Medical Imaging Data Marketplace Survey
Frequently Asked Questions

Updated as of April 10, 2024

Why do ARPA-H and FDA want to develop a medical imaging data marketplace?

With the increase in the development of artificial intelligence and machine learning in software as a medical device (SaMD), access to high-quality, regulatory-ready medical imaging data for developing and testing new technology is one of the most significant challenges faced by medical device and software developers. This limited access not only hinders innovation but also delays FDA clearance or premarket authorization due to data quality deficiencies, ultimately limiting the performance and availability of AI/ML software and medical devices.

ARPA-H and FDA seek to develop a medical imaging data marketplace to organize and manage medical imaging data at scale by connecting existing databases, marketplaces, and data providers to a trusted, streamlined platform that researchers and customers can use to find and affordably access the data needed to develop and test new algorithms. This initiative will include developing new tools that will give users of the marketplace greater confidence that the data accessed will be consistent with the regulatory requirements for future premarket authorizations.

What types of data will be included in the marketplace?

The envisioned marketplace aims to include a wide range of medical imaging data related to prostate and breast cancer, screening and diagnostic imaging, digital pathology or radiology as well as imaging recommended by the community. Imaging data such as DICOM images, X-rays, MRI scans, CT scans, ultrasound images, PET scans, and more will be included. Additionally, metadata related to patient demographics, imaging modalities, collection and acquisition parameters, and clinical annotations may also be included to enrich the dataset. To ensure compatibility and interoperability of datasets, the marketplace aims to support standard data formats such as Digital Imaging and Communications in Medicine (DICOM) and Neuroimaging Informatics Technology Initiative (NIfTI). We are eager to support other relevant formats based on community feedback and requirements.

What volume of data do you anticipate will be available?

We don't have a defined number yet, and it will depend on the modality but expect 10s to 100s of thousands of images per modality. We assume additional types of images may become part of the marketplace over time, that the marketplace will be across regions, and that the volume of data available will grow over time. We seek input from the research community on how many images (and what level of annotation) are necessary for their specific research applications.

How will data be used?

The data available on the marketplace is intended to serve a multitude of purposes. Researchers can leverage the data for developing net diagnostic tools, treatment algorithms, and predictive

models. Medical device manufacturers can access data for product development, validation, and regulatory submissions. The goal is for the data to be used to drive advancements in medical imaging technology, improve patient outcomes, and overall healthcare delivery.

Is there a plan to include other types of data? (Clinical data, treatments, genomic information, etc.)

Yes. We believe the program will expand overtime to meet additional needs from the research community.

Will metadata be included in consideration for the marketplace?

Yes, the intention is to have robust metadata included in the marketplace. We are looking for input on the specific pieces of metadata that should be provided. *(E.g. For Whole-slide images (WSI) would a half page of pathology data be enough?)*

Is this a dedicated medical imaging marketplace?

As the marketplace evolves over time, it may add new data types. The first acute use case identified as a high need for research is medical imaging data; specifically, medical whole-slide pathology and mammography because this is where the FDA is receiving the highest number of submissions. Based on survey response data, we will build a prioritized list of which indications the marketplace should expand upon after the initial modalities.

What regulatory approvals will be required to use this data?

Users will be responsible for obtaining any necessary approvals or permissions for data access, including, but not limited to, institutional review board (IRB) approval for research studies for submission for medical device development.

Are you looking to leverage existing cloud marketplace?

Yes. We don't want to start from scratch. Performers can reuse things they have already built or add onto solutions previously available.

How will business models be evaluated as part of the proposals?

To be determined. The network survey asks questions about the community's experience with database, exchanges and marketplaces. Using the survey responses, we are interested in the advantages and disadvantages of different business models. In particular, we are looking for business models that-meet to the goals of the medical imaging data marketplace and ensure its long-term self-sustainment.

Can an existing product be proposed as the solution?

Yes, an applicant can suggest expanding their existing solution to create the medical imaging data marketplace for any potential programs launched at a future date, however it must meet the needs which have yet to be defined and will incorporate feedback from this survey.

Could this be a public marketplace or be part of a public marketplace?

We are looking for a sustainable long-term business model that will enable the marketplace to continue long-term. That could come in the form of a non-profit or for-profit solution.

Who will have access to the data?

Access to the data is anticipated to be granted to authorized users, including researchers, healthcare providers, device manufacturers and software developers and other appropriate stakeholders. We are looking for your input on how to ensure access privileges are determined, and how user roles, permissions, and compliance are maintained while still allowing for an easy-to-use highly accessible site.

How will using data from this marketplace impact regulatory review?

The medical imaging marketplace will provide a variety of tools that will assess data for common regulatory compliance issues that impede premarket authorization. We envision that this will give users greater confidence that the data they use will meet FDA's requirements.
The medical imaging marketplace will provide a central site for data quality, annotation, and truthing and sequestration of imaging data. By consolidating data from various sources, regulatory bodies can more efficiently review and access data for bias, diversity, integrity, ethical considerations, and requirements for pre-approval. We envision this will support an expedited data approval process for research and medical device pre-approval submission.

What FDA/CDRH tools will be available?

ARPA-H and the FDA envision the marketplace will provide users with a range of tools and resources to facilitate the exploration, query, analysis, and utilization of medical imaging data. Our goal is to empower users with the tools they need to leverage the data effectively and drive innovation. To that end, we are considering tools that evaluate data quality, annotation, and truthing, sequestration, bias and diversity, support metrics, and analysis. Please provide input on what tools would be most valuable to you and your organization in the survey.

What does "Regulatory Ready Data" mean?

While the FDA cannot guarantee certification for products developed using data from the Marketplace, we want to give users confidence that the data they obtain is unlikely to have significant underlying data-quality issues that will impede their ability to obtain FDA premarket authorization. We characterize this data as "Regulatory Ready." The evaluation mechanisms, test methods, and standards supported will be developed or implemented as part of the marketplace.

How does this activity interact with the NIH Data Sharing initiatives?

The Medical Imaging Data Marketplace is intended to complement the work across the NIH. For example, NIH-funded studies would be able to connect their data to this marketplace.

What is the purpose of the medical data imaging marketplace Survey?

APRA-H is conducting the Medical Data Imaging Marketplace Survey to solicit feedback from relevant stakeholders to inform the design of a potential program to create a medical imaging data marketplace.

Who should complete the survey?

We invite all interested parties to participate in giving feedback on the medical imaging data marketplace survey. Teams can give their feedback as Data Users, Data Managers, and Providers. You may select multiple archetypes if you/your organization contains multiple types of stakeholders.

How do we define a medical imaging data user, data manager, and data provider?

- **Data Users** are researchers (academic or private sector) that need/use medical imaging data to train and test AI/ML models. We are specifically interested in Data Users working on both radiology and digital pathology applications.
- **Data Managers** are entities that manage, sell, or aggregate medical imaging data for researchers. These organizations can be databases, platforms, or aggregators. They do not necessarily produce this data themselves, but they make it available to Data Users.
- **Data Providers** create the medical imaging data needed by Data Users. They do not necessarily provide or sell their data to other parties. We are highly interested in providers of radiology and pathology images.

Will my responses be kept confidential?

This survey is administered by Investor Catalyst Hub, VentureWell, part of the Advanced Research Projects Agency for Health (ARPA-H) nationwide health innovation network, ARPANET-H. Contact information provided to complete the survey is confidential and will only be used for survey follow-up. Individual responses to survey questions are confidential and will only be accessible to select personnel in ARPA-H, FDA, and ARPA-H's Investor Catalyst Hub, VentureWell. Please do not include medical imaging data, patient data, or personally identifiable information of others in your response. Survey findings and key insights will only be presented in an aggregated format with no identifying information. Findings will be made available to all respondents through a public findings document. This survey does not constitute an endorsement of any initiative from ARPA-H, FDA or VentureWell.

When are responses due?

April 16th, 2024 at 5:00 p.m. ET.

We filled out the Survey once but we realized there's a bit more info we can provide that might be helpful. Will filling it out again erase our original response?

Yes, you can edit your responses after you've submitted your survey. If you have additional thoughts or comments to add in addition to the survey questions, you can email midm@arpa-h.gov.

What is the link? How do I provide input to the marketplace?

Responses can be submitted here.

When will the public report be published?

We expect to release the public report in late June 2024. This is subject to change.