



Research Opportunity Announcement

OTA-21-018

BioData Catalyst Data Management Core Program

Introduction

[NHLBI BioData Catalyst](#) is a cloud-based ecosystem providing tools, applications, and workflows in secure workspaces (**see Figure in Addendum 1**). By increasing access to NHLBI datasets and innovative data analysis capabilities, BioData Catalyst accelerates efficient biomedical research that drives discovery and scientific advancement, leading to novel diagnostic tools, therapeutics, and prevention strategies for heart, lung, blood, and sleep disorders. BioData Catalyst supports the overall [Strategic Vision](#) of NHLBI and, in particular, [Objective 7](#).

Though the primary goal of the BioData Catalyst program is to build a data science ecosystem, at its core, this is a people-centric endeavor. BioData Catalyst is also building a community of practice working collaboratively to solve technical and scientific challenges. The community includes developers, computer scientists and bioinformaticians who are building and operating the cloud-based platform; the data generators who collect and generate the data from NHLBI-funded studies; and platform users who access and compute on the data within the platform.

To make the BioData Catalyst endeavor successful, data must become a first-class research object. Investing time and resources into making research data FAIR - findable, accessible, interoperable, and reusable - is critical for ensuring maximum return of the Institute's investment in generating the data. NHLBI wants its research community to be able to easily find and use its data assets. NHLBI also wants to ensure that its data can be integrated and aggregated and optimally reused by its investigators.

The Institute faces several challenges in making data a first-class research object. These include:

- Prioritizing and sequencing NHLBI datasets for ingest into BioData Catalyst
- Standardizing (meta)data
 - Retrofitting existing datasets by working with data generators
 - Promulgating standards for prospective/future data collection
 - Using technology to facilitate standardization
- Understanding governance of datasets and translating for BioData Catalyst developers

The Institute has several immediate data management needs. These include ingesting multiple data sets of varied data types (electronic health record, research clinic surveys and assessments, genotypes, molecular and laboratory, imaging) into BioData Catalyst and the number of data sets has increased as a result of the COVID-19 pandemic. To efficiently ingest these data sets, a shared understanding between the data generators and the BioData Catalyst

developers is required of data formats, structure, standards, and governance. Minimal (meta)data standards are required across the BioData Catalyst ecosystem and these standards need to be communicated to data generators and NHLBI staff. Standard operating procedures for getting data ready for ingest into BioData Catalyst need to be developed.

This solicitation invites applicants with appropriate expertise and capacity to submit a proposal to serve as the BioData Catalyst Data Management Core, outlining the functions, roles, and responsibilities of such a Core in ensuring the efficient and effective management of data before and during ingest into BioData Catalyst.

Authority

This Research Opportunity Announcement (ROA) is issued with the goal of establishing an “Other Transactions” (OT) agreement pursuant to 42 U.S.C. § 285b-3.

Objectives

The NHLBI seeks to establish a Data Management Core (DMC) to support the overall research mission of the NHLBI. The overall goals of the DMC will be to:

- Identify and prioritize, in collaboration with NHLBI, data for onboarding into BioData Catalyst
- Standardize data and streamline the data onboarding process
- Maximize value of data through harmonization prior to onboarding
- Identify and develop, as needed, technology to innovate and create efficiencies in the above areas

The DMC will interact closely with stakeholders across the NHLBI research enterprise. The DMC will work with domain experts relevant to the NHLBI research community (genomics, imaging, phenotypes of interest to NHLBI (e.g., heart failure, asthma, sickle cell disease, etc.)) to understand what data standards exist and are applicable to the research of the NHLBI community. The DMC will educate data generators about data standards and will subcontract to particular data generators and consortia to configure their data for ingest into BioData Catalyst. The DMC will work with data generators and consortia to facilitate standardization and harmonization of their data including harmonization of data with existing NHLBI datasets. This will be achieved through proactive promulgation of Common Data Elements (CDEs) and common measurement methods. Where necessary, the DMC will also support the data generator in retrospectively harmonizing phenotypes to enable integration and aggregation of data with existing NHLBI datasets such as NHLBI’s cohort studies. It is anticipated that the DMC will develop and/or collate a suite of tools to facilitate phenotype harmonization ranging from clearly documented methods to software tools and Jupyter notebooks to artificial intelligence and machine learning methods. The DMC will catalog similar tools and guidance for standardization of molecular data, for example alignment and variant calling pipelines for genomics.

The DMC will work with BioData Catalyst developers to understand BioData Catalyst architecture, ingest process, and requirements for ingest. The DMC will also work with BioData Catalyst developers to develop cloud-based work spaces for data preparation. The DMC will also work collaboratively with NHLBI staff on understanding data governance at NHLBI, developing data standards, identifying content experts, and recommending guidance to the research community.

It is expected the DMC will take an agile and strategic approach to accomplishing its objectives. Nimble subcontracting will be required to bring in subject matter experts as needed and to support data generators as they prepare their data. It is expected that the DMC will not have all required subject matter experts on their team at project initiation.

Statement of Objectives and Milestones (SOM)

The activities of the Data Management Core may be grouped into 4 categories: operations, data management, data standards, and data technology and these activities are reflected across the objectives below.

1.0 Operations

- 1.1 Plan and implement an NHLBI-inclusive Data Management Core that works collaboratively with the NHLBI BioData Catalyst program team, NHLBI leadership, data generators, data stewards, coordinating centers, the NHLBI Data Access Committee, and the BioData Catalyst Consortium.

Under the guidance of the NHLBI BioData Catalyst (BDC) team, the Awardee will:

Provide basic administrative support for coordinating and planning meetings, drafting agendas, taking notes, providing meeting minutes, and recording / monitoring / follow up action items, relevant to DMC activities. This includes administrative meeting support for the NHLBI, the Core and its subcontractors, and working groups. Meetings may take the form of teleconferences, webinars and up to 4 in-person meetings per year.

- 1.2 Establish and enable well defined and documented data management practices that adhere to FAIR principles (Findable, Accessible, Interoperable, Reusable/Reproducible) within NHLBI and its broad research community. Where possible, the DMC should leverage existing practices and standards.
- 1.3 Implement a comprehensive program and project management practice with the expertise and experience to coordinate, monitor, and implement a diverse and complex array of activities across multiple research programs. The program must incorporate the development of methods for prioritization of curation and ingest of data sets into BioData Catalyst that is cognizant of NHLBI leadership priorities and the needs of the research community and other relevant stakeholders and is consistent with informed consent. The DMC should work with NHLBI staff to

establish pathways and processes for seeking NHLBI leadership approval where needed.

Under the guidance of the NHLBI BDC team, the Awardee will:

Provide day to day project management services, including project plan tracking and schedule management through the use of appropriate project management software and other electronic tools in order to coordinate and integrate the milestone-driven activities, as well as budget monitoring, and risk identification and mitigation for all consortium-based projects.

Implement project modifications as necessary, after consultation with NHLBI. Serve as the key point of contact with parties at NHLBI and the BioData Catalyst consortium awardees to help coordinate efforts in close collaboration with all technical staff. Prepare and distribute detailed project plan status reports and other project plan related documentation to the NHLBI and the consortium. Oversee the various components of the Core activities with appropriate milestones and metrics.

- 1.3.1 Monitor and report program activity status, risk, cost and schedule against approved plans. Reporting should be progress relative to metrics which the Core to identifies and develops, in consultation with NHLBI leadership, the BioData Catalyst consortium, and other stakeholders.
- 1.4 The Awardee must embrace the speed and flexibilities of the Other Transaction mechanism by taking a more active and direct role than may be usual in helping the government foster the success of the initiative. As such, the Awardee cannot passively monitor progress but must be responsible for ensuring strategic alignment, programmatic oversight and ensuring the government's interests are met. The awardee should quickly identify issues and making recommendations to address them. Given changing priorities and/or inadequate progress, the DMC must be able to stop a project and pivot efforts and resources toward more compelling scientific challenges/needs.
- 1.5 Coordinate with and provide financial support to research institutions in order to bring in needed outside expertise and to support data generators/data stewards in their work to prepare data for ingest. Awardee must have efficient and flexible subcontracting processes in place.
- 1.6 Regularly communicate and report on activities to stakeholders including NHLBI staff and leadership, the NHLBI BioData Catalyst consortium, the BioData Catalyst External Expert Panel (EEP), research programs etc. to ensure they have an accurate and up to date picture of the core's activities and plans, and for the NHLBI, financial status. Communicate plans and priorities for data curation and ingest with time horizons.

2.0 *Data management*

- 2.1 Publish and maintain a compendium of high value NHLBI data sets to enable the research community to understand the available data and assess the utility of the data to their research goals including a program to evaluate and maintain an assessment of the FAIRness of the data sets consistent with other NIH data management and sharing efforts to assess and promote FAIRness.
- 2.2 Develop a prioritization scheme of how to order data sets for ingest by soliciting external experts, EEP and NHLBI leadership input that may also include input from the research community.
 - 2.2.1 Develop a multi-year road map which includes the timelines and estimated scope, level of effort, and cost needed to on-board the data sets.
 - 2.2.2 Provide regular updates to enable NHLBI, the BioData Catalyst Consortium and the research community to understand plans and progress on data ingestion including dashboards relating to program activities and status.
- 2.3 Support data generators/stewards in preparing data for ingest into BioData Catalyst, in a consultant type role. This will involve ingestion of varied data types, including, but not necessarily limited to, data from electronic health records, research clinic surveys and assessments, genotypes, molecular and laboratory assays, and imaging procedures. This will involve some combination of deidentification, digitization, alignment to data standards, QA/QC checks, governance translation, etc. This might involve subcontracting with data generators/stewards to accomplish this work. This may involve use to tools/workflows developed by BioData Catalyst and/or the DMC. This may involve supporting/educating data generators/stewards in working in cloud-based workspaces to accomplish this work.
 - 2.3.1 The DMC should demonstrate expertise in the data types highly relevant to the NHLBI research community, including genomics, cardiac and lung imaging, and relevant phenotypes, such as heart failure, sickle cell anemia, and chronic obstructive pulmonary disease. It is not expected that all experts will be at a single institution nor is it expected that all expertise will be present in the initial application. The DMC must be able to identify the appropriate expertise and be facile at bringing that expertise into the DMC.
 - 2.3.2 The DMC should develop policies and procedures for interacting with data generators and their respective coordinating centers (if relevant). An example is the Administrative Coordinating Center for the TOPMed program. These should include methods for communication and should

describe how the DMC will work with data generators to register studies in dbGaP, support data harmonization, and support data ingestion into BDC.

- 2.3.3 The DMC should develop policies and procedures for interacting with the BioData Catalyst consortium. These should include methods of communication, representation on relevant BDC committees (such as the Data Release and Management Working Group (DRMWG)), and how the hand-off of data to relevant BDC components will occur.

3.0 *Data standards*

- 3.1 Solicit external/internal expert advice, via working groups or requests for information (RFIs), to define enterprise-wide approaches to data standardization and publish appropriate guidance to the research community.
- 3.2 Develop and maintain guidance on data formatting, meta-data model, common data elements, ontologies, mappings, etc. This guidance may be data type and/or disease specific. The DMC should not develop de novo standards where well-accepted, reasonable standards exist. The guidance should be developed in consultation with external experts, NHLBI, NLM and other NIH staff and the research community. These standards shall be published for wide access by the NHLBI research community.
- 3.3 Ensure that appropriate processes and procedures exist for quality assurance and quality control for data being ingested to BioData Catalyst by working with BioData Catalyst developers, data generators/stewards, and external/internal experts. Provide regular assessments and reporting on quality factors to the NHLBI.
- 3.4 Publish and maintain data governance procedures/policies for BioData Catalyst. Work with BioData Catalyst developers and data generators to implement alternative approaches as needed.

4.0 *Data technology*

- 4.1 Collate a suite of tools that enable and facilitate ingest of data into BioData Catalyst including tools for assessing data readiness, standardizing and harmonizing data and preparing data for ingest and validating data including compliance with relevant policies and consistency with study participant consents following ingest. This objective may require ongoing research into data preparation methods using a variety of techniques including artificial intelligence and machine learning methods.

5.0 *Reporting*

The Awardee will engage with the NHLBI to assess progress and will be required to provide reporting updates on a regular schedule.

5.1 Monitor, assess, and provide reports on the overall status of Core activities

Under the guidance of the NHLBI PMO, the Awardee will track and monitor progress under a unified milestone-based project management framework to keep the NHLBI informed of consortium-based projects and overall strategic progress. The awardee includes metrics identification, collection and analyses to determine if the projects and program are meeting intended goals. Awardee will monitor and report on cost, schedule, and performance relative to projections and address anomalies and actions which are being taken to address.

5.1.1 Monthly reports

On a monthly basis, the Awardee Business Official (ABO) will submit a single report to the Agreements Officer (AO) Scientific Program Director (SPD), and the PMO assigned Program Manager consisting of financial and technical information, that provides the government with sufficient detail to monitor schedule, risk, budget, and overall progress of each milestone appearing in this SOM. The NHLBI/NIH will provide a suggested template to the Awardee for reporting purposes. Monthly reports are due to the NHLBI OT Mail Box (NHLBI_OTA@mail.nih.gov) and to the NHLBI/NIH Points of Contact 15 calendar days following the end of each calendar month.

5.3 Monthly meetings

On at least a monthly basis, the Awardee will make the Principal Investigator and key personnel available and hold a telephone, WebEx/zoom, and/or in-person meeting with the NHLBI/NIH. In addition, the Awardee's ABO and PI may need to be available for direct discussions with the NHLBI SPD, PM and AO as needed.

5.4 Semi-annual reports

Provide semi-annual evaluation reports to NHLBI with respect to metrics for the overall program and individual projects. Reports may be used for the BioData Catalyst external advisory panel.

6.0 *Phasing and potential milestones*

6.1 It is expected that the activities to achieve the above Objectives of the DMC will be conducted in phases: Initiation, Operationalization, and Maintenance. Initiation phase will involve start-up activities, landscape analyses, and preliminary work. Operationalization phase will involve establishment of standard operating procedures (SOP), policies, and procedures. Maintenance phase will involve implementation of plans, ongoing work, and continuous monitoring and improvement of processes. Overlap of some activities across phases is expected.

6.2 A list of potential deliverables by phase are provided below. These are just examples and the lists are not exhaustive.

Initiation phase:

- Landscape analysis of work to date and proposed strategy to leverage existing work
- Evaluation of high value NHLBI data sets (**see list of released and backlogged data sets in Addendum 2**)
- Prioritization schema for ordering data sets for ingestion (approved by DSLT) and implication to NHLBI data sets
- SOPs for ingesting data into BDC
- Solution for providing stakeholders data ingest status and beta implementation
- Recommendations for enterprise-wide approach for data standardization
- Compendium of existing data standards used by NHLBI studies
- Landscape analysis of available tools
- Review and evaluation of data standardization/harmonization work to date
- OT management plan that outlines how the DMC will manage within the Biodata Catalyze eco systems and the various stakeholders.
- Subcontracts for needed expertise or support transitional datasets

Operationalization phase:

- SOP for continuous landscape analysis of datasets
- SOP for periodic data prioritization
- Subcontracts to data generators, as needed, for pre-ingest data standardization
- SOP for maintaining and updating data standards guidance
- Test tools to support data standardization

Maintenance phase:

- Continuous landscape analysis
- Periodic data prioritization
- Develop durable materials (e.g., guidance documents, checklists, etc.) to support data generators
- Maintenance and updating of data standards guidance
- Supports users working with data standardization/readiness tools

Eligibility

Organizations

The following entities are eligible to apply under this ROA:

- Higher Education Institutions

- Public/State Controlled Institutions of Higher Education
- Private Institutions of Higher Education

Nonprofits Other Than Institutions of Higher Education

- Nonprofits with 501(c)(3) IRS Status (Other than Institutions of Higher Education)
- Nonprofits without 501(c)(3) IRS Status (Other than Institutions of Higher Education)
- For-Profit Organizations
- Small Businesses
- For-Profit Organizations (Other than Small Businesses)

Special Award Terms

The complete terms and conditions of each Other Transaction award issued under this ROA are subject to negotiation and will be contained in the OT Agreement entered between the NHLBI and the Awardee. This Special Award Terms section is provided for informational purposes only in order to provide prospective applicants with an understanding of key expectations and terms that may differ from traditional NIH award mechanisms.

Award Criteria and Selection Information

The awardee will be selected through an objective review process. The level of funding for the award made under this ROA has not been predetermined but will depend on (1) the objectives proposed by the applicant and how well they fit with the goals of the host targeting therapies initiative, (2) quality of the proposals received, and (3) budget evaluation of value for dollar. The agreement for the award will be negotiated with the eligible entity whose proposals is determined to be the most meritorious and provide the best value to the NHLBI toward achieving the goals of the DMC and in accordance with the NHLBI priorities.

The NHLBI reserves the right to:

- Select for negotiation all, some, one, or none of the proposals received in response to this ROA;
- Segregate portions of resulting awards into components and their associated budget and/or milestones that differ from those that have been proposed;
- Accept proposals in their entirety or to select only portions of proposals for award;
- Fund proposals in increments and/or with options for continued work at the end of one or more phases, which can consist of more than one milestone;
- Fund proposals of two or more applicant entities as part of a reorganized, consolidated consortium operating under an article of collaboration, teaming arrangement, or other means acceptable to the NHLBI;
- Fund proposers as sub-awardees of a separate Coordinating Center entity to be established by the NHLBI;
- Request additional documentation (certifications, etc.); and

- Remove proposers from award consideration should the parties fail to reach a finalized, fully executed agreement, or the proposer fails to provide requested additional information in a timely manner.

Proposal Process

Submission in response to this ROA occurs in two stages. Stage 1 requires submitting a Letter of Request (see [>\\$500k process on NHLBI web site](#)). If invited by NHLBI, the applicant shall proceed to Stage 2, which requires submitting a full proposal using eRA ASSIST.

STAGE 1 - NHLBI Biodata Catalyst Data Management Core Program Initial Eligibility and Preliminary Review

NHLBI will review and determine whether the applicant should proceed with completing the full proposal submission. The NHLBI may request additional information be provided by the applicant to complete their initial eligibility and preliminary review. These requests will be sent to the applicant via email. Applicants are strongly encouraged to provide the requested information in a timely manner to prevent any potential delays in the review process. Proposals that do not meet the initial Biodata Catalyst Data Management Core Program ROA program and eligibility criteria will be rejected.

Stage 1 will be patterned after the NHLBI >\$500K process (<https://www.nhlbi.nih.gov/grants-and-training/policies-and-guidelines/applications-with-direct-costs-of-500000-or-more-in-any-one-year>) that includes a Letter of Request addressed to Dr. David Goff, DCVS Director, Chair of the Data Science Leadership Team, outlining major elements of the proposed core, personnel, budget, and feasibility.

In order for the NHLBI to make an informed decision about whether to accept a proposed application for review, the Letter of Request (see link) should not exceed five (5) pages and should include:

- The proposed project title
- The anticipated solicitation (Funding Opportunity Announcement/Research Opportunity Announcement)
- The anticipated application receipt date
- The expected start date to launch the DMC
- The key personnel (the eRA Commons *user id* must be included for the PI or contact PI)
- The submitting organization or institution
- Brief description of the overall goals of the DMC, their rationale, and how they will support the overall NHLBI research mission
- Brief description of the main DMC activities and how they will be accomplished
- Brief description of any prior work or preliminary work in this area including data management across the varied data types described above

- Brief description of program and project management approach and plan, with particular attention to how milestones will be planned, monitored, reported, and, when appropriate, shared with stakeholders
- Brief description of DMC personnel including their relevant expertise and history in conducting the type of activities necessary for establishment and implementation of the DMC including data management, data harmonization, across the varied data types described above
- Demonstrated ability to efficiently issue subcontracts
- Demonstrated ability to operate in a highly collaborative and flexible manner in a complex program as required by the Other Transaction mechanism
- Direct and total costs by phase of the project – generally a one-paragraph description of major sources of costs in the study using one of the two tables provided via the following link: https://www.nhlbi.nih.gov/files/docs/500K%20Budget%20Tables_2018-09-17.xlsx
 - Specify (1) any funding provided by other entities (federal agencies, foundations, companies), and (2) any goods or services (and their value) provided by any of these parties.
- A description of any anticipated agreements with third-parties relevant to the proposed project, including details about any provisions or restrictions related to intellectual property, publication, and data sharing

STAGE 2 - Full Proposal

If upon review of the Letter of Request the NHLBI determines the proposed research to be in scope and hold significant promise for achieving the objectives of this ROA, the applicant will be invited to submit a full proposal. The full proposal will be submitted via eRA ASSIST following instructions that will be provided to the applicant.

The full proposal should include information in the following areas:

- Additional administrative information about the applicant and institution or organization (name, address, entity and Principal Investigator NIH Commons Registration information), including SAM information and DUN and Bradstreet number, human and animal assurance approvals as appropriate.
- Project Plan uploaded as searchable PDF format in a font size of 11 or 12 point and font type of Arial or Times New Romans. Margins must be 1-inch wide (top, bottom, left, and right). The project plan must not exceed 30 pages in length. Biosketches must not exceed 4 pages in length and are not counted in the page limit. Also excluded from the page limitation are cover sheets, letters from collaborators and consultants, and representation and certification documents.
- Budget reflecting the total cost proposed, accounting for cost share amounts offered by the applicant. (If proposing F&A include a negotiated federal rate approval.)

Project Plan

The Project Plan should generally include the following elements:

- Project Summary: Description of the overall goals of the DMC and the activities that will support its implementation.
- Detailed description of how each Objective will be met and how progress will be monitored. Applicant should describe how innovative practices and approaches will be used to accomplish the Objectives.
- Operational Milestone Based Plan: The plan should describe all proposed Operational Milestones by Objective and by each Phase. Each Operational Milestone should include objective completion criteria and an anticipated completion date, as well as a 9-12 month timeline showing each milestone in a Gantt chart like format.
- Financial Contingency Plan: Potential risks, mitigation strategies, and associated costs, including a description of a viable source to cover these costs (other than NHLBI and not including co-funding).
- Team Organization: Team structure, leadership and communications plan, including biosketches of individuals identified as the principal investigator and all key personnel.
- Resources and Environment: Resources available to the project and environment in which the activities will be performed.
- References

Budget

The Budget section of the application must provide a realistic, fully justified budget and cost proposal for performing the work over a specified period of performance needed to accomplish project objectives.

Funding will be obligated annually, with bi-annual restrictions in place. An initial amount of funding, covering the execution of the first half of the activity period, will be made available (unrestricted) upon award. Subject to NHLBI program staff review and approval, additional bi-annual funding will be unrestricted based on achievement of planned project schedule objectives.

Provide the overall expected cost for each of the following categories:

- Personnel
- Equipment
- Travel
- Subawards/subcontracts/consultants
- Other direct costs
- Total cost (with indirect costs included)
- Proposed Cost Share contribution

Submission and Contact Information

For best consideration, initial Stage 1 Letters of Request should be submitted via email by **Tuesday, February 22, 2022, by 5 PM EDT** to NHLBI_OTA@mail.nih.gov. This mailbox will also be used to answer questions. If invited to submit a Stage 2 full proposal, the applicant will be provided specific instructions for how to submit via eRA ASSIST.

Financial and administrative questions should be addressed to Benjamin Sakovich, NHLBI Agreements Officer, Benjamin.Sakovich@nih.gov.

Technical questions should be addressed to the NHLBI Division that best aligns with the study: Christopher Miller at christopher.miller4@nih.gov

A note about eRA Registration

NHLBI uses the eRA Commons system to administer OT awards. If you are invited to submit a full proposal in eRA ASSIST, you will need to be registered in eRA Commons, which can take some time to complete – as many as several weeks in some cases. Therefore, if you are considering submitting a proposal and are not yet registered in eRA, it is highly recommended that you begin the process of registering your organization, Program Director/Principal Investigator (PD/PI) and Signing Official (SO) in eRA Commons as soon as possible to avoid possible award processing delays. To register, please follow the instructions via this website: <https://public.era.nih.gov/commons/public/registration/registrationInstructions.jsp>.

1. Complete the online Institution Registration Form and click Submit.
2. The NIH database will send you an email with the link to confirm your email address.
3. Once your email address is verified, the NIH will review your request and let you know of the result via email.
4. If your request is denied, you will get an email notifying you of the reason.
5. If your request is approved, you will get an email with your Commons User ID and temporary password.
6. Log into Commons with the temporary password and the system will prompt you to change temporary password to a permanent one. Your SO will be prompted to electronically sign your registration request. (Please review your registration information carefully.)
7. Once your SO has electronically signed the request, your organization will be active in Commons and you may create and maintain additional accounts for your institution staff.

To complete the registration above, you may need to register for the following if you haven't done so already:

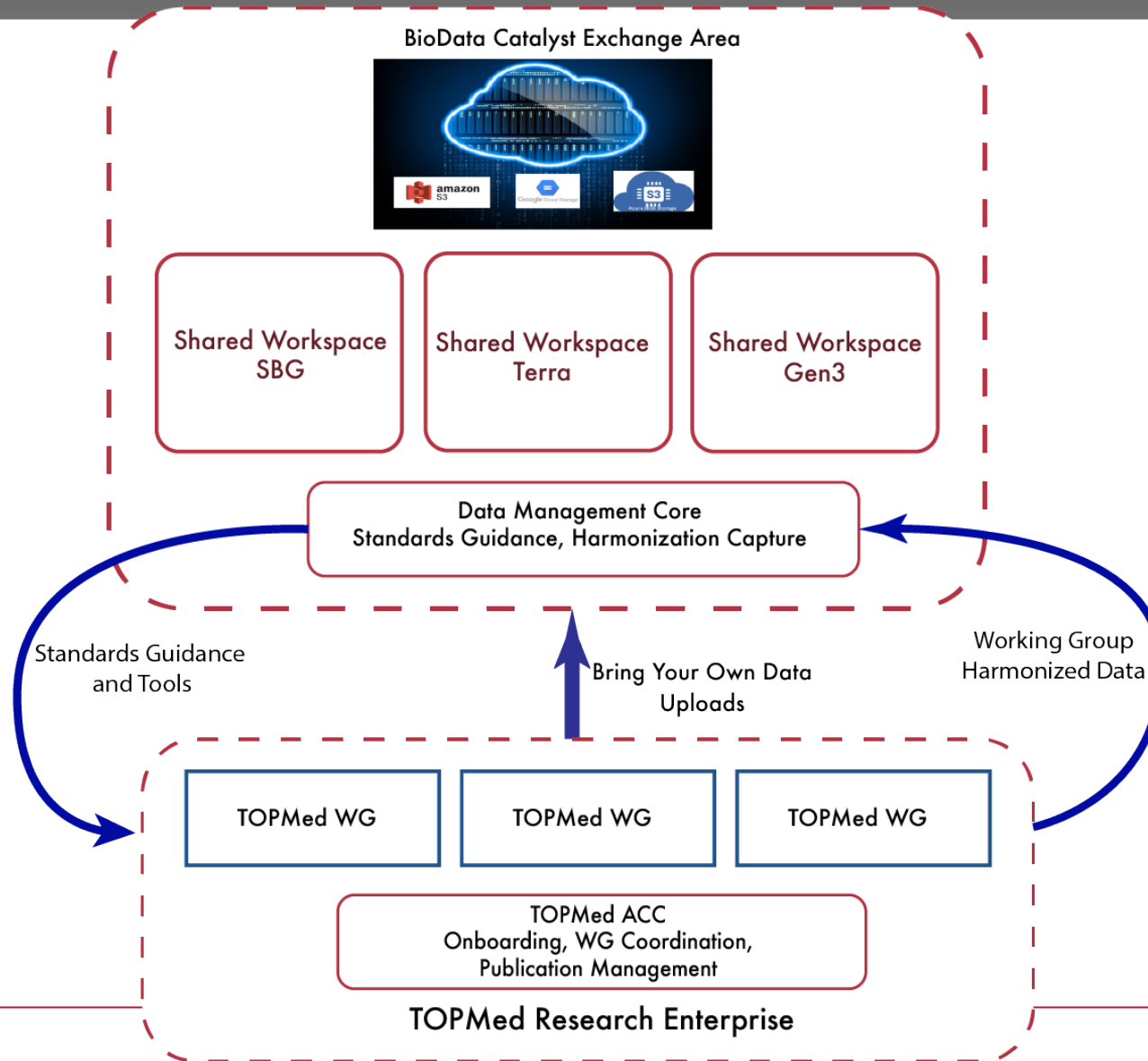
Dun & Bradstreet Number (DUNS) - <https://fedgov.dnb.com/webform/>

Employer Identification Number (EIN)- <https://www.irs.gov/businesses/small-businesses-self-employed/apply-for-an-employer-identification-number-ein-online>

Small Business Administration (SBA) - <https://www.sbir.gov/registration>

System for Award Management (SAM) - <https://www.sam.gov/SAM/>

HLBS Scientific Community Served by Enterprise



Jira Ticket Number	Study/Data Name	Status	Comments	COVID-19	Sickle Cell	Genomic	Phenotypic/ Clinical	Imaging	EHR
MONITOR-715	1000 Genomes	BDCatalyst Released	"BDCatalyst Released" Indicates availability across the entire ecosystem	-	-	X	X	-	-
MONITOR-716	TOPMed Freeze 8 -- Batch 3	BDCatalyst Released	"BDCatalyst Released" Indicates availability across the entire ecosystem	-	-	X	X	-	-
MONITOR-712	TOPMed Freeze 8 -- Batch 2	BDCatalyst Released	"BDCatalyst Released" Indicates availability across the entire ecosystem	-	-	X	X	-	-
MONITOR-710	Freeze 8 Combined Exchange Area	BDCatalyst Released	"BDCatalyst Released" Indicates availability across the entire ecosystem	-	-	X	X	-	-
MONITOR-723	TOPMed Freeze 8 -- Batch 1	BDCatalyst Released	"BDCatalyst Released" Indicates availability across the entire ecosystem	-	-	X	X	-	-
MONITOR-722	TOPMed Freeze 5b	BDCatalyst Released	"BDCatalyst Released" Indicates availability across the entire ecosystem	-	-	X	X	-	-
MONITOR-708	Modification of Metadata Fields	BDCatalyst Released	"BDCatalyst Released" Indicates availability across the entire ecosystem	-	-	-	-	-	-
MONITOR-818	Additional TOPMed Studies	BDCatalyst Released	"BDCatalyst Released" Indicates availability across the entire ecosystem	-	-	X	X	-	-
MONITOR-755	CureSCI - MSH (via BioLINCC)	BDCatalyst Released	"BDCatalyst Released" Indicates availability across the entire ecosystem	-	X	-	X	-	-
MONITOR-756	CureSCI - CSSCD (via BioLINCC)	BDCatalyst Released	"BDCatalyst Released" Indicates availability across the entire ecosystem	-	X	-	X	-	-
MONITOR-742	PCGC (CMG/Wagner)	BDCatalyst Released	"BDCatalyst Released" Indicates availability across the entire ecosystem	-	-	X	-	-	-
MONITOR-849	New data from TOPMed combined EA	BDCatalyst Released	"BDCatalyst Released" Indicates availability across the entire ecosystem	-	-	X	X	-	-
MONITOR-810	TOPMed Freeze 8 Corrected Zip Files	BDCatalyst Released	"BDCatalyst Released" Indicates availability across the entire ecosystem	-	-	X	X	-	-
MONITOR-753	CureSCI - CIBMTR (via BioLINCC) - aka HCT for SCD	BDCatalyst Released	"BDCatalyst Released" Indicates availability across the entire ecosystem	-	X	-	X	-	-
MONITOR-764	TOPMed Parent Study Updates	BDCatalyst Released	"BDCatalyst Released" Indicates availability across the entire ecosystem	-	-	-	X	-	-
MONITOR-759	PETAL - ORCHID	BDCatalyst Released	"BDCatalyst Released" Indicates availability across the entire ecosystem	X	-	-	-	-	-
MONITOR-906	Additional Parent Freeze8 Study (CARDIA)	BDCatalyst Released	"BDCatalyst Released" Indicates availability across the entire ecosystem	-	-	X	X	-	-
MONITOR-920	PETAL - ORCHID (re-ingestion)	BDCatalyst Released	"BDCatalyst Released" Indicates availability across the entire ecosystem	X	-	-	-	-	-
MONITOR-760	PETAL - RED CORAL (COVID-19)	BDCatalyst Released	"BDCatalyst Released" Indicates availability across the entire ecosystem	X	-	-	-	X	-
MONITOR-935	PCGC SRA Data	Pre-Ingestion	"Pre-Ingestion" - Data is available and undergoing QA/QC in NHLBI buckets before ingestion.	-	-	X	-	-	-
MONITOR-1051	Immune response to Life-Threatening Respiratory Infections in Children and Young Adults (COVID-19)	Prioritized	"Prioritized" - Data is prioritized by the BDC CCB for ingestion and is in the process of being uploaded to NHLBI buckets.	X	-	-	X	-	-
MONITOR-971	TOPMed Freeze 9 - Batch 3	Prioritized	"Prioritized" - Data is prioritized by the BDC CCB for ingestion and is in the process of being uploaded to NHLBI buckets.	-	-	X	X	-	-
MONITOR-852	BioLINCC Phase 2 (16 TOPMed Studies)	Prioritized	"Prioritized" - Data is prioritized by the BDC CCB for ingestion and is in the process of being uploaded to NHLBI buckets.	-	-	-	X	-	-
MONITOR-831	C3PO (COVID-19)	Prioritized	"Prioritized" - Data is prioritized by the BDC CCB for ingestion and is in the process of being uploaded to NHLBI buckets.	X	-	-	X	-	-
MONITOR-851	NuMoM2b Heart Health Study	Prioritized	"Prioritized" - Data is prioritized by the BDC CCB for ingestion and is in the process of being uploaded to NHLBI buckets.	-	-	X	X	-	-
MONITOR-832	ACTIV4a (COVID-19)	Prioritized	"Prioritized" - Data is prioritized by the BDC CCB for ingestion and is in the process of being uploaded to NHLBI buckets.	X	-	-	-	-	-
MONITOR-1054	ACTIV4b (COVID-19)	Prioritized	"Prioritized" - Data is prioritized by the BDC CCB for ingestion and is in the process of being uploaded to NHLBI buckets.	X	-	-	-	-	-
MONITOR-758	CureSCI - BabyHug (via BioLINCC)	Released by Gen3	"Released by Gen3" - Data is indexed by Gen3 and available for downstream processing by the remaining BDC platform teams prior to BDC-wide release	-	X	-	X	-	-
MONITOR-919	Additional TOPMed Freeze 8 Studies (CATHGen)	Released by Gen3	"Released by Gen3" - Data is indexed by Gen3 and available for downstream processing by the remaining BDC platform teams prior to BDC-wide release	-	-	X	X	-	-
MONITOR-887	Additional TOPMed Freeze 8 Studies (PMBB)	Released by Gen3	"Released by Gen3" - Data is indexed by Gen3 and available for downstream processing by the remaining BDC platform teams prior to BDC-wide release	-	-	X	X	-	-
MONITOR-1014	TOPMed Freeze 9 - Batch 2	Released by Gen3	"Released by Gen3" - Data is indexed by Gen3 and available for downstream processing by the remaining BDC platform teams prior to BDC-wide release	-	-	X	X	-	-
MONITOR-757	CureSCI - STOP-II (via BioLINCC)	Released by Gen3	"Released by Gen3" - Data is indexed by Gen3 and available for downstream processing by the remaining BDC platform teams prior to BDC-wide release	-	X	-	X	-	-
MONITOR-754	CureSCI - Walk-PhaSST (via BioLINCC)	Released by Gen3	"Released by Gen3" - Data is indexed by Gen3 and available for downstream processing by the remaining BDC platform teams prior to BDC-wide release	-	X	-	X	-	-
MONITOR-817	TOPMed Freeze 9 - Batch 1	Released by Gen3	"Released by Gen3" - Data is indexed by Gen3 and available for downstream processing by the remaining BDC platform teams prior to BDC-wide release	-	-	X	X	-	-
MONITOR-707	BioLINCC (Phase I) - Training Data (CAMP, Framingham)	Released by Gen3	"Released by Gen3" - Data is indexed by Gen3 and available for downstream processing by the remaining BDC platform teams prior to BDC-wide release	-	-	-	X	-	-
MONITOR-888	BioLINCC (Phase I) - Training Data (Digitalis)	Released by Gen3	"Released by Gen3" - Data is indexed by Gen3 and available for downstream processing by the remaining BDC platform teams prior to BDC-wide release	-	-	-	X	-	-
MONITOR-709	COPDGene Image Data (Phase I)	Backlog	"Backlog" - Data is in our queue but not ready for ingestion due to administrative or other actions required (e.g. - Institutional Certification, dhCaP Registration, etc.)	-	-	-	-	X	-
MONITOR-713	COPDGene Image Data (Phase II)	Backlog	"Backlog" - Data is in our queue but not ready for ingestion due to administrative or other actions required (e.g. - Institutional Certification, dhCaP Registration, etc.)	-	-	-	-	X	-
MONITOR-1029	TOPMed Freeze 9 - CARE_PACT	Backlog	This is not a complete list of TOPMed Freeze 9 studies, just a list of the studies we were alerted were ready for ingestion by the former TOPMed DCC. More will be added as they are ready.	-	-	X	X	-	-
MONITOR-1028	TOPMed Freeze 9 - CARE_BADGER	Backlog	This is not a complete list of TOPMed Freeze 9 studies, just a list of the studies we were alerted were ready for ingestion by the former TOPMed DCC. More will be added as they are ready.	-	-	X	X	-	-
MONITOR-1027	TOPMed Freeze 9 - GGAF	Backlog	This is not a complete list of TOPMed Freeze 9 studies, just a list of the studies we were alerted were ready for ingestion by the former TOPMed DCC. More will be added as they are ready.	-	-	X	X	-	-

MONITOR-798	CHS (Echo)	Backlog	This is not a complete list of TOPMed Freeze 9 studies, just a list of the studies we were alerted were ready for ingestion by the former TOPMed DCC. More will be added as they are ready.	-	-	-	-	X	-
MONITOR-797	HeartShare (Echo)	Backlog	This is not a complete list of TOPMed Freeze 9 studies, just a list of the studies we were alerted were ready for ingestion by the former TOPMed DCC. More will be added as they are ready.	-	-	-	-	X	-
MONITOR-796	INDIE (Echo)	Backlog	This is not a complete list of TOPMed Freeze 9 studies, just a list of the studies we were alerted were ready for ingestion by the former TOPMed DCC. More will be added as they are ready.	-	-	-	-	X	-
MONITOR-795	NEAT (Echo)	Backlog	This is not a complete list of TOPMed Freeze 9 studies, just a list of the studies we were alerted were ready for ingestion by the former TOPMed DCC. More will be added as they are ready.	-	-	-	-	X	-
MONITOR-794	RELAX (Echo)	Backlog	This is not a complete list of TOPMed Freeze 9 studies, just a list of the studies we were alerted were ready for ingestion by the former TOPMed DCC. More will be added as they are ready.	-	-	-	-	X	-
MONITOR-793	Framingham (Echo)	Backlog	This is not a complete list of TOPMed Freeze 9 studies, just a list of the studies we were alerted were ready for ingestion by the former TOPMed DCC. More will be added as they are ready.	-	-	-	-	X	-
MONITOR-792	CARDIA (Echo)	Backlog	This is not a complete list of TOPMed Freeze 9 studies, just a list of the studies we were alerted were ready for ingestion by the former TOPMed DCC. More will be added as they are ready.	-	-	-	-	X	-
MONITOR-791	SOL (Echo)	Backlog	This is not a complete list of TOPMed Freeze 9 studies, just a list of the studies we were alerted were ready for ingestion by the former TOPMed DCC. More will be added as they are ready.	-	-	-	-	X	-
MONITOR-790	ARIC (Echo)	Backlog	This is not a complete list of TOPMed Freeze 9 studies, just a list of the studies we were alerted were ready for ingestion by the former TOPMed DCC. More will be added as they are ready.	-	-	-	-	X	-
MONITOR-789	MESA (Echo)	Backlog	This is not a complete list of TOPMed Freeze 9 studies, just a list of the studies we were alerted were ready for ingestion by the former TOPMed DCC. More will be added as they are ready.	-	-	-	-	X	-
MONITOR-743	HeartShare	Backlog	This is not a complete list of TOPMed Freeze 9 studies, just a list of the studies we were alerted were ready for ingestion by the former TOPMed DCC. More will be added as they are ready.	-	-	X	X	X	X