

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

Accessible

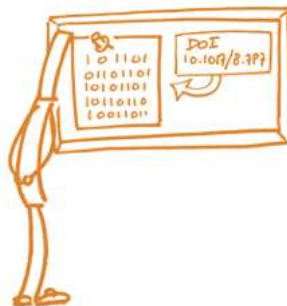
Interoperable

Reusable /
Reproducible

FAIR DATA PRINCIPLES



FINDABLE



ACCESSIBLE



INTEROPERABLE

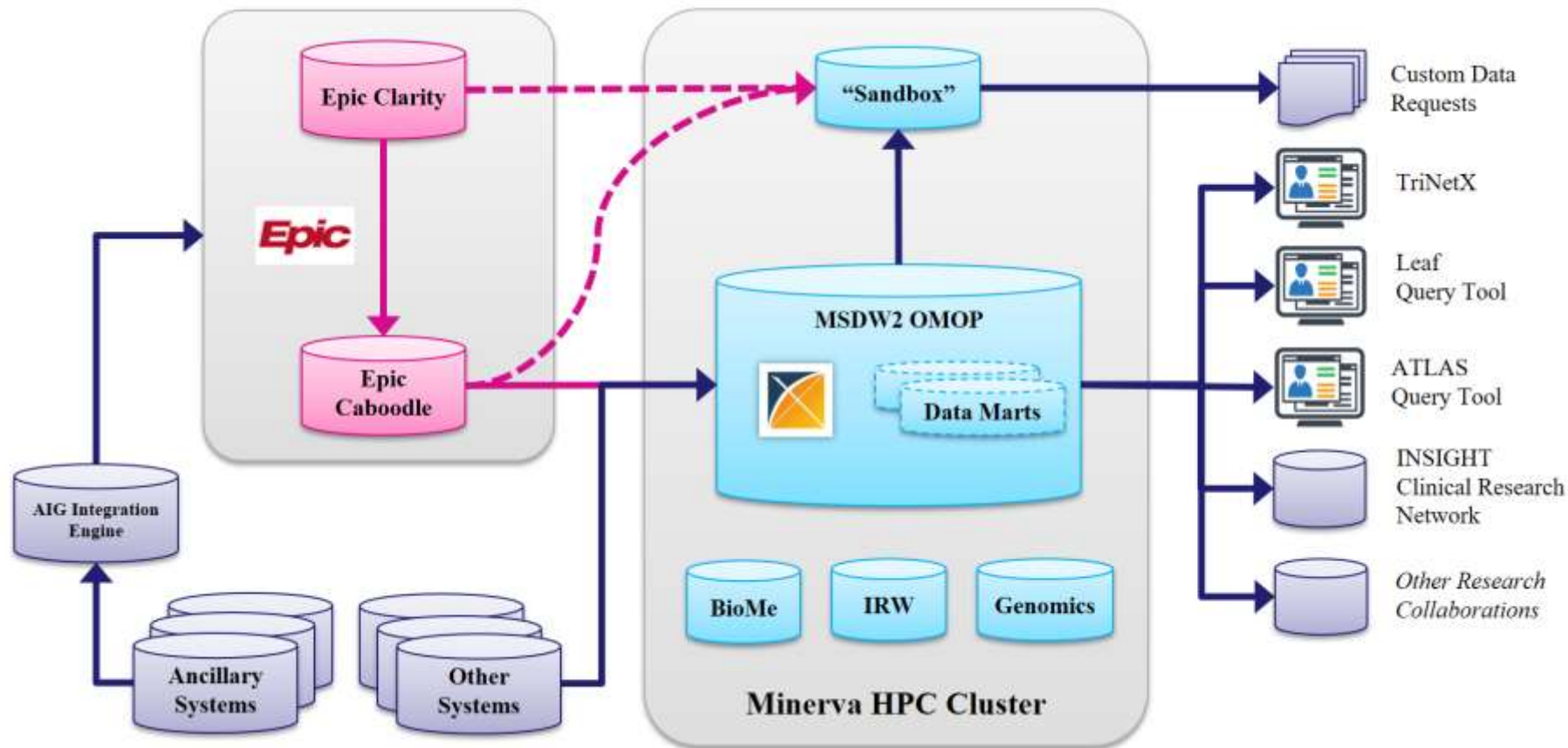


REUSABLE

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

All Concepts ▾ chest pain

- Conditions (ICD-10-CM)
 - Symptoms, signs and abnormal clinical and laboratory findings, not elsewhere classified (R00-R99) (ICD10CM:R00.0-R99)
 - Symptoms and signs involving the circulatory and respiratory systems (R00-R09) (ICD10CM:R00.0-R09.89)
 - Pain in throat and chest (ICD10CM:R07.0-R07.9)
 - Chest pain on breathing (ICD10CM:R07.1)
 - Chest pain, unspecified (ICD10CM:R07.9)
 - Other chest pain (ICD10CM:R07.81-R07.89)
 - Other chest pain (ICD10CM:R07.89) 90.562
- Lab Results (LOINC)
 - LOINCCLASSTYPES (LNC:10154-3-MTHU854759)
 - Clinical Class (LNC:10154-3-MTHU854759)
 - PhenX (LNC:56057-3-67789-8)
 - In the past Mo, have you had chest pain when you were not doing physical activity (LNC:66235-3)
 - Pain in the chest, left arm or shoulder or jaw within 72H of death (LNC:58326-0)
 - Pain included chest (LNC:58327-8)
 - Surveys (LNC:28079-2-95414-9)
 - General Healthy survey (LNC:70301-7-95400-8)
 - How many times have you had chest pain, chest tightness or angina on average over the past 4W (LNC:88476-7)
 - How many times have you had to take nitroglycerin (nitroglycerin)

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

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

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

  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

All Concepts ▾ myocardial infarction ✕

▾ ⌚ Conditions (ICD-10-CM)

▾ ⌚ Diseases of the circulatory system (I00-I99) (ICD10CM:I00-I99.9) 📊 Learn More

▾ ⌚ Ischemic heart diseases (I20-I25) (ICD10CM:I20.0-I25.9)


- ⌚ Acute myocardial infarction (ICD10CM:I21.01-I21.A9)
- ⌚ Certain current complications following ST elevation (STEMI) and non-ST elevation (NSTEMI) myocardial infarction (within the 28 day period) (ICD10CM:I23.0-I23.8)
- ⌚ Other acute ischemic heart diseases (ICD10CM:I24.0-I24.9)
- ⌚ Subsequent ST elevation (STEMI) and non-ST elevation (NSTEMI) myocardial infarction (ICD10CM:I22.0-I22.9)

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.ica hn.mssm.edu/msdw/>



Scientific Computing / Cohort Selection Tools

Query Tool Access Request Form

Requestor Name:

Network ID:

Email:

Department Name:

Name of PI, if applicable (optional):

Platform:

Leaf

Data Source:

- ☐ MSDW 2.0 Anal Cancer
- ☐ MSDW 2.0 BioMe
- ☒ MSDW 2.0 De-identified
- ☐ MSDW 2.0 Liver
- ☐ MSDW 2.0 PHI

Brief Description of Data Request Purpose (optional):

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

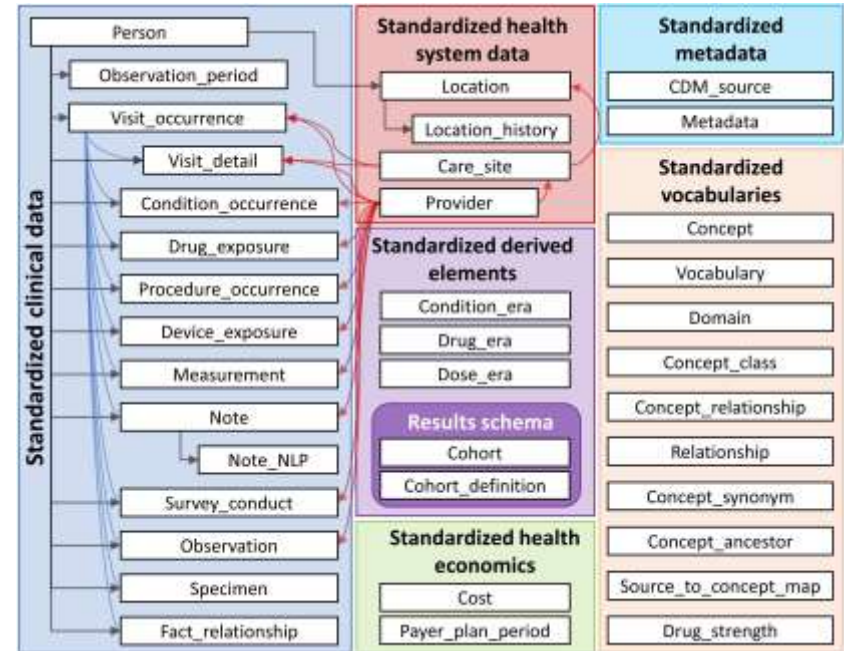


► Observational Medical Outcomes Partnership (OMOP)

- 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.ica hn.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

medabp01

Search

type 2 Diabetes

Advanced Options

Filter: Search...

Column visibility Copy CSV Show 15 entries

Showing 1 to 15 of 6,422 entries

Vocabulary

EPIC EDG .1 (5892)

SNOMED (147)

ICD10CM (125)

Nebraska Lexicon (108)

Class

Epic Concept (5892)

Clinical Finding (236)

7-char billing code (52)

Read (48)

Domain

Condition (6374)

Observation (41)

Procedure (4)

Measurement (2)

Standard Concept

Non-Standard (5892)

Unknown (404)

Standard (124)

Classification (2)

Invalid Reason

Valid (5892)

Unknown (483)

Invalid (47)

	Id	Code	Name	Class	RC	DRC	Domain	Vocabulary
<input checked="" type="checkbox"/>	4193704	313436004	Type 2 diabetes mellitus without complication	Clinical Finding	42,485	42,869	Condition	SNOMED
<input checked="" type="checkbox"/>	2001259621	1321962	Type 2 diabetes mellitus without complication, without long-term current use of insulin	Epic Concept	16,988	16,988	Condition	EPIC EDG .1
<input checked="" type="checkbox"/>	2000602205	521601	Type 2 diabetes mellitus without complications	Epic Concept	16,983	16,983	Condition	EPIC EDG .1
<input checked="" type="checkbox"/>	37016349	368051000119109	Hyperglycemia due to type 2 diabetes mellitus	Clinical Finding	12,144	12,144	Condition	SNOMED
<input checked="" type="checkbox"/>	201826	44054006	Type 2 diabetes mellitus	Clinical Finding	11,123	62,947	Condition	SNOMED
<input checked="" type="checkbox"/>	2000599869	523781	Type 2 diabetes mellitus with hyperglycemia	Epic Concept	6,409	6,409	Condition	EPIC EDG .1
<input checked="" type="checkbox"/>	443729	422166005	Peripheral circulatory disorder due to type 2 diabetes mellitus	Clinical Finding	2,761	4,465	Condition	SNOMED
<input checked="" type="checkbox"/>	2000599917	527480	Type 2 diabetes mellitus with other specified complication	Epic Concept	2,420	2,420	Condition	EPIC EDG .1
<input checked="" type="checkbox"/>	43531578	771000119108	Chronic kidney disease due to type 2 diabetes mellitus	Clinical Finding	2,400	2,612	Condition	SNOMED
<input checked="" type="checkbox"/>	2000602088	528363	Type 2 diabetes mellitus with unspecified complications	Epic Concept	2,391	2,391	Condition	EPIC EDG .1
<input checked="" type="checkbox"/>	2000691482	1322061	Type 2 diabetes mellitus without complication, with long-term current use of insulin	Epic Concept	2,326	2,326	Condition	EPIC EDG .1
<input checked="" type="checkbox"/>	2000480410	1321789	Uncontrolled type 2 diabetes mellitus with hyperglycemia	Epic Concept	2,295	2,295	Condition	EPIC EDG .1
<input checked="" type="checkbox"/>	443731	420279001	Renal disorder due to type 2 diabetes mellitus	Clinical Finding	2,207	4,819	Condition	SNOMED
<input checked="" type="checkbox"/>	37017432	713706002	Polyneuropathy due to type 2 diabetes mellitus	Clinical Finding	2,139	2,139	Condition	SNOMED
<input checked="" type="checkbox"/>	2000692533	1321828	Type 2 diabetes mellitus with hyperglycemia, without long-term current use of insulin	Epic Concept	1,853	1,853	Condition	EPIC EDG .1

Showing 1 to 15 of 6,422 entries

Previous 1 2 3 4 5 ... 429 Next

Concept Sets

[New Concept Set](#)

Column visibility

Copy

CSV

Show 15 entries

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

Home

Data Sources

Search

Concept Sets

Cohort Definitions

Characterizations

Cohort Pathways

Incidence Rates

Profiles

Estimation

Prediction

Jobs

Configuration

Feedback

Cohort Definitions

New Cohort

Column visibility Copy CSV Show 15 entries

Filter Search

Showing 1 to 15 of 41 entries

	Id	Name	Created	Updated	Author
Created					
2+ Weeks Ago (34)	55	DM Poor Control	05/02/2022 6:23 PM	05/02/2022 7:01 PM	medabp01
Last Week (3)	53	New users of ACE inhibitors as First line of Monotherapy	05/02/2022 5:06 PM	05/02/2022 5:06 PM	medabp01
Within 24 Hours (3)	52	High Myopia	05/02/2022 3:44 PM	05/02/2022 3:44 PM	banasb01
This Week (1)					
Updated					
2+ Weeks Ago (34)	51	Demo Cohort	05/01/2022 6:05 PM	05/01/2022 6:05 PM	medabp01
Within 24 Hours (3)	48	Thassos	04/20/2022 11:45 AM	04/29/2022 3:00 PM	lees152
This Week (2)	50	aa	04/22/2022 2:33 PM	04/22/2022 2:33 PM	minerg01
Last Week (2)	49	Externa for Nasal Polyps	04/22/2022 9:32 AM	04/22/2022 9:32 AM	donbok01
Author					
sampa01 (8)	47	Depression	04/17/2022 1:59 PM	04/17/2022 2:01 PM	medabp01
medabp01 (4)	46	Down 2	03/09/2022 11:08 AM	03/09/2022 11:08 AM	bogund01
konie01 (4)	45	Down syndrome test	03/09/2022 10:25 AM	03/09/2022 10:25 AM	bogund01
dant01 (2)	44	tbc demographics	01/26/2022 2:49 PM	01/26/2022 4:08 PM	owja01
Designs					
Other designs (34)	42	test rhino	01/25/2022 6:35 PM	01/25/2022 6:35 PM	hernam13
My designs (7)	41	affiliated	01/18/2022 4:51 PM	01/18/2022 4:52 PM	johnsk26
	40	Sexual assault encounter	01/11/2022 1:11 AM	01/11/2022 1:38 AM	lepowl01
	38	DRF Rehab	12/22/2021 7:33 AM	12/22/2021 7:39 AM	sternb06

Apache 2.0
open source software

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

- 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



Demographics



Diagnoses



Procedures



Medications



Lab Results



Vitals



Oncology



Genomics



NLP



Patient Linking

Search Term... 

ADD TO QUERY



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

- 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

all changes saved

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

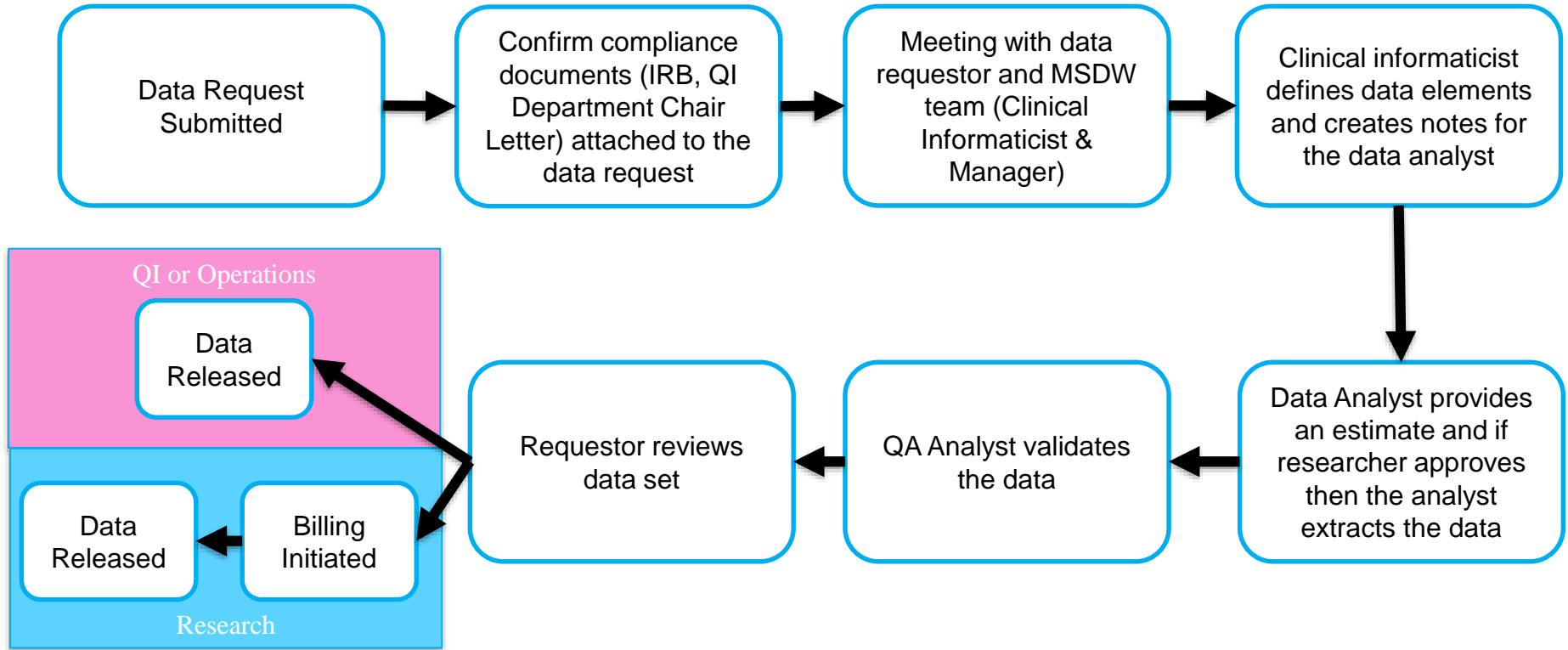
1

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.icaohn.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**