Mount Sinai Data Warehouse Town Hall

Mark Green, PhD
Farhan Mahmood
Sharon Nirenberg, MD, MS
Scientific Computing and Data
Icahn School of Medicine at Mount Sinai
May 1, 2024

Icahn School of Medicine at Mount Sinai

Agenda

1. MSDW Team Updates
2. MSDW Operations
3. MSDW Major Accomplishments
    - Leaf improvements
    - Epic for research
    - Access to new data sources
4. MSDW Roadmap April - October 2024

MSDW Team Updates

The MSDW Team

Patricia Kovatch
Professor & Dean for Scientific Computing

Mark Green PhD
Executive Director

Sharon Nirenberg MD
Physician Informaticist

Naomi So MD
Physician Informaticist

Farhan Mahmood
Director Scientific Computing

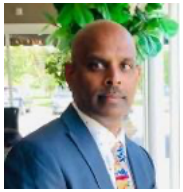
Teja Ganta MD
Physician Informaticist

Raj Bose PhD
Director Research Engagement

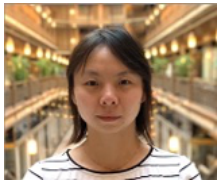
Jiani Xiang
Clinical Data Specialist

Priyal Mehta
Healthcare Data Analyst

Timothy Quinn PhD
Principal Data Architect



Shivaji Punukollu
Healthcare Data Engineer



Jing Yang PhD
Research Engagement Specialist

Jacob Weiser
Healthcare Data Engineer

Darius Boopal
Healthcare Data Engineer

Rupan Hossain
Database Administrator

MSDW Operations

MSDW Data Sets Delivered

From Apr 2023 until Mar 2024, MSDW team has delivered **253** custom data sets to **180** distinct users and **37%** are repeat users

30 (~15%) more custom data sets delivered from April 2023 to March 2024 compared to April 2022 to March 2023

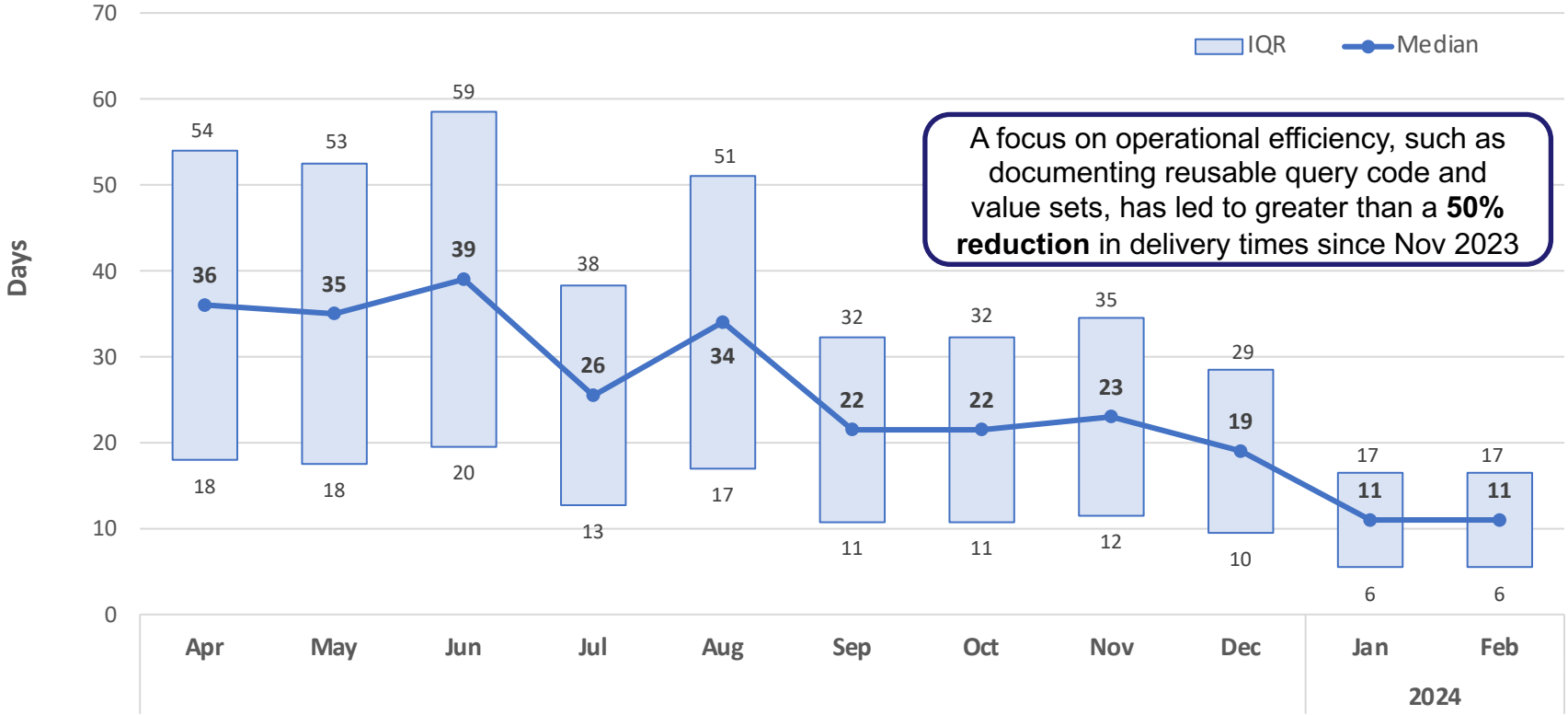
Top 5 MSDW Users – Apr 2023 until Mar 2024

User	Department	Data Sets Delivered
Harm Van Bakel	Genetics/Genomic Sci - ISM	11
Amy Kontorovich	Cardiology - ISM	7
Juan Wisnivesky	Div Gen Int Medicine	6
Dusan Bogunovic	Clinical Immunology	5
Robert Hirten	Medicine Gastroenterology	4

Top 5 MSDW Departments – Apr 2023 until Mar 2024

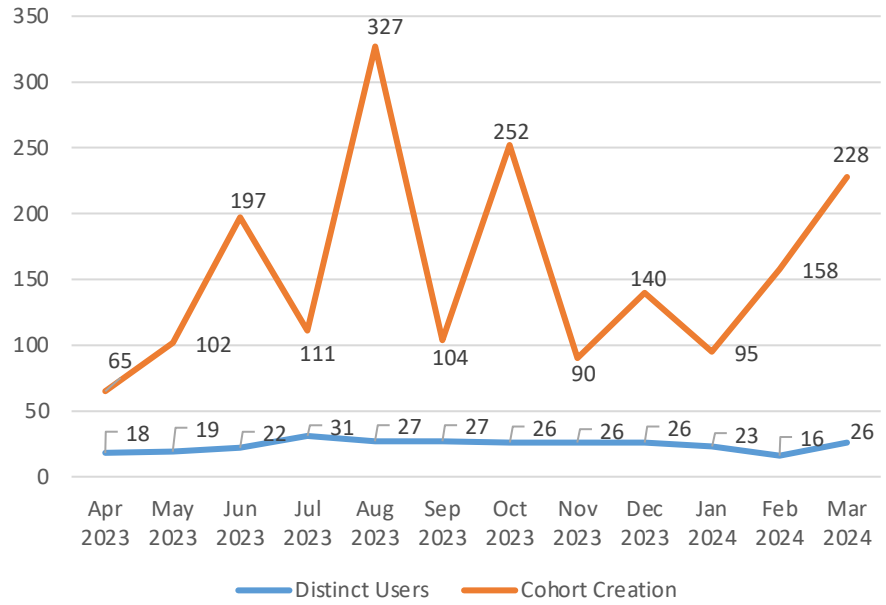
Department	Data Sets Delivered
Medicine - General Internal Medicine	27
Genetics and Genomic Sciences	24
Medicine - Cardiology	18
Population Health Science and Policy	16
Medicine - Gastroenterology	14

Custom Data Set Delivery Times Improved in 2024

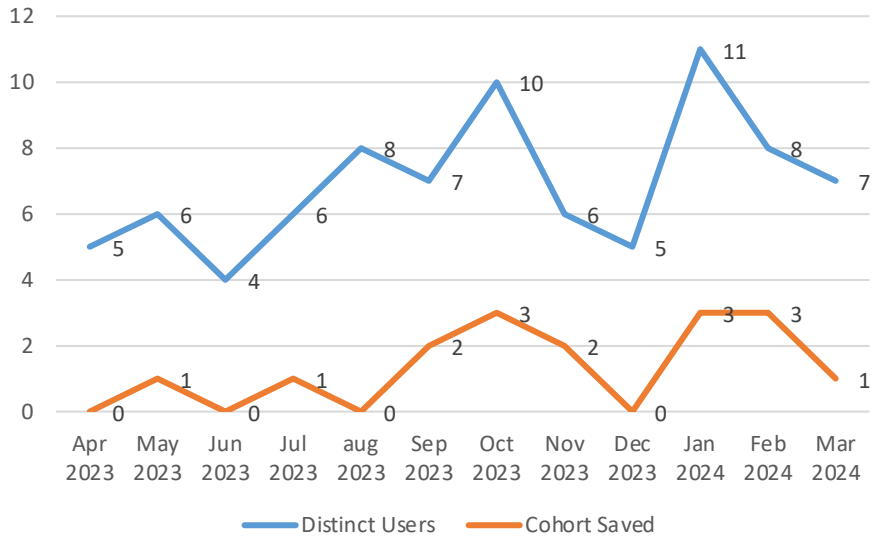


Utilization of Leaf and ATLAS (April 2023 to March 2024)

Leaf Utilization



ATLAS Utilization



The number of distinct individuals using Leaf has increased more than 40% over the last 12 months with a Month over Month (MoM) cohort creation increase of 65% in 2024

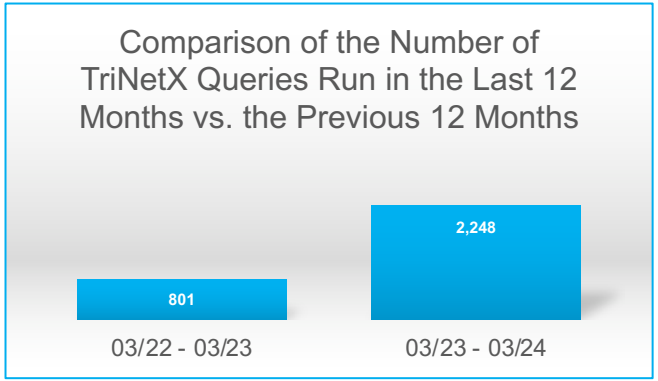
TriNetX Utilization Has Increased Over the Last Year

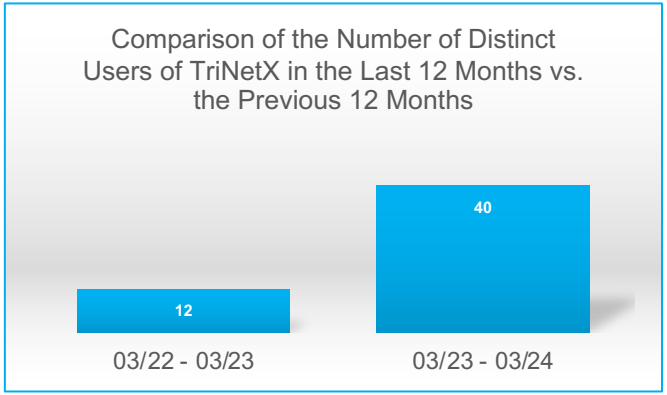
181% increase in the number of queries run on TriNetX in the last year

> 3X TriNetX users in the last year

Reason for increased utilization:

1. Quarterly data updates in TriNetX from MSDW OMOP database over the last 12 months
2. Increased TriNetX utilization to assess clinical trial feasibility





Outreach to MSDW Users: September 2023 – March 2024

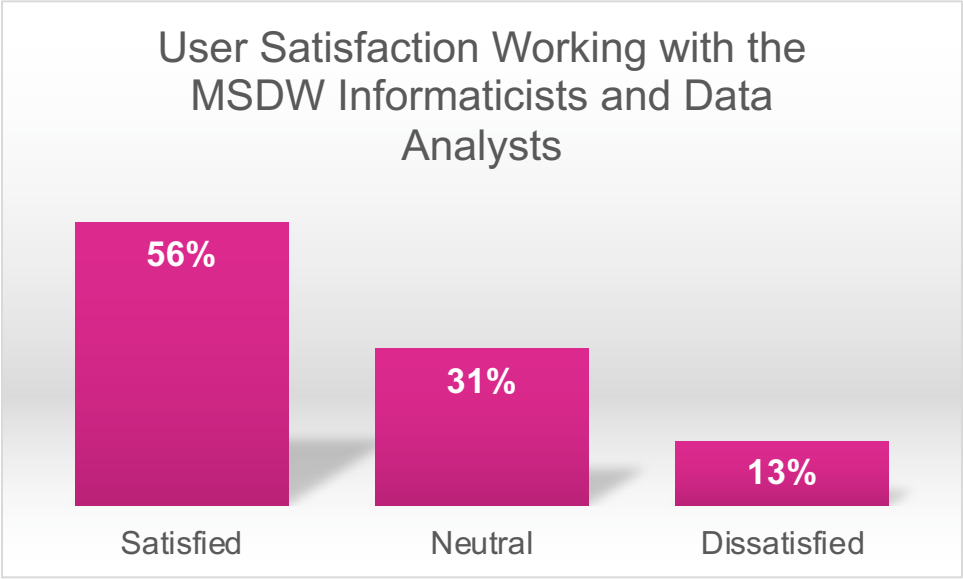
Date	Event	Participants
October 24, 2023	Town Hall	4
November 7, 2023	Epic for Research Training	150
January 10, 2023	TriNetX Training Session	13
January 17, 2024	Leaf and ATLAS Training Session	11
March 15, 2024	Presentation to Infectious Disease Researchers	8
March 21, 2024	Presentation to Oncology Fellows on Leaf	14
Every Wednesday	Digital Concierge	515
	TOTAL	**715**

Annual MSDW User Survey

Survey distributed in January 2024 to 616 active users of MSDW services and resources

19 users responded to the survey (3% response rate)

More than 56% of respondents had a very satisfying or satisfying experience working the MSDW informaticists and data analysts



Free Texts Comments from the User Survey

Positive:

”I consulted with the MSDW team recently. They were extremely helpful in answering all of my questions.”

“Farhan was great and worked hard to get us what we needed. Also, Naomi So was very good.”

Needs improvement:

”Terrible turn around time.”

“There is no availability of the major of the data I need within MSDW. I work in neonatology and the BirthFact table from Caboodle isn’t integrated.”

“MSDW has a lot of missing values, even for data from the EMR, but somehow not in the correct fields.”

Action items:

Reassessing strategies to speed up the data delivery process after payment process begins

Adding the BirthFact table to the MSDW integration wish list

Require MSDW Users to Agree to Acknowledge the CTSA

As of January 1, 2024, anyone who requests a custom data set from the Mount Sinai Data Warehouse must agree to cite the CTSA in any publications resulting from the requested data set.

The agreement is part of the ticket intake process.

This is because of the support provided by the CTSA for the MSDW.

Supported by the Clinical and Translational Science Awards (CTSA) grant UL1TR004419 from the National Center for Advancing Translational Sciences, National Institutes of Health.

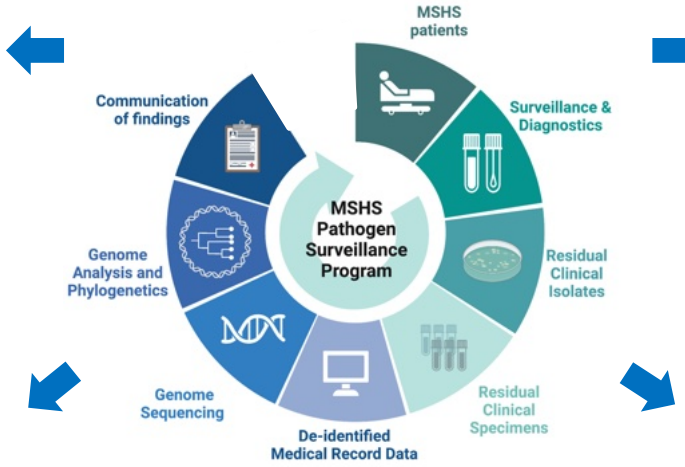
Winner of the CTSA Team Science Award:
The Mount Sinai Pathogen Surveillance Program

Leadership Team


Harm van Bakel

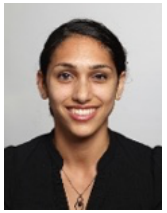

Viviana Simon


Mia Sordillo



Departments of Medicine and Microbiology


Komal Srivastava


Deena Altman

Department of Genetics and Genomic Sciences


Ana Silvia Gonzalez-Reiche


Bobby Sebra

Scientific Computing and Data


Farhan Mahmood


Sharon Nirenberg


Alona Rabin

Department of Pathology, Molecular and Cell Based Medicine


Alberto Paniz-Mondolfi


Juan Ramirez Gonzalez


Melissa Gitman

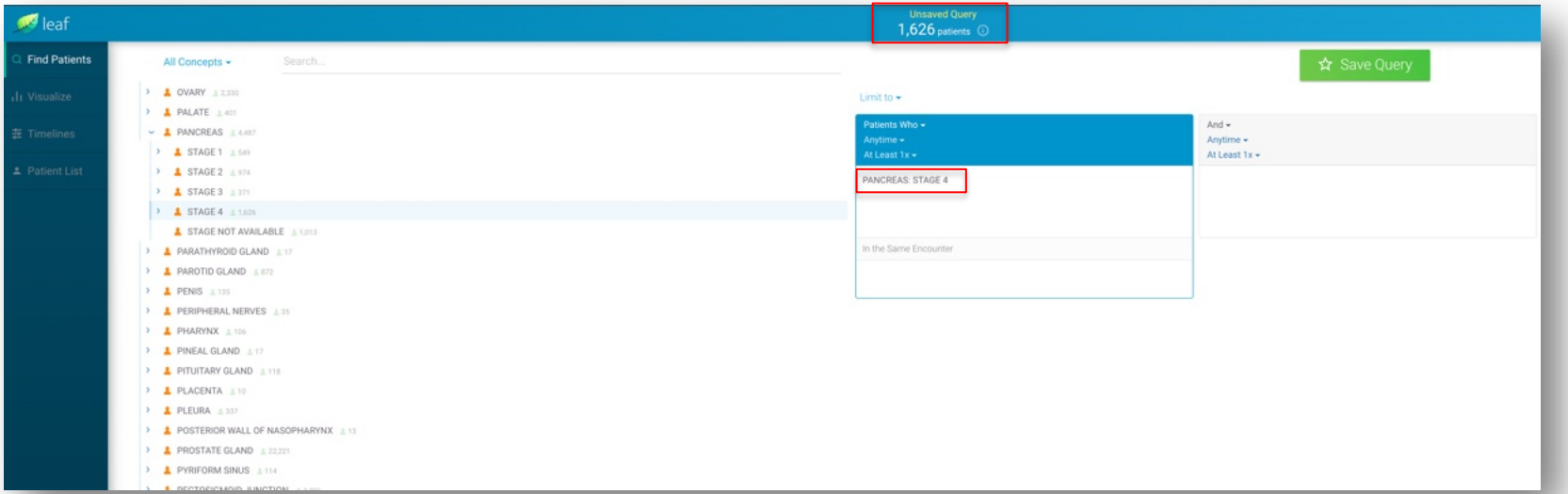
MSDW Major Accomplishments

Cancer Stage Now Searchable in Leaf

Cancer Stage data from Mount Sinai's Cancer Registry

- Categorized by primary tumor site
- Next step to augment with cancer stage data available in Epic

Improves assessment of oncology clinical trial feasibility



Tumor histology coming in summer 2024

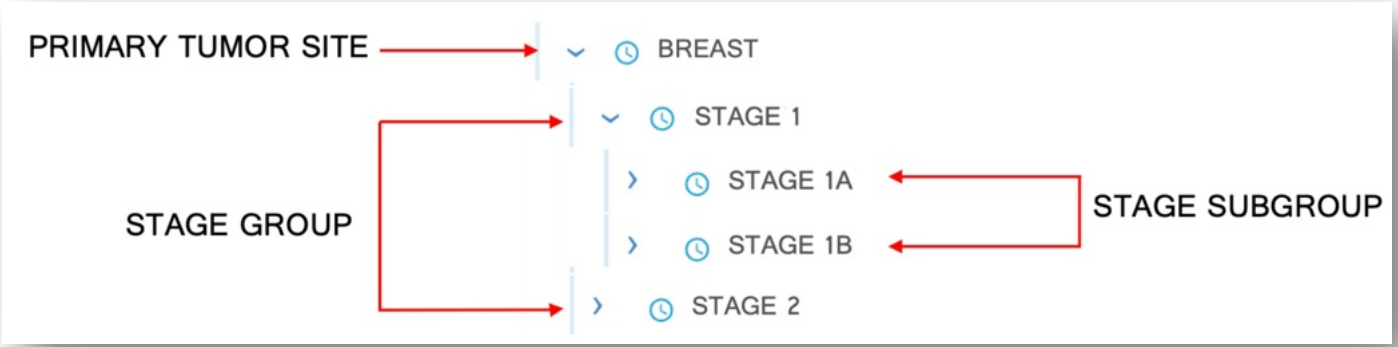
Deployed New Algorithm to Clean Staging Data from Cancer Registry

Four cancer stage summary fields in the cancer registry

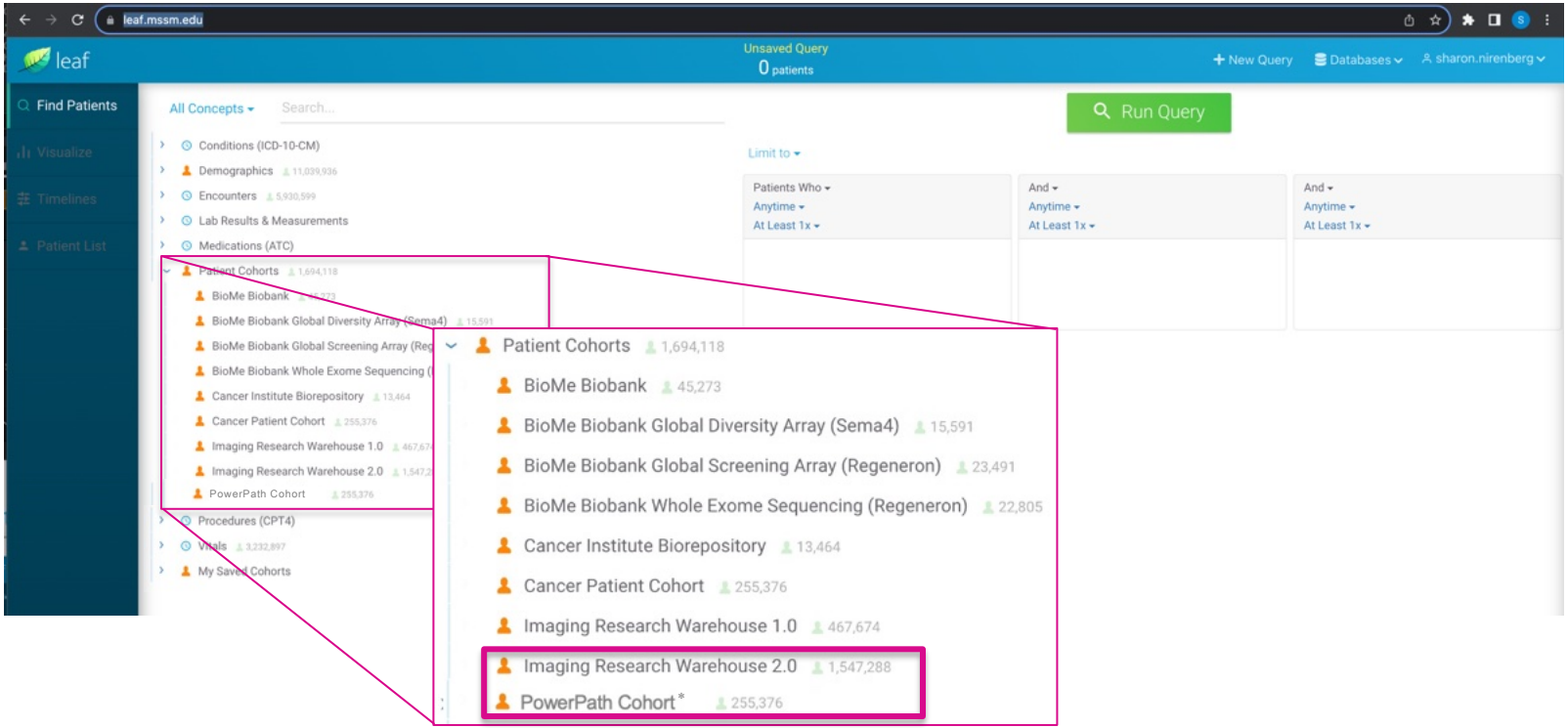
- PATH_STG - TNM Path Stage Group
- CLIN_STG - TNM Clinical Stage Group
- AJCPTHSG - AJCC TNM Path Stage Group
- AJCCLNSG - AJCC TNM Clinical Stage Group

Subject matter expertise provided by Dr. Teja Ganta to prioritize the 4 stage summary fields in the cancer registry

- **PATH_STG > AJCPTHSG > CLIN_STG > AJCCLNSG**



Added New Institutional Patient Cohorts to Leaf



Imaging Research Warehouse 2.0 and PowerPath cohorts searchable in Leaf

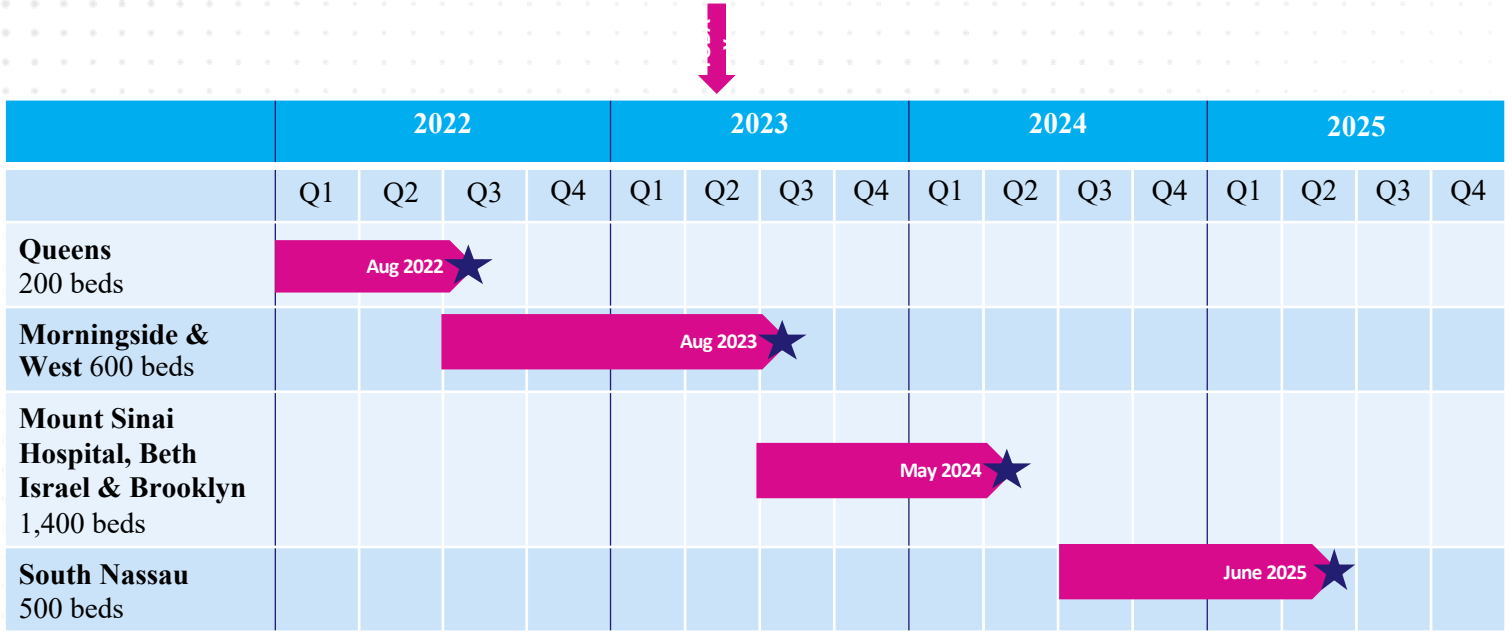
Improved Leaf Query Speed

Across all domains, average query performance increased by ~50%!

Domain	Query	Pre-Optimization Run Time (seconds)	Post-Optimization Run Time (seconds)	Improvement
Demographic + Encounter	Age >=18 AND (Inpatient Visit Or Emergency Room Visit in past 12 months)	88.6	29.9	66 %
Patient Cohorts	IRW 2.0	143	10.2	93 %
Labs	Hemoglobin [mass/volume] blood	Query timed out	63.4	100%
All Domains	All queries	N/A	N/A	~50%

Optimizing the speed of “Conditions” (ICD-10 CM) this summer

Mount Sinai Health System (MSHS): one Epic
Increasing opportunities for patient recruitment



2,700 beds in total

Patients are actively engaged with Epic MyMountSinai Helping us bring research studies to all!

Epic EHR

- Our Epic EHR is a common element across all sites and clinics
- Our investment in the Epic EHR will grow
- Our goal is to bring clinical trials to all

MyMountSinai (Epic MyChart) website & app

- 1.7M total MyMountSinai accounts
- ~370k patients use MyMountSinai every month
- 61.4% of patients seen in the last month have an account

Patients on Clinical Trials

- 69.1% of patients participating in clinical trials have an account
    - Reference: 60% average across Epic academic organizations

Mount Sinai Community Adopting Epic Research

Successfully using MyChart to identify eligible and interested study participants

- PRISM study – PI: Dr. Rosalind Wright
- About 30 new participants identified weekly
- Second study to begin using MyChart to recruit patients in April 2024

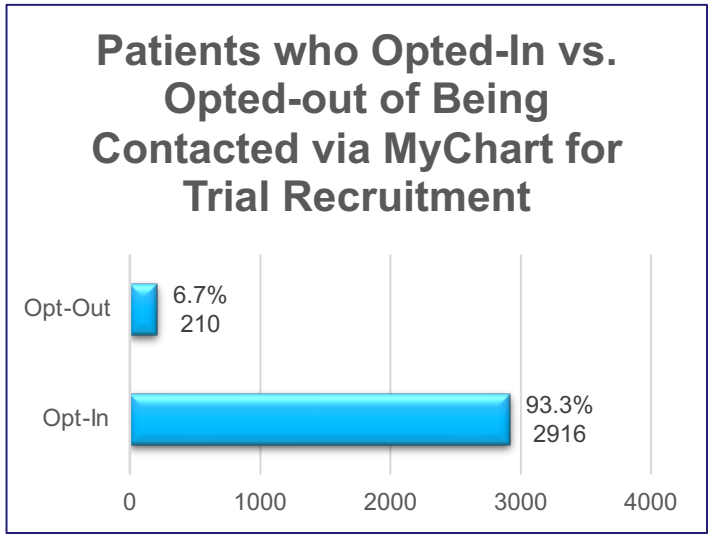
Created New Epic Research Advisory Group

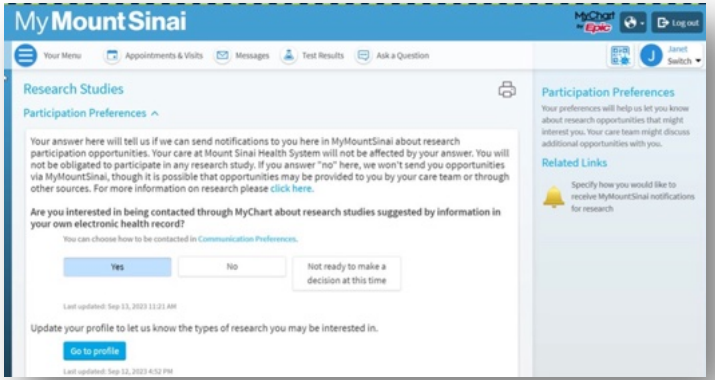
- Participation from clinical informatics, Epic, PPHS, and Scientific Computing and Data
- Review requests received from ServiceNow and MSDW to use Epic for research

Training Sessions on current Epic Research features

- Collaborative effort between Epic and Scientific Computing & Data Division
- Initial training session held in November 2023 with > **150 attendees**
- Spring Epic for Research training scheduled for May 14, 2024

Launched MyChart Research Opt-Out in June 2023





In total 3,126 patients have responded to the Epic Research Consent

210 (6.7%) of respondents have opted-out of being contacted via MyChart for research studies suggested by information in their electronic health record

Linked Geocoded Patient Addresses to the American Community Survey in MSDW

Geocoded patient addresses are now stored in the Mount Sinai Data Warehouse

- Used the Decentralized Geomarker Assessment for Multi-Site Studies (DeGAUSS) application
- Patient home address recorded in Epic converted to latitude and longitude points
- Processes conducted on current and historic patient home addresses

The 2022 American Community Survey (ACS) available in MSDW

- Yearly survey conducted by the United States Census Bureau
- ACS results link demographic, social, economic and housing data to geographic points
- Mechanism to link Social Determinants of Health (SDoH) data to geocoded patient addresses

Linked the Geocoded Addresses with the ACS

American Community Survey (ACS)

Yearly survey of 3.5 million households across the US conducted by the United States Census Bureau

Current and historic geocoded patient addresses in MSDW linked to the ACS

SDoH and Health Equity Domain	ACS Content
DEMOGRAPHIC	**Race/Ethnicity** **Citizenship** – Place of birth, ancestry, year of entry **Language** – Spoken at home, English proficiency, Linguistic isolation **Household Makeup** – Single-parent family, Multifamily household
SOCIAL	**Disability** – Type (cognitive, vision, hearing, other physical, etc.) **Educational Attainment** **Health Insurance**
ECONOMIC	**Income/Poverty Status** **Employment** – Status, Labor force participation **Public Programs** – SNAP, Income support **Transportation** – Commute to work, Vehicles available
HOUSING	**Type and Occupancy** – Owner/renter, Type, Time at address **Housing Costs and Conditions** – Monthly rent, Annual heating costs **Technology** – Internet connectivity Computers

Non-Epic Clinical Data Sources Now Available for Custom Data Requests

PowerPath metadata including the results in the pathology reports

CNExT Cancer Registry

ASOB metadata from obstetric ultrasounds

MOSAIQ – Radiation oncology electronic health record

DENTRIX – electronic health record used at Mount Sinai dental clinics

MSDW Roadmap April-October 2024

MSDW Projects in Progress

	Project	Target Date	New Capabilities for Researchers
1.	**Cancer Registry** • Self-service querying of select data elements from Cancer Registry	2024-Q2	Ability to query Leaf and TriNetX for cancer stage and tumor histology
2.	**Image Research Warehouse (IRW 2.0)** • Identify key radiology data elements	2024-Q3	Enable self-service cohort identification combining key imaging study metadata and EHR data
3.	**Digital Pathology** • Pathologic diagnosis	2024-Q4	Enable self-service cohort identification combining pathology metadata and EHR data Enable access to 10 million de-identified pathology images on Minerva
4.	**Expand Access to new Data Sources** • Structured genomic results from external vendors • MOSAIQ radiation oncology	2024-Q4	Expand the scope of data elements available to include in custom data sets

Obtaining Genomic Data Directly from the Vendors

Project objectives

- To make genomic data and genetic results accessible to researchers in a single central location
- To link the phenotypic and genomic data on Minerva for researchers

Partnering with MSIP on this initiative

- MSIP ensuring contracts address Mount Sinai's best interest for use of the genomic data

Genomic results to be stored on Minerva

- Results include both structured and raw genomic data
- File formats received include BAM, VCF, PDF, XML, JSON and CSV
  - File types available vary by vendor

Foundation Medicine

Foundation Medicine transferred the results of 7,404 ordered tests to Mount Sinai

- Transferred on March 27, 2024

Result files types transferred include:

- XML
- PDF
- VCF (some orders placed before February 2020 do not have VCF results)
- JSON (orders placed after January 2021)

Backload of BAM files to be delivered by May 22, 2024

Somatic Genomic Testing Vendors

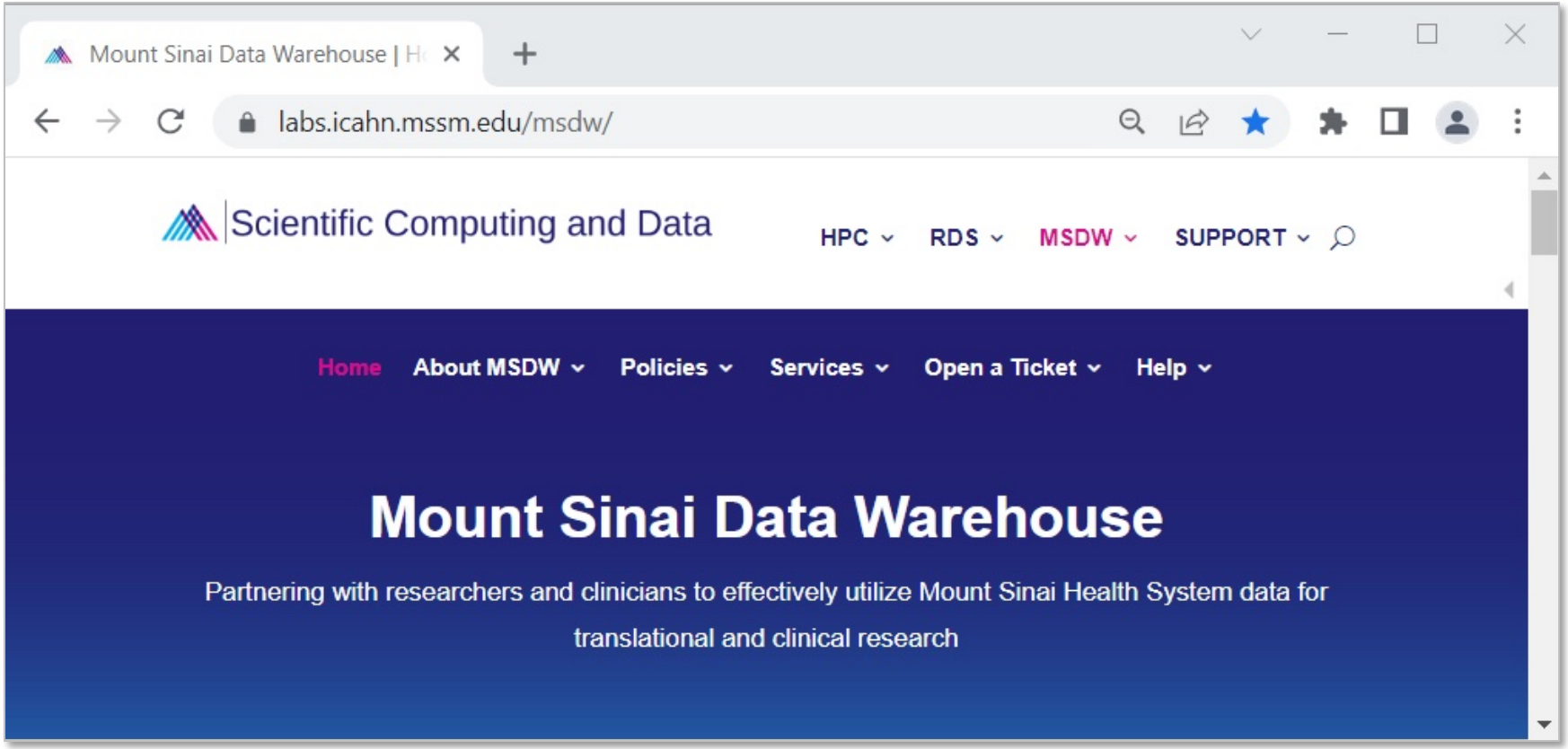
Vendor	Status
Foundation Medicine	Data transfer in progress
Caris	Contract signed
NeoGenomics	MSIP reviewing existing agreement
Tempus	Identifying contact at company
Guardant	Identifying contact at company
Natera	Identifying contact at company
Sema4 (historic testing)	MSIP discussing with Mount Sinai legal

Researcher Engagement: Upcoming Events Spring 2024

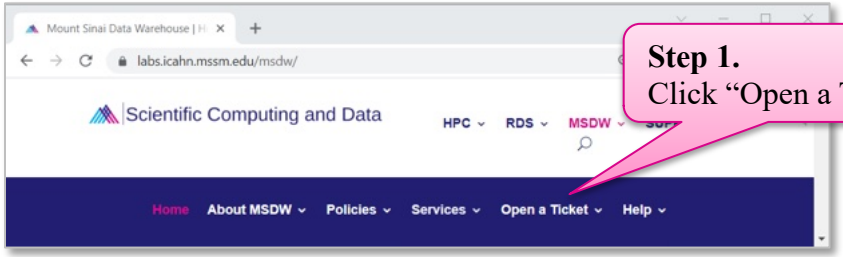
Date	Event	Topics
Tuesday, April 15 9AM - 5PM	**Epic Research Training for Clinical Research Coordinators**	• Requesting a new research study to be in Epic • Enrolling patients to a research study • New patient registration
Tuesday, May 14 1-2PM	**Leveraging Epic for Research**	• Using MyChart for Trial recruitment • Identifying patients in research studies in Epic • Patient opt-out of clinical trial recruitment via MyChart
Wednesday, May 1 3 pm–4 pm	**Mount Sinai Data Warehouse (MSDW) Town Hall**	• 2023 MSDW Usage • 2023 MSDW Accomplishments and Updates • MSDW Initiatives and Roadmap
Monday, June 3	**CTSA Translational Science Research Day**	• Representing Leaf, TriNetX and MSDW at Speed*Connections* round
June 2024	**Leaf/ATLAS Training**	• Overview of the Mount Sinai Data Warehouse • Introduction to Leaf and ATLAS query tools
June 2024	**TriNetX Training**	• Overview of the Mount Sinai Data Warehouse • Introduction to the TriNetX query tool

How to Contact Us

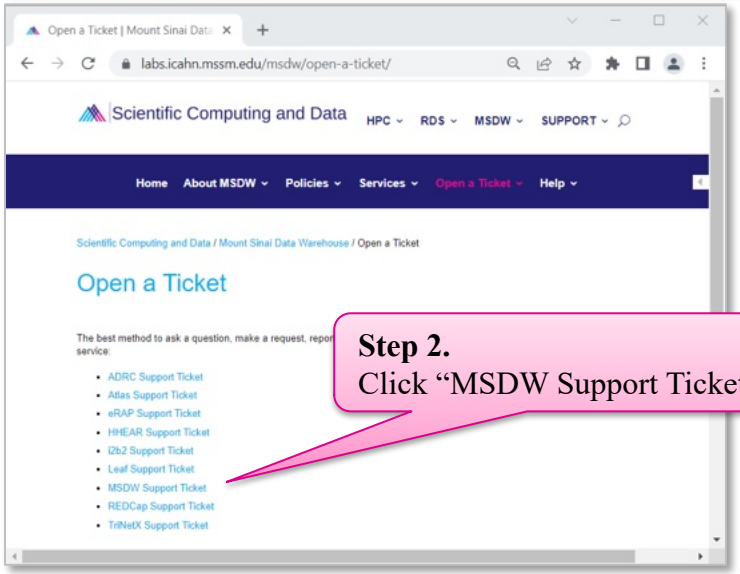
Our Website: msdw.mountsinai.org



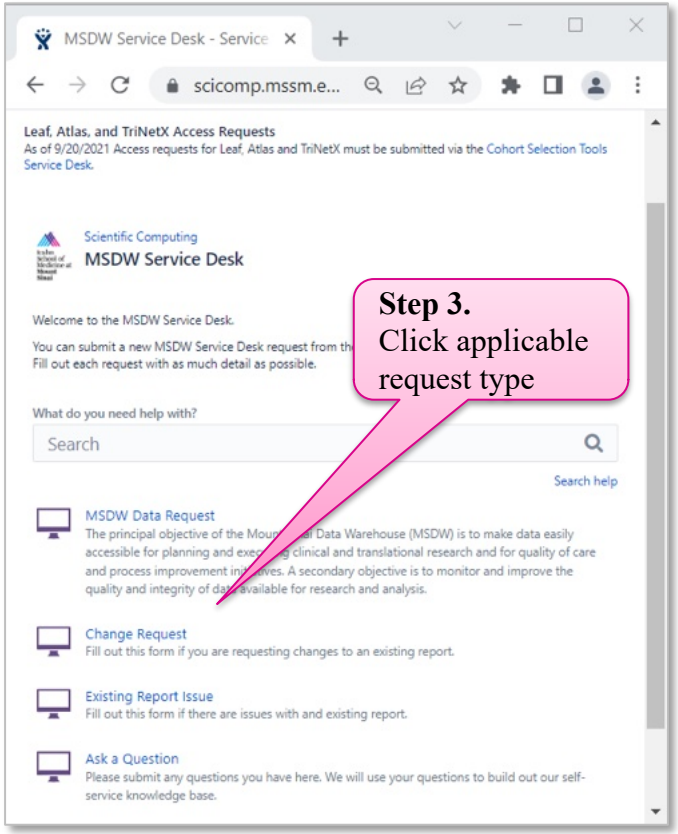
How to Open an MSDW Request Ticket



Step 1.
Click "Open a Ticket"



Step 2.
Click "MSDW Support Ticket"

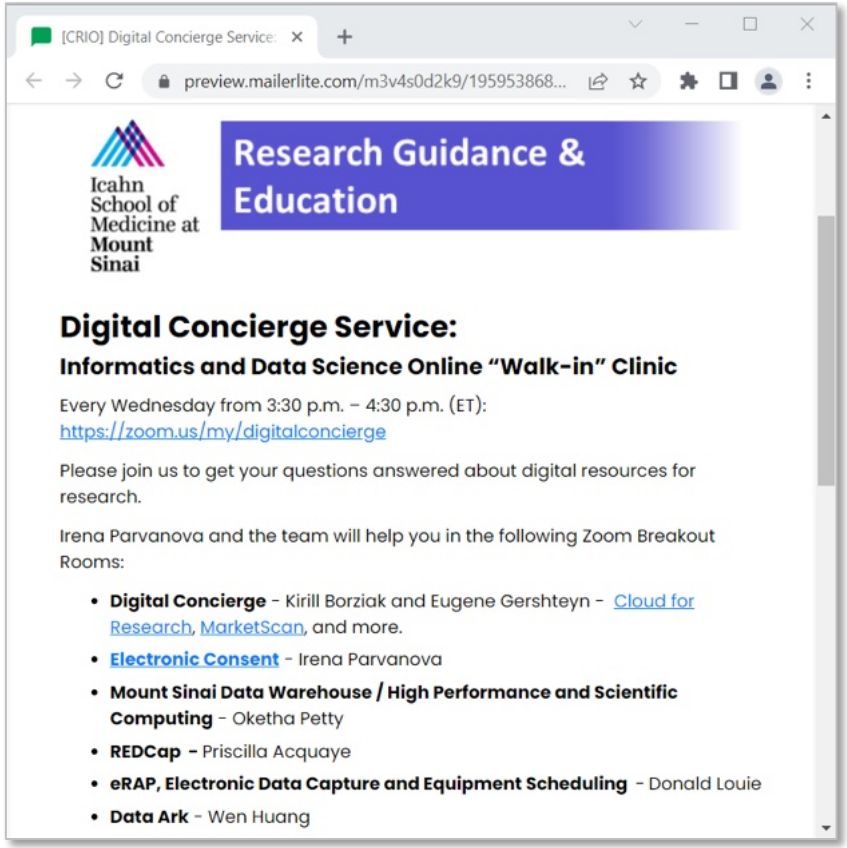


Step 3.
Click applicable request type

Digital Concierge – Walk in Clinic

"Walk-in" **Digital Concierge** service hosted by the MSDW team **every Wednesday from 3:30 PM to 4:30 PM**

Reminder notifications sent by the ORS team every week



Thank You!
For help and more information:
msdw.mountsinai.org