on CDMs and mapping



FHIR

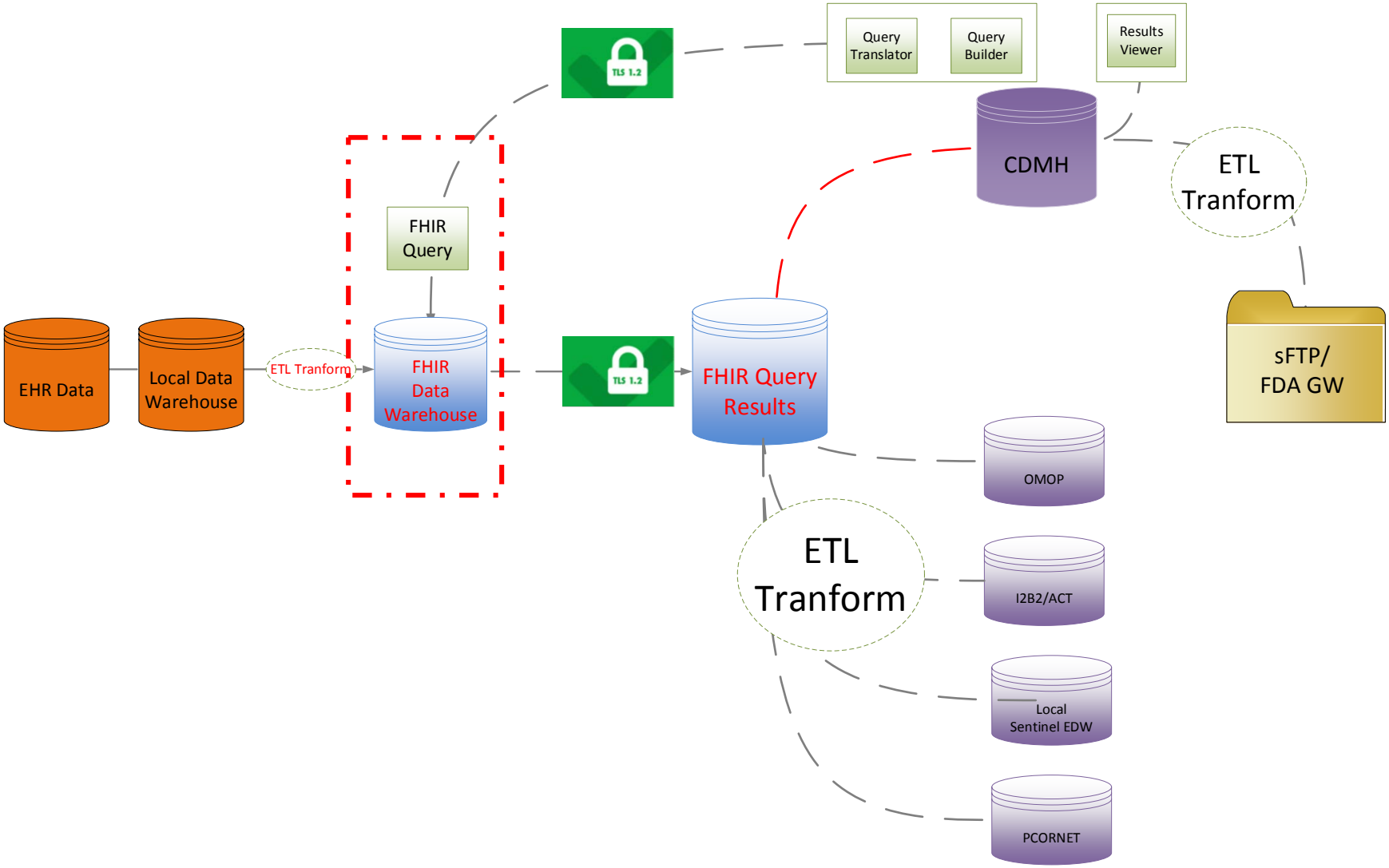
Phase II - FHIR Model: Common Query

Leveraged Query Translator to transform the CDMH query into a “FHIR Query”



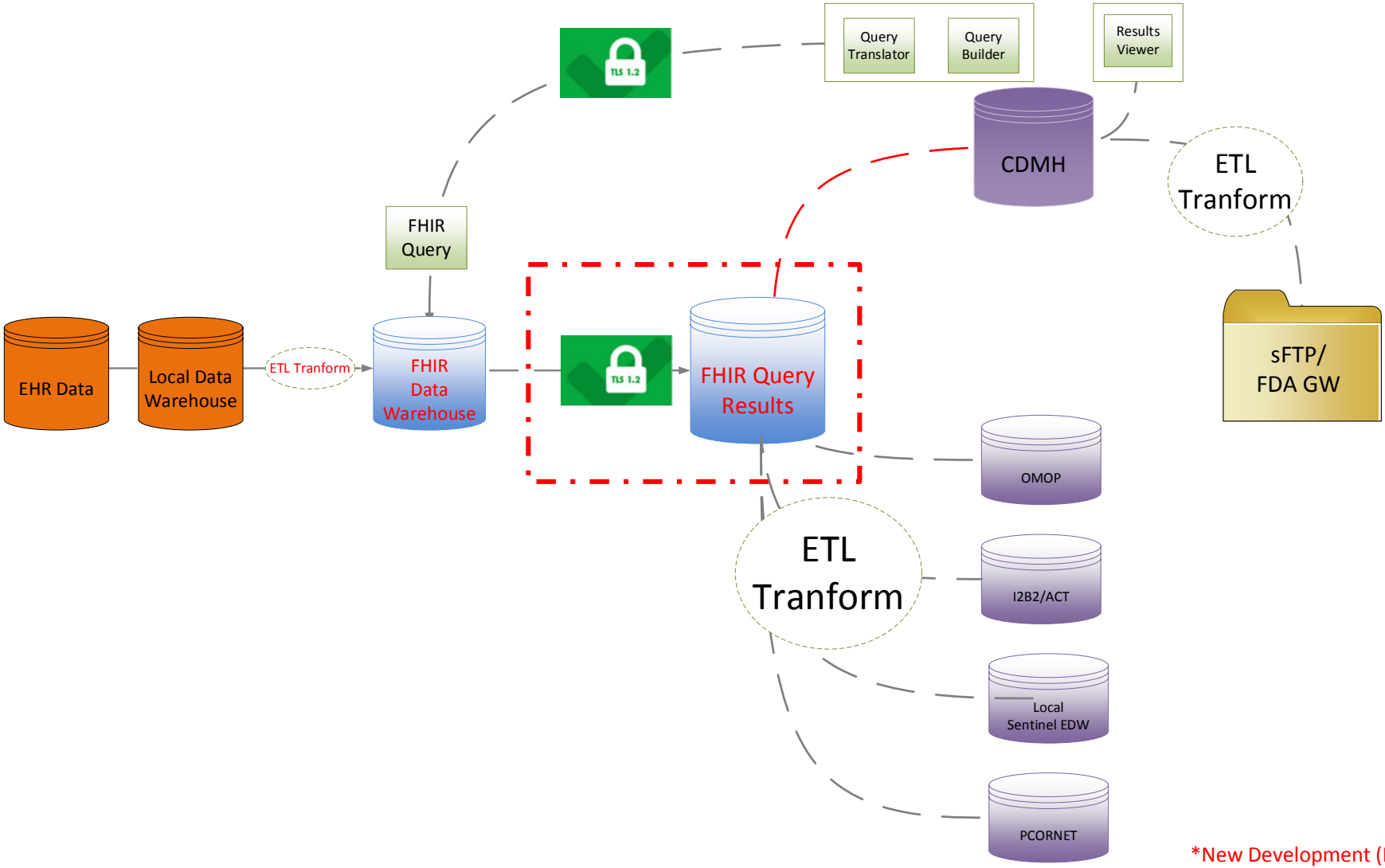
Phase II - FHIR Model: FHIR Server

*Institution runs FHIR query against de-identified FHIR Server



Phase II - FHIR Model: Results

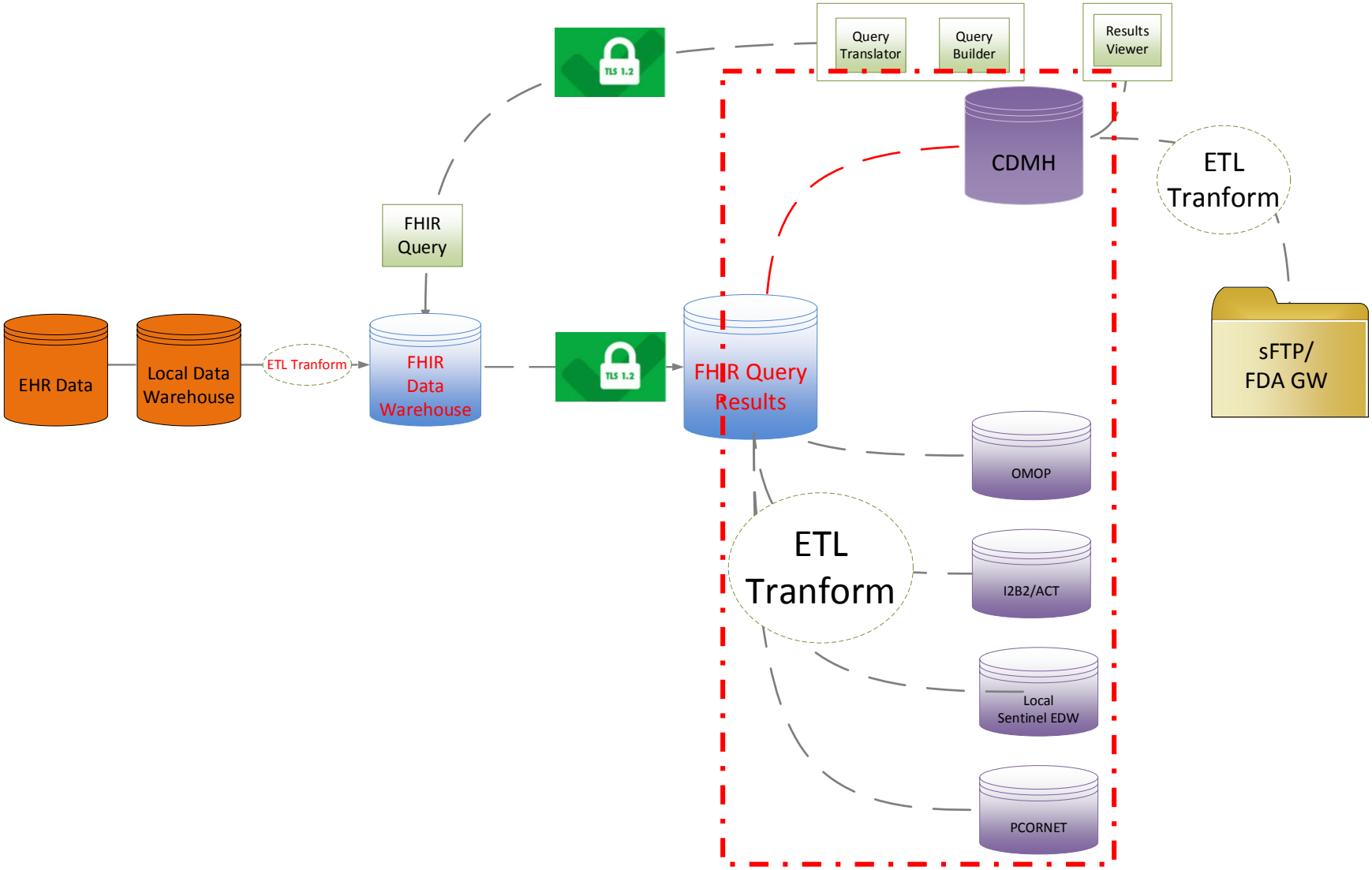
Results Loaded into a staging database



*New Development (FDA)

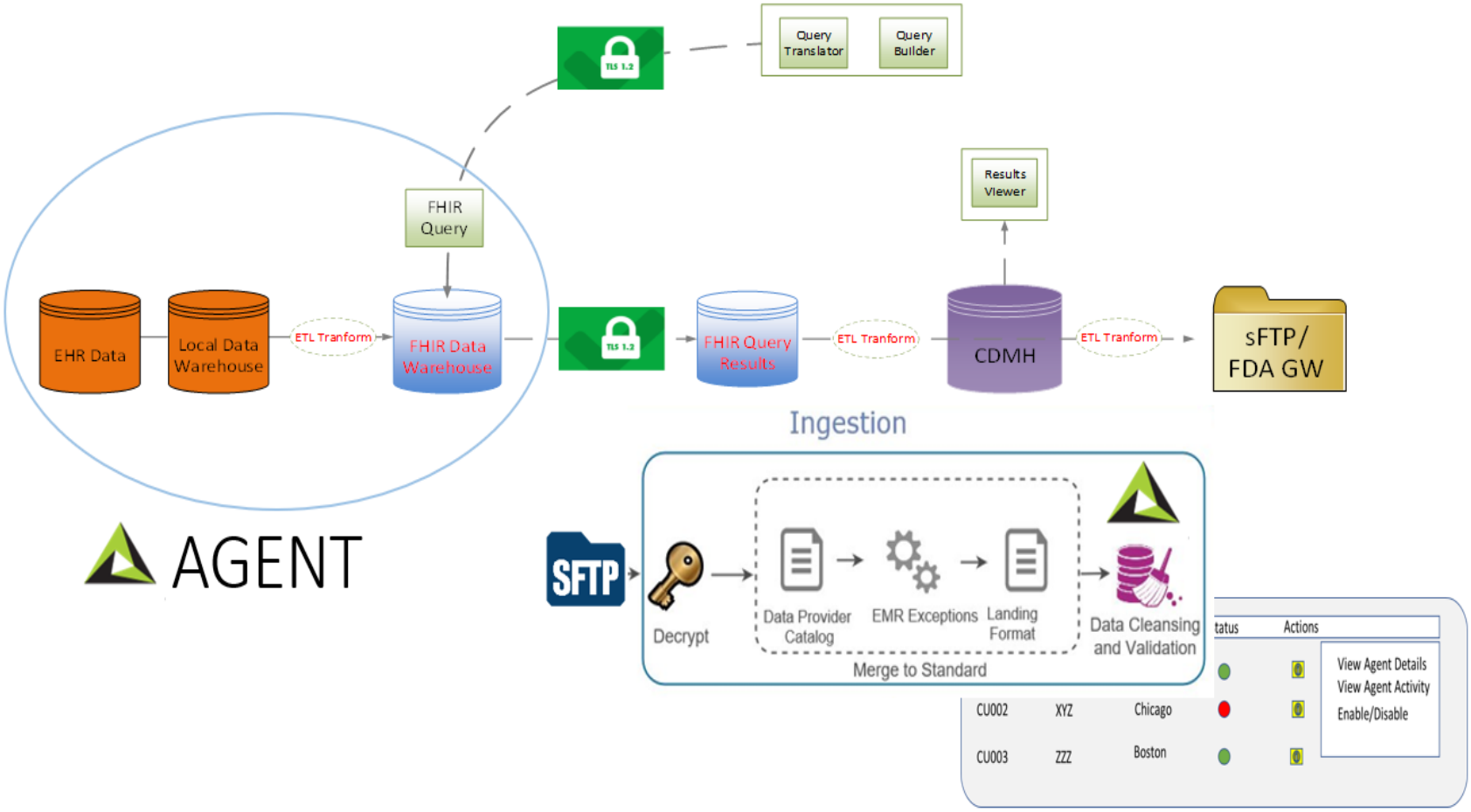
Phase II - FHIR Model: Transformation

FHIR Results → PCORI, OMOP, i2b2/ACT Sentinel and BRIDG



Phase II - FHIR Model: Agent

Agent includes: local ETL, de-identification process, FHIR Server, management oversight



END