
Data Storage, Toolspace, Access, and Analytics for biG-data Empowerment (DataSTAGE) Project Management Plan

V1.0 - 20190611

Data**STAGE**

DataSTAGE Project Management Plan

DRAFT - 20190412

Document Status

Version

V0.4

Approvals

Signatures presented below denote review and approval of the DataSTAGE Project Management Plan (PM Plan). These approvals are given based on the understanding that the PM Plan, and the information herein, will be revised at regular periods over the course of the program. It is the responsibility of the Principal Investigator of each funded team and select NHLBI program staff to sign in the indicated space below. Agreements listed on this page do not constitute an agreement to text linked outside of this document. PM approval constitutes proxy for their respective PI's.

Approved Date

N/A

PI Approvals:

PI	Team	Approval Date
Stefan Kuhn for Benedict Paten (+U/Chi & Broad PIs)	Calcium+	04/05/2019
Jessica Lyons for Paul Avillach	Carbon+	04/04/2019

Sandy Skipper for Ashok Krishnamurthy	Helium+	04/03/2019
Alison Leaf for Brandi Davis-Dusenbery	Xenon+	04/05/2019

NIH Approvals:

	NIH NHLBI Data STAGE Role	Approval Date
Jonathan Kaltman, MD	Program Manager	6/11/2019
Alastair Thomson, NHLBI CIO	Information Security	4/17/2019

Next Review Date

1 year from most recent approved date

Document Owner

DataSTAGE Coordinating Center

Revision History

Date (YYYYMMDD)	Version Number	Revision Reviewed / Approved By	Brief Description of Change
20190412	V0.4	Marcie Rathbun	Updated Stakeholder Management section, adding the link to a draft engagement plan; updated RFC section to include the Decision Log
20190405	V0.3	Team PM's/PI's	Element team consensus on text content (not all linked docs)
20190327	V0.2	Marcie Rathbun	Incorporated PM comments

20190306	V0.1	Marcie Rathbun	Incorporated NHLBI comments
20190222	V0	Jon Kaltman	Draft document submitted to NHLBI for review.

TABLE OF CONTENTS

INTRODUCTION	6
PURPOSE OF PROJECT MANAGEMENT PLAN	6
BACKGROUND INFORMATION	6
EXECUTIVE SUMMARY OF PROJECT CHARTER	6
ASSUMPTIONS AND CONSTRAINTS	7
Assumptions	7
Constraints	7
SCOPE MANAGEMENT	7
CONSORTIUM GOVERNANCE	8
Summary of Roles and Responsibilities	8
REQUEST FOR COMMENT (RFC) PROCESS	9
WORK PLAN MANAGEMENT	10
DEPLOYMENT AND TESTING PLAN	10
Initial Operationalization for Alpha and Beta Users	11
DECISION LOG	11
SCHEDULE MANAGEMENT	12
PROJECT SCHEDULE	12
PROJECT METRICS	13
FINANCIAL MANAGEMENT	13
QUALITY MANAGEMENT	13
QUALITY TABLE	13
RESOURCES MANAGEMENT	15
ORGANIZATIONAL CHART	15

STAKEHOLDER MANAGEMENT	15
DATASTAGE OPERATIONAL RESPONSIBILITY MATRIX	15
STAFFING	16
CLOUD RESOURCE MANAGEMENT	16
COMMUNICATIONS MANAGEMENT	16
MEETINGS TABLE	16
PUSH, PULL, AND INTERACTIVE WRITTEN COMMUNICATIONS TABLE	19
COMMUNICATION GUIDELINES	19
RISK MANAGEMENT	19
RISK REGISTER	19
COMPLIANCE-RELATED PLANNING	20
INSTITUTIONAL REVIEW BOARD	20
SYSTEMS SECURITY	21
APPENDIX A: REFERENCES	26
APPENDIX B: NOMENCLATURE	27
APPENDIX C: SUMMARY OF SPENDING	28

INTRODUCTION

PURPOSE OF PROJECT MANAGEMENT PLAN

The DataSTAGE Project Management Plan will provide a definition of the project, including the project's goals and objectives. Additionally, the plan will serve as an agreement between the Project Sponsor (NHLBI), Steering Committee, External Expert Panel, Element teams, Data Stewards, and other stakeholders associated with the project.

The intended audience of the DataSTAGE Project Management Plan is all DataSTAGE Consortium members.

The Project Management Plan defines the following:

- Project purpose
- Goals and objectives
- Scope and expectations
- Roles and responsibilities
- Assumptions and constraints
- Project management approach
- Ground rules for execution of the project
- Project timeline
- Conceptual design of the Platform solution

BACKGROUND INFORMATION

The DataSTAGE project was initiated in April 2018 as a customized implementation of the NIH Data Commons. The project has since shifted focus to become an independent, yet coordinated, development effort to leverage existing cloud systems to serve the NHLBI research community. This requires careful coordination to create a system of systems, henceforth referred to as the DataSTAGE Platform of Platforms, that can be extended to incorporate emerging solutions to address HLBS researcher's needs with an immediate focus on servicing the TOPMed research community.

To expedite the development of the Platform, the initial cyberinfrastructure will be based upon FireCloud and Seven Bridges with Interoperability Service Agreements (ISA) with the Data Commons Framework (DCF) Services of Gen3 to provide critical infrastructure, common security and the data gold master. The i2b2/transSMART platform will be the clinical data gold master database leveraging the PIC-SURE metaAPI. The functionality of the Platform will be extended through a number of options to integrate third-party applications.

EXECUTIVE SUMMARY OF PROJECT CHARTER

NHLBI has funded the creation of large, high-value data resources that are currently underused due to constraints in data identification, storage, computation, and access. NHLBI has tasked the

DataSTAGE project to develop a cloud-based platform for tools, applications, and workflows to provide secure workspaces to share, store, cross-link, and compute large sets of data generated from biomedical and behavioral research. DataSTAGE will be developed such that it can exchange information with other NIH-relevant cyberinfrastructure.

The project management approach described in this document aims to coordinate and align the efforts across DataSTAGE to execute, monitor, and control work towards deliverables within the program. While the individual software development teams apply an Agile approach to executing deliverables, the project management across the Consortium will be a hybrid approach of Waterfall and Agile to manage, monitor, and control the various team plans and outputs. The full DataSTAGE Charter can be found [here](#).

ASSUMPTIONS AND CONSTRAINTS

Assumptions

The following assumptions were made in preparation of the Project Management Plan:

- HLBS investigators are willing to learn a new system to advance their research if the proper functionality and support is provided
- The Consortium will work together to ensure that the needed resources are made available
- The teams will participate in Working Groups (WG) and Tiger Teams (TT) to ensure close coordination of related activities
- The Steering Committee and teams will participate in the timely execution of the Project Management Plan
- Failure to identify changes in team milestones within the timelines for the Features may result in project delays
- Element team members will adhere to the data access and communication plans
- DataSTAGE leadership will foster support and “buy-in” of project goals and objectives
- NHLBI, as the project sponsor, will ensure the approval of use for STRIDES cloud providers
- NHLBI, as the project sponsor, will identify a process to enable data access by the DataSTAGE team members and for research users
- All Consortium participants will abide by the guidelines within this Plan
- The Project Management Plan may change as new information and issues are revealed

Constraints

The following represent known project constraints:

- Project funding availability and limits
- Separate and distinct Other Transaction Agreement (OTA) awards to performers

SCOPE MANAGEMENT

The DataSTAGE project will introduce a new cyberinfrastructure that aligns with the Work Streams as defined in the Strategic Framework and Implementation Plan that includes the following:

- A cloud-hosted production platform that utilizes existing Authority to Operate (ATO) systems to:
 - Provide access to a suite of analysis tools needed by HLBS researchers;
 - Store and manage access to TOPMed data by an approved NIH process;
 - Import novel tools and datasets to complement the core resources; and
 - Support tools and processes for data management according to FAIR standards.
- A process for streamlining the current data access process for software developers as well as researchers for the TOPMed datasets.
- Activities to support user engagement throughout the development process, as well as training to use the Platform.

The mechanism of award, Other Transaction Agreement (OTA), provides a degree of flexibility in the scope of the work that is needed to advance this type of high risk/high reward project. Through the DataSTAGE governance processes, the Consortium will manage these changes to balance new opportunities and priorities with progress towards the vision of the DataSTAGE program.

CONSORTIUM GOVERNANCE

Governance is the process by which the Consortium makes collaborative decisions related to the activities, strategies, and direction of the work to realize the DataSTAGE vision.

DataSTAGE Consortium decision-making rests largely with the **Steering Committee (SC)**. The NHLBI Director, through the NHLBI Program Team (consisting of all of the NHLBI representation on the Steering Committee), retains final decision-making authority on strategic direction of the NHLBI DataSTAGE effort.

The **External Expert Panel (EEP)**, appointed by NHLBI, provides relevant advice to NHLBI and the Consortium.

Day-to-day development decisions rest with the development teams following the guidance of the SC and other relevant Consortium policies and procedures.

Detailed information regarding composition and functioning of governance activities can be found in the following governance documents:

- DataSTAGE Consortium Charter
- DataSTAGE Code of Conduct
- DataSTAGE Publications Guidelines
- DataSTAGE Working Group and Tiger Team Quick Start Guide

Summary of Roles and Responsibilities

	EEP	SC	Team Collaboration Meeting (TCM)	WGs/TTs	NHLBI
Members	NHLBI invitation	Internal and external stakeholders	Principal Investigators, technical staff, PMs, other	Self-identified Consortium members	NHLBI staff
Membership Type	Fixed	Fixed (for voting purposes)	Fluid	Fluid membership; cross-teamed	Fixed
Participation	Per invite	Open	Open	Open	Restricted
Scope	Consortium-wide	Consortium-wide	Surface team issues	Themed	Scoping expectations, stewardship of funds
Role	Advisory, particularly regarding conflicting priorities	Decisions	Reconcile work from WGs/TTs	Solution-focused	Fiscal and programmatic decisions
Responsibility	Advising	Direction-setting	Cross-WGs/TTs communication	Finding a path forward through tech solutions	Approval and feedback
Purpose	Advisory	Facilitation of directions setting and decisions	Information sharing and consensus building	Agreements on technical, collaborative approaches	Oversight
Meeting Notes	Yes	Yes	Yes Summary minutes presented to SC	Yes Summary minutes presented to TCM	N/A
Meeting Frequency	Bi-weekly (SC)	Bi-weekly	Bi-weekly	Periodical	N/A

REQUEST FOR COMMENT (RFC) PROCESS

The activities of the NHLBI DataSTAGE teams play a vital role in the NHLBI DataSTAGE, and may also be a critical part of the wider NIH data ecosystem. NHLBI DataSTAGE decisions are informed by feedback from individual contributors, the DataSTAGE teams, NHLBI stakeholders, DataSTAGE Data Stewards, and other members of DataSTAGE via the Requests

for Comments (RFCs) process. The RFC process helps to create, confirm, and communicate Consortium consensus on relevant topics.

Generally, an RFC should be created when:

- We seek to reach agreement.
- We seek to establish an agreement amongst the consortium.
- We are defining conventions, e.g., interfaces, APIs, data models, etc.
- There is a need for transparency and inclusion on a necessary decision.
- A decision could impact more than one system component, team, or stakeholders.
- We are adding dependencies that can affect more than one team.

For more information and guidelines, refer to the full version of the DataSTAGE RFC Implementation Process.

Decision Log

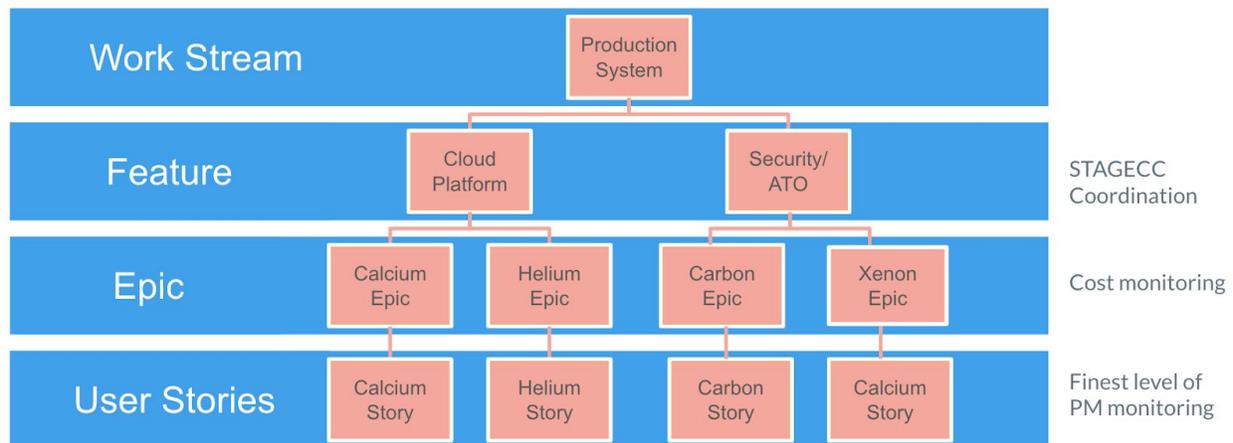
The Coordinating Center will assume ownership and manage any changes to Consortium-wide documents and processes that will include governance documents, planning documents, RFCs, SOPs, and WG and TT products by maintaining a Decision Log.

Planning documents, including the Strategic Framework Plan, Implementation Plan, and [this] Project Management Plan, will undergo periodic reviews by NHLBI, the Steering Committee, and the Coordinating Center to assess any high-level changes in scope and/or direction of the project. Proposed reviews will occur on an annual basis, per the Annual Strategic and Team Work Plan Cycle SOP. Once a document is approved, the document is locked for editing by the Coordinating Center and a PDF is placed on the password-protected website. Changes to these approved documents can be made as suggestions in the google drive documents themselves.

If the Stakeholder determines that a proposed change requires review for document amendment prior to the next formal review cycle, then they must request that it be added as an agenda item to the next Steering Committee meeting. At the SC meeting, the proposed change will be reviewed and the decision on acceptance or change to documentation will be captured in the Decision Log, further described here in STAGE-DRAFT-1.

Each OTA is responsible for managing and documenting any changes to Epics, User Stories, or subtasks within their own Work Plans. Preferably, this is done through versioning in their Work Plans and tracking Epics in Jira. Each team will utilize their software development or project management tool of choice for tracking User Story changes.

WORK PLAN MANAGEMENT



During the 2018-2019 phases of DataSTAGE, teams worked according to individual Work Plans that organized and aggregated activities differently. Beginning in the Spring of 2019, the Coordinating Center will work with the individual teams to align their activities to the overall Work Stream, Feature, Epic, and User Story framework. This alignment will take time and reflect a change in the project management across the Consortium. We will take a phased approach to this alignment, beginning with aligning the current Work Plans 2.0 activities to the Features through the use of labels in the current GitHub repositories. The Coordinating Center will work with the teams as they prepare Work Plans 3.0 to create a more direct alignment of work to the hierarchy and facilitate cost reporting by teams at the Epic level.

Work Plans 2.0:

- Calcium+ (Ca+) Work Plan
- Carbon+ (C+) Work Plan
- Helium+ (He+) Work Plan, Helium+ (He+) Work Plan Extended
- Xenon+ (Xe+) Work Plan

A breakdown of the User Narratives into Features and Epics, can be found in the STAGE User narratives - Decomposed . This document has been reviewed and approved by all PI’s for the June 2019 User Narratives.

DEPLOYMENT AND TESTING PLAN

The deployment and testing plans are being worked on through the Operationalization Tiger Team.

The working folder for that team is here:

https://drive.google.com/open?id=1li8D_-wS2q8zGcnVXUD16ttNRLlGey3j

DataSTAGE will be extended through the integration of third-party applications. There are a number of possible models in which a third-party application can operate within the DataSTAGE

Platform. The terms of operation for these applications are being developed collaboratively between the Tools and Applications Working Group and the Operationalization Tiger Team.

The Coordinating Center utilizes the *nhlbidatastage* GitHub for code sharing, with specific, existing open source projects also using GitHub repositories for source control. The GitHub application for Jira will allow syncing of activities to the DataSTAGE Jira instance.

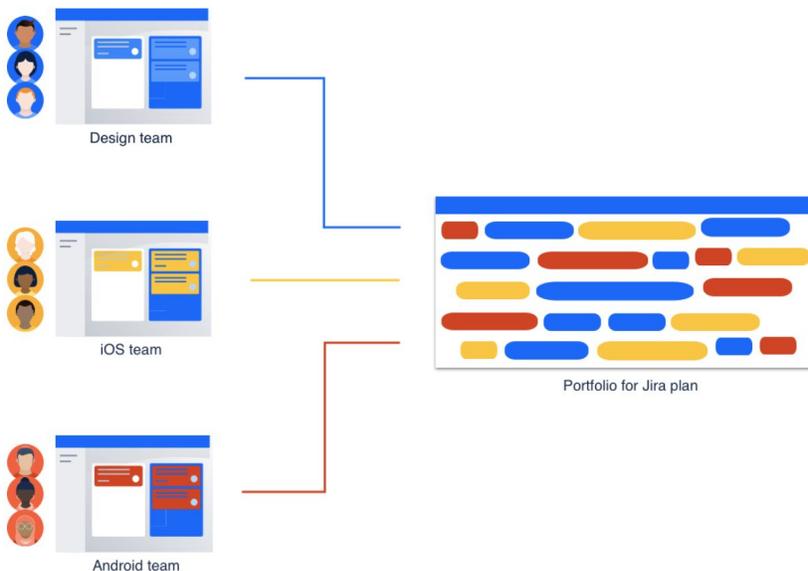
The Coordinating Center will facilitate the creation of a Testing SOP. The CC will coordinate with all teams to finalize this document.

Initial Operationalization for Alpha and Beta Users

The Operationalization Tiger Team has put together a Phase 1 plan to operationalize DataSTAGE in 60 days. This DRAFT plan can be found here.

SCHEDULE MANAGEMENT

Project schedules for DataSTAGE will be monitored using Portfolio for Jira starting with the Work Streams and Features identified in the project’s DataSTAGE User Narratives- Decomposed. Project schedules will be managed by each Element team and rolled-up to the Feature level for tracking and reporting purposes.



Activity sequencing will be used to determine the order of Features and Epics and assign relationships between project activities. In Jira, this will be implemented by identifying blockers between Jira issues (team Epics). Activity duration estimates will be used to calculate the number of work periods required to complete the Features. Element teams are responsible for identifying the resource estimates for each of their Epics. A PM Standards Document <<<Placeholder for link>> accompanied by a PM Dictionary <<<Placeholder for link>> will be

developed to standardize items such as resource estimation so that the Epics from the different teams can consistently be rolled up into cross-team Features.

If there are any scheduling delays that may impact the release of a Feature, they must be communicated amongst the PMs at the earliest possible time so that collaborative proactive measures may be taken to mitigate slips in delivery dates. The Coordinating Center will monitor Jira and coordinate any conversations that must occur to understand and/or accommodate the delivery date change. NHLBI will participate in periodic reviews of the schedule progress for progress and metric tracking.

PROJECT SCHEDULE

The overall project schedule will be managed by the coordinating center by utilizing Portfolio for Jira to enable cross-team collaboration that will address the priority-ranked backlog, release dates, and blockers. The CC will manage the cross-team Features in Jira as a release or collection of Jira issues (made up of element team Epics). The overall project schedule is fed by each element team's Epics that are rolled up into features (and releases). Each team is responsible for maintaining their Epics in the NHLBI demilitarized zone (DMZ) Jira Cloud instance. Element teams are highly encouraged to also maintain a more detailed project schedule or PM software release management tool to plan and track activities down to the user story level. Goals and deadlines will be centrally communicated and managed by the CC.

>>Placeholder: The CC will eventually add a link to the schedule/portfolio dashboard<<

PROJECT METRICS

The Standard Operating Procedure for Project Metrics.

FINANCIAL MANAGEMENT

NHLBI is responsible for compiling financial reports from the DataSTAGE Element teams. The Coordinating Center will facilitate this process by working with the teams to develop reporting templates (see example in Appendix C). The DataSTAGE teams will complete these templates and send them directly to NHLBI.

Each team is responsible for tracking their finances based upon the award conditions and for providing status updates as requested to NHLBI.

The goal for implementing this financial reporting template will be to align it with the Work Plan 3.0 structures.

QUALITY MANAGEMENT

Quality is paramount to the DataSTAGE project and having clear best practices for quality management will ensure that the project, and its individual products, are consistent. The leadership team commits to the development and implementation of this Project Management Plan to guide team leaders and evaluate progress toward DataSTAGE-specific aims and

milestones. The Quality Table below highlights the DataSTAGE approach to the four main components of quality management: quality planning, quality assurance, quality control, and quality improvement.

Governance and policy documentation, along with a collection of Standard Operating Procedures, have been developed to establish a framework for Quality Assurance (QA). These will be expanded over time to encompass additional SOPs. Part of the Coordinating Center focus will be to provide clear guidelines for adding QA processes and metrics to new SOPs. Furthermore, as Element teams develop Features through the realization of Epics and User Stories, the four main components of quality management will be documented in Work Plans and applied to solution development, testing, and production deployment.

While the applications of quality management practices will be an ongoing process for DataSTAGE and expand over time, some initial QA processes and activities can be identified.

QUALITY TABLE

QA Process	QA Activity	Responsible Person(s)	Frequency/Interval
Quality Planning	Inclusion in Strategic Framework Plan, Implementation Plan, Project Management Plan, and SOPs	Coordinating Center	As needed
	Inclusion of quality planning in Epic/User Story development	Element teams	As part of Agile development process
Quality Assurance	Continuous Integration (CI) testing of software components from Epics/User Stories	Element teams	Triggered as part of development process
	Validation of the accuracy of data/tools onboarded in the DataSTAGE system	Data Stewards, Element teams, and Working Groups (Data Access/UX-UI Working Group)	Triggered as part of data and tool onboarding
Quality Control	Periodic testing and validation of systems to ensure uptime and documented functionality	Coordinating Center, Element teams, Working Groups, and Tiger Teams (Tool and Apps Working Group, Operationalization Tiger Team)	Triggered on a schedule, such as nightly
Quality Improvement	Solicit feedback from researchers using the DataSTAGE Platform	Coordinating Center and Data Access/UX-UI Working Group	As part of a regular assessment process

RESOURCES MANAGEMENT

ORGANIZATIONAL CHART

National Heart, Lung, and Blood Institute

NHLBI provides global leadership for a research, training, and education program to promote the prevention and treatment of heart, lung, and blood disorders and enhance the health of all individuals so that they can live longer and more fulfilling lives.

Director Gary Gibbons, M.D.	CIO Alastair Thomson	Program Officer Jon Kaltman, M.D.
---------------------------------------	--------------------------------	---

Steering Committee

Responsible for decision-making and communication in DataSTAGE

Chairperson Ingrid Borecki STAGE PIs STAGECC PIs	User Community Representative Tasha Fingerlin NHLBI Working Group
---	--

External Expert Panel

An independent body of experts that informs and advises the work of the DataSTAGE Consortium

Donna Arnett David Mendelson	Mark Craven Jason Williams	Warrant Kibbe
---------------------------------	-------------------------------	---------------

 Ca⁺ Calcium	 C⁺ Carbon	 He⁺ Helium	 Xe⁺ Xenon	 DS Data Stewards	 O16N Operationalization Tiger Team	 UI/UX Data Access / UI-UX	 DH Data Harmonization	 T/A Tools & Applications				
Grossman, PI Paten, PI Philippakis, PI The Broad Institute University of Chicago University of Southern California Vanderbilt University Medical Center					Kohane, PI Avillach, PI Harvard Medical School		Krishnamurthy, PI Berkeley Lab Oregon State University RTI International University of New Mexico UNC-CH/RENCI		Davis-Dusenbery, PI Elsevier Reproductive Seven Bridges Genomics Veterans Affairs TOPMed COPDGene		Partnering with element teams on data accessibility and interoperability	
Co-Chairs Davis-Dusenbery Grossman					Co-Chairs Bradford Lyons		Chair Thessen		Co-Chairs Leaf Sloan			
C+, Ca+, He+, Xe+, CC, NHLBI					C+, Ca+, He+, Xe+, CC, NHLBI, SC, TOPMed		C+, Ca+, He+, Xe+, CC, NHLBI, SC, TOPMed		C+, Ca+, He+, Xe+, CC, NHLBI, TOPMed, COPD, TOPMed IRC			
Element Teams					Tiger Teams/Working Groups							

STAGECC
 Stan Ahalt, PI
 Rebecca Boyles, co-PI

Project management, coordination support, communications platform, project reporting, and standards for collaboration
 More information at <https://www.nhlbidatastage.org> and <http://bit.ly/nhlbiSTAGEsite>

STAKEHOLDER MANAGEMENT

All members of the DataSTAGE Consortium are project stakeholders. Consortium stakeholders are represented in the Steering Committee (SC). Each SC member is given one vote. Considerations of the stakeholders that are not Consortium members are brought to the Consortium by the EEP. Quarterly, stakeholders convene at in-person meetings, to which the invitation is open to all Consortium members.

The DataSTAGE Consortium Directory serves as the **Stakeholder Registry** contains names, contact information, team affiliation, roles, and other data for all Consortium members. It is maintained by the Coordinating Center on the password-protected website.

A Stakeholder Engagement Plan has been drafted to include the roles and engagement strategies for key stakeholders.

DATASTAGE OPERATIONAL RESPONSIBILITY MATRIX

The DataSTAGE Operational Responsibility Matrix is the RACI, developed by NHLBI, for the

key STAGE operational functions.

STAFFING

Each OTA is responsible for the management and staffing of their Element teams. The Coordinating Center will notify NHLBI of any changes in key personnel for any of the OTAs.

Project Managers are responsible for completing the Request a New Member Form to initiate the STAGECC to complete consortium member onboarding as well as submitting off-boarding requests for members leaving the teams. Details on this process can be found on the website under Resources: <https://nhlbidatastage.org/stage-resources/>.

CLOUD RESOURCE MANAGEMENT

NHLBI will need to weigh in on the construction of the cloud resources.

The DataSTAGE consortium will need to be able to:

- 1) Understand the available amount of cloud resources.
- 2) Track use to determine possible places for efficiencies.
- 3) Liaise with STRIDES

COMMUNICATIONS MANAGEMENT

Communications Management on DataSTAGE is paramount to the project's success. While each Element team is responsible for communications within their Element team, the Coordinating Center is primarily responsible for communications across the Consortium. This includes top-down communications that come from NHLBI and must be disseminated through the Steering Committee and down to each Element team, and as necessary, the EEP, Data Stewards, and other Consortium members. Cross-Consortium communications that the Coordinating Center facilitate also include coordinating with the EEP, SC, Data Stewards, and outreach (user community). It is also part of the Coordinating Center's mission for cross-Consortium collaboration to facilitate communication amongst Element teams and test users.

The Coordinating Center approach to communication is broken up into the two main modes of communication: written and verbal. Verbal communication, the most effective mode, is conducted across the Consortium by means of video conference and face-to-face meetings. A summary of the current meetings is listed in the Meetings Table below.

Written communications range from the informal media of Slack chats to the more formal Governance documents, RFCs, and other project deliverable documentation. The RFC process is of note since it provides a clear mechanism to establish contracts/agreements between systems and consensus on key technical decisions (see the Request for Comment (RFC) Process section). Written communications are further detailed in the Push, Pull, and Interactive Written Communications Table below.

MEETINGS TABLE

Communication Type	Description	Format	Frequency	Participants	Owner
Project Manager Monthly Meeting	PM updates on team progress, issues, and blockers	Zoom	First Monday of the month, 4-5pm ET	Coordinating Center, He+, Ca+, C+, and Xe+	Coordinating Center
NHLBI/CC Planning Meeting	Plan topics for the next SC meeting	Zoom	Bi-weekly, Thursdays, 2-3pm ET	Coordinating Center, NHLBI Project team, SC Chair	Coordinating Center
Steering Committee Meeting	Discuss progress, issues, and concerns	Zoom	Bi-weekly Thursdays, 10:30-11:30am ET	SC Chair, PIs, NHLBI representatives, user community, Coordinating Center	SC Chair
Team Collaboration Meeting	Discuss cross-element team related collaboration and concerns	Zoom	Bi-weekly Thursdays, 1-2pm ET	All Element teams and Coordinating Center members	Coordinating Center
Quarterly Face-to-Face (F2F) Meeting	Status overview, presentations, demos, collaborative efforts, etc.	In-person (location rotates, 2x/yr. in Bethesda, MD)	Quarterly (1-2 day sessions)	All Consortium	Coordinating Center
Operationalization Tiger Team	Develop overall framework, supporting plans, SOPs, and the minimum supporting documents to operate DataSTAGE	Zoom	Weekly Wednesdays, 11am-12pm ET	Coordinating Center, NHLBI, Xe+, Ca+, He+, C+	Co-Chairs: Brandi Davis-Dusenbery and Bob Grossman
Tools and Applications Working Group	Produce a suggested list of alpha & beta test users for SC approval; maintain the official list of 10-20 tools, workflows, or “apps” desired by the community; produce definitions to be approved by SC	Zoom	Bi-weekly Fridays, 12-1pm ET	Coordinating Center, NHLBI, Xe+, Ca+, C+, He+, Xe+, TOPMed Data Coordinating Center (DCC), TOPMed Informatics Research Center (IRC)	Co-Chairs: Cricket Sloan and Alison Leaf
Data Harmonization Working Group	Identify opportunities to support phenotype harmonization efforts and maximize efficiency across	Zoom	First Tuesday of the month, 12-1pm ET	Coordinating Center, NHLBI, He+, C+, Ca+, Xe+, TOPMed	Chair: Anne Thessen

	DataSTAGE. Collectively define requirements for search and analytics across TOPMed.				
Data Access/UX-UI Working Group	TBD	Zoom	TBD	Coordinating Center, NHLBI, C+, Ca+, He+, Xe+, TOPMed, SC	Co-Chairs: Kira Bradford and Jessica Lyons

PUSH, PULL, AND INTERACTIVE WRITTEN COMMUNICATIONS TABLE

Communication Type	Communication Vehicle	Examples	Additional notes
Push	Email	<ul style="list-style-type: none"> Meeting/due date reminders Call for document reviews (Governance docs, matrices, etc.) 	<ul style="list-style-type: none"> Email distributions are determined by the audience. Lists are formed using the Consortium Directory
Pull	DataSTAGE website	<ul style="list-style-type: none"> Calendar (hosts all relevant project- and Consortium-level meetings and events) Collaboration pages Resources page Reports Page Meeting materials Governance documents and SOPs 	<ul style="list-style-type: none"> Meeting notes Consortium members are added to the private website during the onboarding process
Interactive	<ul style="list-style-type: none"> Slack Google Drive RFCs 	<ul style="list-style-type: none"> Slack: channels and direct messaging Google Drive: working documents 	Consortium members are added to these shared platforms during the onboarding process

COMMUNICATION GUIDELINES

Additionally, the Coordinating Center has collected a set of DataSTAGE Communication Guidelines that overviews meeting best practices and etiquette as well as details on using the DataSTAGE website’s Consortium Directory for mailing lists.

RISK MANAGEMENT

The Coordinating Center will implement Risk Management through a methodical process by which the Element teams identify, score, and rank the various risks. Every effort will be made to proactively identify risks ahead of time in order to implement a mitigation strategy as early as possible. The most likely and highest impact risks will be added to the project schedule to ensure that the assigned Risk Owner take the necessary steps to implement the mitigation response at

the appropriate time during the schedule.

RISK REGISTER

The Coordinating Center will work with the Steering Committee to identify high-level Consortium-wide risks, as well as work with the individual team Project Managers to identify risks, which can relate to Epics, Consortium interaction, security concerns, etc. These risks will be documented and categorized in the Risk Register linked below.

Under the guidance of NHLBI, the Coordinating Center and the Steering Committee will meet to review the Risk Register and approve the risk response plans. The SC may advise on amendments needed to the Risk Register and the Element teams will update the risk plans, as necessary. Risks may be elevated to NHLBI for review based on impact probability and severity. Proposed informal reviews of the Risk Register will occur quarterly during the Face-to-face meetings and formal reviews of the Risk Register will occur on a bi-annual basis.

Note: Risk review and mitigation will be an ongoing and collaborative effort undertaken by the Element teams throughout the project lifecycle.

DataSTAGE Risk Register

COMPLIANCE-RELATED PLANNING

INSTITUTIONAL REVIEW BOARD

Process here will be superseded by the following Data Access section:

Some data are subject to restricted access and may require additional documentation, such as documentation of local IRB approval and/or letters of collaboration with the primary study PI(s).

Consortium members are responsible for obtaining relevant approvals from their respective Institutional Review Boards and submitting them to their dbGaP applications as needed. Consortium members are responsible for their compliance with the IRB requirements.

As of Feb 15, 2019, Consortium members listed below are covered by the smartIRB Reliance Determination, submitted and approved under the Harvard Medical School, and are eligible to access 29 datasets located in dbGaP, for the purpose described in the Research Statement of the dbGaP applications.

- Boston Children's Hospital
- Broad Institute
- Harvard Medical School and Harvard School of Dental Medicine
- The Jackson Laboratory
- University of California, Santa Cruz
- University of Chicago BSD IRB - Biological Sciences/Medical
- University of North Carolina at Chapel Hill

RTI has entered a separate reliance agreement with Harvard Medical School and have received approval.

The University of Washington and Seven Bridges have submitted their own separate IRB protocol applications and received IRB approval for this work.

DATA ACCESS

NHLBI is working on a solution for Data Access that will result in the amendment to each OTA. The document linked here is an example of Data Access Guidelines borrowed from the DCPPC: [DCPPC Guidelines for Appropriate Data Use for Tier 1 Data Access](#)

SYSTEMS SECURITY

The DataSTAGE Platform will be developed to utilize existing production systems that have attained an NIH Authority to Operate (ATO) that attests to the system's compliance with the NIST 800-53 standards for moderate security.

Of specific relevance to the DataSTAGE Platform planning are the below NIST 800-53 control families and associated controls.

Name	Control Description	ID
Access Control		
Access Control Policy and Procedures	Have policies and procedures around access control? Documented and up to date?	AC-1
Account Management	Define roles and access Someone responsible for overseeing accounts Notice events such as when users are added or need to be removed Moderate only: Automatically disable inactive accounts after a X days Automatically audit and notify when account related events occur, e.g., create/delete, enable/disable	AC-2
Access Enforcement	The system enforces rules around who can access what. Users cannot simply get around them or change them.	AC-3
Information Flow Enforcement	Moderate only: The system ensures that sensitive data only goes where it should	AC-4
Separation of Duties	Moderate only: Different people have different roles (everyone is not an admin)	AC-5
Least Privilege	Moderate only: People only have access to what they need; audit and review; special private accounts for private actions	AC-6
Unsuccessful Login Attempts	Enforce a limit on login failures	AC-7
System Use Notification	Show a banner before login	AC-8
Session Lock	Moderate only: Lock after X idle minutes and require re-auth; Conceal data while locked	AC-11
Session Termination	Moderate only: Auto-logout after Y minutes	AC-12
Permitted Actions without Identification or Authentication	Document the rationale for actions users can take without authenticating	AC-14
Remote Access/ Wireless Access/ Access Control for Mobile Devices	Authorize before allowing remote access Have restrictions and documentation Moderate only: Encryption Limit remote access points and privileged operations	AC-17/ 18/ 19

	Authenticate wireless	
Use of External Information Systems	<p>Before exchanging data with other systems, be aware of their security measures Have conditions for access from private devices</p> <p>Moderate only: ATO and Connection Agreements Limit use of portable storage</p>	AC-20
Information Sharing	<p>Moderate only: Provide users guidance on deciding whether specific data can be shared with a given user</p>	AC-21
Publicly Accessible Content	Control and review what is posted publicly; have an authorized role for it with training	AC-22
Identification and Authentication		
Identification and Authentication Policy and Procedures	<p>Have policies and procedures around identification and authentication Write them down and keep them up to date</p>	IA-1
Identification and Authentication (Organizational Users)	<p>Uniquely identify and authenticate internal users Multifactor for network access to privileged accounts Implement PIV if required.</p> <p>Moderate only: Multifactor for local access to private accounts and network access to non-private accounts Resist replay for network access to private accounts (can be TLS) Multifactor with separate device for remote access</p>	IA-2
Device Identification and Authentication	<p>Moderate only: Uniquely identify and authenticate devices Strength of authentication is up to us Only apply where truly needed</p>	IA-3
Identifier Management	<p>Usernames are unique and are not reused; examples are device IDs and role names</p>	IA-4
Authenticator Management	<p>Verify identity on issuing passwords, tokens, and PKI certs Select hardware carefully and update if needed Password rules</p> <p>Moderate only: For PKI-based auth, follow chain to trusted root, have a local store of revocation data, etc. Define policies for distributing authenticators either in person or by a trusted third party</p>	IA-5
Authenticator Feedback	Display asterisks for passwords	IA-6
Cryptographic Module Authentication	<p>Use strong encryption An approved module implementing an approved algorithm (usually AES; at least SHA-256 for hashes)</p>	IA-7

	https://csrc.nist.gov/projects/cryptographic-module-validation-program	
Identification and Authentication (Non-Organizational Users)	Uniquely identify and authenticate external users Comply with Federal Identity, Credential, and Access Management Architecture (FICAM) Implement PIV, if required	IA-8
Audit and Accountability		
Audit and Accountability Policy and Procedures	Have a policy and procedures around audit and accountability Write them down and keep them up to date	AU-1
Audit Events	Have a list the auditable events with justification Moderate only: Review and update every X months	AU-2
Content of Audit Records	Log type, time, location, source, outcome, and users Moderate only: Log additional data points	AU-3
Audit Storage Capacity	Ensure sufficient storage for logs	AU-4
Response to Audit Processing Failures	Alert someone if logging fails Decide what else to do and implement it	AU-5
Audit Review, Analysis, and Reporting	Review/analyze and report anomalies Moderate only: Automate analysis and correlate across audit repositories	AU-6
Audit Reduction and Report Generation	Moderate only: Create the capability to summarize logs on demand for investigation of incidents without altering content or sequencing	AU-7
Time Stamps	Include system time stamps Moderate only: Synchronize with a time authority	AU-8
Protection of Audit Information	Prevent logs from being altered Moderate only: Only a subset of privileged users can manage audit functionality	AU-9
Audit Record Retention	Keep logs for X days	AU-11
Audit Generation	Implement the logging as claimed in AU-2	AU-12
Additional Controls		
Developer Configuration Management	Moderate only: Use configuration management Get approval for all document changes Track security flaws and fixes	SA-10
Developer Security Testing and Evaluation	Moderate only: Create and implement a security assessment plan, including testing, remediation, and recording of evidence Recommend reading	SA-11

Malicious Code Protection	Guard against malicious code at entry and exit points. Perform scans with current tools. Moderate only: Manage centrally and auto-update	SI-3
---------------------------	---	------

APPENDIX A: REFERENCES

The following table summarizes the documents referenced in this document.

Document Name/Version	Description	PM Plan Section
Strategic Framework Plan	Identifies the mission and vision of the DataSTAGE program and describes how the program will align across stakeholders to execute on common goals and how that performance will be measured.	Scope Management
Implementation Plan	Describes the process by which the DataSTAGE Consortium will incrementally progress towards the vision of the program described in the DataSTAGE Strategic Framework Plan.	Scope Management
DataSTAGE Consortium Charter	The charge for the entire DataSTAGE consortium aligning the visions and goals for multiple OTAs.	Introduction - Executive Summary of Project Charter; Consortium Governance
DataSTAGE Code of Conduct	Summary of member expectations for a safe environment.	Consortium Governance
DataSTAGE Publications Guidelines	Set of requirements for publishing content related to DataSTAGE.	Consortium Governance
DataSTAGE Working Group Quick Start Guide	A set of parameters for initiating a DataSTAGE Working Group or Tiger Team.	Consortium Governance
Proposed RFC Implementation Process	Outlines the proposed steps for cross-consortium agreement documentation.	Request For Comment (RFC) Process
Calcium+ Work Plan	Work Plan 2.0	Work Plans
Carbon+ Work Plan	Work Plan 2.0	Work Plans
Helium+ Work Plan	Work Plan 2.0	Work Plans
Xenon+ Work Plan	Work Plan 2.0	Work Plans
STAGE User narratives - Decomposed	Detailed breakdown of user narratives, features, and epics.	Work Plans; Schedule Management

NHLBI DataSTAGE 60 Day Operationalization Phase 1 Plan	A plan to operate DataSTAGE Phase 1 in the next 60 days that includes incident response plans, a concept of operations, whitelist, training, and a rough timeline.	Deployment Plan - Initial Operationalization for Alpha and Beta Users
Portfolio for Jira	Atlassian Portfolio for Jira product overview	Schedule Management
Governance and Policy Documentation [Folder]	Folder containing DataSTAGE Governance & Policy Documents.	Quality Management
SOPs [Folder]	Folder containing a collection of internal SOPs utilized by the Coordinating Center.	Quality Management
DataSTAGE Operational Responsibility Matrix	DataSTAGE RACI matrix.	Stakeholder Management
Consortium Directory	DataSTAGE website hosted directory containing consortium contact information.	Stakeholder Management
Coordinating Center Calendar	DataSTAGE website hosted calendar containing consortium meeting details and events.	Communications Management
Coordinating Center Communication Guidelines	Provides communication guidance on email distribution, meeting access, and meeting conduct.	Communications Management
Coordinating Center Risk Register	Tool for documenting, identifying, and managing risks.	Risk Management
DCPPC Guidelines for Appropriate Data Use for Tier 1 Data Access	Provides guidance to the NIH Data Commons Pilot Phase Consortium (DCPPC) investigators for the proper use of the controlled access data of GTEx and selected TOPMed and NHLBI studies.	Compliance Related Planning

APPENDIX B: NOMENCLATURE

The Glossary provides definitions and examples for terms relevant to this document.



APPENDIX C: SUMMARY OF SPENDING

Budget Item	PY-1	PY	CY	BY	BY + 1	BY + 2	BY + 3	BY + 4	Total
Planning:									
Budgetary Resources					\$ -	\$ -	\$ -	\$ -	\$0.00
Outlays					\$ -	\$ -	\$ -	\$ -	\$0.00
Development & Implementation of Project:									
Budgetary Resources					\$ -	\$ -	\$ -	\$ -	\$0.00
Outlays					\$ -	\$ -	\$ -	\$ -	\$0.00
Total, sum of stages:									
Budgetary Resources					\$ -	\$ -	\$ -	\$ -	\$0.00
Outlays					\$ -	\$ -	\$ -	\$ -	\$0.00
Operations & Maintenance:									
Budgetary Resources					\$ -	\$ -	\$ -	\$ -	\$0.00
Outlays					\$ -	\$ -	\$ -	\$ -	\$0.00
Total, all stages:									
Budgetary Resources					\$ -	\$ -	\$ -	\$ -	\$0.00
Outlays					\$ -	\$ -	\$ -	\$ -	\$0.00
Government FTE cost					\$ -	\$ -	\$ -	\$ -	\$0.00

PY: Previous Year; CY: Current Year; BY: Budget Year