Mount Sinai Data Warehouse Town Hall

Patricia Kovatch
Farhan Mahmood
Sharon Nirenberg, MD
Timothy Quinn, PhD
Scientific Computing
Icahn School of Medicine at Mount Sinai
May 2nd, 2023

Icahn
School of
Medicine at
Mount
Sinai

Agenda

1. MSDW Team Updates
2. MSDW Operations Since Q4 2022
3. MSDW 2022 User Survey Results
4. New Services: Interactive Dashboards & Geocoding
5. Epic Research Module
6. Improving OMOP Data Content
7. MSDW Roadmap

MSDW Team Updates

The MSDW Team

Patricia Kovatch
Professor & Dean for Scientific Computing

Sharon Nirenberg MD
Physician Informaticist

Timothy Quinn PhD
Principal Data Architect

Farhan Mahmood
Director Scientific Computing

Naomi So MD
Physician Informaticist

Teja Ganta MD
Physician Informaticist

Shivaji Punukollu
Healthcare Data Engineer

Manoj Chekuri
Healthcare Data Engineer

Rupan Hossain
Database Administrator

Eric Rosenberg
Sr. System Administrator

AJ Caberto
Clinical Data Specialist

Jiani Xiang
Clinical Data Specialist

Darius Boopal
Healthcare Data Engineer

Jacob Weiser
Healthcare Data Engineer

Priyal Mehta
Healthcare Data Analyst

Vidhya Venkatesan
Healthcare Data Analyst

New Scientific Computing Team Members

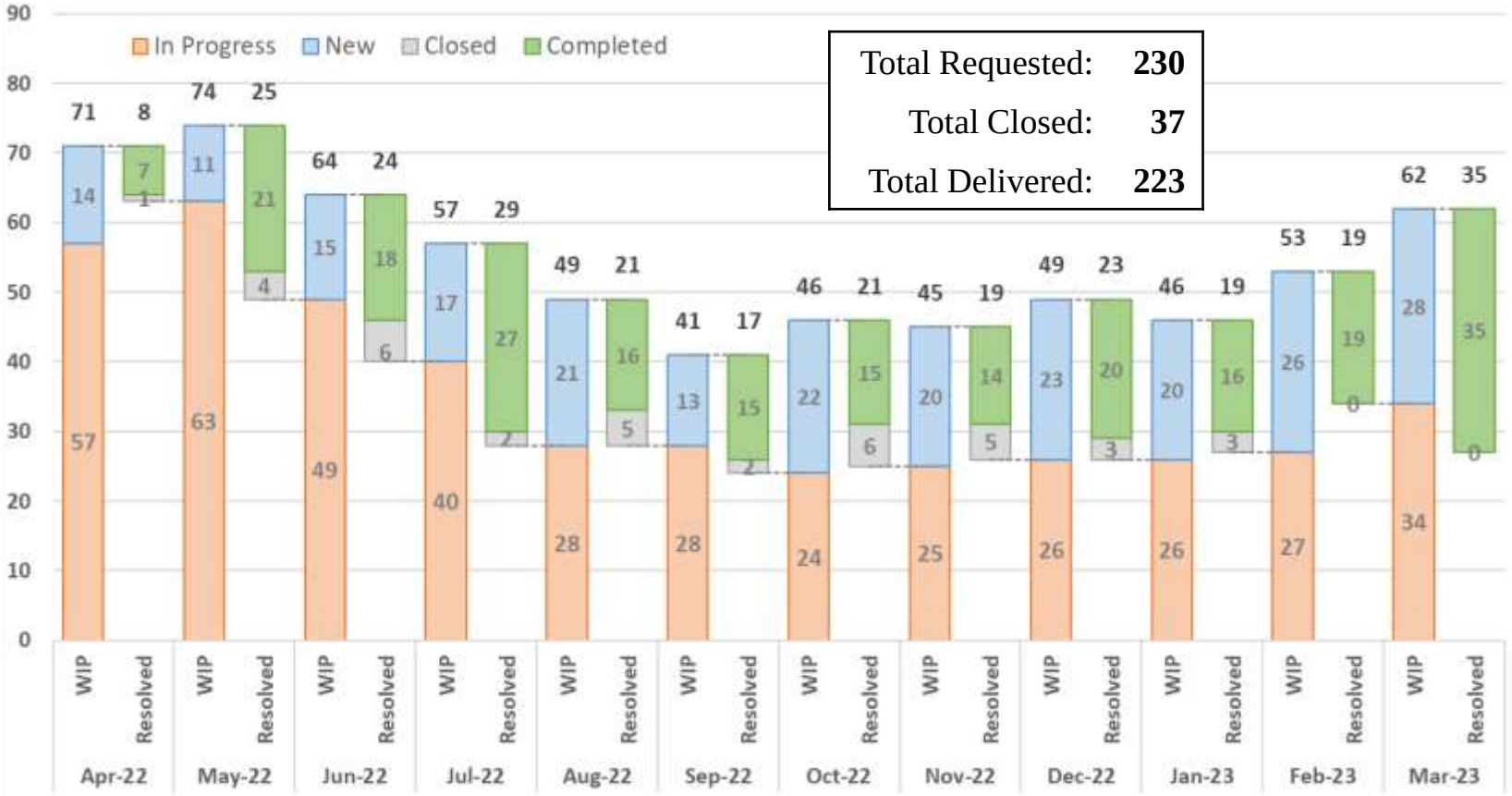
- **Naomi So MD**
  - Physician Informaticist
  - Joins us from Cedars-Sinai Medical Center, Los Angeles

- **Teja Ganta MD** (30% from Aug 2023)
  - Physician Informaticist
  - Joint appointments with Division of Hematology & Medical Oncology, the Tisch Cancer Institute, and the Office of the CMIO
  - Joins us from ISMMS's fellowship programs in Hematology & Medical Oncology and Clinical Informatics

- **Rajendra Bose PhD**
  - Director, Researcher Engagement
  - Joins us from the Chan Zuckerberg Biohub (UCSF, UC Berkeley, Stanford)

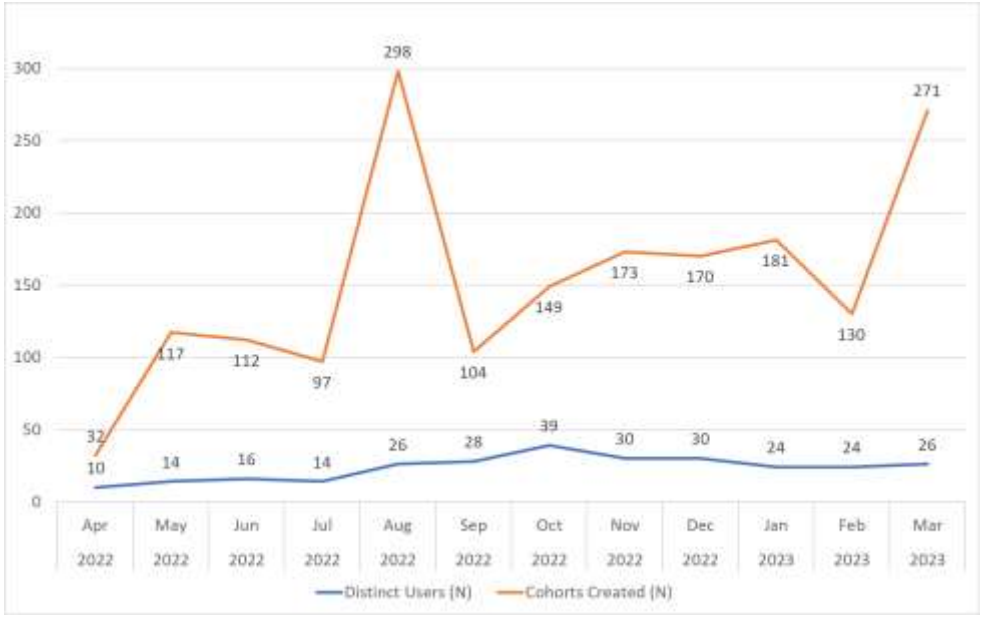
MSDW Operations Since Q4 2022

Custom Data Sets Requested & Delivered (Last 12 Months)

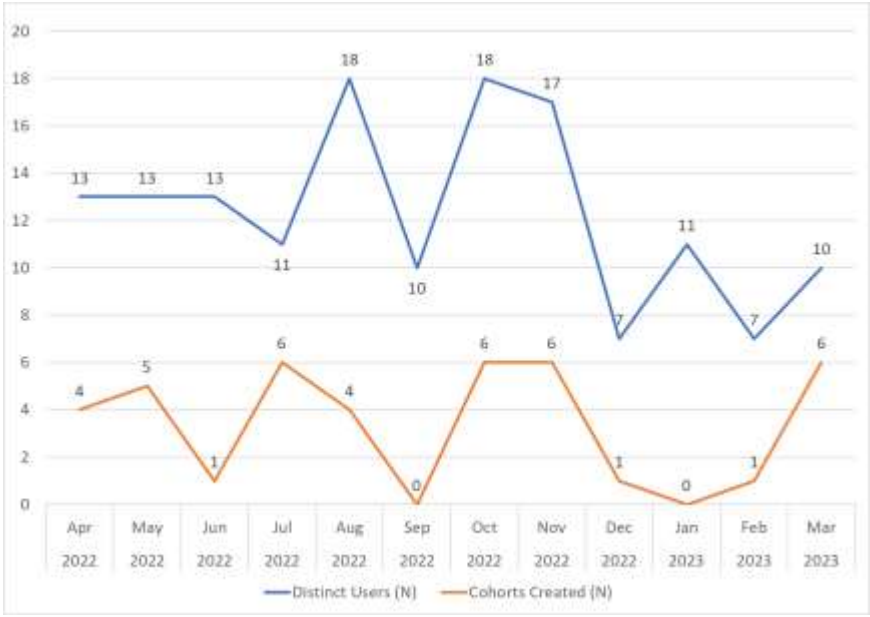


Utilization of Leaf and ATLAS (Apr 2022 to Mar 2023)

Leaf Utilization

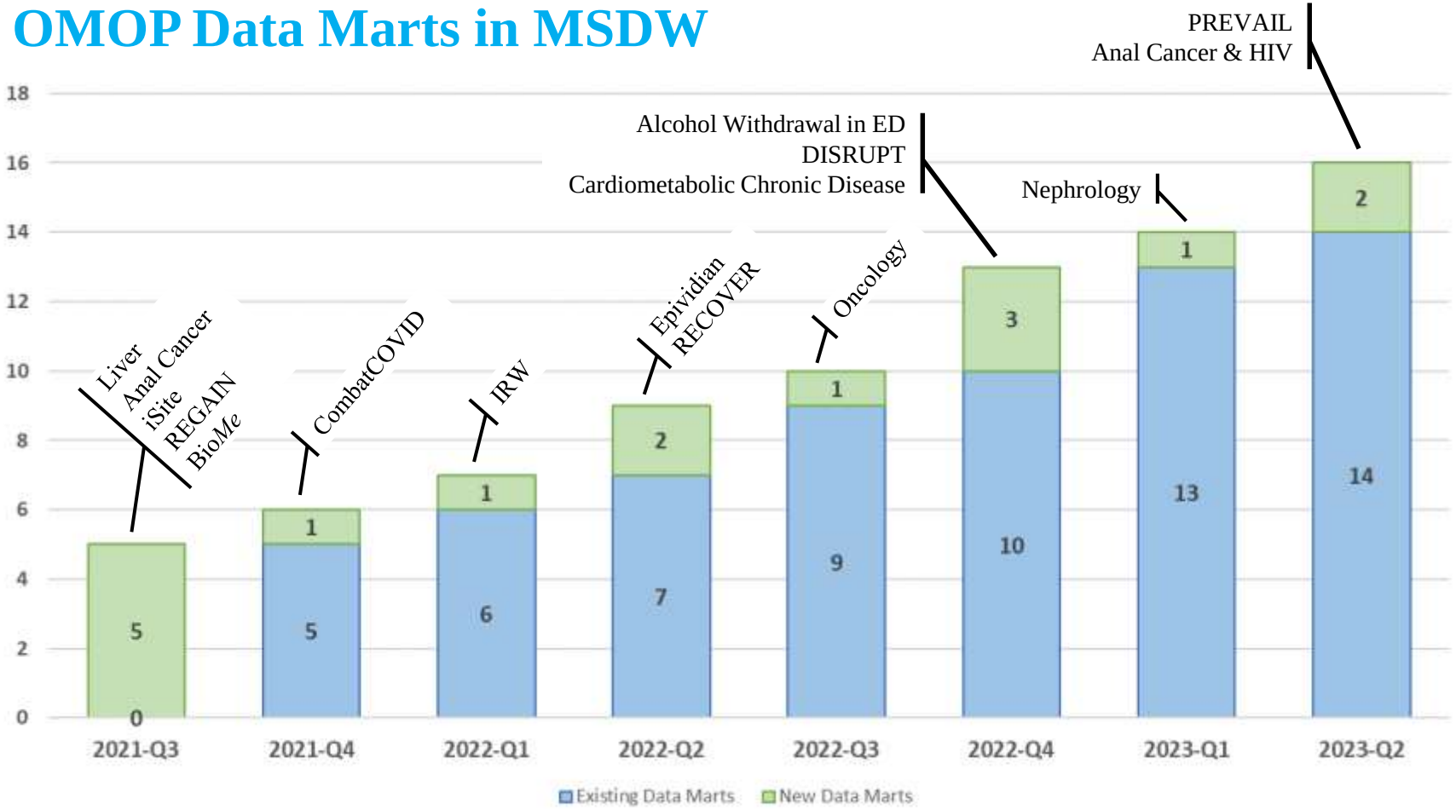


ATLAS Utilization

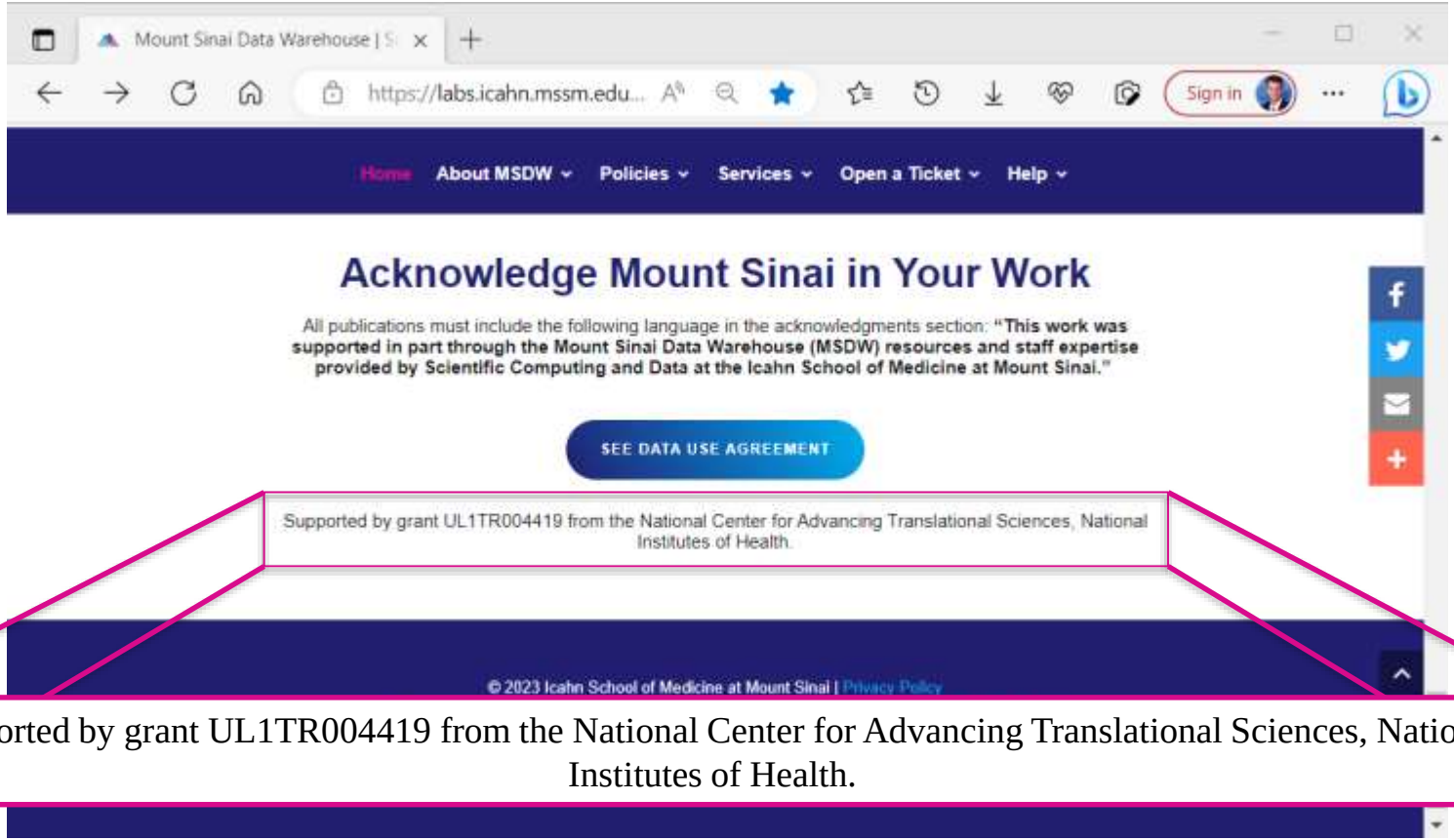


Leaf utilization has remained steady at about 30 users per month over the year

OMOP Data Marts in MSDW

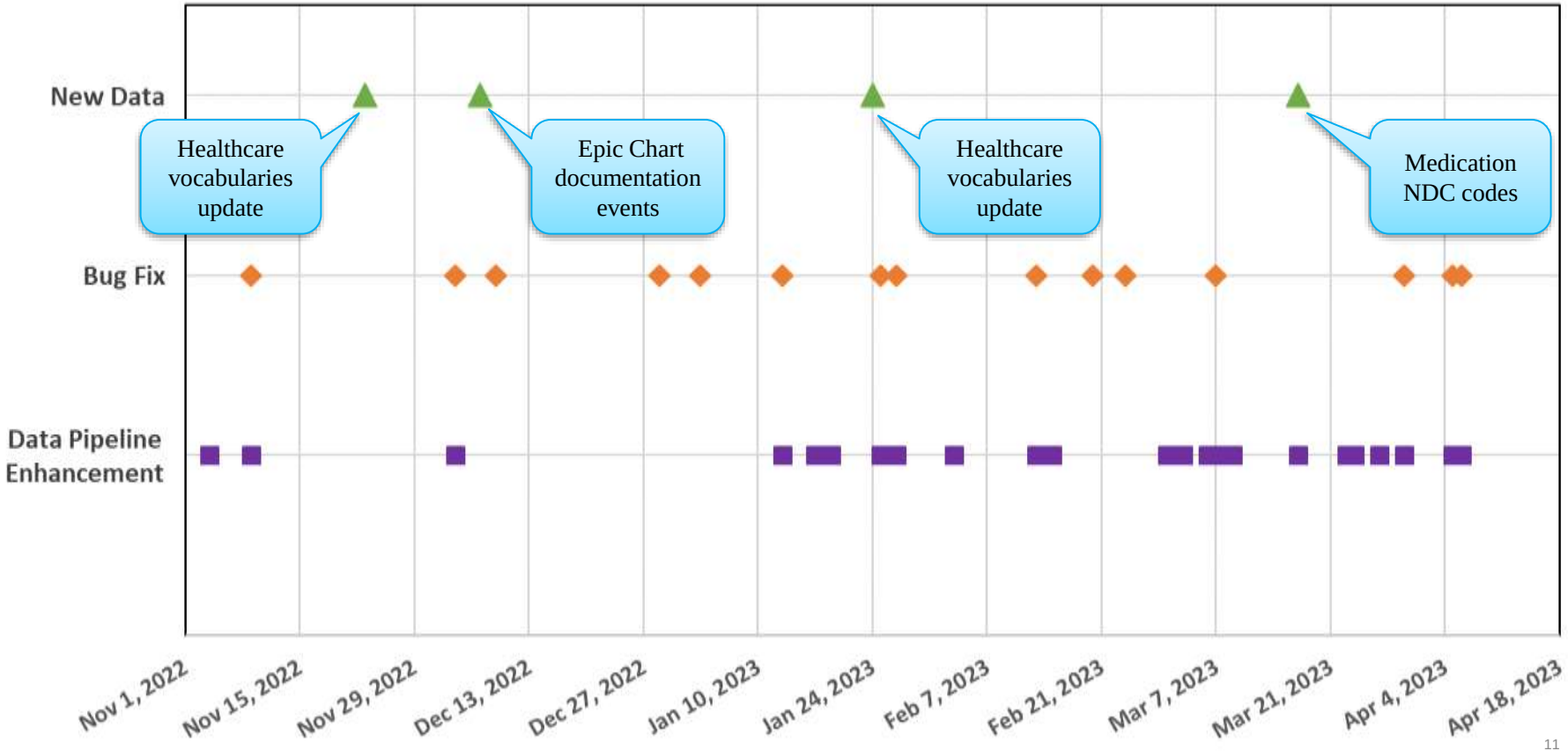


Encourage MSDW Users to Acknowledge CTSA



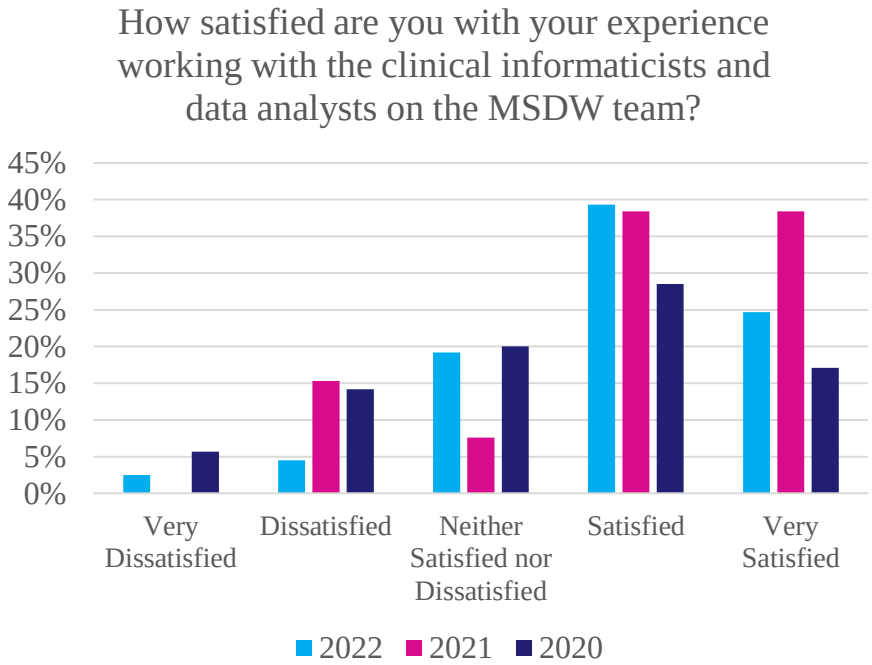
Supported by grant UL1TR004419 from the National Center for Advancing Translational Sciences, National Institutes of Health.

32 Releases Since November 1, 2022 (~1.4 per week)

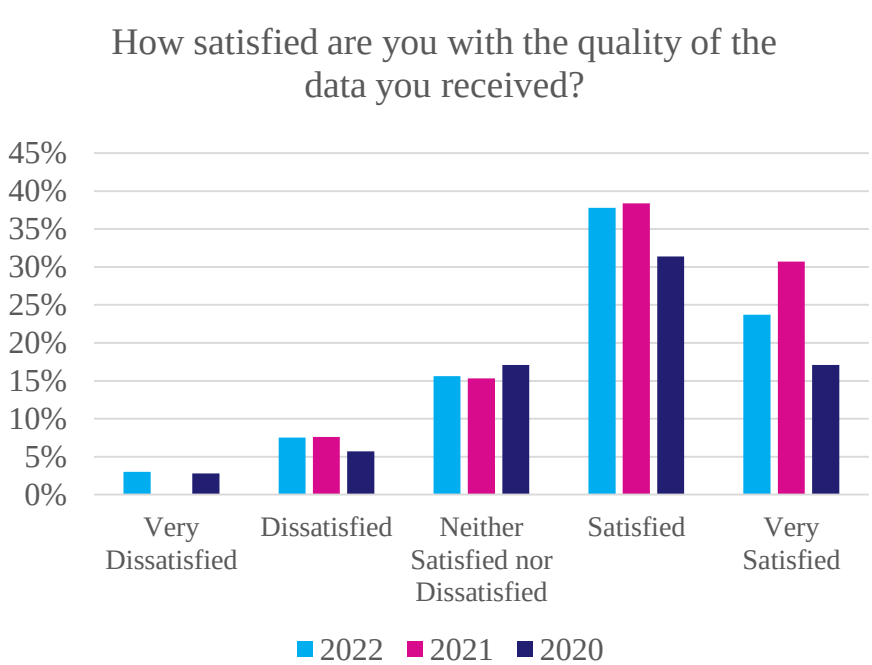


MSDW 2022 User Survey Results

MSDW 2022 User Satisfaction Survey Results

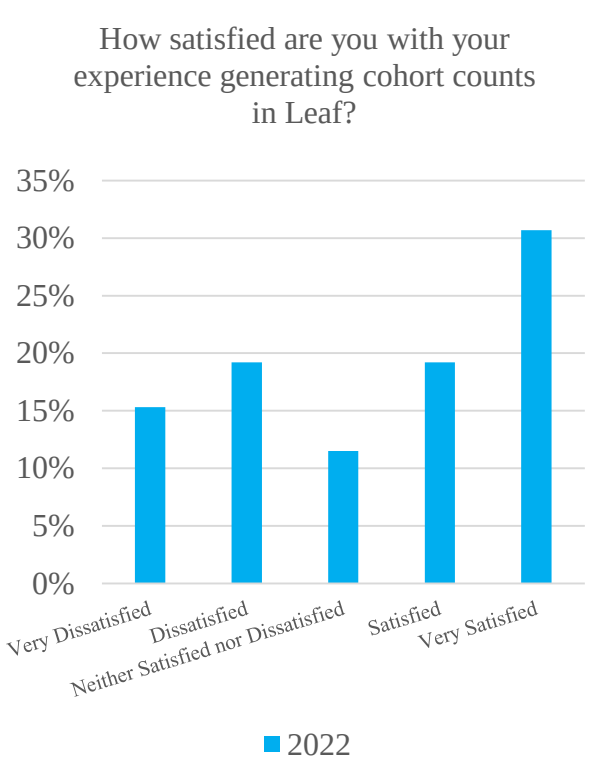


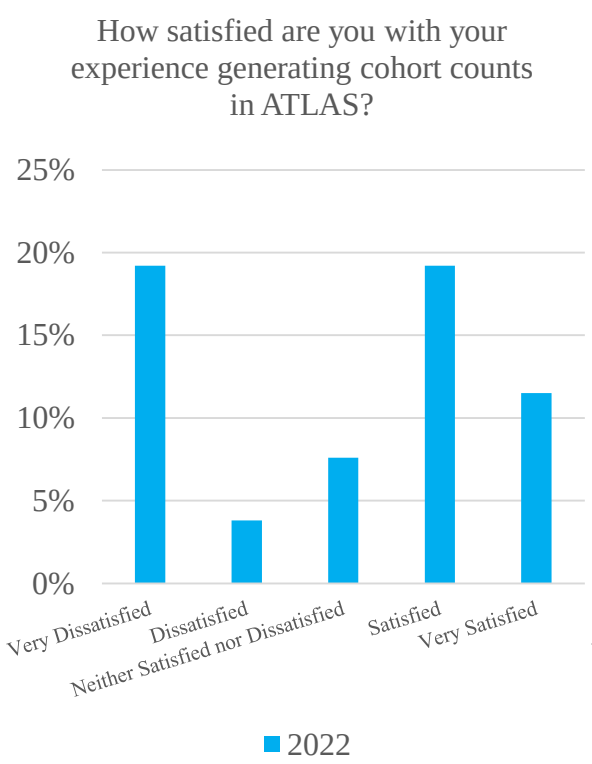
In 2020, n=35, 45% of respondents reported satisfied or very satisfied
In 2021, n=13, 76% of respondents reported satisfied or very satisfied
In 2022, n=198; 64% of respondents reported satisfied or very satisfied

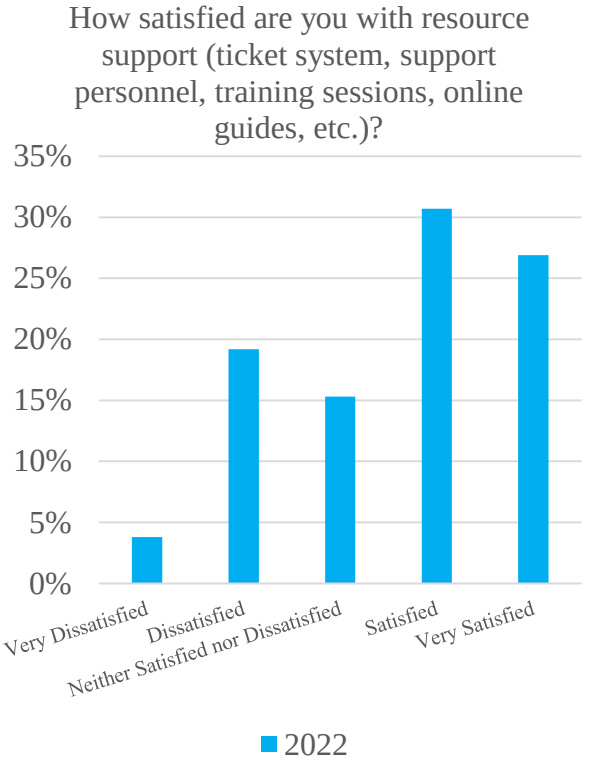


In 2020, n=35, 48% of respondents reported satisfied or very satisfied
In 2021, n=13, 69% of respondents reported satisfied or very satisfied
In 2022, n=198; 61% of respondents reported satisfied or very satisfied

Leaf & ATLAS 2022 User Satisfaction Survey Results







In 2022, n=26; 13% of MSDW users reported to using either Leaf or ATLAS in 2022

User Satisfaction Survey Comments

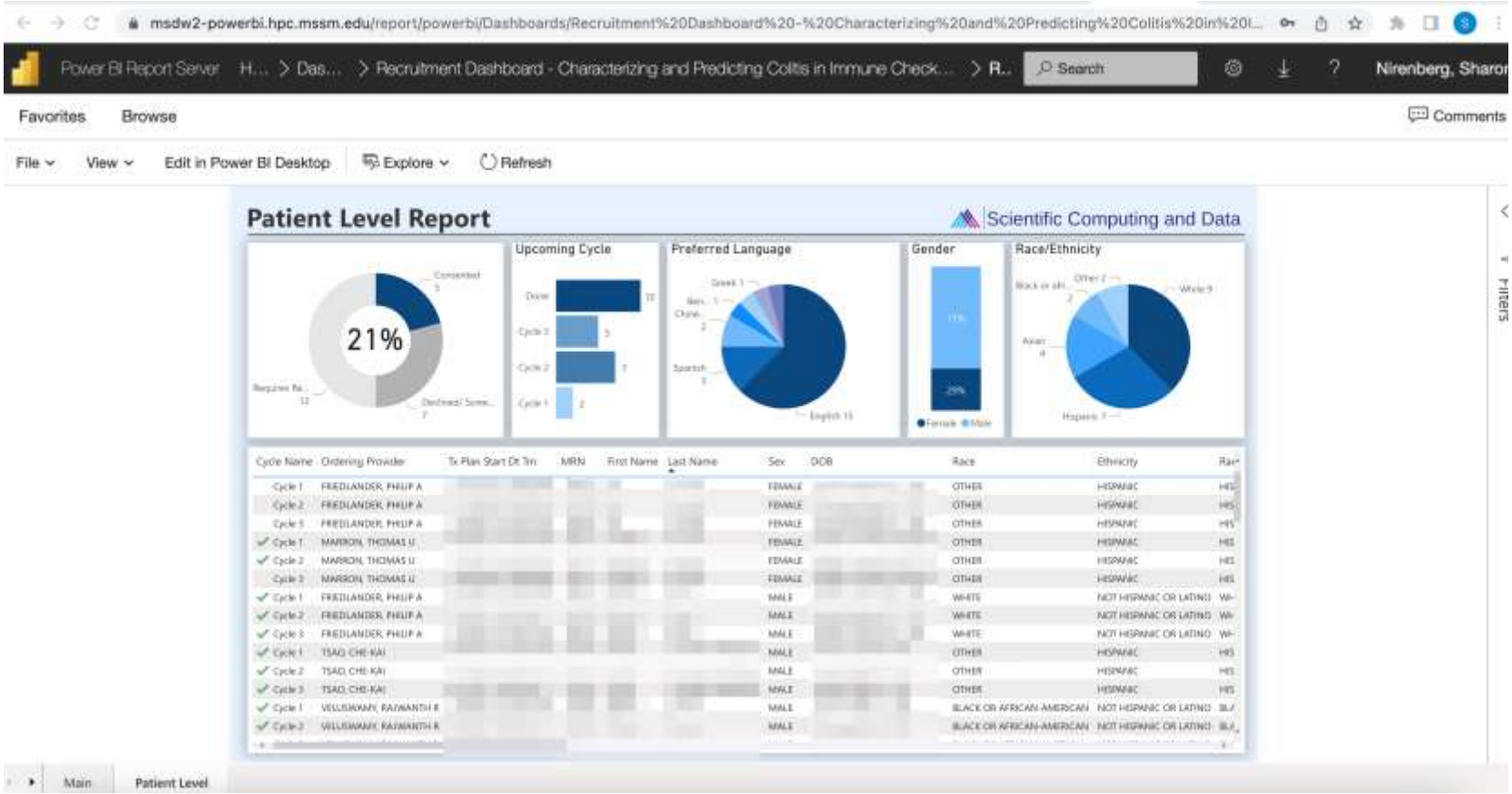
1. **Positive comments**
   - "Extraordinarily professional, competent, and collaborative"
   - "I am having such a great experience working with MSDW team. Very informative environment."
   - "I enjoy working with this team. They are hard-working, professional and very responsive"
   - "First, let me say that I'm extremely thankful for the MSDW, and that it exists as a service to researchers. …[W]hen the MSDW [team] and I did come to a shared understanding of what I needed, the data were always well-labeled and easy to understand."
2. **Opportunity for improvement: Turnaround time**
   - We are expanding our capacity by hiring another Informaticist, Naomi So MD
3. **Opportunity for improvement: Payment processing delays data release**
   - With new projects we will underscore attention to finance/payment procedure and communications

New Services:

Interactive Dashboards & Geocoding

Trial Recruitment Dashboard

Trial Recruitment Dashboard – Patient-Level Details



Role of MSDW in Clinical Trial Recruitment

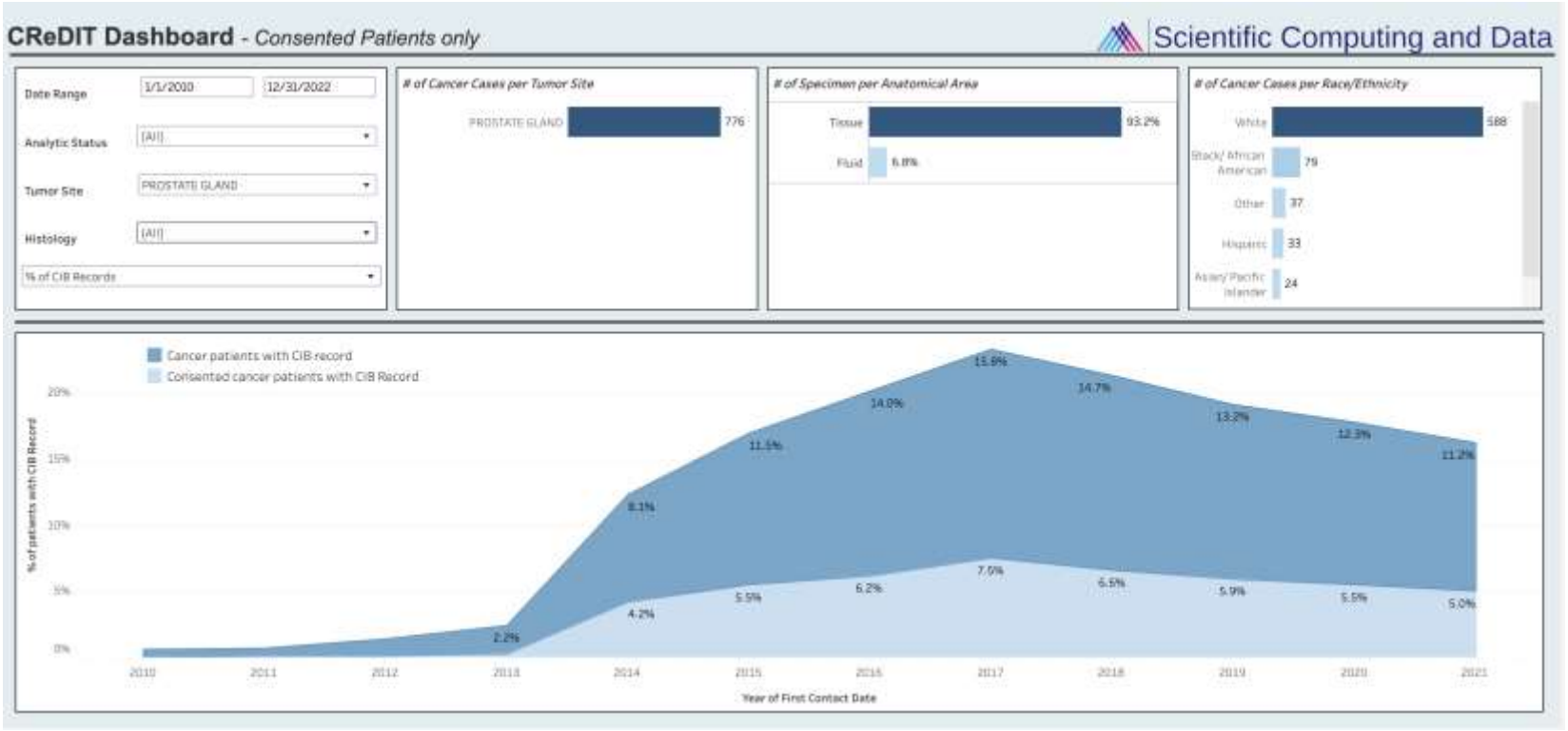
Using electronic health record data to provide focused lists of eligible patients to clinical research coordinators

1. Define patient cohort **inclusion/exclusion criteria** for studies
   - Operationalized with Epic EHR data
   - Obtain as narrow a list as possible

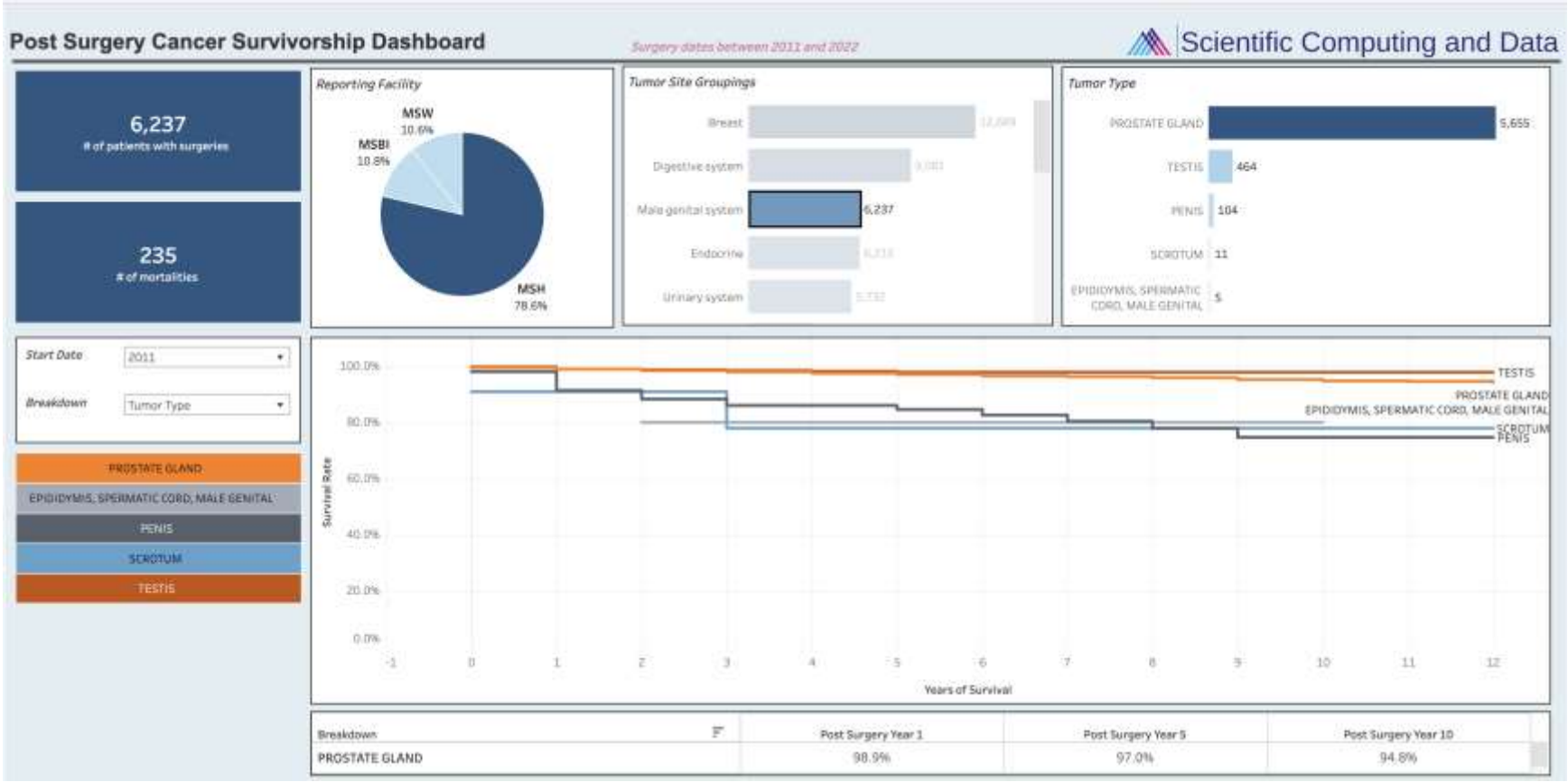
2. Identify new patients **potentially eligible** to enroll in each clinical trial
   - Provide automated, regular reports
   - Facilitate easy chart review by clinical research coordinators

3. CRCs review charts of short-listed patients to **ensure enrollment eligibility**

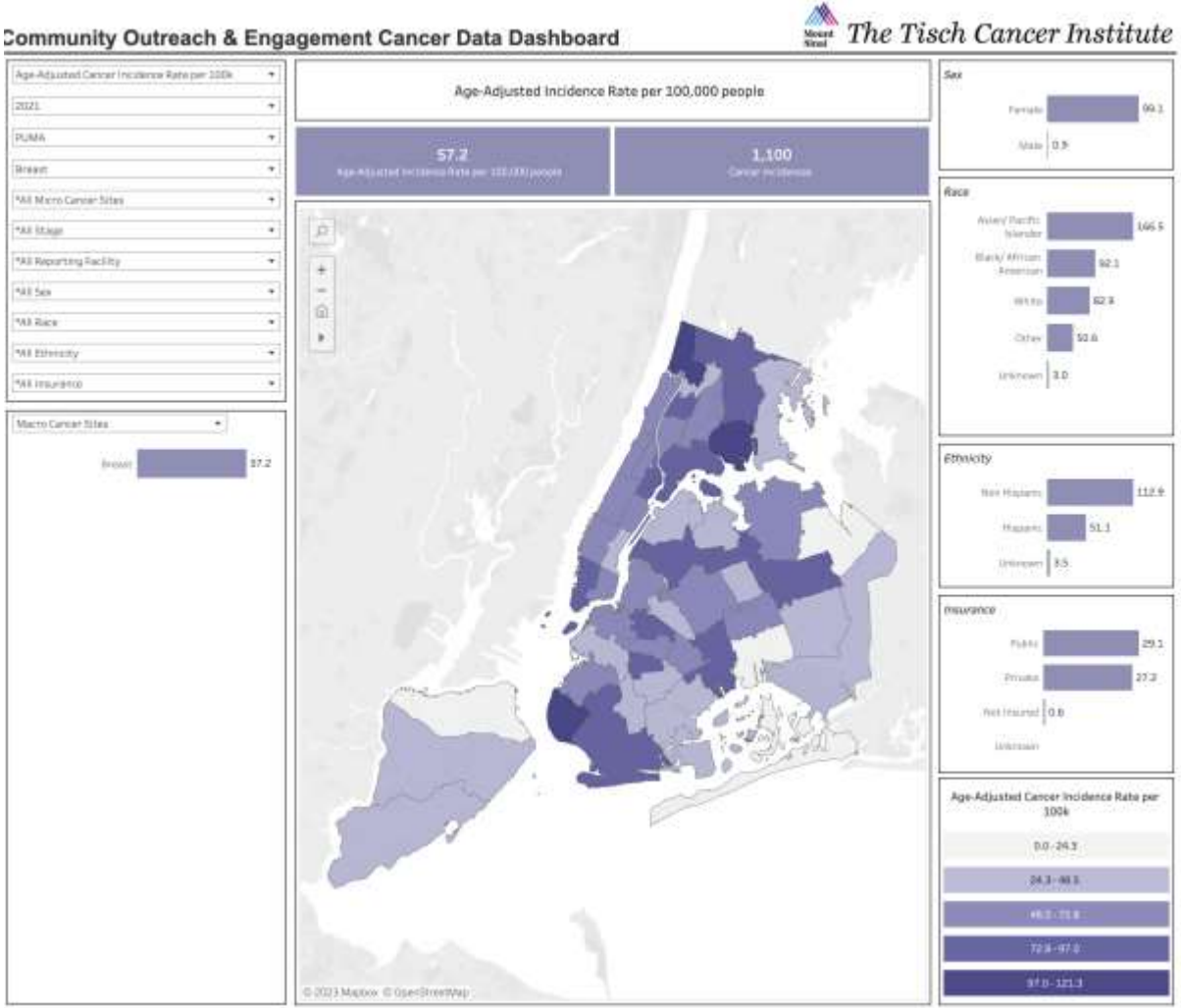
CReDIT Dashboard



Post Surgery Cancer Survivor Dashboard



Example of a Dashboard Using Geocoding



Epic Research Module

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

- 1.4M 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

CTSA-Supported Initiative: Epic Research Module

2023 initiative to raise awareness of Epic's research-related capabilities via ORS and increased training

Epic Research Functionality	Currently Available at Mount Sinai
Automatically process research-specific claims	✓
Designate an encounter as a research-specific encounter	✓
Notify research staff of research patient upcoming appointment, ED or hospital admission, demise, etc.	✓
Enable research billing review	✓
Designate an order as research related	✓
Link research studies with study protocols	✓
Enable best practice advisories (BPAs) to identify potential research participants	✓
Send research project recruitment messages to patients via MyMountSinai	
Display current research studies in MyMountSinai	
Patient opt in/out of research recruitment via MyMountSinai	
Consent study participants via MyMountSinai	
Document and manage adverse events in Epic	

Improving OMOP Data Content

National COVID Cohort Collaborative (N3C) OMOP Data Quality Scorecard

- NIH's National Center for Advancing Translational Sciences (NCATS)
- New data quality scorecards launched in March 2023
- Mount Sinai benchmarked against 30 sites contributing in OMOP format

	Where We Did Well
1.	No implausible dates or lengths of stay
2.	High concept mapping rates for 5/6 domains
3.	More SDoH variables than 80% of sites

	Opportunities for Improvement
1.	Fill rates for race and ethnicity
2.	Map COVID-19 tests to more specific LOINC codes
3.	Map more procedure codes
4.	Map more lab test units of measure

Terminology Mapping Services Vendor

Clinical Architecture®

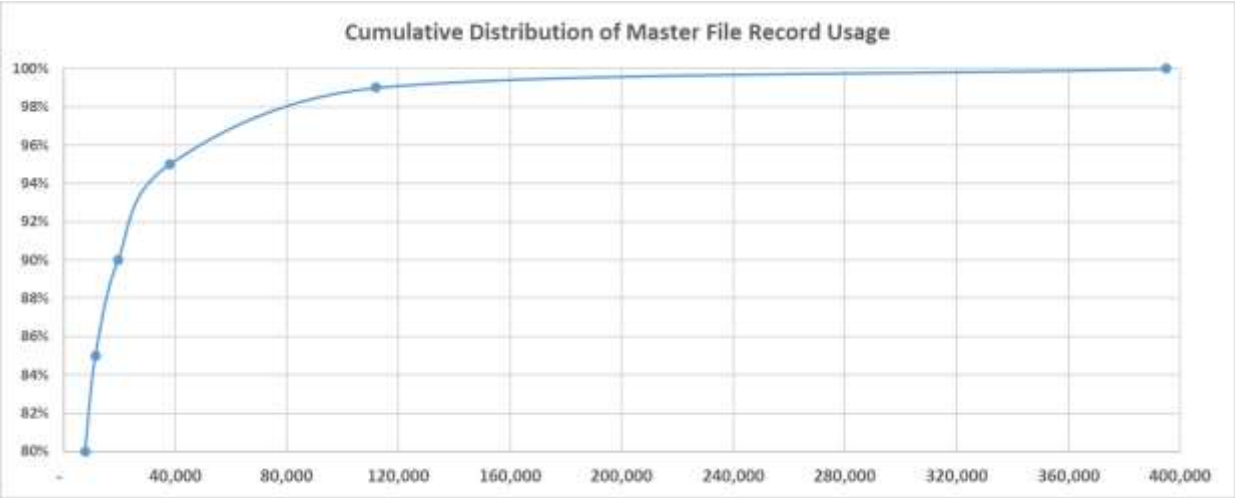
11611 North Meridian St
Suite 500
Carmel, IN 46032

Leading provider of healthcare terminology mapping services

Symedical mapping software

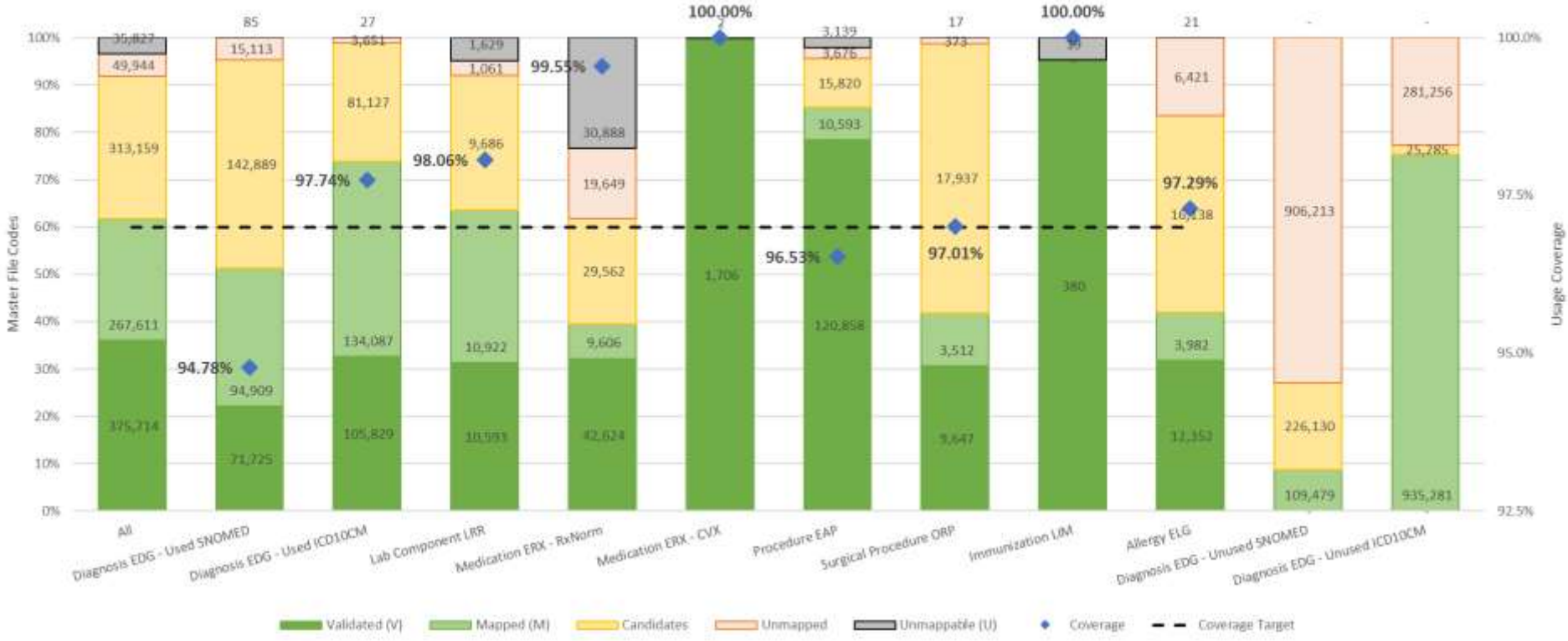
- Multiple matching algorithms
- Matching algorithms tailored to domain
- Each algorithm tuned for domain and customer data
- Underlying data model, dictionaries, thesauri for each domain

Most Epic Codes Used Infrequently (“Long Tail”)

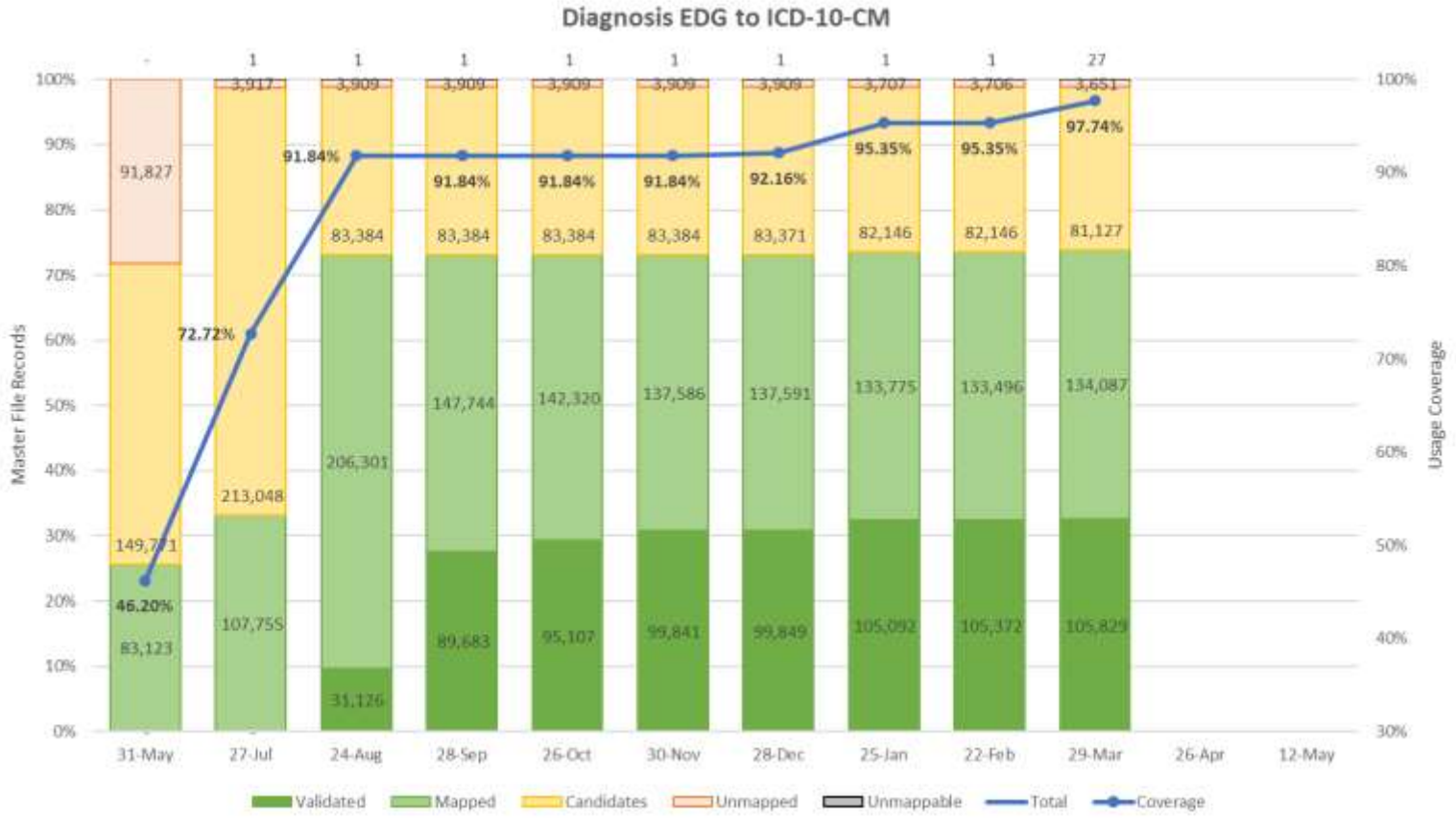


Master File	Target Vocabulary	Rows for 80%	Rows for 85%	Rows for 90%	Rows for 95%	Rows for 99%	Used Rows	Total Rows	Current Mapped
Diagnosis EDG	SNOMED-CT	6,000	9,000	16,000	32,000	96,000	324,721	1,566,543	276,198
Diagnosis EDG	ICD10CM	6,000	9,000	16,000	32,000	96,000	324,721	1,566,543	1,175,224
Lab Component LRR	LOINC / SNOMED	200	300	400	600	1,800	21,866	33,891	23,144
Medication ERX	RxNorm	800	1,000	1,400	2,400	6,200	53,376	134,037	83,118
Procedure EAP	CPT / HCPCS	200	400	600	1,000	3,000	26,159	154,086	134,590
Surgical Procedure ORP	CPT / HCPCS	400	600	800	1,400	2,600	4,712	31,486	13,176
Immunization LIM	CVX / RxNorm	30	40	50	80	130	333	399	399
Allergy ELG	SNOMED-CT / RxNorm	150	200	350	700	2,300	8,249	38,914	16,355
	TOTAL:	**7,780**	**11,540**	**19,600**	**38,180**	**112,030**	**439,416**	**1,959,356**	**1,446,006**
		80%	85%	90%	95%	99%	100%		73.8%

Mapping Status: 7 of 9 Domains Above 97%



Mapping Example: Epic Diagnosis (EDG) to ICD-10-CM



MSDW Roadmap

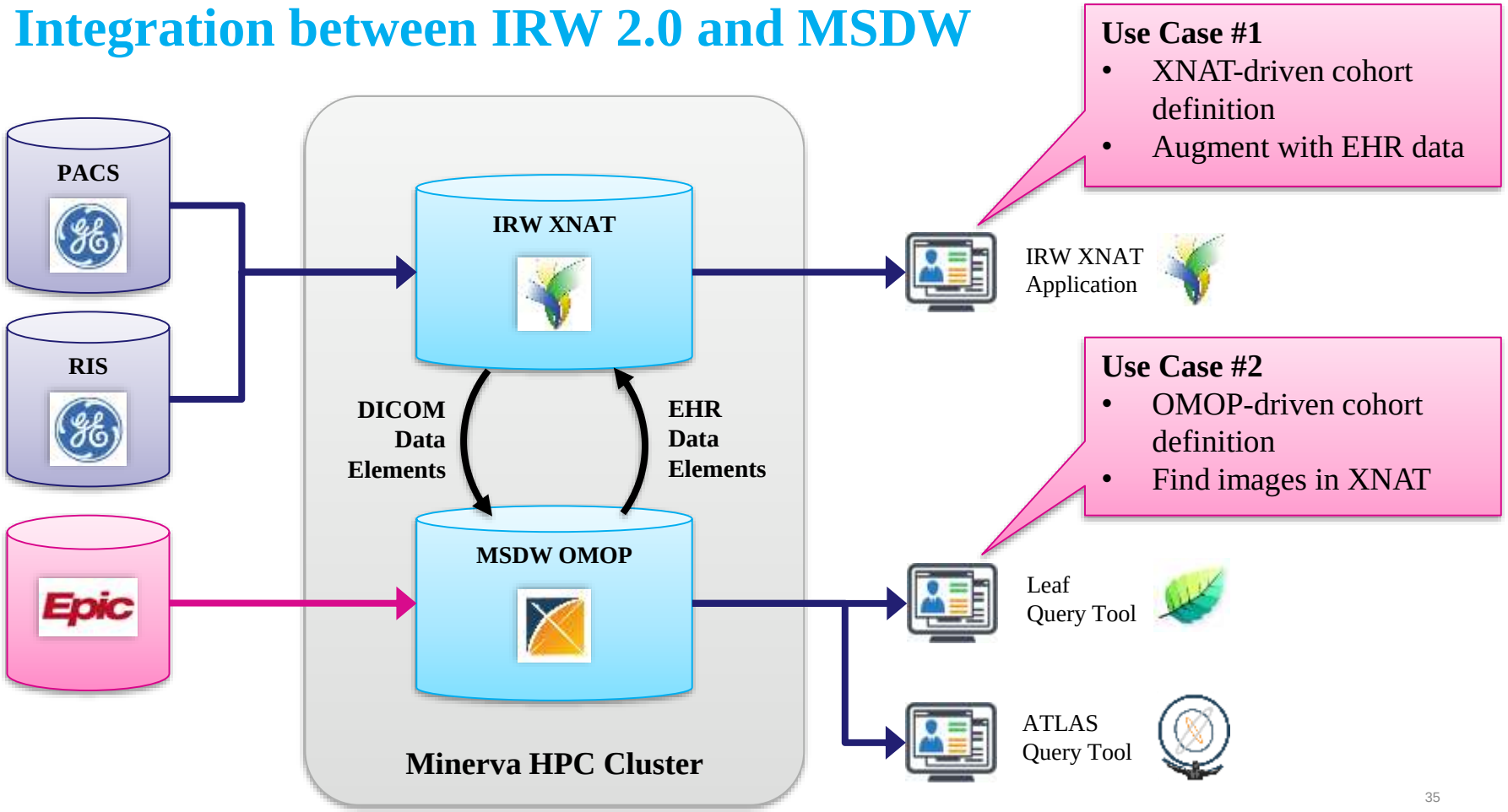
CTSA Data Science Survey (Feb 2023)

Data Collection Method	Participants	Topics
Electronic Survey • distributed via ORS Research email listserv	146	1. Access to expert help 2. New data sources 3. Access to Informatics tools 4. Training/workshops
1:1 Interviews • with Thought Leaders	12	

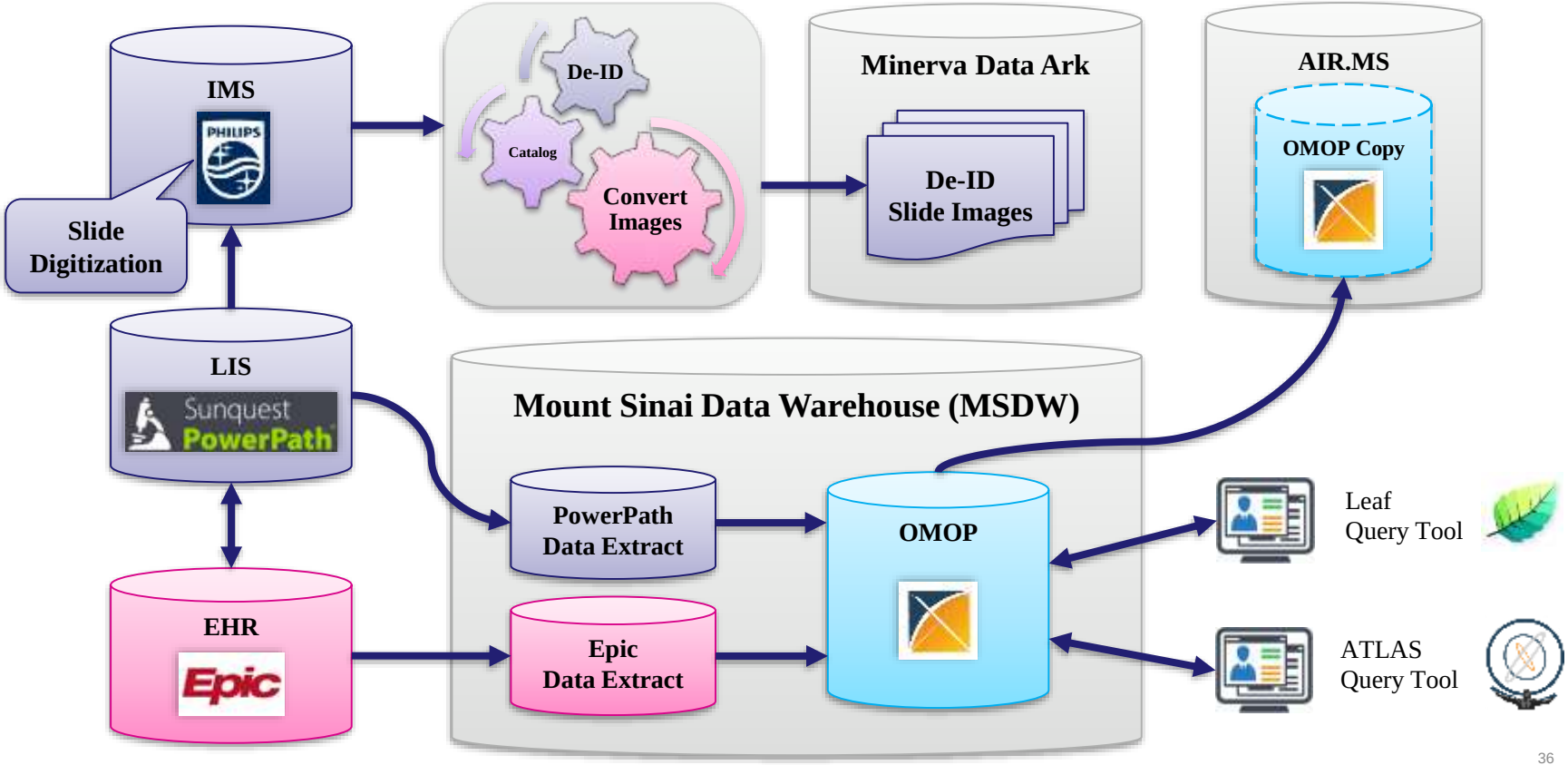
CTSA Data Science Survey: Highest-Priority Needs

	Need	Potential Action Item
1.	SDoH and NLP-extracted terms from unstructured clinical notes	• Extract SDoH from structured/unstructured data, link to clinical data • Create NLP task force to choose best-in-class software to map notes to SNOMED terms
2.	Hospital & professional billing data	• Integrate Mt Sinai's billing data from MSX
3.	Radiology imaging studies data	• Integration with IRW 2.0 and Data Ark
4.	Ongoing access to genomics data	• (*Partially solved*) De-identified BioMe exome/genotyping data now on Data Ark
5.	Long-term and hands-on support for EHR and multi-modal analysis	• Create a taskforce to assess how best to address this need

Integration between IRW 2.0 and MSDW



Digital Pathology Data Integration

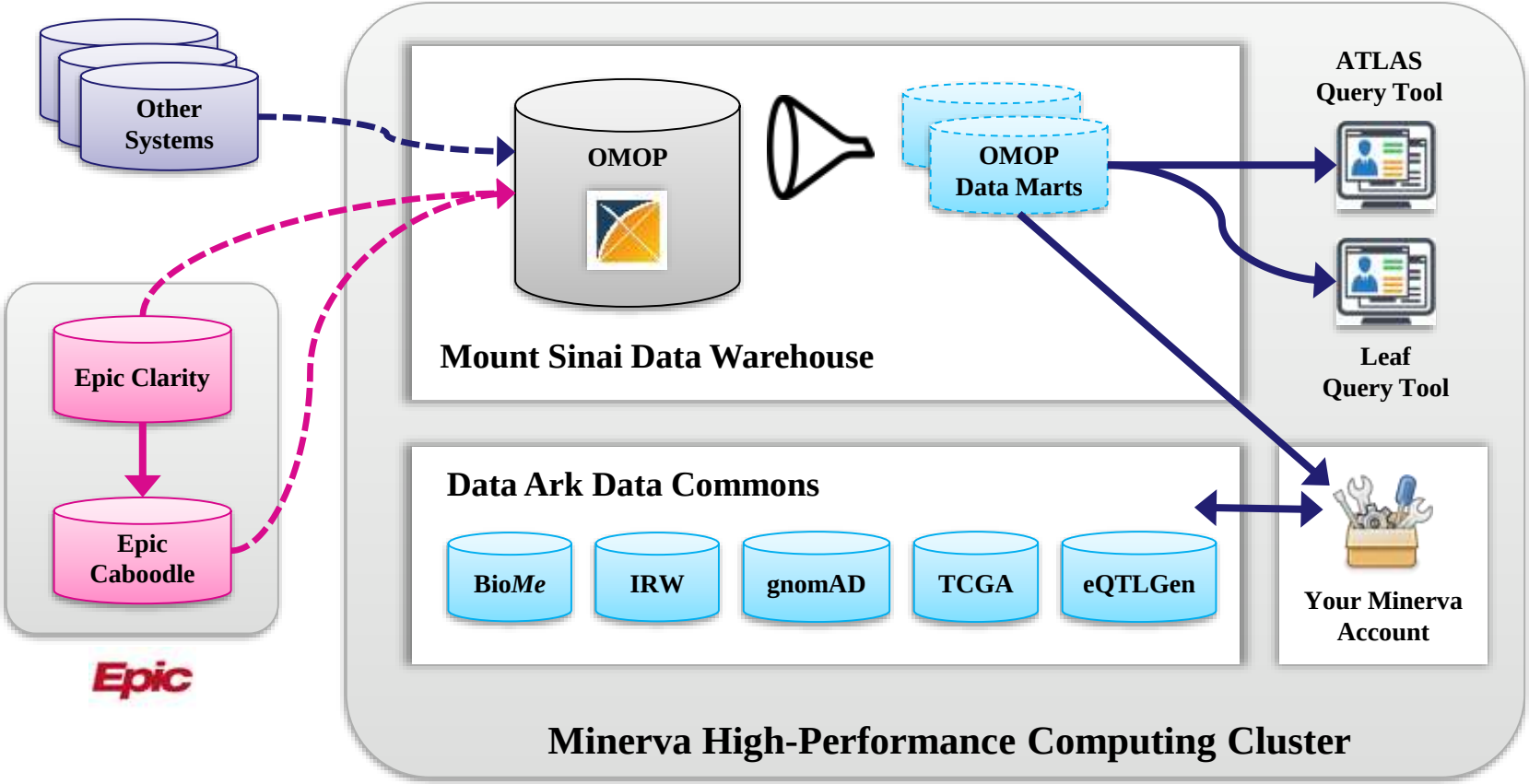


MSDW Development Roadmap for 2023

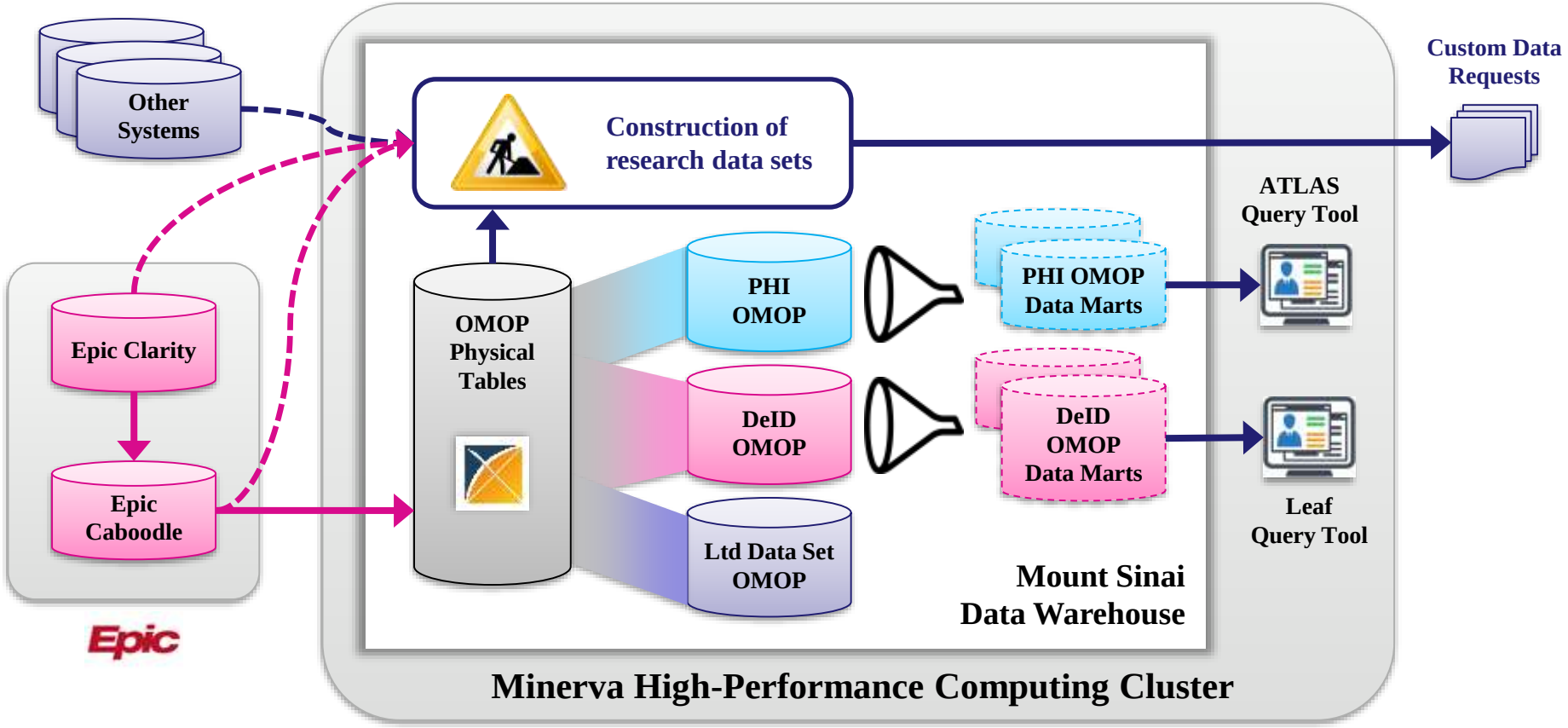
Quarter	Integrate New Data into MSDW	Upgrade MSDW Infrastructure
Q2 2023	• Radiology reports • Pathology specimen and slide data from PowerPath • Patients with imaging studies in IRW 2.0 • Searchable in Leaf • De-identified OMOP and Bio*Me* on Data Ark	• Relauch of TriNetX cohort query tool • Automated data quality checks
Q3 2023	• Links to digitized pathology slide images on Minerva	• Bi-directional interface with IRW 2.0
Q4 2023	• Imaging study data elements from IRW 2.0	• Framework for Epic upgrades • Next upgrade on Sun, Nov 5 2023

Appendix

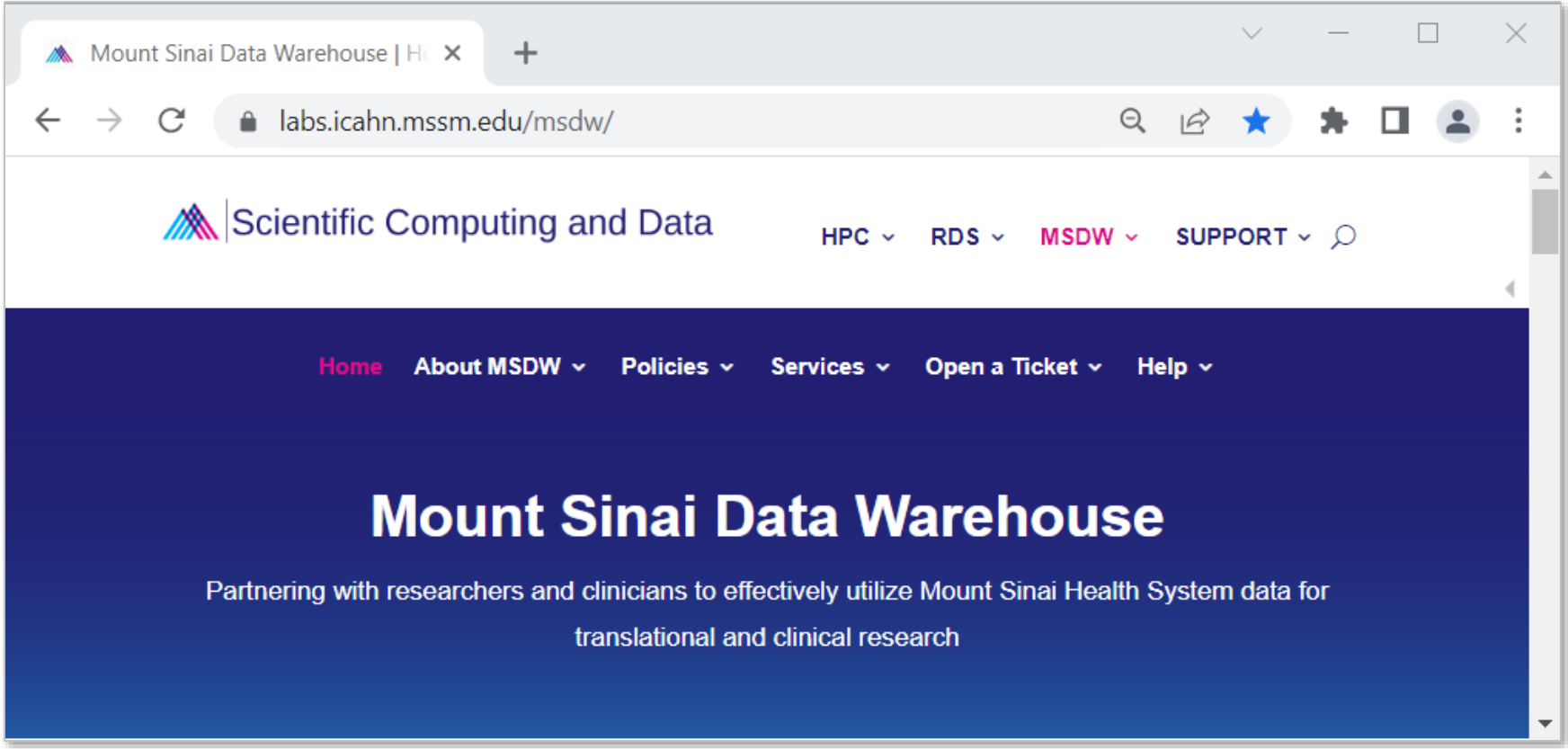
Computational & Data Science Ecosystem



Mount Sinai Data Warehouse on Minerva HPC Cluster



Our Website: msdw.mountsinai.org



MSDW Website Resources (a subset)

Website Resource	URL Link
MSDW Homepage	https://labs.icahn.mssm.edu/msdw/
Open a Ticket for Assistance	https://labs.icahn.mssm.edu/msdw/open-a-ticket/
Data Contents & Record Statistics	https://labs.icahn.mssm.edu/msdw/data-sources/
Protected Patient Categories & Data Exclusions	https://labs.icahn.mssm.edu/msdw/about-us/protected-patient-categories/
MSDW Release Notes	https://labs.icahn.mssm.edu/msdw/msdw-release-notes/
Presentations	https://labs.icahn.mssm.edu/msdw/presentations/
Training & Tutorials	https://labs.icahn.mssm.edu/msdw/training/
FAQs	https://labs.icahn.mssm.edu/msdw/faqs/