Leaf and Atlas Query Tools

Praveen Medabalmi, MD, MPH
Physician Informaticist
Scientific Computing
Icahn School of Medicine at Mount Sinai

May 4, 2022

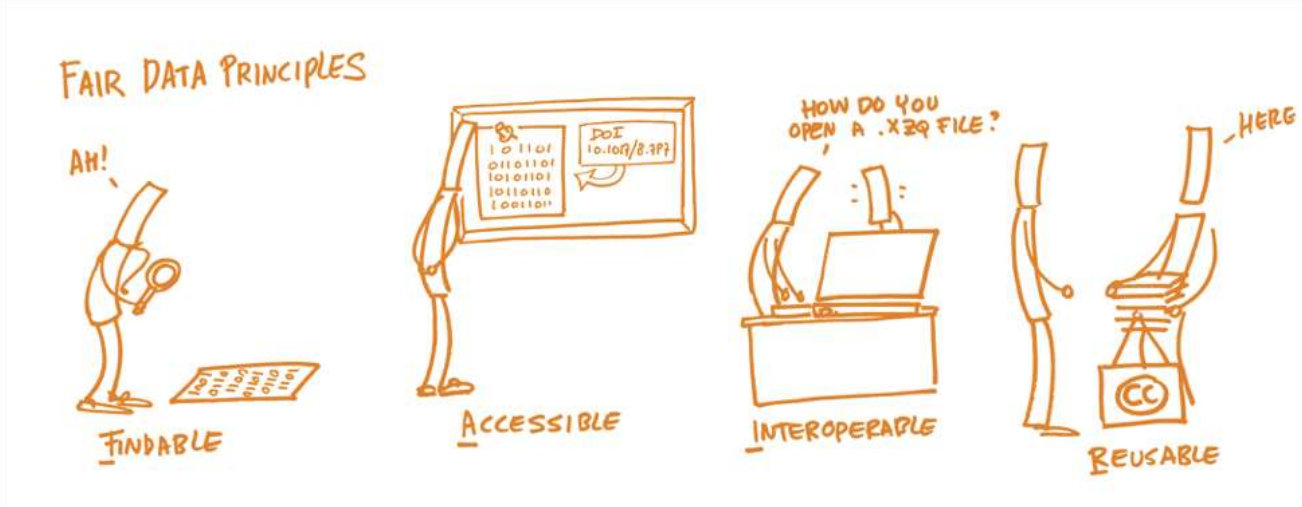
Icahn
School of
Medicine at
Mount
Sinai

Objectives:

1. Scientific Computing Fair Principles of Data
2. Data Sources – Epic and OMOP Databases
3. Clinical query tools
    - Leaf and Atlas
4. How to build a query in each tool
5. When and how to request a custom dataset

Scientific Computing FAIR Principles for Data

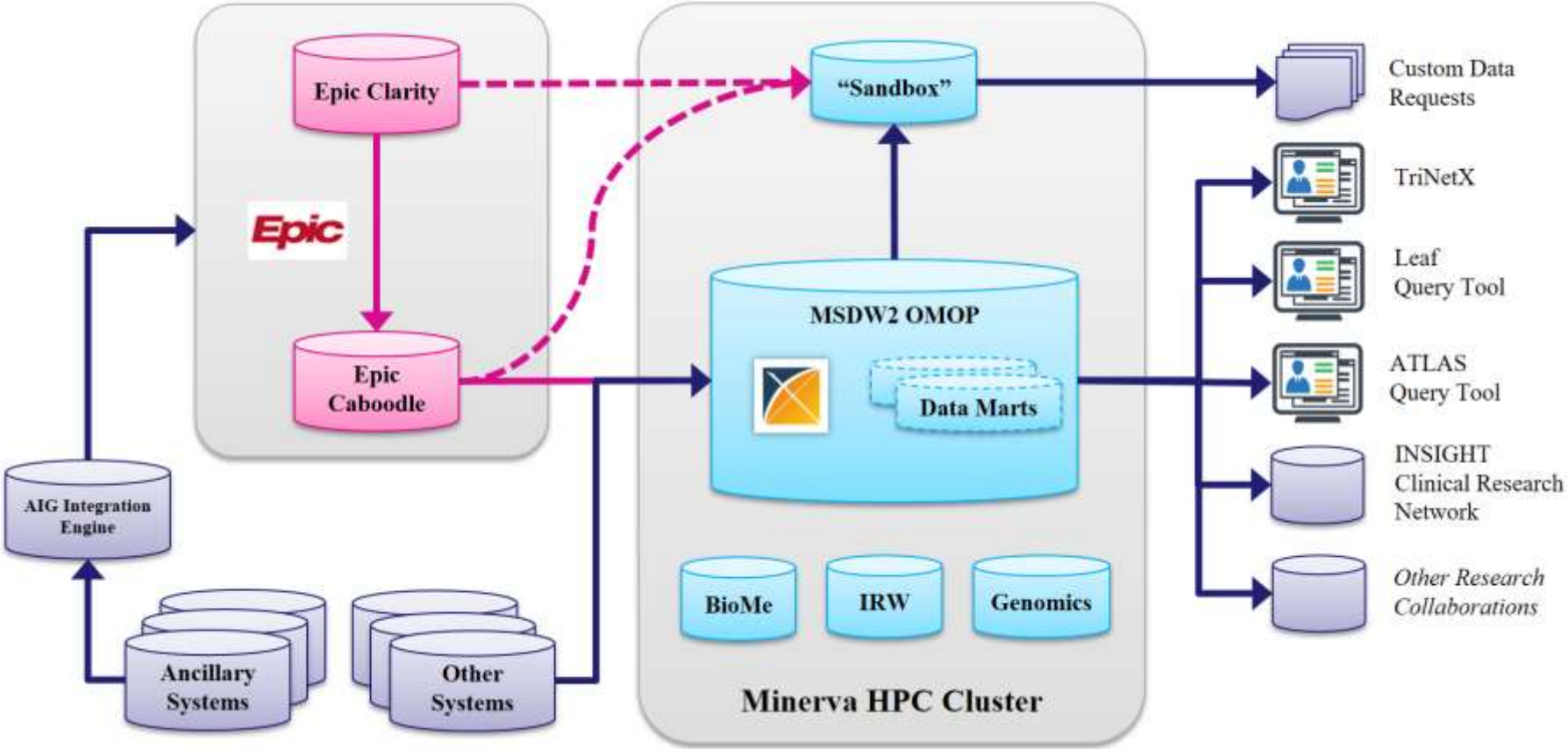
Findable | **A**ccessible | **I**nteroperable | **R**eusable / Reproducible



Source: NIH's Big Data to Knowledge (BD2K) Initiative (https://commonfund.nih.gov/bd2k)

Image Source: https://book.fosteropenscience.eu/

MSDW2 Ecosystem



Leaf Query Examples

1. Patients 18 years and older with diagnosis of IVD or MI – Aspirin or Antiplatelet Agent.
2. Patients diagnosed with Diabetes Mellitus and HBA1c greater than 8 and less than equal to 9.
3. Female Patients between age 50-74 and screened for mammogram
4. Patients diagnosed with COVID-19 and were intubated
5. Patients diagnosed with Essential hypertension and most recent blood pressure was adequately controlled (less than 140/90)

Leaf Query Domains

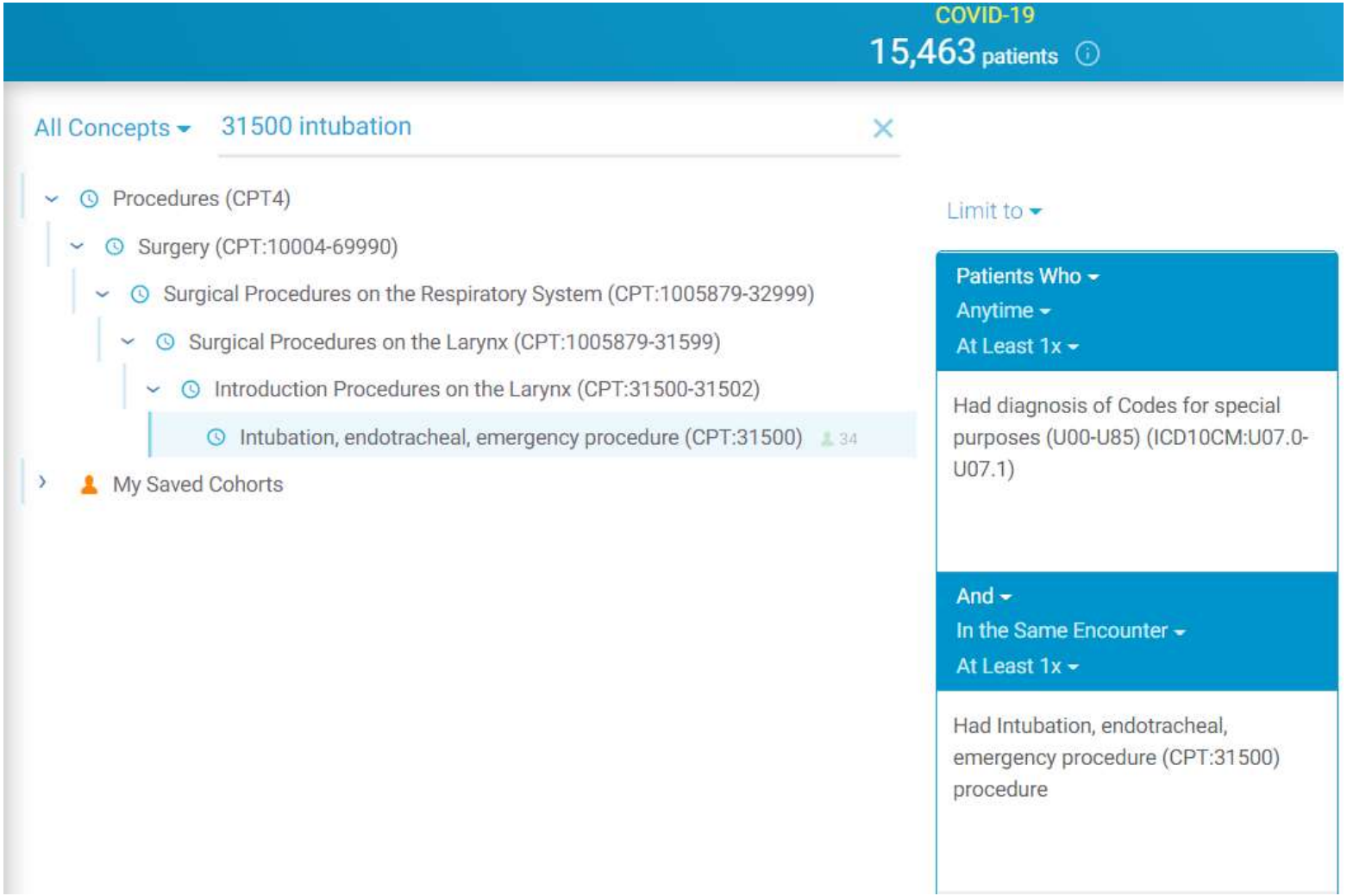
- Conditions – ICD10 Codes
- Lab Results
- Medications
- Procedures
- Demographics, such as current age, ethnicity, gender, race
- Vitals
- Patient cohorts in Sinai repositories – BioMe, Cancer Institute Biorepository and Imaging Research Warehouse

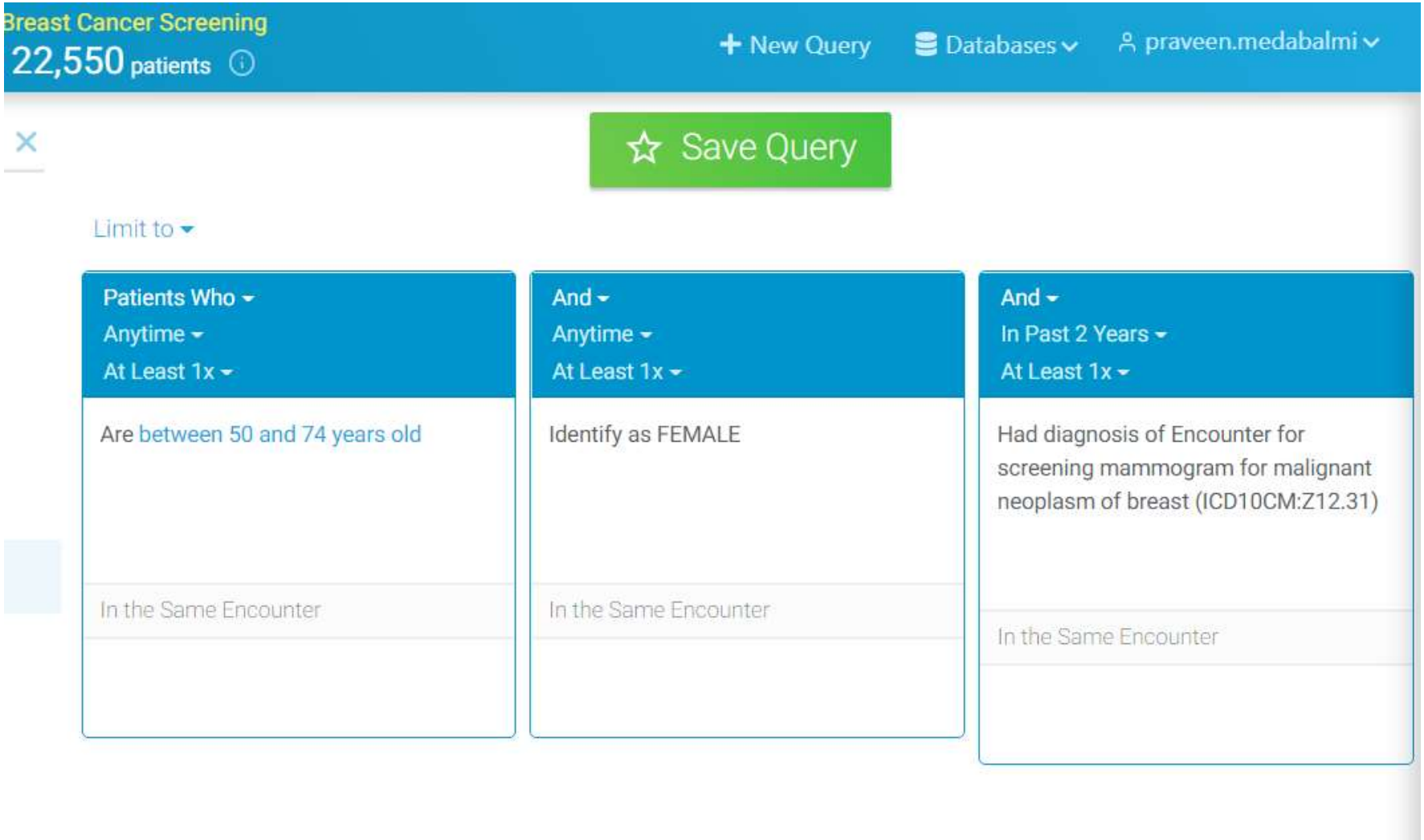
De-Identified Data Elements

- Name
- Street Address, city, county, zip code (the first three digits of the zip code may be used if there are more than 20,000 people in the zip code)
- All element of dates (except year), including dates of birth, admission, discharge or death
- All ages over 89
- All telephone numbers
- Fax number
- E-mail addresses
- Social Security Number (SSN)
- Medical Record Number (MRN)
- Health plan beneficiary number
- Account numbers
- Certificate/License number
- Vehicle identifiers, including license plate numbers
- Device identification and/or serial number
- Uniform Resource Locator (URL)
- Internet Protocol (IP) address
- Biometric identifiers, including finger and voiceprints
- Full face photographic images and other comparable images
- Any other unique identifying number, characteristic, or code

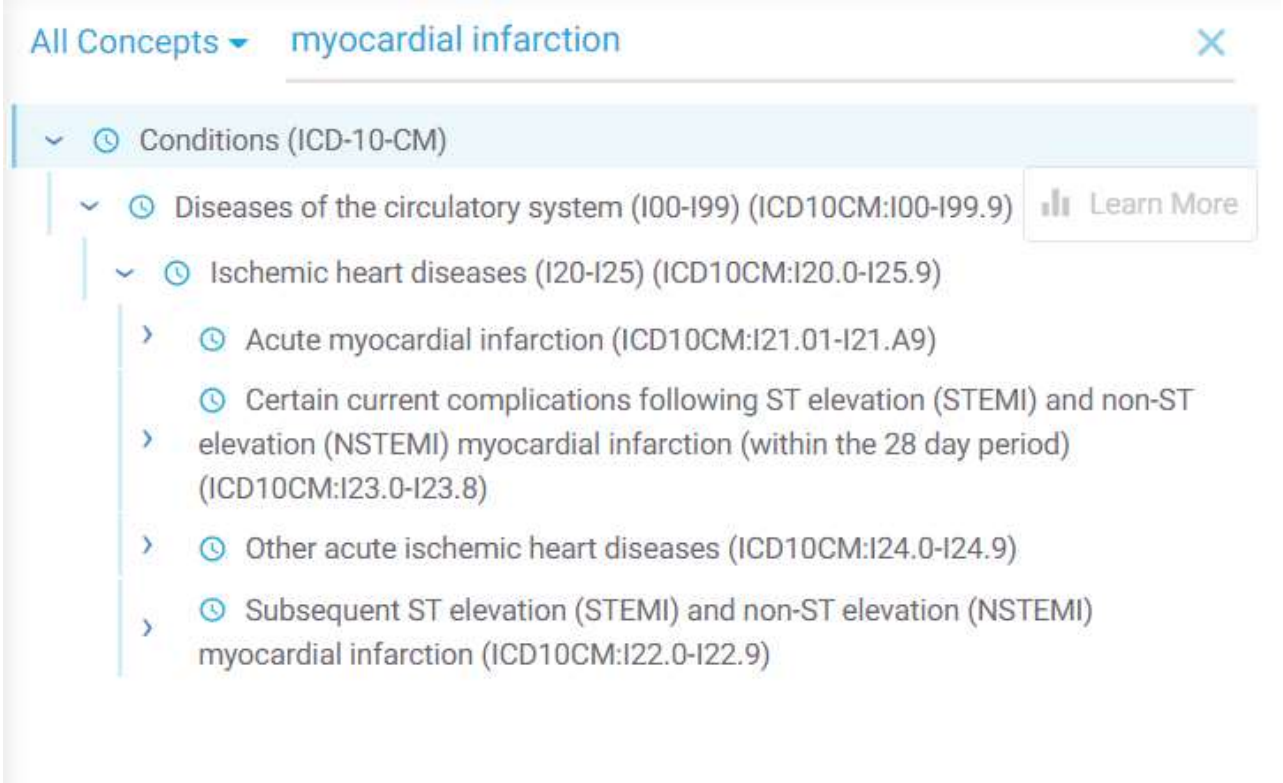
HIPAA Safe Harbor Hripcsak's Shift-and-Truncate (SANT) Method "Elapsed Days" Approach

- Searches on keywords in concept descriptions
- Shows nodes in the hierarchy that satisfy the query


COVID-19
15,463 patients
All Concepts
31500 intubation
Procedures (CPT4)
Surgery (CPT:10004-69990)
Surgical Procedures on the Respiratory System (CPT:1005879-32999)
Surgical Procedures on the Larynx (CPT:1005879-31599)
Introduction Procedures on the Larynx (CPT:31500-31502)
Intubation, endotracheal, emergency procedure (CPT:31500) 34
My Saved Cohorts
Limit to
Patients Who
Anytime
At Least 1x
Had diagnosis of Codes for special purposes (U00-U85) (ICD10CM:U07.0-U07.1)
And
In the Same Encounter
At Least 1x
Had Intubation, endotracheal, emergency procedure (CPT:31500) procedure


Breast Cancer Screening
22,550 patients
+ New Query
Databases
praveen.medabalmi
Save Query
Limit to
Patients Who
Anytime
At Least 1x
Are between 50 and 74 years old
In the Same Encounter
And
Anytime
At Least 1x
Identify as FEMALE
In the Same Encounter
And
In Past 2 Years
At Least 1x
Had diagnosis of Encounter for screening mammogram for malignant neoplasm of breast (ICD10CM:Z12.31)
In the Same Encounter

Searching Domains/ Concepts

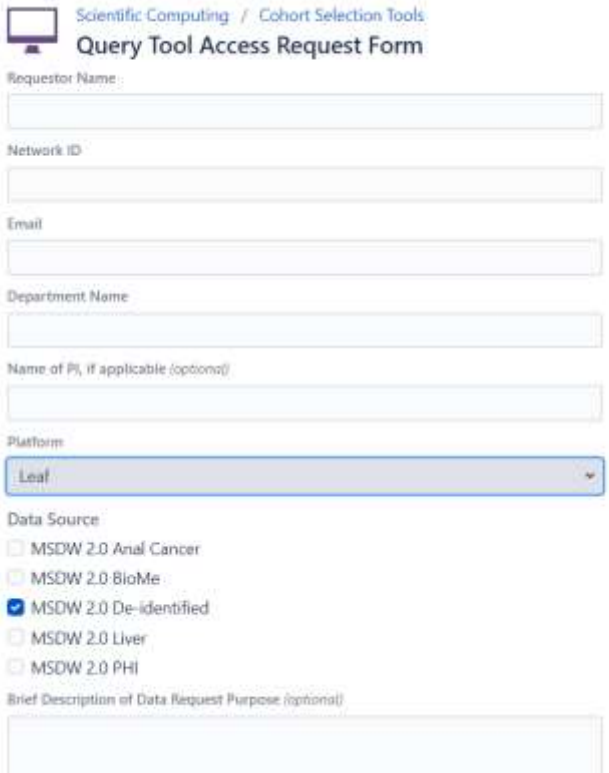


Clinical Query Tools – Leaf and Atlas

	Leaf	ATLAS
Development	Nic Dobbins, University of Washington Open-source collaborators (including Mount Sinai)	OHDSI Community
Similarities	Open-source software Extensive use by academic researchers Specify cohort inclusion criteria Visualize cohort characteristics Export cohort patient list	
Ease of Use	User Friendly	Learning curve
Cohort Inclusion Criteria	Easy drag-and-drop	Complex rules, temporal dependencies
Analytics Capabilities	Minimal	Built-in
For Data Marts	Extensive setup	Easy Setup

Query Tool Access Request Form - Cohort Selection Tools

https://labs.icahn.mssm.edu/msdw/



Leaf Application Status and Roadmap

Last Updated: April 18, 2022

Status

Quality improvement and research users can use Leaf to conveniently query the Mount Sinai Data Warehouse. Leaf only accesses de-identified data. Documentation and user service are available.

Leaf will be offline for scheduled maintenance on Thursday, April 21, from 3:00-7:00pm to improve usage reporting tools and to test query facets for external patient cohorts. Please log out of Leaf during scheduled maintenance.

On Wednesday, May 4, an application training for Leaf and ATLAS functionality will be hosted online, followed by a brief Q&A. Click here to register for the application training on May 4, 2022, from 12:00-12:30 pm.

Leaf supports these query domains:

- Conditions (diagnoses) using ICD-10-CM
- Visit (encounter) locations
- Lab results using LOINC
- Medications using ATC
- Procedures using CPT4
- Demographics, such as current age, ethnicity, gender, race and vital status
- Vitals
- Patient cohorts in Sinai repositories, starting with BioMe

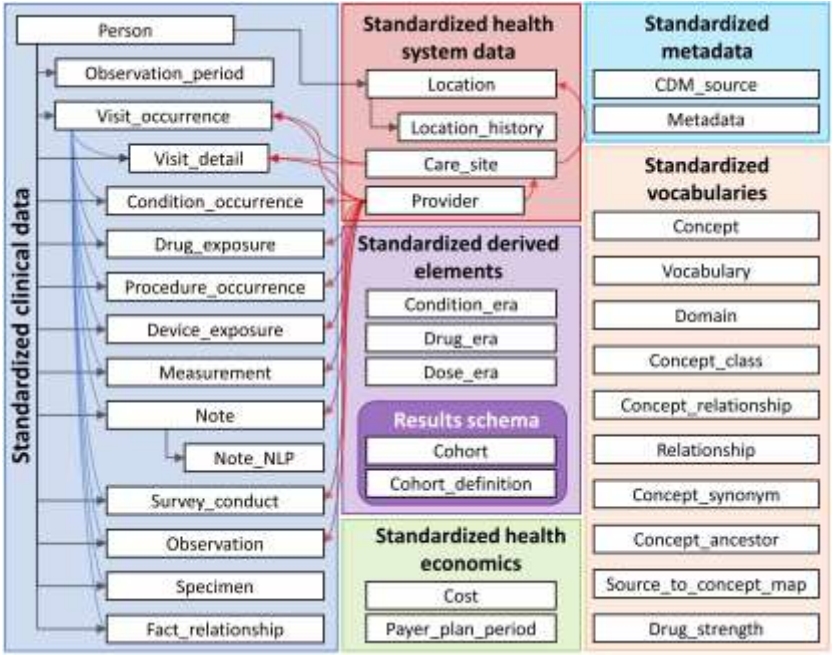
Leaf contains these known bugs:

- Patient List temporarily disabled
- Queries that last longer than 1 minute terminate
- Leaf is updated every 2 weeks (rather than nightly)

OMOP Common Data Model

- **Observational Medical Outcomes Partnership (OMOP)**
  - Clinical Data
  - Health System Data
  - Standardized derived elements
  - Standardized Metadata
  - Standardized Vocabularies
  - Standardized Health Economics data

- **Common Data Model**
  - Standardizes both structure & content for interoperability
  - Produce meaningfully comparable and reproducible results



MSDW Data Standardized though OMOP CDM

OMOP Table	Record Type
person	Patient Demographics
death	Patient Date of Death
visit_occurrence	Type of Visit - Inpatient Hospitalization
condition_occurrence	Encounter Diagnosis
measurement	Lab Component Result
measurement	Vital Signs
drug_exposure	Immunization Administration
drug_exposure	Medication Order
drug_exposure	Medication Order with Administration
drug_exposure	Medication Order with Dispense
procedure_occurrence	General/ Surgical Procedure
observation	Family History
note	Clinical Note
note	Lab Component Result Note

https://labs.icahn.mssm.edu/msdw/data-sources/

ATLAS Query Tool

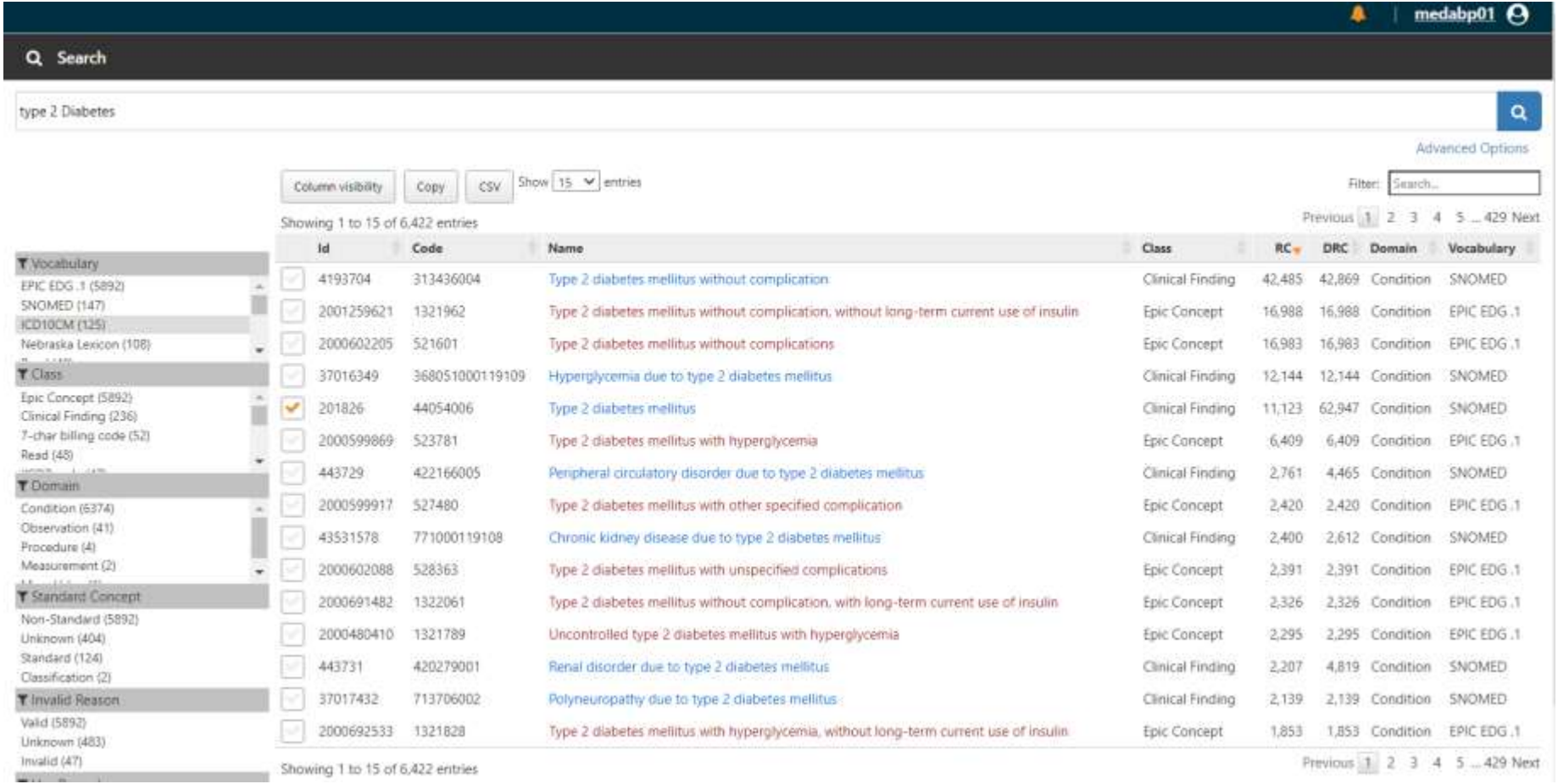
- Web based platform developed by OHDSI Community
- Facilitates design and execution of Analyses
- Vocabularies for Diagnosis – EPIC, SNOMED
- SNOMED Code Mapped to OMOP CDM concept id
- Cohort Definition – Criteria

ATLAS Query

- Concept Set – list of Concepts that can be reused
- Attributes – Descendants, Mapped, Excluded
- Concept Set Expression – 2 Concept sets

Concept ID	Concept Name	Descendants	Mapped	Excluded
4329847	Myocardial Infarction	YES	YES	NO
314666	Old Myocardial Infarction	YES	YES	YES

Concepts and Vocabularies



Concept Sets

New Concept Set

Column visibility | Copy | CSV | Show 15 entries

Search: Search...

Showing 1 to 15 of 52 entries

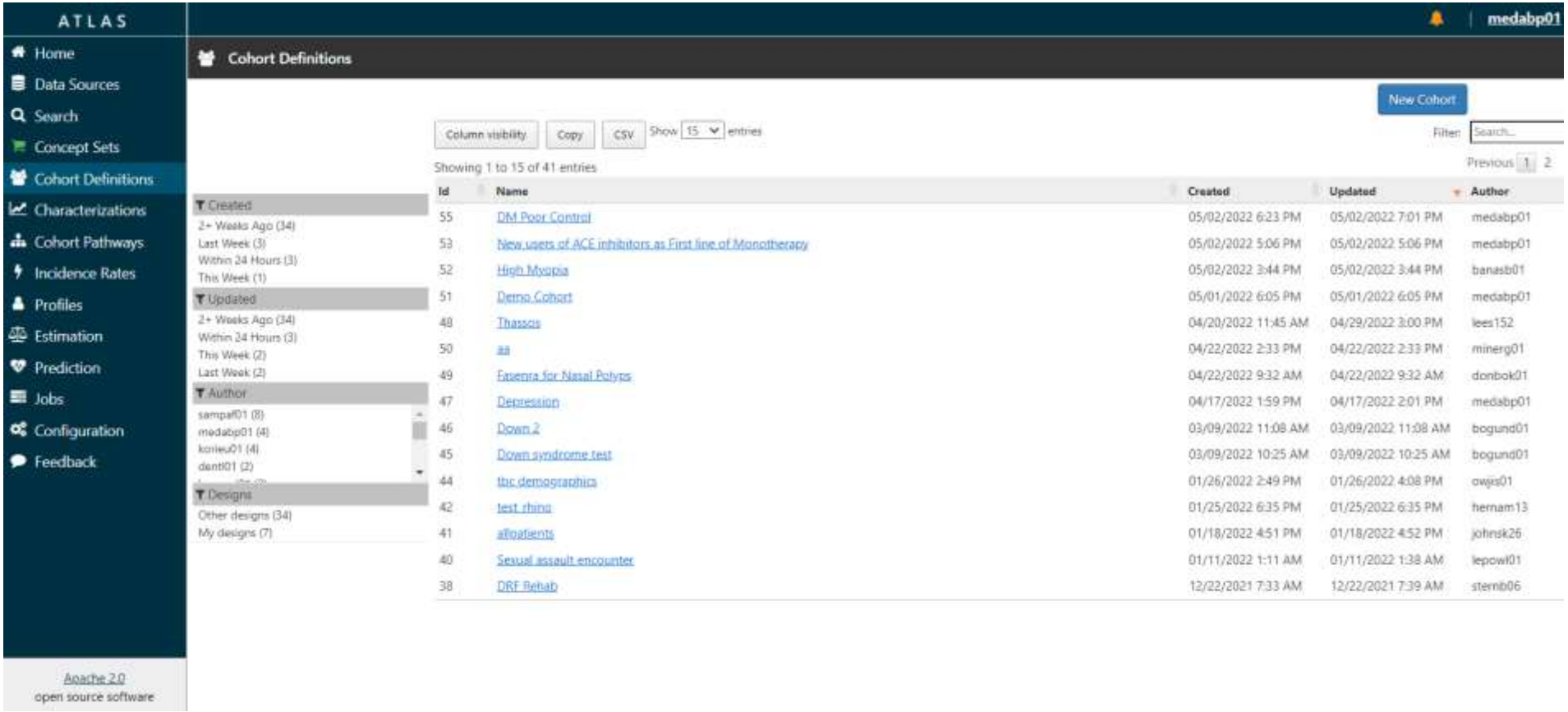
Previous 1 2 3 4 Next

Id	Name	Created	Modified	Author
68	Essential hypertension	05/02/2022 6:51 PM	05/03/2022 4:30 PM	medabp01
70	Colitis	05/03/2022 4:08 PM	05/03/2022 4:08 PM	medabp01
69	GI Bleed	05/03/2022 2:49 PM	05/03/2022 2:49 PM	medabp01
65	COVID-19 and MI	05/02/2022 3:48 PM	05/02/2022 5:46 PM	banasb01
64	Diabetes Mellitus	05/01/2022 5:53 PM	05/01/2022 5:53 PM	medabp01
63	THASSOS_p_link	04/29/2022 1:33 PM	04/29/2022 1:33 PM	lees152
62	COPY OF aa	04/22/2022 1:55 PM	04/22/2022 1:55 PM	minerg01
61	aa	04/22/2022 12:59 PM	04/22/2022 1:42 PM	minerg01
59	Depression	04/17/2022 1:09 PM	04/17/2022 1:56 PM	medabp01
17	covid	10/19/2021 12:55 PM	04/07/2022 9:31 PM	nirens01
56	Migraines	04/01/2022 10:16 AM	04/01/2022 10:19 AM	dentl01
54	MSI Test	03/15/2022 1:49 PM	03/24/2022 2:27 PM	grubec05
53	Product of Conception	03/09/2022 11:01 AM	03/09/2022 11:08 AM	bogund01
52	Down syndrome	03/09/2022 10:13 AM	03/09/2022 10:58 AM	bogund01
51	d22.9	01/26/2022 2:50 PM	01/26/2022 2:50 PM	owjis01

Showing 1 to 15 of 52 entries

Previous 1 2 3 4 Next

Cohort Definitions



Atlas Application Status and Roadmap

Last Updated: March 10, 2022

Status

ATLAS application is currently in software version 1 at Mount Sinai and is available for user testing. Features and functionality are regularly updated, and scheduled maintenance is announced in advance through msdw.mountsinai.org.

On Wednesday May 4, an application training for Leaf and ATLAS functionality will be hosted online, followed by a brief Q&A. Click here to register for the application training on May 4, 2022, from 12:00-12:30 pm.

As of December 16, 2021, all ATLAS functions (left-hand tabs) are available to users.

ATLAS functionality is currently limited to include:

- De-identified data source only
- OMOP standard concept IDs are contained within only the following clinical domains: conditions, procedures, visits, vitals measurements
- All other domains can be queried using concepts in Epic vocabularies

ATLAS beta contains the following limitations:

- Data containing PHI, including datamarts, are not yet available for query
- Most ATLAS data uses OMOP standard concept IDs, but some live data uses non-standard vocabularies

Roadmap

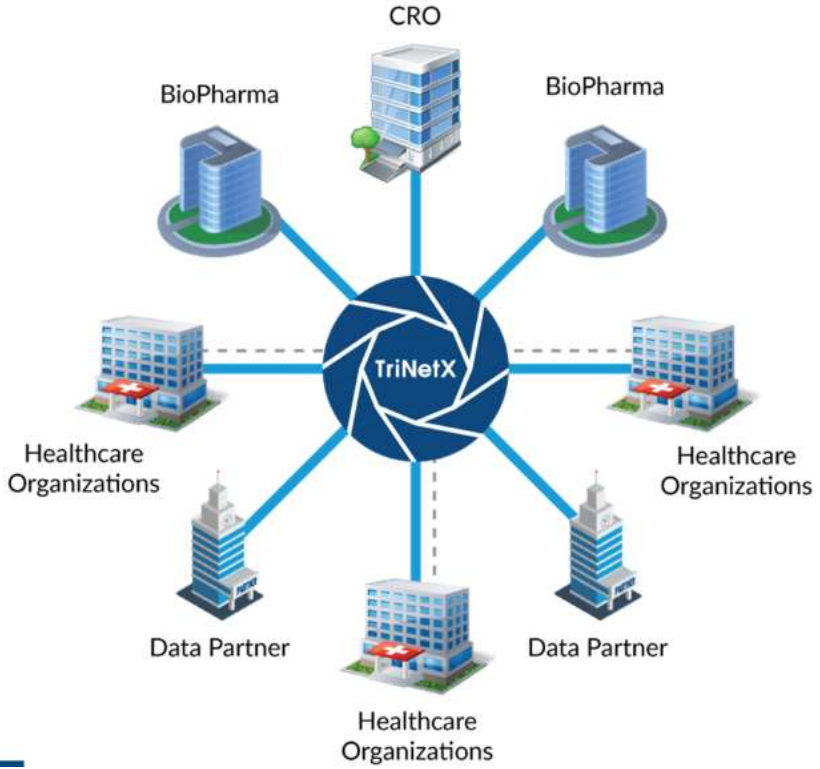
The following improvements to ATLAS are planned for the 4th quarter of 2021:

- Add four datamarts containing PHI: BioMe, Anal Cancer, Liver, and iSite
- Enhance ATLAS deployment process to support ATLAS users

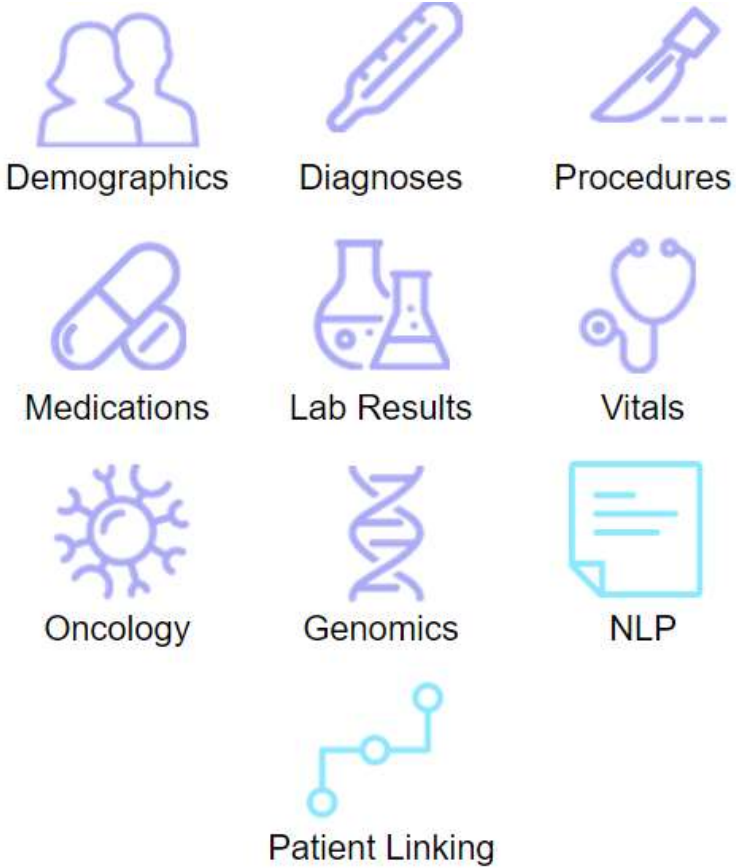
TrinetX

- A global research network that optimizes clinical research for data at Mount Sinai

- TrinetX platform utilizes real-time modeling and simulation of patient populations based on inclusion and exclusion criteria

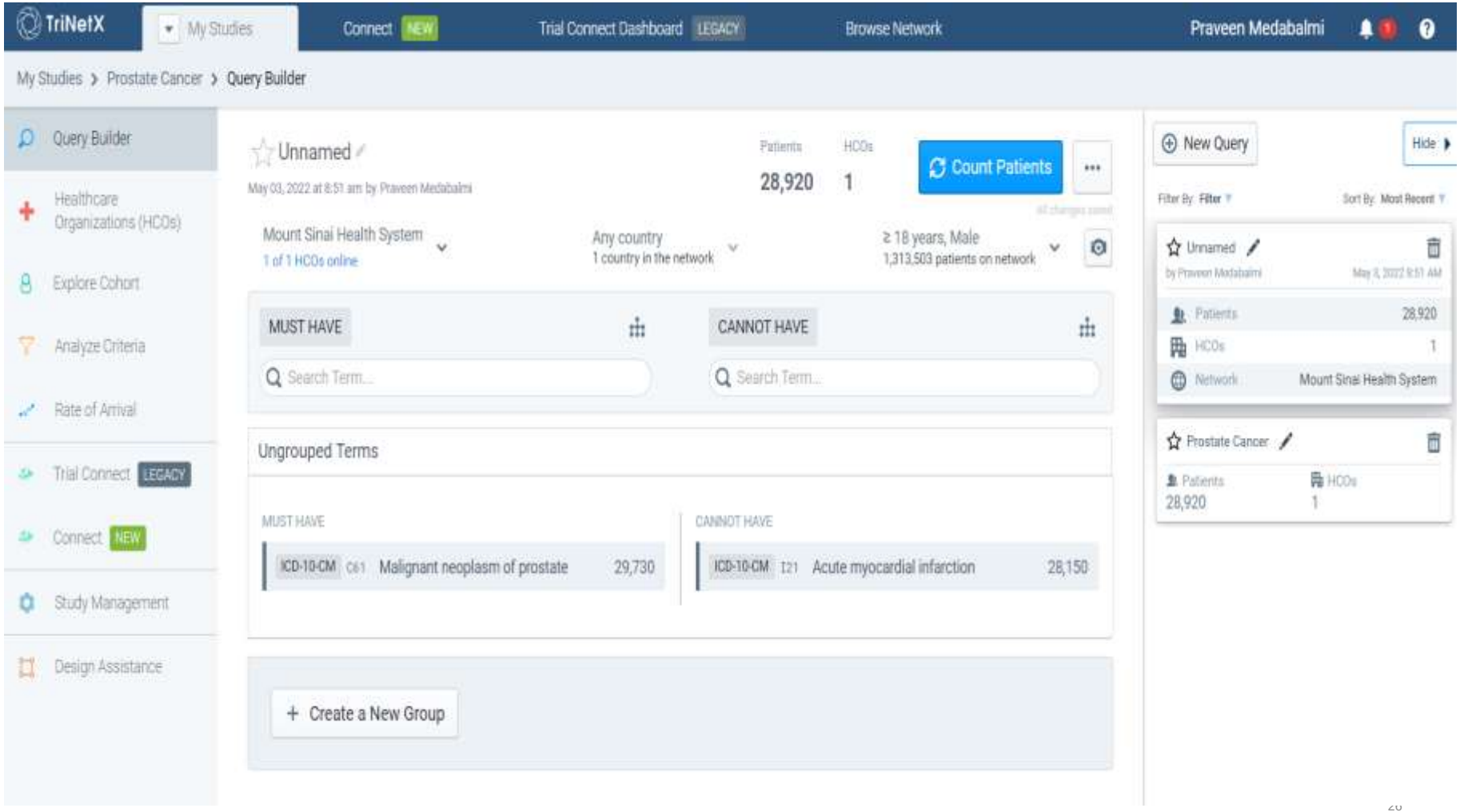


Query Real World Data



Protocol Design and Feasibility

Real-World Data	Query Builder	Explore Cohort	Criteria Analysis
Self-service access to continuously-refreshed patient data	Determine if a sufficient patient population matches a protocol	Investigate all attributes and comorbidities of the eligible cohort	Analyze inclusion / exclusion criteria and the impact of protocol changes

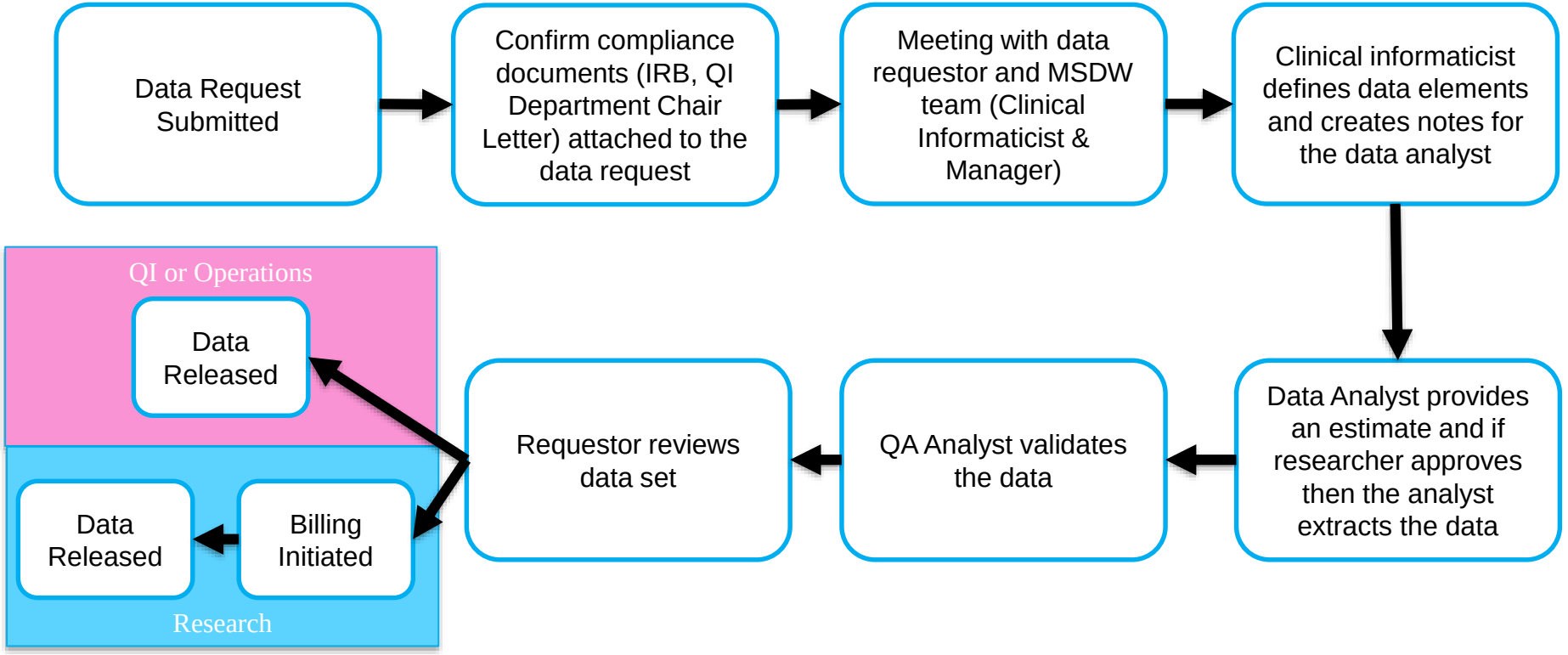

TriNetX
My Studies
Connect NEW
Trial Connect Dashboard LEGACY
Browse Network
Praveen Medabalmi
My Studies › Prostate Cancer › Query Builder
Query Builder
Healthcare Organizations (HCOs)
Explore Cohort
Analyze Criteria
Rate of Arrival
Trial Connect LEGACY
Connect NEW
Study Management
Design Assistance
Unnamed
May 03, 2022 at 8:51 am by Praveen Medabalmi
Patients
28,920
HCOs
1
Count Patients
Mount Sinai Health System
1 of 1 HCOs online
Any country
1 country in the network
≥ 18 years, Male
1,313,503 patients on network
MUST HAVE
CANNOT HAVE
Search Term...
Search Term...
Ungrouped Terms
MUST HAVE
ICD-10-CM C61 Malignant neoplasm of prostate 29,730
CANNOT HAVE
ICD-10-CM I21 Acute myocardial infarction 28,150
+ Create a New Group
New Query
Hide
Filter By: Filter
Sort By: Most Recent
Unnamed
by Praveen Medabalmi
May 3, 2022 8:51 AM
Patients 28,920
HCOs 1
Network Mount Sinai Health System
Prostate Cancer
Patients 28,920
HCOs 1

Custom Data Request Process

When You Need Custom Data

- Complex question that cannot be answered with one of the self-service query tools
- Need additional data that is not included in a de-identified data set
- Need PHI data for your analysis

Workflow Once Data Request Submitted



JIRA ticketing system used to monitor the status of data requests

Leaf, Atlas and TrinetX

	Leaf	OMOP/ATLAS	TriNetX
Description	Web-based, lightweight drag-and-drop query tool that quickly analyzes population demographics using ICD-10 codes.	A web-based integrated platform for database exploration, standardized vocabulary browsing, cohort definition, and population-level analysis.	A global research network that optimizes clinical research for data at Mount Sinai.
Access	Request access here. You can then use your Sinai network username/password to login.	Login to the ATLAS System using your Sinai School network username/password. For users with only a Hospital account, request a School account through Sailpoint to access ATLAS.	Request access here. Log in to the TriNetX system using your email address and password.
Training	Tutorial	Tutorial, Videos	PEAK Tutorial
Data Types	Conditions (diagnoses), procedures, demographics, encounters (patient encounter locations) and vitals	Facilities, diagnoses, procedures, medications, labs, orders, patient demographics	Demographics, diagnoses, procedures, medications, labs, vitals
PHI	No	Yes, if IRB Approved	
Cost	Free	Free	Free
Application Status	Leaf Status and Roadmap	ATLAS Status and Roadmap	
Turnaround	Seconds	Seconds	Seconds to minutes
Advantages/Disadvantages	Can visualize demographic details of cohorts, drag-and-drop query feature	Utilizes common data model and queries; numerous query analysis tools available in github	Platform utilizes real-time modeling and simulation of patient populations based on inclusion and exclusion criteria; contains MSDW patient data through June 2021; can search BioMe patients; can search Image Research Warehouse patients; can be slow to operate

https://labs.icahn.mssm.edu/msdw/services/

Learn more about MSDW2 and Clinical Query tools from the links below:

https://labs.icahn.mssm.edu/msdw/

https://labs.icahn.mssm.edu/msdw/services/

https://labs.icahn.mssm.edu/msdw/data-sources/

Praveen Medabalmi, MD, MPH
Physician Informaticist
Scientific Computing
Icahn School of Medicine at Mount Sinai

May 4, 2022

Icahn
School of
Medicine at
Mount
Sinai