Leaf and Atlas Query Tools

Patricia Kovatch, Dean
Sharon Nirenberg, MD
Praveen Medabalmi, MD
Physician Informaticist
Scientific Computing and Data
Icahn School of Medicine at Mount Sinai

October 12, 2022
Objectives:

1. Scientific Computing FAIR Principles of Data
2. Data Sources and MSDW2 Ecosystem
3. Clinical query tools - Leaf and ATLAS
4. How to build a query using Leaf and ATLAS tools
5. When and how to request a custom dataset
Scientific Computing FAIR Principles for Data

**F**indable

**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/
### The FAIR Principles

<table>
<thead>
<tr>
<th><strong>Findability:</strong></th>
<th><strong>Accessibility:</strong></th>
<th><strong>Interoperability:</strong></th>
<th><strong>Reusability:</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>F1. (meta)data are assigned a globally unique and persistent identifier;</td>
<td>A1. (meta)data are retrievable by their identifier using a standardized communication protocol;</td>
<td>I1. (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.</td>
<td>R1. (meta)data are richly described with a plurality of accurate and relevant attributes;</td>
</tr>
<tr>
<td>F2. data are described with rich metadata;</td>
<td>A1.1. the protocol is open, free, and universally implementable;</td>
<td>I2. (meta)data use vocabularies that follow FAIR principles;</td>
<td>R1.1. (meta)data are released with a clear and accessible data usage license;</td>
</tr>
<tr>
<td>F3. metadata clearly and explicitly include the identifier of the data it describes;</td>
<td>A1.2. the protocol allows for an authentication and authorization procedure, where necessary;</td>
<td>I3. (meta)data include qualified references to other (meta)data;</td>
<td>R1.2. (meta)data are associated with detailed provenance;</td>
</tr>
<tr>
<td>F4. (meta)data are registered or indexed in a searchable resource;</td>
<td>A2. metadata are accessible, even when the data are no longer available;</td>
<td></td>
<td>R1.3. (meta)data meet domain-relevant community standards;</td>
</tr>
</tbody>
</table>

Improving reproducibility by using high-throughput observational studies with empirical calibration - [https://doi.org/10.1098/rsta.2017.0356](https://doi.org/10.1098/rsta.2017.0356)
MSDW2 Ecosystem
Leaf Query Tool

• Open-source, model-agnostic and data-driven web application for cohort discovery

• Lightweight self-service web application for querying data from different sources.

• Simple drag-and-drop user interface

• Research, Clinical Quality improvement

• Generate and test hypotheses

• Cohort size estimates before submitting Institutional Review Board (IRB) requests

• Access and Analyze clinical data after IRB approval as well

• De-Identified data
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
Leaf Data Types

1. Conditions: coded and descriptive diagnoses – ICD10
2. Procedures: procedure, service and events - CPT
3. Demographics: Age, Ethnicity, Gender, Race, and Vital Status
4. Encounters: ER visit, Inpatient, Outpatient, Telehealth
5. Vitals: Height, weight, BMI, O2 Sat, pulse, RR, Body Temp
6. Lab Results: clinical, laboratory, and survey measurements - LOINC
7. Medications: Medications prescribed and administered - ATC

LOINC - Logical Observation Identifiers Names and Codes
ICD-10 - International Classification of Diseases
ATC - In the Anatomical Therapeutic Chemical (ATC) classification system, the active substances are divided into different groups according to the organ or system on which they act and their therapeutic, pharmacological and chemical properties.
CPT – Current Procedural Terminology
# Leaf – Patient Cohorts

<table>
<thead>
<tr>
<th>Patient Cohorts on Leaf</th>
<th>Description</th>
<th>Number of patients</th>
</tr>
</thead>
<tbody>
<tr>
<td>BioMe Biobank</td>
<td>Patients who submitted tissue samples to Mount Sinai’s BioMe Biobank</td>
<td>61,541</td>
</tr>
<tr>
<td>BioMe Biobank Global Diversity Array - Sem4</td>
<td>Patients who submitted tissue samples to Mount Sinai’s BioMe Biobank and have had their DNA analyzed with Illumina’s Global Diversity Array by Sema4</td>
<td>20,521</td>
</tr>
<tr>
<td>BioMe Biobank Global Screening Array – Regeneron</td>
<td>Patients who submitted tissue samples to Mount Sinai’s BioMe Biobank and have had their DNA analyzed with Illumina’s Infinium Global Screening Array by Regeneron</td>
<td>31,304</td>
</tr>
<tr>
<td>BioMe Biobank whole Exome Sequencing – Regeneron</td>
<td>Patients who submitted tissue samples to Mount Sinai’s BioMe Biobank with whole exome sequence (WES) data generated by Regeneron</td>
<td>30,656</td>
</tr>
<tr>
<td>Cancer Institute Biorepository</td>
<td></td>
<td>14,831</td>
</tr>
<tr>
<td>Cancer Patient Cohort</td>
<td>Patients who have been diagnosed with cancer, refreshed on a monthly basis around the 15th of every month</td>
<td>254,041</td>
</tr>
<tr>
<td>Imaging Research Warehouse 1.0</td>
<td>Patients who have image data in version 1.0 of the Imaging Research Warehouse (IRW)</td>
<td>528,865</td>
</tr>
</tbody>
</table>
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)
• Searches on keywords in concept descriptions

• Shows nodes in the hierarchy that satisfy the query
Leaf Query – Lab Results
Leaf Query – Medications
Leaf Query - Diagnosis
Leaf Query - Patients diagnosed with COVID-19 and Intubated

[Diagram showing the process of selecting criteria for patients intubated due to COVID-19, including procedures and codes.]

- 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
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
Age, Race, Gender, and Ethnicity: these are considered immutable patient characteristics. Immutable patient characteristics are denoted by an orange human icon: ⚪. These characteristics cannot be combined with Encounter concepts in one query box. Doing so produces a search error.
Leaf Application Status and Roadmap

**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 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, BioBank, the Imaging Research Warehouse 1.0, the Cancer Institute Biorepository, and Cancer Patients

Leaf contains these known bugs:
- Patient List temporarily disabled
- Queries that last longer than 1 minute terminate

**Roadmap**

These improvements to Leaf are ongoing for the first half of 2022:
- Fix high-priority known bugs
- Enhance the Leaf development and configuration process to make it easier to support Leaf’s users
- In addition, Leaf will access all improvements that are made to MSDW2.
ATLAS

• Open source web-based application developed by OHDSI - Observational Health Data Sciences and Informatics Community
• Generate real world evidence from patient level clinical data
• Data Sources, Concepts, Cohorts
• **Common Data Model** - A convention for representing healthcare data that allows portability of analysis
• **Concept** - A term (with a code) defined in a medical terminology, all clinical events in the OMOP CDM are expressed as concepts
• **Concept set** is an expression representing a list of concepts that can be used as a reusable component in various analyses
• **Cohort** is a set of persons who satisfy one or more inclusion criteria for a duration of time
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
## Concepts

<table>
<thead>
<tr>
<th>Concept</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>CONCEPT_ID</td>
<td>313217</td>
</tr>
<tr>
<td>CONCEPT_NAME</td>
<td>Atrial fibrillation</td>
</tr>
<tr>
<td>DOMAIN_ID</td>
<td>Condition</td>
</tr>
<tr>
<td>VOCABULARY_ID</td>
<td>SNOMED</td>
</tr>
<tr>
<td>CONCEPT_CLASS_ID</td>
<td>Clinical Finding</td>
</tr>
<tr>
<td>STANDARD_CONCEPT</td>
<td>S</td>
</tr>
<tr>
<td>CONCEPT_CODE</td>
<td>49436004</td>
</tr>
<tr>
<td>VALID_START_DATE</td>
<td>01-Jan-1970</td>
</tr>
<tr>
<td>VALID_END_DATE</td>
<td>31-Dec-2099</td>
</tr>
<tr>
<td>INVALID_REASON</td>
<td></td>
</tr>
</tbody>
</table>

- **Primary key**: CONCEPT_ID
- **English description**: Atrial fibrillation
- **Domain**: Condition
- **Vocabulary**: SNOMED
- **Class in vocabulary**: Clinical Finding
- **Standard, Source of Classification**: S
- **Code in vocabulary**: 49436004
- **Valid during time interval**: 01-Jan-1970 to 31-Dec-2099
Disease Hierarchy (SNOMED)
ATLAS Data Sources – Person

- One unique ID for each Person

- Person table includes all the demographics - year of birth, gender, race, and ethnicity

- Overview of demographic distribution and observe any potential data outliers.

- Foreign key to the LOCATION, PROVIDER, and CARE_SITE table that contains one record
Visit Occurrence

- Visits are Encounters
- Span of times for Medical Services
- Visit Types
  - Emergency room
  - Inpatient
  - Inpatient/Emergency
  - Outpatient
  - Long-term care
Condition Occurrence

- Diagnoses, signs, or symptoms of a condition either observed by a Provider or reported by the patient

- Condition Concept prevalence over time, age information within the population, and frequency of diagnosis

- Examples:
  - Billing Diagnosis
  - Problem List
Drug Exposure

• Records about the intent or actual introduction of a drug into the body of the patient.

• Prescription and over-the-counter medicines, vaccines, and large-molecule biologic therapies.

• Clinical events associated with orders, prescriptions written, pharmacy dispensing, procedural administrations, patient-reported information.
Procedure Occurrence

- Contains records of activities or processes ordered or carried out by a healthcare Provider on the patient with a diagnostic or therapeutic purpose.

- Vocabularies include CPT-4, HCPCS, ICD-9 Procedures, ICD-10 Procedures, LOINC, SNOMED
Measurement

- Contains records of:
  - Measurement
  - Structured values (numerical or categorical) obtained through systematic and standardized examination or testing of a Person or Person's sample
- Data sources:
  - Structured
  - Quantitative measures, such as laboratory tests
- Measures have associated units
Observation

• Captures clinical facts about a Person obtained in the context of examination, questioning or a procedure

• Any data that cannot be represented by any other domains, such as social and lifestyle facts, medical history, family history, etc. are recorded here

• Instrument for CDM extension
ATLAS - Cohort

• Review the dashboard – Data sources

• Search Concepts – Hypertension, Diabetes, Asthma, Obesity

• Review the Concept hierarchy – Parent and Children Concepts

• Add Descendants – Concepts that are children to the main concept

• Exclude concepts – Example Gestational Hypertension

• Build the Cohort using the concept sets

• Add Initial event, Dates and Attributes
<table>
<thead>
<tr>
<th>Id</th>
<th>Code</th>
<th>Name</th>
<th>Class</th>
<th>RCoV</th>
<th>DRC</th>
<th>Domain</th>
<th>Vocabulary</th>
</tr>
</thead>
<tbody>
<tr>
<td>4193704</td>
<td>313436904</td>
<td>Type 2 diabetes mellitus without complication</td>
<td>Clinical Finding</td>
<td>42.485</td>
<td>42.869</td>
<td>Condition</td>
<td>SNOMED</td>
</tr>
<tr>
<td>20012569621</td>
<td>1321962</td>
<td>Type 2 diabetes mellitus without complication, without long-term current use of insulin</td>
<td>Epic Concept</td>
<td>16.988</td>
<td>16.988</td>
<td>Condition</td>
<td>EPIC EDG.1</td>
</tr>
<tr>
<td>2000602205</td>
<td>521601</td>
<td>Type 2 diabetes mellitus without complications</td>
<td>Epic Concept</td>
<td>16.983</td>
<td>16.983</td>
<td>Condition</td>
<td>EPIC EDG.1</td>
</tr>
<tr>
<td>37016349</td>
<td>368051000119109</td>
<td>Hyperglycemia due to type 2 diabetes mellitus</td>
<td>Clinical Finding</td>
<td>12.144</td>
<td>12.144</td>
<td>Condition</td>
<td>SNOMED</td>
</tr>
<tr>
<td>201826</td>
<td>44054006</td>
<td>Type 2 diabetes mellitus</td>
<td>Clinical Finding</td>
<td>11.123</td>
<td>62.947</td>
<td>Condition</td>
<td>SNOMED</td>
</tr>
<tr>
<td>2000599869</td>
<td>523781</td>
<td>Type 2 diabetes mellitus with hyperglycemia</td>
<td>Epic Concept</td>
<td>6.409</td>
<td>6.409</td>
<td>Condition</td>
<td>EPIC EDG.1</td>
</tr>
<tr>
<td>443729</td>
<td>422160005</td>
<td>Peripheral circulatory disorder due to type 2 diabetes mellitus</td>
<td>Clinical Finding</td>
<td>2.761</td>
<td>4.465</td>
<td>Condition</td>
<td>SNOMED</td>
</tr>
<tr>
<td>2000599917</td>
<td>527480</td>
<td>Type 2 diabetes mellitus with other specified complication</td>
<td>Epic Concept</td>
<td>2.420</td>
<td>2.420</td>
<td>Condition</td>
<td>EPIC EDG.1</td>
</tr>
<tr>
<td>43531578</td>
<td>771000119108</td>
<td>Chronic kidney disease due to type 2 diabetes mellitus</td>
<td>Clinical Finding</td>
<td>2.400</td>
<td>2.612</td>
<td>Condition</td>
<td>SNOMED</td>
</tr>
<tr>
<td>2000602088</td>
<td>528363</td>
<td>Type 2 diabetes mellitus with unspecified complications</td>
<td>Epic Concept</td>
<td>2.391</td>
<td>2.391</td>
<td>Condition</td>
<td>EPIC EDG.1</td>
</tr>
<tr>
<td>2000480410</td>
<td>1322061</td>
<td>Type 2 diabetes mellitus without complication, with long-term current use of insulin</td>
<td>Epic Concept</td>
<td>2.326</td>
<td>2.326</td>
<td>Condition</td>
<td>EPIC EDG.1</td>
</tr>
<tr>
<td>2000691482</td>
<td>1322061</td>
<td>Type 2 diabetes mellitus without complication, with long-term current use of insulin</td>
<td>Epic Concept</td>
<td>2.295</td>
<td>2.295</td>
<td>Condition</td>
<td>EPIC EDG.1</td>
</tr>
<tr>
<td>2000480410</td>
<td>1321789</td>
<td>Uncontrolled type 2 diabetes mellitus with hyperglycemia</td>
<td>Clinical Finding</td>
<td>2.207</td>
<td>4.819</td>
<td>Condition</td>
<td>SNOMED</td>
</tr>
<tr>
<td>443731</td>
<td>420279001</td>
<td>Renal disorder due to type 2 diabetes mellitus</td>
<td>Clinical Finding</td>
<td>2.139</td>
<td>2.139</td>
<td>Condition</td>
<td>SNOMED</td>
</tr>
<tr>
<td>37017432</td>
<td>713706002</td>
<td>Polyneuropathy due to type 2 diabetes mellitus</td>
<td>Epic Concept</td>
<td>1.853</td>
<td>1.853</td>
<td>Condition</td>
<td>EPIC EDG.1</td>
</tr>
</tbody>
</table>
Concept set – Inclusion and Exclusion Criteria
ATLAS Concept Sets

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

<table>
<thead>
<tr>
<th>Concept ID</th>
<th>Concept Name</th>
<th>Descendants</th>
<th>Mapped</th>
<th>Excluded</th>
</tr>
</thead>
<tbody>
<tr>
<td>4329847</td>
<td>Myocardial Infarction</td>
<td>YES</td>
<td>YES</td>
<td>NO</td>
</tr>
<tr>
<td>314666</td>
<td>Old Myocardial Infarction</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Id</td>
<td>Name</td>
<td>Created</td>
<td>Modified</td>
<td>Author</td>
</tr>
<tr>
<td>----</td>
<td>-----------------------</td>
<td>------------------</td>
<td>------------------</td>
<td>---------</td>
</tr>
<tr>
<td>68</td>
<td>Essential hypertension</td>
<td>05/02/2022 6:51 PM</td>
<td>05/03/2022 4:30 PM</td>
<td>medabp01</td>
</tr>
<tr>
<td>70</td>
<td>Colitis</td>
<td>05/03/2022 4:08 PM</td>
<td>05/03/2022 4:06 PM</td>
<td>medabp01</td>
</tr>
<tr>
<td>69</td>
<td>GI Bleed</td>
<td>05/03/2022 2:49 PM</td>
<td>05/03/2022 2:49 PM</td>
<td>medabp01</td>
</tr>
<tr>
<td>65</td>
<td>COVID-19 and MI</td>
<td>05/02/2022 3:48 PM</td>
<td>05/02/2022 5:46 PM</td>
<td>banasb01</td>
</tr>
<tr>
<td>64</td>
<td>Diabetes Mellitus</td>
<td>05/01/2022 5:53 PM</td>
<td>05/01/2022 5:53 PM</td>
<td>medabp01</td>
</tr>
<tr>
<td>63</td>
<td>THASSOS_p_link</td>
<td>04/29/2022 1:33 PM</td>
<td>04/29/2022 1:33 PM</td>
<td>lees152</td>
</tr>
<tr>
<td>62</td>
<td>COPY OF aa</td>
<td>04/22/2022 1:55 PM</td>
<td>04/22/2022 1:55 PM</td>
<td>minerg01</td>
</tr>
<tr>
<td>61</td>
<td>aa</td>
<td>04/22/2022 12:59 PM</td>
<td>04/22/2022 1:42 PM</td>
<td>minerg01</td>
</tr>
<tr>
<td>59</td>
<td>Depression</td>
<td>04/17/2022 1:09 PM</td>
<td>04/17/2022 1:56 PM</td>
<td>medabp01</td>
</tr>
<tr>
<td>17</td>
<td>covid</td>
<td>10/19/2021 12:55 PM</td>
<td>04/07/2022 9:31 PM</td>
<td>nirens01</td>
</tr>
<tr>
<td>56</td>
<td>Migraines</td>
<td>04/01/2022 10:15 AM</td>
<td>04/01/2022 10:19 AM</td>
<td>dent01</td>
</tr>
<tr>
<td>54</td>
<td>MSI Test</td>
<td>03/15/2022 1:49 PM</td>
<td>03/24/2022 2:27 PM</td>
<td>grubec05</td>
</tr>
<tr>
<td>53</td>
<td>Product of Conception</td>
<td>03/09/2022 11:01 AM</td>
<td>03/09/2022 11:08 AM</td>
<td>bogund01</td>
</tr>
<tr>
<td>52</td>
<td>Down syndrome</td>
<td>03/09/2022 10:13 AM</td>
<td>03/09/2022 10:58 AM</td>
<td>bogund01</td>
</tr>
<tr>
<td>51</td>
<td>d22.9</td>
<td>01/26/2022 2:50 PM</td>
<td>01/26/2022 2:50 PM</td>
<td>owjis01</td>
</tr>
</tbody>
</table>
COHORT Definition

A cohort definition defines when an individual qualifies (or does not qualify) for a cohort during this span of time.

**Cohort entry event** which happens in a moment in time to qualify an individual for cohort entry.

**Inclusion criteria**

**Cohort exit** is the event in time at which an individual does not qualify anymore for a cohort and leaves it.
Defining a Cohort
### Cohort Definitions

<table>
<thead>
<tr>
<th>Id</th>
<th>Name</th>
<th>Created</th>
<th>Updated</th>
<th>Author</th>
</tr>
</thead>
<tbody>
<tr>
<td>34</td>
<td>2+ Weeks Ago</td>
<td>05/02/2022 6:23 PM</td>
<td>05/02/2022 7:01 PM</td>
<td>medabp01</td>
</tr>
<tr>
<td>34</td>
<td>Last Week</td>
<td>05/02/2022 5:06 PM</td>
<td>05/02/2022 5:06 PM</td>
<td>medabp01</td>
</tr>
<tr>
<td>34</td>
<td>Within 24 Hours</td>
<td>05/02/2022 3:44 PM</td>
<td>05/02/2022 3:44 PM</td>
<td>medabp01</td>
</tr>
<tr>
<td>34</td>
<td>This Week</td>
<td>05/01/2022 6:05 PM</td>
<td>05/01/2022 6:05 PM</td>
<td>medabp01</td>
</tr>
<tr>
<td>34</td>
<td>Demo Cohort</td>
<td>04/20/2022 11:45 AM</td>
<td>04/29/2022 3:00 PM</td>
<td>lees152</td>
</tr>
<tr>
<td>34</td>
<td>Thessos</td>
<td>04/22/2022 2:33 PM</td>
<td>04/22/2022 2:33 PM</td>
<td>minerg01</td>
</tr>
<tr>
<td>34</td>
<td>aa</td>
<td>04/22/2022 9:32 AM</td>
<td>04/22/2022 9:32 AM</td>
<td>donbak01</td>
</tr>
<tr>
<td>34</td>
<td>Fasenza for Nasal Polyps</td>
<td>04/17/2022 1:59 PM</td>
<td>04/17/2022 2:01 PM</td>
<td>medabp01</td>
</tr>
<tr>
<td>34</td>
<td>Depression</td>
<td>03/09/2022 11:08 AM</td>
<td>03/09/2022 11:08 AM</td>
<td>bogund01</td>
</tr>
<tr>
<td>34</td>
<td>Down 2</td>
<td>03/09/2022 10:25 AM</td>
<td>03/09/2022 10:25 AM</td>
<td>bogund01</td>
</tr>
<tr>
<td>34</td>
<td>Down syndrome test</td>
<td>01/26/2022 4:08 PM</td>
<td>01/26/2022 4:08 PM</td>
<td>ogwjd01</td>
</tr>
<tr>
<td>34</td>
<td>tbc demographics</td>
<td>01/25/2022 6:35 PM</td>
<td>01/25/2022 6:35 PM</td>
<td>hemam13</td>
</tr>
<tr>
<td>34</td>
<td>test_ching</td>
<td>01/18/2022 4:52 PM</td>
<td>01/18/2022 4:52 PM</td>
<td>johnok26</td>
</tr>
<tr>
<td>34</td>
<td>alyparents</td>
<td>01/11/2022 1:38 AM</td>
<td>01/11/2022 1:38 AM</td>
<td>leprow01</td>
</tr>
<tr>
<td>34</td>
<td>Sexual assault encounter</td>
<td>12/22/2021 7:39 AM</td>
<td>12/22/2021 7:39 AM</td>
<td>sternd06</td>
</tr>
<tr>
<td>34</td>
<td>DRF Rehab</td>
<td>12/22/2021 7:39 AM</td>
<td>12/22/2021 7:39 AM</td>
<td>sternd06</td>
</tr>
</tbody>
</table>
ATLAS Application Status and Roadmap

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.

As of December 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 version 1 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
# Clinical Query Tools – Leaf and Atlas

<table>
<thead>
<tr>
<th></th>
<th>Leaf</th>
<th>OMOP/ATLAS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Description</strong></td>
<td>Web-based, lightweight drag-and-drop query tool that quickly analyzes population demographics</td>
<td>Web-based integrated platform for database exploration, standardized vocabulary browsing, cohort definition, and population-level analysis</td>
</tr>
<tr>
<td><strong>Access</strong></td>
<td>Use your Mount Sinai network username/password to login</td>
<td>Use your Mount Sini School network username/password. For users with only a Hospital account, request a school account through Sailpoint to access ATLAS.</td>
</tr>
<tr>
<td><strong>Data Types</strong></td>
<td>Conditions (diagnoses), procedures, demographics, encounters (patient encounter locations), lab results, vitals, patient cohorts</td>
<td>Facilities, diagnoses, procedures, medications, labs, orders, patient demographics</td>
</tr>
<tr>
<td><strong>PHI</strong></td>
<td>No</td>
<td>Yes, if IRB approved</td>
</tr>
<tr>
<td><strong>Turnaround</strong></td>
<td>Seconds</td>
<td>Seconds</td>
</tr>
<tr>
<td><strong>Advantages/Disadvantages</strong></td>
<td>Can visualize demographic details of cohorts, drag-and-drop query feature</td>
<td>Utilizes common data model and queries; numerous query analysis tools available in github</td>
</tr>
</tbody>
</table>

See more details at [https://labs.icahn.mssm.edu/msdw/services/](https://labs.icahn.mssm.edu/msdw/services/)
Query Tool Access Request Form - Cohort Selection Tools
https://labs.icahn.mssm.edu/msdw/
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

Data Request Submitted

Confirm compliance documents (IRB, QI Department Chair Letter) attached to the data request

Meeting with data requestor and MSDW team (Clinical Informaticist & Manager)

Clinical informaticist defines data elements and creates notes for the data analyst

Clinical Informaticist

Data Requestor

Clinical Informaticist

Data Analyst

Clinical Informaticist

Billing Initiated

Research

JIRA ticketing system used to monitor the status of data requests
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/

Office Hours onsite – Nov 1st week

Survey - [https://redcap.link/alt8n8dm](https://redcap.link/alt8n8dm) or