

2022 MSDW User Survey Comments and Responses

Scientific Computing and Data

May 12, 2023

The 2022 MSDW (Mount Sinai Data Warehouse) user survey—distributed in January 2023—solicited feedback from all active users of MSDW. From all MSDW users, 192 users opened a new ticket in 2022. 198 users responded anonymously.

Survey Questions:

1. **How satisfied are you with your experience working with the MSDW Clinical Data Informaticists and Data Analysts?**
 - a. Very Satisfied: 25%
 - b. Satisfied: 39%
 - c. Neither Satisfied Nor Dissatisfied: 19%
 - d. Dissatisfied: 5%
 - e. Very Dissatisfied: 3%
 - f. Decline to Answer: 10%
2. **How satisfied are you with the quality of the data you received?**
 - a. Very Satisfied: 24%
 - b. Satisfied: 38%
 - c. Neither Satisfied Nor Dissatisfied: 16%
 - d. Dissatisfied: 8%
 - e. Very Dissatisfied: 3%
 - f. Decline to Answer: 12%
3. **Did you use one of the self-service query tools Leaf or ATLAS in 2022?**
 - a. Yes: 13%
 - b. No: 83%
 - c. Decline to Answer: 3%
4. **[Of the 13% of users who answered “Yes” to #3] How satisfied are you with your experience generating cohort counts in Leaf?**
 - a. Very Satisfied: 31%
 - b. Satisfied: 19%
 - c. Neither Satisfied Nor Dissatisfied: 12%
 - d. Dissatisfied: 19%
 - e. Very Dissatisfied: 15%
 - f. Decline to Answer: 4%
5. **[Of the 13% of users who answered “Yes” to #3] How satisfied are you with your experience generating cohort counts in ATLAS?**
 - a. Very Satisfied: 19%
 - b. Satisfied: 4%
 - c. Neither Satisfied Nor Dissatisfied: 8%
 - d. Dissatisfied: 19%
 - e. Very Dissatisfied: 12%
 - f. Decline to Answer: 38%
6. **[Of the 13% of users who answered “Yes” to #3] How satisfied are you with resource support (ticket system, support personnel, training sessions, online guides, etc.)?**
 - a. Very Satisfied: 27%
 - b. Satisfied: 31%

- c. Neither Satisfied Nor Dissatisfied: 15%
- d. Dissatisfied: 19%
- e. Very Dissatisfied: 4%
- f. Decline to Answer: 4%

7. Please share any additional comments:

Comments regarding service:

- Extraordinarily professional, competent, and collaborative.
- I am having such a great experience working with MSDW team. Very informative environment.
- I am very satisfied with the service
- I enjoy working with this team. They are hard-working, professional and very responsive.
- I felt there was tension / passive aggressiveness in terms of customer service for first time users.

- The team is friendly to work with whenever we meet. The time it takes to have a request fulfilled can be quite long, though.
- The turnaround time is too long,
- Response time has slowed
 - Thank you for providing feedback regarding your experience. We have recently expanded our team by adding more staff.

- When emailing with MSDW to clarify questions about data that will be received, their responses are incomplete or vague. They do not provide complete and clear answers in a timely manner.
- Related to the satisfaction of my experience working with the clinical informaticists, I'll try to explain why I answered "Dissatisfied." First, let me say that I'm extremely thankful for the MSDW, and that it exists as a service to researchers. The reason my experience was dissatisfying is that communicating which data I wanted from MSDW (specific diagnosis, prescriptions, inclusion/exclusion criteria, control groups) was a major difficulty in working with the clinical informaticists, as they repeatedly did not understand what I was asking for. Now, I don't blame MSDW for this because the kinds of studies that researchers regularly conduct, myself included, are complicated! It's not easy to communicate exactly what you want to someone who is in no way, shape, or form involved in the study. This led to many, many rounds of back-and-forth meetings between myself and the clinical informaticists over several months, and it got very frustrating, for two reasons. The main reason is the obvious; that this kept delaying work getting done. The second reason is that MSDW kept denying my proposed solution to the first issue, which was to simply have a full or more general export of the MSDW, so that I could filter/subset the data as needed without having to have multiple meetings with MSDW so they could understand what I needed. I don't know if others have had the same difficulties I have in communicating about which data they want, but I think it would serve MSDW to have a full export of the EMR so that researchers can do whatever exploratory analyses they like, instead of having to fill out more forms, schedule more meetings, and review the data gathered by MSDW before finally receiving the data. When I proposed this, I was told that this kind of export would take a very long time and

couldn't be done, but I think this needs to be an option in the case there are difficult misunderstandings between researchers and MSDW about which data are needed for a study. And if exporting the entire MSDW is not an option, a version with minimal filtering should be, an example being an export of all diagnoses (or prescriptions) between certain years or above certain patient ages. Apologies if I'm an outlier and these kinds of miscommunications that end up consuming a lot of time are not common, but if others are raising this issue, it needs to be understood that this is consuming valuable research time, and more efficient solutions are needed. In regards to the quality of the data I received, I answered "Satisfied" because when the MSDW and I did come to a shared understanding of what I needed, the data were always well-labeled and easy to understand. There were 1-2 occasions where a column was missing, but aside from that it's always been great.

- [Thank you for providing feedback regarding your experience. We have recently expanded our team by adding more staff.](#)
 - [MSDW is available to discuss concerns regarding data and to answer general questions about data sets during the open office hours at the Digital Concierge, every Wednesday from 3:30-4:30 pm plus one additional weekday each month. \[Click here\]\(#\) to learn more about the Digital Concierge.](#)
- Additionally, the information on costing of different data options on the website is inaccurate.
 - [Custom data requests are processed at a rate of \\$180/hour as described on our FAQ page on our website under "\[Q: How much are the Data Warehouse services?\]\(#\)" and our \[Chargeback Policy page\]\(#\)](#)
- Additionally, the address I was told to send the fee to was not accurate and it took months to get it to be received.
- Finance team communication needs to be better and more proactive, and finance needs to communicate better with requester and data scientists. My data request was hung up on intermediate steps with finance that I had to facilitate individually.
- Turnaround time for data release has been delayed because of slow payment processing. When the data are available but are not being released because of the time taken to go through the payment process, the requester should be informed, so that attempts to expedite can be made if possible.
 - [We understand your concern with delays due to the billing process. We can make attempts to follow up on payment processing through Sinai, but data cannot be released until the payment has been processed. If you have a hard deadline for data delivery, please let our team know so that we can work with you to ensure a timely delivery.](#)
 - [We are now able to pull data from the MSX billing database in order to integrate CPT codes.](#)
- The ticketing system is also a little confusing. When I receive email requests/updates, it appears as though I should respond via email, however if I do, the email goes nowhere. It would be great if you could clarify the email text to say **DO NOT REPLY TO THIS EMAIL** and direct people to use the service website.

- Thank you for your comment. Currently the unmonitored noreply email communicates this message, and all users are encouraged to use our ticketing/tracking system submit a ticket [here](#)
- We are investigating improved ways for interacting with our ticketing/tracking system via email

Comments regarding data:

- I received data with mixed patient population which included patients who are not relevant to my study.
- Patient site is mixed up - requested Morningside only and got patients from main campus, Roosevelt and Queens. Need more curation of these fields. Data request is very difficult to amend - hard to get access to original query to straighten out patient site. Integration of CPT codes would be helpful
 - The [MSDW Custom Data Request Process](#) is explicitly designed to fully review and confirm details of the data request and, once developed, schedule a data approval meeting to confirm your satisfaction with the data before going to billing and data delivery. Any data elements missing should be identified in the final approval meeting for a data set.
- some of the MSDW data from EMRs is not reliable
 - MSDW data pulls directly from Epic EHR. EHR data is inherently complex, and the quality of data in records can vary. Please contact us [here](#) with more details and we will do our best to assist.
- Improved documentation of data sources included in the data provided by MSDW would be great
 - Please contact us [here](#) so we can learn more about your experience and improve our service.
 - Our current list of MSDW Data Contents is located [here](#)
- I've started working with colleagues in other health systems to use their records and their cohort samples because the Sinai system is restricted and prohibitive above and beyond the needs of HIPPA and data integrity. I am disappointed that I was unable to access and use the data that I needed for my presentations through Mount Sinai systems.
 - Please contact us [here](#) so we can learn more from your experience and improve our services to researchers. Some restrictions may be related to security policies set forth by the IRB team to ensure the appropriate protection of human subjects data for research.
- MSDW should not charge for data searches if the data is being used for resident quality assurance projects.
 - MSDW's policy is not to charge for QA/QI projects if approved by the QA Board. As you may already know, it is possible to query de-identified data within MSDW utilizing one of the free self-service cohort tools: [Leaf](#) or [ATLAS](#). Custom data requests are carefully constructed and vetted for quality, which incurs a cost of \$180/hour.

- Sometimes we have had to work to organize the data so we could use the information
- The data pulls are expensive and often come back with several errors (which at that price shouldn't happen). It is also not possible for the clinical informaticists and data analysts to pull data from PDFs and written records.
 - A vast majority of MSDW data is drawn from Mount Sinai Health System's Epic EHR. Please note that the MSDW team does not offer de-identified health information in free text fields (such as progress notes). Our Data Analysts can help you to determine what will work best for your research and ensure your requested data set is accurate and comprehensible.
- Data delivery schedule could be more accurate, even it is it a range of dates
- Two of the variables I asked for were not delivered with the data. I discovered it only months later when I finally got to analyzing the data and then had to submit new tickets to get it. In addition to the data, I suggest delivering a document summarizing the variables sent. This way the PI can review and make sure all the needed information is there.
- The process was a significant expense and our results were not initially completed as we requested. We had to reach out about an error in the report and have results recalculated and resent once they were correct. Overall we ended up with the report we wanted, but this kind of oversight should not have happened with the level of service we thought we were getting.
 - Please contact us [here](#) so we can learn more about your experience and improve our service.
 - The [MSDW Custom Data Request Process](#) is explicitly designed to fully review and confirm details of the data request and, once developed, schedule a data approval meeting to confirm your satisfaction with the data before going to billing and data delivery. Any data elements missing should be identified in the final approval meeting.

Comments regarding Leaf and/or ATLAS:

- I would like to learn how to use the leaf tool or atlas tool for clinical data for qi projects on my unit
 - To get started with Leaf, review the [online tutorial document](#), or visit the [Leaf Introduction Training Module in Peak](#). For ATLAS, review the [tutorial document](#), or visit the [ATLAS Introduction Training Module in Peak](#). Revisit session recordings and slide decks from past presentations on Atlas [on our website](#).
- The interface in LEAF is relatively easy to use, but some additional details on the timeframe of the numbers in LEAF would be helpful - are numbers for all time, for a year, etc. Some additional training videos on how to build cohorts in ATLAS would be helpful.
 - Regarding your question on timeframes: In Leaf, patient age always refers to the patient's age at the time a query runs. This means the same patients may not consistently appear in age-related searches that you run on different days, and that age may not correlate to the date of patient's vitals. For example, a record

for a patient born on January 1, 2000 will not appear in searches of Age: > 18 when queried on January 1, 2017, but will appear in the same search when queried after January 1, 2019.

- To get started with Leaf, review the [online tutorial document](#), or visit the [Leaf Introduction Training Module in Peak](#).
For ATLAS, review the [tutorial document](#), or visit the [ATLAS Introduction Training Module in Peak](#).
Revisit session recordings and slide decks from past presentations on Atlas [on our website](#).

- Steep learning curve to use the system. I would be interested in a more programmatic access to the data, using APIs
- Earlier this year I was looking for statistics on rates of NICU admission for infants born in the Sinai health system. Most of my queries were yielding results that were obviously wrong (very small Ns like a few tens to few hundreds while Sinai has 14000 deliveries/year, many patient ages over 70 y where it seemed vanishingly unlikely that data from their births would be in the system, etc). I did submit a ticket and was supposed to be connected with a specialist but that person never answered my emails so I eventually gave up on the question. That was earlier in 2022, not sure if there have been improvements since that time as I haven't tried again recently.
- Many criteria exist in both leaf and atlas that return null data. For example, searching various criteria to show women who delivered vaginally for a certain time period came back with a handful of results when the real number is in the thousands.
- I am not familiar with this program and this was first time trying to find data. I'm not sure this is the venue for me to gather data on imaging studies. thank you