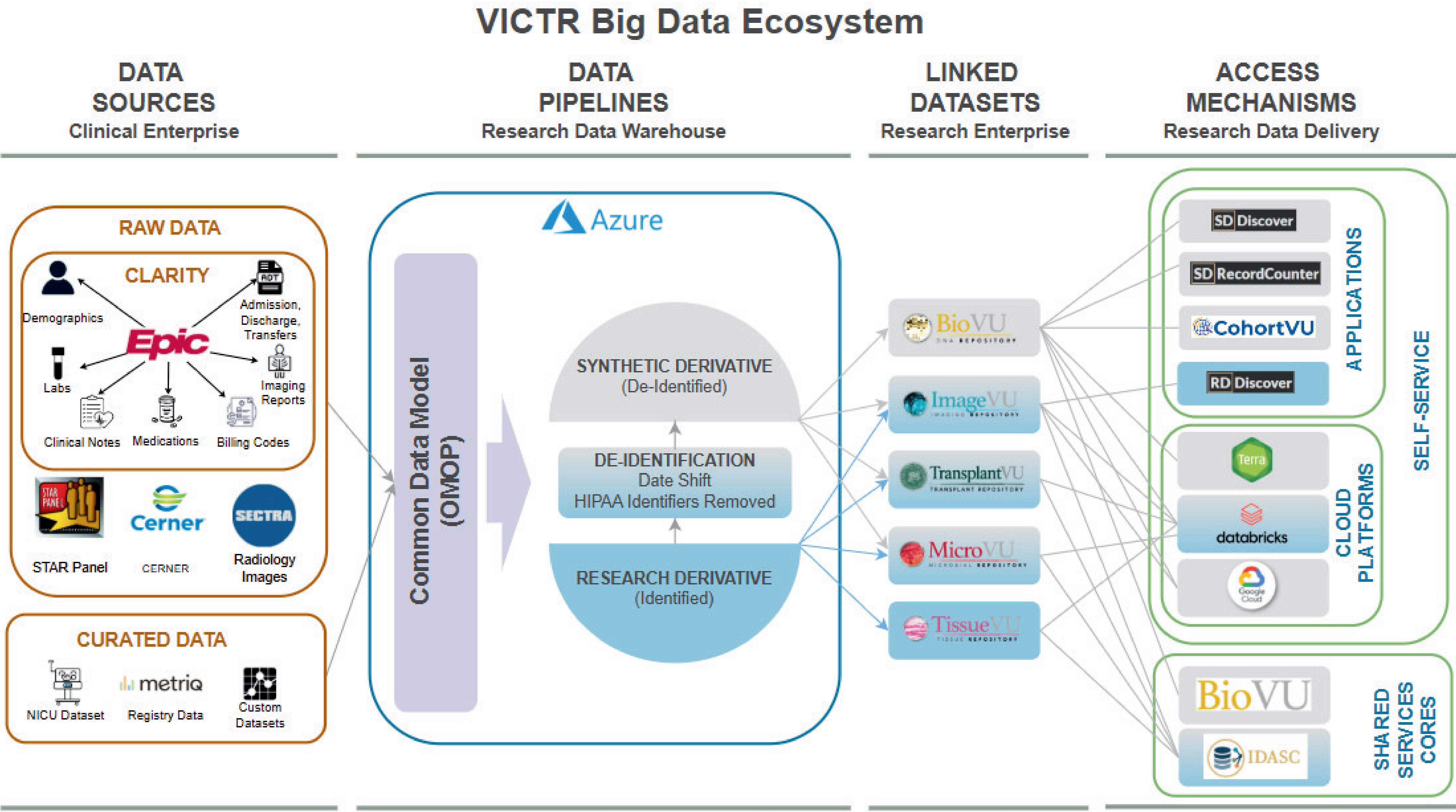


Visualizing the Research and Synthetic Derivative Pipeline:

Leveraging Clinical Data for Research Insights

A SECURE, SCALABLE PIPELINE TURNING RAW CLINICAL DATA INTO IDENTIFIED AND DE-IDENTIFIED DATASETS FOR RESEARCH AND DISCOVERY.

- 5.4+ Million Patient Records in the ecosystem
- 30+ Years of Data available for longitudinal research
- Hundreds of Active Research Projects supported annually

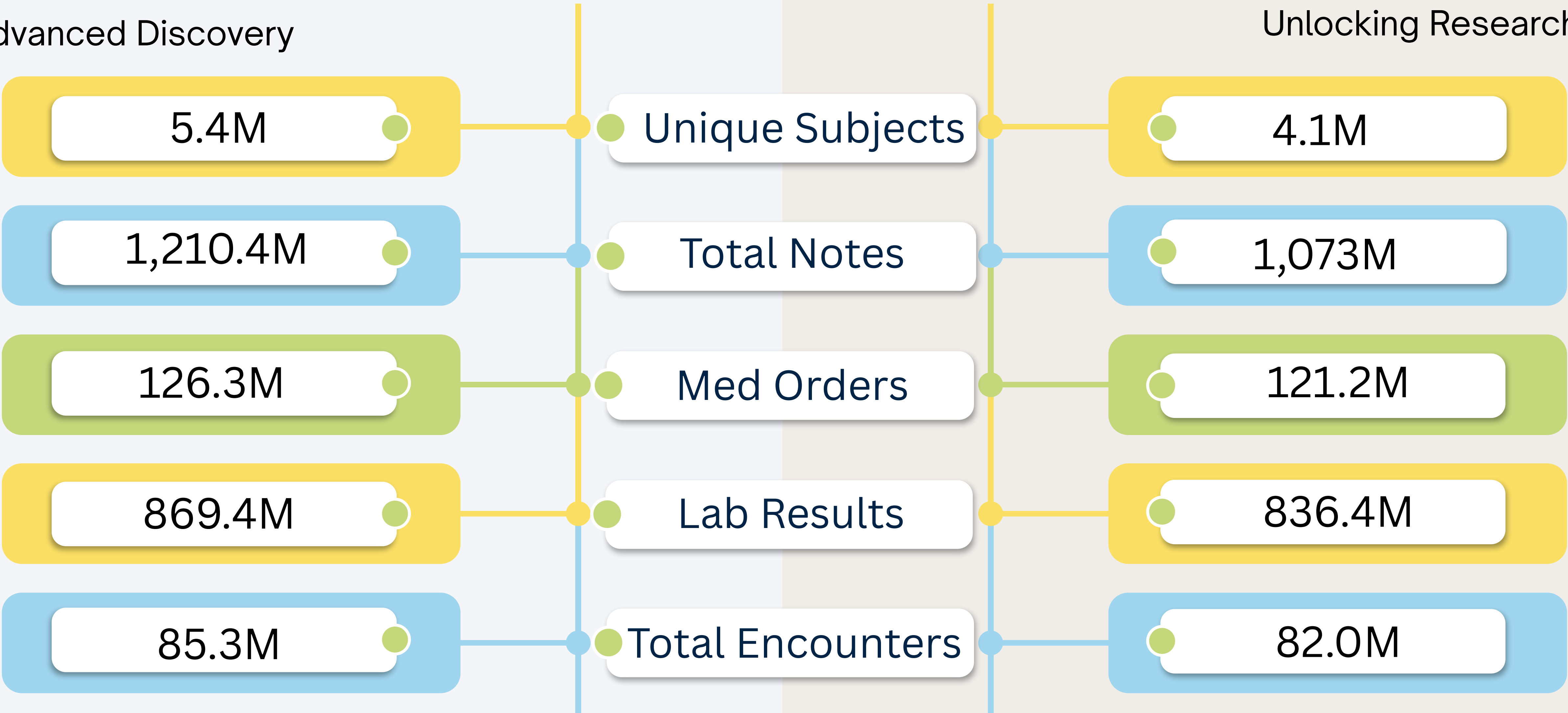


TOGETHER, THESE DATABASES REPRESENT A CORNERSTONE OF VUMC'S COMMITMENT TO ADVANCING BIOMEDICAL RESEARCH THROUGH DATA INNOVATION.

RESEARCH DERIVATIVE (RD)

Full-Fidelity Clinical Data for Advanced Discovery

- **Comprehensive clinical data** from VUMC enterprise systems
- **Identifiers retained** for cross-dataset linkage
- **Structured and unstructured data**, including ICD codes, CPT codes, labs, medications, notes and visit details
- **Research-ready formats** for advanced analysis



VOLUME OF RECORDS (IN MILLIONS) AND TYPES OF DATA AVAILABLE

SYNTHETIC DERIVATIVE (SD)

Unlocking Research While Safeguarding Privacy

- **Fully de-identified** version of medical record data
- **Privacy-preserving** via automated de-identification
- **Semantic tags** maintain meaning for dates, names, and identifiers
- **Structured and unstructured data**, including ICD codes, CPT codes, labs, medications, notes and visit details