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:

<table>
<thead>
<tr>
<th>PI</th>
<th>Team</th>
<th>Approval Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stefan Kuhn for Benedict Paten (+U/Chi &amp; Broad PIs)</td>
<td>Calcium+</td>
<td>04/05/2019</td>
</tr>
<tr>
<td>Jessica Lyons for Paul Avillach</td>
<td>Carbon+</td>
<td>04/04/2019</td>
</tr>
</tbody>
</table>
NIH Approvals:

<table>
<thead>
<tr>
<th>NIH NHLBI Data STAGE Role</th>
<th>Approval Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jonathan Kaltman, MD</td>
<td>6/11/2019</td>
</tr>
<tr>
<td>Alastair Thomson, NHLBI CIO</td>
<td>4/17/2019</td>
</tr>
</tbody>
</table>

Next Review Date

1 year from most recent approved date

Document Owner

DataSTAGE Coordinating Center

Revision History

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

Sandy Skipper for Ashok Krishnamurthy
Helium+ 04/03/2019
Alison Leaf for Brandi Davis-Dusenbery
Xenon+ 04/05/2019
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
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>STAKEHOLDER MANAGEMENT</td>
<td>15</td>
</tr>
<tr>
<td>DATASTAGE OPERATIONAL RESPONSIBILITY MATRIX</td>
<td>15</td>
</tr>
<tr>
<td>STAFFING</td>
<td>16</td>
</tr>
<tr>
<td>CLOUD RESOURCE MANAGEMENT</td>
<td>16</td>
</tr>
<tr>
<td>COMMUNICATIONS MANAGEMENT</td>
<td>16</td>
</tr>
<tr>
<td>MEETINGS TABLE</td>
<td>16</td>
</tr>
<tr>
<td>PUSH, PULL, AND INTERACTIVE WRITTEN COMMUNICATIONS TABLE</td>
<td>19</td>
</tr>
<tr>
<td>COMMUNICATION GUIDELINES</td>
<td>19</td>
</tr>
<tr>
<td>RISK MANAGEMENT</td>
<td>19</td>
</tr>
<tr>
<td>RISK REGISTER</td>
<td>19</td>
</tr>
<tr>
<td>COMPLIANCE-RELATED PLANNING</td>
<td>20</td>
</tr>
<tr>
<td>INSTITUTIONAL REVIEW BOARD</td>
<td>20</td>
</tr>
<tr>
<td>SYSTEMS SECURITY</td>
<td>21</td>
</tr>
<tr>
<td>APPENDIX A: REFERENCES</td>
<td>26</td>
</tr>
<tr>
<td>APPENDIX B: NOMENCLATURE</td>
<td>27</td>
</tr>
<tr>
<td>APPENDIX C: SUMMARY OF SPENDING</td>
<td>28</td>
</tr>
</tbody>
</table>
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/tranSMART 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

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

## 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 nhlbidastage 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**

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

ORGANIZATIONAL CHART

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

<table>
<thead>
<tr>
<th>Communication Type</th>
<th>Description</th>
<th>Format</th>
<th>Frequency</th>
<th>Participants</th>
<th>Owner</th>
</tr>
</thead>
<tbody>
<tr>
<td>Project Manager Monthly Meeting</td>
<td>PM updates on team progress, issues, and blockers</td>
<td>Zoom</td>
<td>First Monday of the month, 4-5pm ET</td>
<td>Coordinating Center, He+, Ca+, C+, and Xe+</td>
<td>Coordinating Center</td>
</tr>
<tr>
<td>NHLBI/CC Planning Meeting</td>
<td>Plan topics for the next SC meeting</td>
<td>Zoom</td>
<td>Bi-weekly, Thursdays, 2-3pm ET</td>
<td>Coordinating Center, NHLBI Project team, SC Chair</td>
<td>Coordinating Center</td>
</tr>
<tr>
<td>Steering Committee Meeting</td>
<td>Discuss progress, issues, and concerns</td>
<td>Zoom</td>
<td>Bi-weekly Thursdays, 10:30-11:30am ET</td>
<td>SC Chair, PIs, NHLBI representatives, user community, Coordinating Center</td>
<td>SC Chair</td>
</tr>
<tr>
<td>Team Collaboration Meeting</td>
<td>Discuss cross-element team related collaboration and concerns</td>
<td>Zoom</td>
<td>Bi-weekly Thursdays, 1-2pm ET</td>
<td>All Element teams and Coordinating Center</td>
<td>Coordinating Center</td>
</tr>
<tr>
<td>Quarterly Face-to-Face (F2F)</td>
<td>Status overview, presentations, demos, collaborative efforts, etc.</td>
<td>In-person</td>
<td>Quarterly (1-2 day sessions)</td>
<td>All Consortium</td>
<td>Coordinating Center</td>
</tr>
<tr>
<td>Operationalization Tiger Team</td>
<td>Develop overall framework, supporting plans, SOPs, and the minimum supporting documents to operate DataSTAGE</td>
<td>Zoom</td>
<td>Weekly Wednesdays, 11am-12pm ET</td>
<td>Coordinating Center, NHLBI, Xe+, Ca+, He+, C+</td>
<td>Co-Chairs: Brandi Davis-Dusenbery and Bob Grossman</td>
</tr>
<tr>
<td>Tools and Applications Working Group</td>
<td>Produce a suggested list of alpha &amp; 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</td>
<td>Zoom</td>
<td>Bi-weekly Fridays, 12-1pm ET</td>
<td>Coordinating Center, NHLBI, Xe+, Ca+, C+, He+, Xe+, TOPMed Data Coordinating Center (DCC), TOPMed Informatics Research Center (IRC)</td>
<td>Co-Chairs: Cricket Sloan and Alison Leaf</td>
</tr>
<tr>
<td>Data Harmonization Working Group</td>
<td>Identify opportunities to support phenotype harmonization efforts and maximize efficiency across</td>
<td>Zoom</td>
<td>First Tuesday of the month, 12-1pm ET</td>
<td>Coordinating Center, NHLBI, He+, C+, Ca+, Xe+, TOPMed</td>
<td>Chair: Anne Thessen</td>
</tr>
</tbody>
</table>
PUSH, PULL, AND INTERACTIVE WRITTEN COMMUNICATIONS TABLE

<table>
<thead>
<tr>
<th>Communication Type</th>
<th>Communication Vehicle</th>
<th>Examples</th>
<th>Additional notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Push</td>
<td>Email</td>
<td>● Meeting/due date reminders</td>
<td>● Email distributions are determined by the audience.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>● Call for document reviews (Governance docs, matrices, etc.)</td>
<td>● Lists are formed using the Consortium Directory</td>
</tr>
<tr>
<td>Pull</td>
<td>DataSTAGE website</td>
<td>● Calendar (hosts all relevant project- and Consortium-level meetings and events)</td>
<td>● Meeting notes</td>
</tr>
<tr>
<td></td>
<td></td>
<td>● Collaboration pages</td>
<td>● Consortium members are added to the private website during the onboarding process</td>
</tr>
<tr>
<td></td>
<td></td>
<td>● Resources page</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>● Reports Page</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>● Meeting materials</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>● Governance documents and SOPs</td>
<td></td>
</tr>
<tr>
<td>Interactive</td>
<td>Slack, Google Drive, RFCs</td>
<td>● Slack: channels and direct messaging</td>
<td>Consortium members are added to these shared platforms during the onboarding process</td>
</tr>
<tr>
<td></td>
<td></td>
<td>● Google Drive: working documents</td>
<td></td>
</tr>
</tbody>
</table>

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.
<table>
<thead>
<tr>
<th>Name</th>
<th>Control Description</th>
<th>ID</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Access Control</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Access Control Policy and Procedures</td>
<td>Have policies and procedures around access control? Documented and up to date?</td>
<td>AC-1</td>
</tr>
<tr>
<td>Account Management</td>
<td>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</td>
<td>AC-2</td>
</tr>
<tr>
<td>Access Enforcement</td>
<td>The system enforces rules around who can access what. Users cannot simply get around them or change them.</td>
<td>AC-3</td>
</tr>
<tr>
<td>Information Flow Enforcement</td>
<td>Moderate only: The system ensures that sensitive data only goes where it should</td>
<td>AC-4</td>
</tr>
<tr>
<td>Separation of Duties</td>
<td>Moderate only: Different people have different roles (everyone is not an admin)</td>
<td>AC-5</td>
</tr>
<tr>
<td>Least Privilege</td>
<td>Moderate only: People only have access to what they need; audit and review; special private accounts for private actions</td>
<td>AC-6</td>
</tr>
<tr>
<td>Unsuccessful Login Attempts</td>
<td>Enforce a limit on login failures</td>
<td>AC-7</td>
</tr>
<tr>
<td>System Use Notification</td>
<td>Show a banner before login</td>
<td>AC-8</td>
</tr>
<tr>
<td>Session Lock</td>
<td>Moderate only: Lock after X idle minutes and require re-auth; Conceal data while locked</td>
<td>AC-11</td>
</tr>
<tr>
<td>Session Termination</td>
<td>Moderate only: Auto-logout after Y minutes</td>
<td>AC-12</td>
</tr>
<tr>
<td>Permitted Actions without Identification or Authentication</td>
<td>Document the rationale for actions users can take without authenticating</td>
<td>AC-14</td>
</tr>
<tr>
<td>Remote Access/ Wireless Access/ Access Control for Mobile Devices</td>
<td>Authorize before allowing remote access Have restrictions and documentation Moderate only: Encryption Limit remote access points and privileged operations</td>
<td>AC-17/18/19</td>
</tr>
</tbody>
</table>
## Use of External Information Systems

Before exchanging data with other systems, be aware of their security measures.

- Have conditions for access from private devices
- Moderate only: ATO and Connection Agreements
- Limit use of portable storage

| AC-20 |

## Information Sharing

Moderate only:

- Provide users guidance on deciding whether specific data can be shared with a given user

| 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 Policy and Procedures

Have policies and procedures around identification and authentication. Write them down and keep them up to date

| IA-1 |

### Identification and Authentication (Organizational Users)

- Uniquely identify and authenticate internal users
- Multifactor for network access to privileged accounts
- Implement PIV if required.

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

| IA-2 |

### Device Identification and Authentication

Moderate only:

- Uniquely identify and authenticate devices
- Strength of authentication is up to us
- Only apply where truly needed

| IA-3 |

### Identifier Management

Usernames are unique and are not reused; examples are device IDs and role names

| IA-4 |

### Authenticator Management

Verify identity on issuing passwords, tokens, and PKI certs

- Select hardware carefully and update if needed
- Password rules

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

| IA-5 |

### Authenticator Feedback

Display asterisks for passwords

| IA-6 |

### Cryptographic Module Authentication

Use strong encryption

- An approved module implementing an approved algorithm (usually AES, at least SHA-256 for hashes)

| IA-7 |
### 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

### Audit and Accountability

<table>
<thead>
<tr>
<th>Activity</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Audit Accountability Policy and Procedures</td>
<td>Have a policy and procedures around audit and accountability Write them down and keep them up to date</td>
</tr>
<tr>
<td>Audit Events</td>
<td>Have a list the auditable events with justification Moderate only: Review and update every X months</td>
</tr>
<tr>
<td>Content of Audit Records</td>
<td>Log type, time, location, source, outcome, and users Moderate only: Log additional data points</td>
</tr>
<tr>
<td>Audit Storage Capacity</td>
<td>Ensure sufficient storage for logs</td>
</tr>
<tr>
<td>Response to Audit Processing Failures</td>
<td>Alert someone if logging fails Decide what else to do and implement it</td>
</tr>
<tr>
<td>Audit Review, Analysis, and Reporting</td>
<td>Review/analyze and report anomalies Moderate only: Automate analysis and correlate across audit repositories</td>
</tr>
<tr>
<td>Audit Reduction and Report Generation</td>
<td>Moderate only: Create the capability to summarize logs on demand for investigation of incidents without altering content or sequencing</td>
</tr>
<tr>
<td>Time Stamps</td>
<td>Include system time stamps Moderate only: Synchronize with a time authority</td>
</tr>
<tr>
<td>Protection of Audit Information</td>
<td>Prevent logs from being altered Moderate only: Only a subset of privileged users can manage audit functionality</td>
</tr>
<tr>
<td>Audit Record Retention</td>
<td>Keep logs for X days</td>
</tr>
<tr>
<td>Audit Generation</td>
<td>Implement the logging as claimed in AU-2</td>
</tr>
</tbody>
</table>

### Additional Controls

<table>
<thead>
<tr>
<th>Activity</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Developer Configuration Management</td>
<td>Moderate only: Use configuration management Get approval for all document changes Track security flaws and fixes</td>
</tr>
<tr>
<td>Developer Security Testing and Evaluation</td>
<td>Moderate only: Create and implement a security assessment plan, including testing, remediation, and recording of evidence Recommend reading</td>
</tr>
<tr>
<td>Malicious Code Protection</td>
<td>Guard against malicious code at entry and exit points. Perform scans with current tools. Moderate only: Manage centrally and auto-update</td>
</tr>
</tbody>
</table>
## APPENDIX A: REFERENCES

The following table summarizes the documents referenced in this document.

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

<table>
<thead>
<tr>
<th>Portfolio for Jira</th>
<th>Atlassian Portfolio for Jira product overview</th>
<th>Schedule Management</th>
</tr>
</thead>
</table>

|---------------------------------------------|---------------------------------------------------------------|-------------------|

<table>
<thead>
<tr>
<th>SOPs [Folder]</th>
<th>Folder containing a collection of internal SOPs utilized by the Coordinating Center.</th>
<th>Quality Management</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>DataSTAGE Operational Responsibility Matrix</th>
<th>DataSTAGE RACI matrix.</th>
<th>Stakeholder Management</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Consortium Directory</th>
<th>DataSTAGE website hosted directory containing consortium contact information.</th>
<th>Stakeholder Management</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Coordinating Center Calendar</th>
<th>DataSTAGE website hosted calendar containing consortium meeting details and events.</th>
<th>Communications Management</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Coordinating Center Communication Guidelines</th>
<th>Provides communication guidance on email distribution, meeting access, and meeting conduct.</th>
<th>Communications Management</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Coordinating Center Risk Register</th>
<th>Tool for documenting, identifying, and managing risks.</th>
<th>Risk Management</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>DCPCC Guidelines for Appropriate Data Use for Tier 1 Data Access</th>
<th>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.</th>
<th>Compliance Related Planning</th>
</tr>
</thead>
</table>

APPENDIX B: NOMENCLATURE

The Glossary provides definitions and examples for terms relevant to this document.
# APPENDIX C: SUMMARY OF SPENDING

<table>
<thead>
<tr>
<th>Budget Item</th>
<th>PY-1</th>
<th>PY</th>
<th>CY</th>
<th>BY</th>
<th>BY + 1</th>
<th>BY + 2</th>
<th>BY + 3</th>
<th>BY + 4</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Planning:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Budgetary Resources</td>
<td>$</td>
<td>$</td>
<td>$</td>
<td>$</td>
<td>$</td>
<td></td>
<td></td>
<td>$</td>
<td>$0.00</td>
</tr>
<tr>
<td>Outlays</td>
<td>$</td>
<td>$</td>
<td>$</td>
<td>$</td>
<td>$</td>
<td></td>
<td></td>
<td>$</td>
<td>$0.00</td>
</tr>
<tr>
<td><strong>Development &amp; Implementation of Project:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Budgetary Resources</td>
<td>$</td>
<td>$</td>
<td>$</td>
<td>$</td>
<td>$</td>
<td></td>
<td></td>
<td>$</td>
<td>$0.00</td>
</tr>
<tr>
<td>Outlays</td>
<td>$</td>
<td>$</td>
<td>$</td>
<td>$</td>
<td>$</td>
<td></td>
<td></td>
<td>$</td>
<td>$0.00</td>
</tr>
<tr>
<td><strong>Total, sum of stages:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Budgetary Resources</td>
<td>$</td>
<td>$</td>
<td>$</td>
<td>$</td>
<td>$</td>
<td></td>
<td></td>
<td>$</td>
<td>$0.00</td>
</tr>
<tr>
<td>Outlays</td>
<td>$</td>
<td>$</td>
<td>$</td>
<td>$</td>
<td>$</td>
<td></td>
<td></td>
<td>$</td>
<td>$0.00</td>
</tr>
<tr>
<td><strong>Operations &amp; Maintenance:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Budgetary Resources</td>
<td>$</td>
<td>$</td>
<td>$</td>
<td>$</td>
<td>$</td>
<td></td>
<td></td>
<td>$</td>
<td>$0.00</td>
</tr>
<tr>
<td>Outlays</td>
<td>$</td>
<td>$</td>
<td>$</td>
<td>$</td>
<td>$</td>
<td></td>
<td></td>
<td>$</td>
<td>$0.00</td>
</tr>
<tr>
<td><strong>Total, all stages:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Budgetary Resources</td>
<td>$</td>
<td>$</td>
<td>$</td>
<td>$</td>
<td>$</td>
<td></td>
<td></td>
<td>$</td>
<td>$0.00</td>
</tr>
<tr>
<td>Outlays</td>
<td>$</td>
<td>$</td>
<td>$</td>
<td>$</td>
<td>$</td>
<td></td>
<td></td>
<td>$</td>
<td>$0.00</td>
</tr>
</tbody>
</table>

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