BioData Catalyst Project Management Plan

V2.0 - 20210104

Document Status

Version

V2.0

Approvals

Signatures presented below denote review and approval of the BioData Catalyst 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.

PI Approvals (V1.0):

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 BioData Catalyst 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

BioData Catalyst Coordinating Center

Contact: Tom Madden (tmadden@rti.org), BDC3 (bdc3@renci.org)

Revision History

Date (YYYYMMDD)	Version Number	Revision Reviewed / Approved By	Brief Description of Change
20201230	V2.0	Tom Madden (BDC3); Team PM's	General text and file/link updates; process updates; WG/TT updates; additions of WP3.1; CCB; Risk; Data Management; Testing; External Communications; Cloud and Metrics;
20190611	V1.0	Jon Kaltman	NHLBI formal sign-off/approval of V1.0
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

PURPOSE OF PROJECT MANAGEMENT PLAN

The BioData Catalyst Project Management Plan will provide a definition of the project, including the project's goals and objectives, and act as a resource for hosting and linking out to more specific areas of practice and scope. 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 BioData Catalyst Project Management Plan is all BioData Catalyst 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 ecosystem solution

BACKGROUND INFORMATION

The BioData Catalyst (BDCatalyst) 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 BioData Catalyst Ecosystem, 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 BDCatalyst, 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 BioData Catalyst project to develop a cloud-based ecosystem 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. BioData Catalyst 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 BioData Catalyst 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 BioData Catalyst 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
- BioData Catalyst 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 BioData Catalyst team members and for research users
- All teams will use Jira for consortium-wide Work Plan management, monitoring and control; regularly updating and providing visibility into project (Feature/Epic) status
- All Consortium participants will abide by the guidelines within this Plan

- The Project Management Plan will 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 BioData Catalyst 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 ecosystem 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 data by an approved NIH process;
        - *Data as identified and prioritized by the DRMWG
    - 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 datasets.
- Activities to support user engagement throughout the development process, as well as training to use the ecosystem.

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 BioData Catalyst governance processes, the Consortium will manage these changes to balance new opportunities and priorities with progress towards the vision of the BioData Catalyst 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 BioData Catalyst vision.

BioData Catalyst 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 BioData Catalyst effort.

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

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

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

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

Summary of Roles and Responsibilities

	External Expert Panel (EEP)	Steering Committee (SC)	Change Control Board (CCB)	Team Collaboration Meeting (TCM)	Working Groups/Tiger Teams	NHLBI	User Services
Members	NHLBI invitation	Internal and external stakeholders	NHLBI, Principal Investigator,s and other identified consortium participants	Principal Investigators, technical staff, PMs, other	Self-identified Consortium members	NHLBI staff	BDC3
Membership Type	Fixed	Fixed (for voting purposes)	Fixed (for voting purposes)	Fluid	Fluid membership; cross-teamed	Fixed	Fixed
Participation	Per invite	Open	Per invite	Open	Open	Restricted	Restricted
Scope	Consortium-wide	Consortium-wide	Consortium-wide	Surface team issues	Themed	Scoping expectations, stewardship of funds	Consortium-wide with User Service focus (outreach, engagement, coordination, etc.)
Role	Advisory, particularly regarding conflicting priorities	Decisions	Consortium Decisions and Changes	Reconcile work from WGs/TTs	Solution-focused	Fiscal and programmatic decisions	Coordination, advisory, and solution-focused
Responsibility	Advising	Direction-setting	Approval and feedback	Cross-WGs/TTs communication	Finding a path forward through tech solutions	Approval and feedback	Coordination, advising, and user facing solutions
Purpose	Advisory	Facilitation of directions setting and decisions	Make decisions for proposed changes to the BDCatalyst Ecosystem	Information sharing and consensus building	Agreements on technical, collaborative approaches	Oversight	Coordination, advising, and user facing solutions

Meeting Notes	Yes	Yes	Yes; Jira framework	Yes Summary report to SC	Yes Summary report to TCM & SC meetings	N/A	N/A
Meeting Frequency	Bi-weekly (SC)	Bi-weekly	Scheduled Bi-weekly (on-demand occurrences)	Bi-weekly	Periodical	N/A	Weekly (internal); Consortium on-demand

CHANGE MANAGEMENT

A Change Control Board (CCB) was formed in April 2020 and formally approved by the Steering Committee on May 28, 2020. The scope of the CCB is to make decisions for proposed changes to the BDCatalyst Ecosystem which includes data, software or applications which impacts more than a single team within the Ecosystem or the scope, schedule, or budget of a single team.

The CCB Co-chairs, together with the BDCatalyst Coordinating Center (BDC3), have the authority and responsibility to direct the BDCatalyst change management process. As such, the BDC3 has ultimate oversight over the Change Management process.

The BioData Catalyst Coordinating Center will provide PM support to create, update, and maintain the Change Management artifacts and Request for Change Kanban Board in Jira. Specifically, the CCB Co-Chairs, in consultation with the CCB, prioritize change requests, and submit 1 vote (combined) on each Request for Change.

The approved CCB Charter can be found here, while a detailed Procedure for BioData Catalyst Ecosystem Configuration and Change Management SOP is available to further outline and detail the CCB purpose, process and associated workflows.

REQUEST FOR COMMENT (RFC) PROCESS

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

WORK PLAN MANAGEMENT

SCHEDULE MANAGEMENT

A project roadmap for BioData Catalyst will be monitored using Portfolio for Jira, along with a schedule for quarterly releases.

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 releases required to complete the Features. Element teams are responsible for identifying and updating a numerical value for the "User Stories" for each of their Epics.

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. Please see the WP3.1 section for additional details

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. Feature "Responsible Persons" have been assigned to guide Feature progress for WP3.1. 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 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.

Jira:

Jira is to be utilized by the BDCatalyst Consortium to:

- Monitor project Risks
- Manage project Issues
- Mitigate project Blockers
- Track project Action Items
- Approve, Manage and Submit Change Requests
- Manage and Submit Feedback Requests
- Manage project work packages (UN, Features, Epics)
- Manage project Schedule
- Report on project Progress
    - Within the Consortium
    - To External Stakeholders (ie. Dr. Gibbons)
    - *Reporting metrics and requirements to be determined

For additional information, please review the BioData Catalyst Jira SOP. Links to the various Jira projects and resources can be found below:

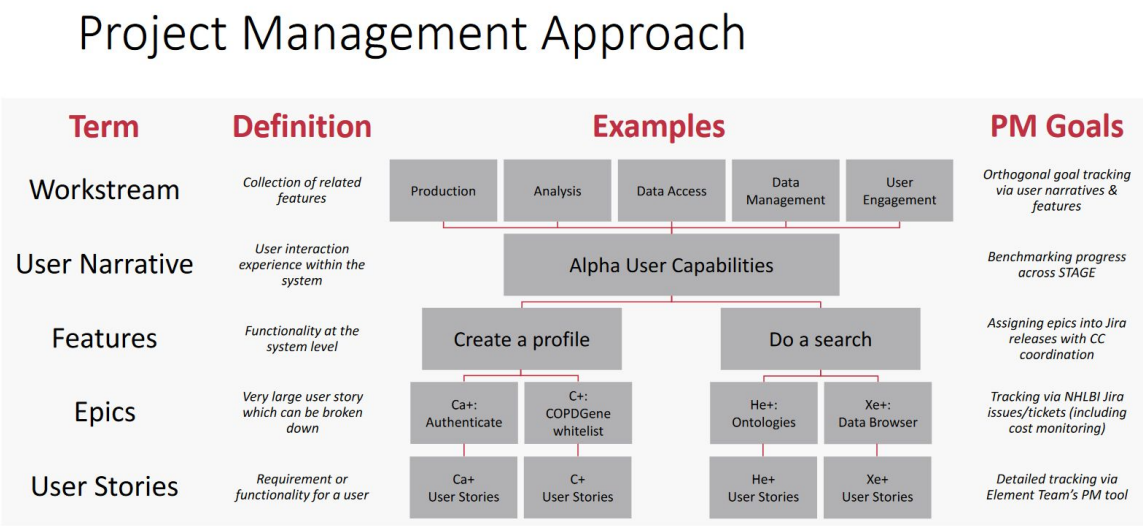
Resource	**URL**
Schedule/portfolio dashboard	https://nhlbijira.nhlbi.nih.gov/secure/PortfolioPlanView.jspa?id=41&sid=41&vid=1#plan/backlog
BDCatalyst Cross-Project Release	https://nhlbijira.nhlbi.nih.gov/secure/PortfolioPlanView.jspa?id=41&sid=41&vid=1#plan/releases
BDCatalyst User Narrative Project	https://nhlbijira.nhlbi.nih.gov/projects/UN/summary
BDCatalyst Feature Project	https://nhlbijira.nhlbi.nih.gov/projects/FEATURE/summary
BDCatalyst Epics Project	https://nhlbijira.nhlbi.nih.gov/projects/EPIC/summary
BDCatalyst Monitoring Project	https://nhlbijira.nhlbi.nih.gov/projects/MONITOR/summary
BDCatalyst Jira SOP	https://docs.google.com/document/d/1GMlY3O2ffqdPqFq_HE8SYU7qYz7k5_CUvuCJOm5hZbA/edit#

WORK PLAN MANAGEMENT

The BDCatalyst Coordinating Center and element teams will work under the direction of NHLBI to develop a unified consortium Workplan for a predefined period of performance. A hierarchical

framework of User Narratives, Features and Epics has been implemented to manage the consortium's work as depicted in Exhibit 1. Further background on WP2.0 and WP3.0, as well as detailed guidance on the current WP3.1 can be found below.

Exhibit 1. Project Management Approach



During the 2018-2019 phases of BioData Catalyst, teams worked according to individual Work Plans that organized and aggregated activities separately resulting in WP2.0.

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

In Spring 2019, the Coordinating Center worked with the individual teams to align their activities to the overall User Narrative, Feature, Epic, and User Story framework which comprised WP3.0.

Work Plan 3.0:

- User Narratives Decomposed
- BioData Catalyst Work Plan 3.0 (Jira Portfolio)
    - BDCatalyst User Narratives
    - BDCatalyst Features
    - BDCatalyst Epics

- WP3.0 retrospective

In Spring 2020, the Coordinating Center worked with teams across the consortium to reassess and realign WP3.0 deliverables in context with WP3.1. For a complete breakdown of WP3.1, see the WorkPlan 3.1 Modified Approach section.

Work Plan 3.1:

- BioData Catalyst Work Plan 3.1: Executive Summary
- BDCatalyst-RFC-8a_Workplan 3.1_Consortium Statement of Work
- BDCatalyst-RFC-8b_WP3.1_Gantt_charts

WORKPLAN 3.1 MODIFIED APPROACH:

As the workplan focus has shifted from mostly development in WP3.0 to a hybrid of development and operations activities in WP3.1, it is necessary for the consortium to take a refreshed approach to managing the workplan into focused quarterly releases with discrete deliverables and continued operations refinement & expansion. The Process below outlines the preparation and review that the Consortium Management team will conduct over a 2 week period (and will then maintain throughout the quarterly releases) ahead of each Quarterly WP3.1 release.

Initial Draft Timeline for conducting the following process for each release:

- ➔ Release 1: 9/14-9/25/2020
- ➔ Release 2: 11/2-11/13/2020
- ➔ Release 3: 1/11-1/22/2021
- ➔ Release 4: 4/12-4/23/2021
- ➔ Release 5: 7/12-7/23/2021
- ➔ Release 6: 11/8-11/19/2021

WP3.1 Proposal baselines / release reports:

How to read these reports: Epic [cells] that are highlighted a pale yellow are the only Epics assigned to that respective release. These reports provide the whole-UN/Feature picture for each release. If a User Narrative or Feature is not included in the Release Report, then there are no Epics from that UN/Feature assigned to that particular release. To view a shorter "list" of all Epics assigned to a release, please refer to the Portfolio Plans in Jira.

- #1
- #2
- #3
- #4

- #5
- #6

WP3.1 Post-NHLBI Award re-baseline release reports:

How to read these reports: Epic [cells] that are highlighted a pale yellow are the only Epics assigned to that respective release. These reports provide the whole-UN/Feature picture for each release. If a User Narrative or Feature is not included in the Release Report, then there are no Epics from that UN/Feature assigned to that particular release. To view a shorter "list" of all Epics assigned to a release, please refer to the Portfolio Plans in Jira.

- #1
- #2
- #3
- #4
- #5
- #6

WP3.1 Post-Awards Re-Baselining Process:

1. **9/28-10/23/2020: BDC3 (Marcie) & NHLBI (Chip)** will meet with **teams (PIs & PMs)** to review WP 3.1 Proposal (filtered by team-specific Epics) and identify which Epics will be de-scoped & "tag" (in Jira) with the component "3.1 Unfunded."
    a. **Teams (PIs & PMs)** should prepare for these meetings by reviewing their respective awards, the WP3.1 proposal, & the NHLBI priorities.
2. **10/26/2020: BDC3 (Ann)** will create a proposed re-baseline SOW (report), color-coding any Epics and Features tagged "3.1 Unfunded" orange.
3. **10/26-11/9/2020: NHLBI (Jon & Alastair)** will review the proposed WP3.1 re-baseline SOW and advise if approved.
4. **11/12-11/20/2020: PMs & Feature RPs** will reflect the NHLBI-approved WP 3.1 re-baseline in Jira by changing any Epics & Features that have been tagged with "3.1 Unfunded" and have been approved by Jon & Alastair to "Abandoned Intentionally" workflow status.
    a. Include the following statement: *Epic has been NHLBI-approved to be descoped from WP 3.1 due to funding restrictions & other priorities.*
5. **11/23-11/24/2020: BDC3 (Ann)** will create re-baselined release reports, and they will be linked above.

Release Preparation / Review PROCESS:

Goals:

- To identify discrete deliverables for the Quarterly Release

- To distribute the PM load across the WP/release
- To ensure cross-team collaboration & accountability
- To focus the work of the WGs
- To identify release/work blockers & interdependencies
- To track, manage, & report through Jira

Steps:

1. **Days 1-3: BDC3** to assign all the WP3.1 Jira Features in the release to an RP. See the section below on the definition and responsibilities of a Feature RP.
    a. This first-pass assignment will likely be a PM or the PM of a UN Champion. **Assignees** can then assess whether or not they want to re-assign their Feature to a Tech Lead or Product Manager or talk to BDC3 about re-assigning to a different team's PM.
2. **Days 1-3: BDC3** to review the Epics, Features, & UNs assigned to the release and comment directly in Jira issues (Features & Epics) by tagging PMs or RPs with questions.
    a. In parallel, **NHLBI & BDC3** are working to review the release content against the NHLBI priorities for WP 3.1.
        i. Part of this process includes Marcie & Chip reviewing UN's, Features, & Epics and tagging PMs in Epics with questions. **PMs** are expected to respond, in Jira, or request a meeting with BDC3/NHLBI to discuss these questions no later than Day 10
3. **Day 3: BDC3** to notify assigned Feature RPs with instructions to review Jira comments and come prepared to the next TCM/PM meeting with questions & concerns and ready to discuss!
4. **Days 3-10: Feature RPs** to:
    a. Review their assigned Feature(s)
        i. Address any comments/questions from NHLBI/BDC3 in any of the Feature or child Epics in Jira
        ii. Prepare for and attend the next TCM/PM meetings (status on: questions, concerns, discussions) → see highlighted "Hot Features" for further detailed instructions in the Feature RPs section below
5. **Days 10-14: PMs & Feature RPs** to meet during the TCMs & PM meetings to discuss discrepancies and re-baseline, as needed. Discuss new Features and decide which need a Feature Planning Document.

Feature RP definition and list of responsibilities:

- The assigned Feature RP has the lead coordinating role in seeing a Feature through from the beginning (brainstorming / requirements gathering) to fruition. Since all Features in WP3.1 include Epics from multiple teams, the Feature RP will be responsible for coordinating work across teams:

    - Leading efforts to gain consortium consensus
    - Shaping the Feature as a cohesive whole
    - Reporting back to the SC on Feature progress/status
        - Blockers
        - WG coordination
        - Contributions to the Quarterly releases
- Responsible for updates & monitoring Jira:
    - For the Feature Jira issue
    - For all of its child Epics
        - NOTE: Project leads are still ultimately responsible for updating their Institution/Team Epics in Jira.
    - For any Jira Issues created in the MONITOR project linked to the Feature or any of the children Epics
        - For "Hot Features" and whenever necessary, the Feature RP should follow up from the Virtual Quarterly meeting [Sept. 2020] with one or more of the following (not mutually exclusive) avenues:
            - Steering Committee
            - TCM
            - Assign to a WG (see further instructions below)
            - Schedule an ancillary meeting
                - ie. 1 more SC meeting to discuss Data Sharing, then move the conversation to the DAWG
        - Including ***creating*** any necessary Jira Issues in the MONITOR project for issues that need to be discussed and worked in the WG's
            - Assign to the appropriate WG chair
            - Add the WG as a component
            - Email or otherwise communicate with the WG chairs to hand off the MONITOR issue to the WG
    - Draft any Change Requests that need to be submitted to the CCB (coordinate amongst consortium)
    - Project Management responsibilities:
        - Helping team's identify interdependencies & blockers
        - Creating Jira issues for any inter-team and NHLBI blockers
        - Linking dependencies & blockers in Epics
- Serving as the Person of Contact (POC) to BDC3 & NHLBI

NOTE: The Feature RP's can be any consortium member that is positioned well to take on the above responsibilities - meaning this role need not be limited to only institution Project Managers.

Feature Planning & Execution PROCESS

Timeline: 6-weeks prior to the start of the Quarterly release cycle, begin immediately after the release review process has completed

Initial Draft Timeline for conducting the planning & execution process:

- ➔ Release 1: 9/28-11/6/2020
- ➔ Release 2: 11/16-12/25/2020
- ➔ Release 3: 2/8 - 3/19/21
- ➔ Release 4: 5/10 - 6/18/21
- ➔ Release 5: 8/9 - 9/17/21
- ➔ Release 6: 11/19 - 12/17/21

Goals:

- To strategize and detail plans for each development feature
  - ***Anything identified as an NHLBI priority requires a meeting with NHLBI to discuss/sign-off on the Feature strategy/process -- ensure everyone is on the same page
  - ***Required if 3 or more teams have an Epic in the Feature, unless noted otherwise
- To document the Feature strategy in one location when too long to be in Jira
- To secure a common understanding amongst the consortium
- To identify release/work blockers & interdependencies
- To identify risks
- To track, manage, & report through Jira

Recommended Workflow Process:

Steps:

1. **Weeks 1-3: Feature RPs** to lead the collaboration with all involved teams to complete the first draft of the Feature Planning document:
   a. Open the template google doc
   b. Save a copy as "Planning Doc: FEATURE-number_Feature name_User Narrative Name"
   c. Link the draft in the description of the Feature in Jira
2. **Weeks 4-6: Feature RPs** work with **BDC3 & NHLBI** to fill in any gaps in the draft and summarize in the Feature in Jira using the Jira Summary Template (below)
   a. ***All Features identified as NHLBI Priority 2021 -- Features RPs to meet with NHLBI to review planning document and strategy -- BDC3 to coordinate
3. ***for Release #1 only* PM meeting on 11/9/2020: Feature RPs, BDC3, & NHLBI** conduct a Lessons Learned

Feature Planning document Template

- template google doc
- example
- Shared Drive Folder

Jira Summary Template [Consideration]

Feature description and purpose:
Approach for implementing:
Teams/individuals involved:
Modules impacted:
Target deadline for implementation:
Risks/Impediments:

PROJECT METRICS

The Coordinating Center is working with NHLBI and the element teams to develop a list of initial BioData Catalyst ecosystem metrics that reflect NHLBI priorities. The metrics, once finalized, will be prioritized for consistent implementation across teams. The targeted timeline for reporting on unified ecosystem metrics aligns with WP3.1 Release #1. Additional information can be found in the Develop and implement v1 of unified cross ecosystem metrics Epic. Version 1 of the ecosystem metrics are lined for reference:

- BioData Catalyst v1 Metrics
- Specs for BioData Catalyst Cross-Platform Metrics (Draft)

Project and user metrics will be managed in WP3.1 by the following Feature and supporting planning doc:

- (User Support: Phase II) Feature D: User Metrics
  - Planning Doc: FEATURE-60_User Metrics_User Support: Phase II

DATA MANAGEMENT

In August 2020 the BDCatalyst Consortium approved the BioData Catalyst Data Management Strategy document/plan which describes the approach taken by the consortium to manage data that the BioData Catalyst Ecosystem is expected to acquire or otherwise interact with. Data management for the BioData Catalyst Ecosystem will continue to evolve to address new challenges as the Ecosystem matures. Additional content regarding BDCatalyst Data Management can be found on the at the following: https://www.biodatacatalyst.org/collaboration/strategic-planning/data-management/

DATA MANAGEMENT RACI

Operational Function	PL	SC	CCB	DRM WG	DH WG	DA WG	BDC3	Element Team DevOps Staff
Identify								
Prioritize ingestion	A	C	C	R	I	I	I	I
Provide a central index service for all data	A	I	C	C	C	C	I	R
Determine data file metadata to identify data assets in the index	C	C	C	R/A	C	C	C	I
Capture and expose provenance information	I	I	C	A	R	I	I	R
Store & Secure								
Plan and manage data storage buckets	R/A	I	C	C	I	I	I	C
Manage platform cloud environments	A	I	C	I	I	I	I	R
Ensure FISMA/FedRAMP compliance	A	I	C	I	I	I	C	R
Report Privacy Concerns and Breaches	A	I	I	C	I	I	R	R
Ensure compliance with consents	A	I	I	C	I	R	I	R
Define quality control procedures at each stage of data transfer	I	I	C	A	R	R	R	R
Run quality control procedures at each stage of data transfer	A	I	C	C	I	I	I	R
Make data available through common identity and access management	A	C	C	C	I	R	I	R
Provision								
Utilize GA4GH and other standards-based interfaces for user's query and interop with other data commons platforms	A	C	C	C	C	R	I	R
Evaluate emerging standards (e.g., PFB, PIC-SURE, FHIR, etc.) to enable data provisioning within the ecosystem	A	C	C	R	C	C	I	I

Craft an Ecosystem-wide search strategy	A	R	C	C	C	C	C	I
Expose								
Enable consistent return of results for user search of data	A	I	C	C	R	I	R	R
Support user's uploading their own data	A	I	C	R	I	I	C	R
Enable users to store and publish data analysis results	A	I	C	C	I	C	I	R
Develop policy for sharing derived data	A	R	C	C	C	C	I	I
Implement policy for sharing derived data	A	I	C	C	I	C	I	R
Develop policy for sharing BYOD back into the ecosystem	A	R	C	C	C	C	I	I
Implement policy for sharing BYOD data	A	I	C	C	I	C	I	R
Governance								
Make recommendations to changes in current data management strategy or processes	A	I	R	R	C	C	I	I
Track status, testing, and changes over time	A	I	C	C	I	I	R	I
Allow users to store and share data using tools in Cloud resources	A	I	C	I	C	C	I	R
Capture and report performance indicators on data management controls	A	I	C	C	C	C	R	R

Legend: R – Responsible; A – Accountable; C – Consulted; I – Informed

Acronyms: PL - Program Leadership; SC- Steering Committee; DRMWG – Data Release Management Working Group; DAWG - Data Access Working Group; DHWG - Data Harmonization Working Group; BDC3 – BioData Catalyst Coordinating Center; CCB – Change Control Board

DRMWG

The scope of the Data Release Management Working Group (DRMWG) is to identify, outline, and make recommendations around:

1) Prioritization of data sets for release
2) Formatting and organization of object files and their directory structure(s) for planned ingested datasets
3) Data ingestion and release-specific metadata, in conjunction with the DataHarmonization WG
4) Data ingestion proposals and desired release timelines
5) Other relevant obstacles related to data ingestion and use within the BioData Catalyst Ecosystem.
6) Collaborate with Data Access Working Group as neededIdentified issues and recommendations will be shared with other BioData Catalyst working groups, teams, stakeholders and Change Control Board, as appropriate.

DRMWG Charter

FINANCIAL MANAGEMENT

NHLBI is responsible for compiling financial reports from the BioData Catalyst Element teams.

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

The Coordinating Center has developed a budget template for consortium-wide use.

QUALITY MANAGEMENT

Quality is paramount to the BioData Catalyst 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 BioData Catalyst-specific aims and milestones. The Quality Table below highlights the BioData Catalyst 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 BioData Catalyst 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, Data Management Strategy Plan, Coordinated DevOps Feature G Planning Doc and various SOPs	Coordinating Center and Feature RPs	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 BioData Catalyst 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	IV&V, Coordinating Center, Element teams, Working Groups, and Tiger Teams (Tool & Apps WG, Integration Testing TT, Load Testing TT)	Triggered on a schedule, such as nightly
Quality Improvement	Solicit feedback from researchers using the BioData Catalyst Ecosystem	Coordinating Center (User Services), User Engagement and Data Access/UX-UI Working Groups	As part of a regular assessment process

DEPLOYMENT AND TESTING PLAN

The Coordinating Center utilizes the *nhlbidatastage GitHub* organization for code sharing, with specific, existing open source projects also using GitHub repositories for source control.

Independent Verification & Validation (IV&V) assessment will analyze and test the ecosystem to:

1. ensure that it performs intended functions correctly
2. ensure that it performs no unintended functions
3. measure its quality and reliability

The IV&V effort will be managed via (Coordinated DevOps) Feature G: IV&V testing. Additional details can be found in the IV&V Feature Planning Document.

The Integration Testing Tiger Team will maintain a working BioData Catalyst ecosystem while simultaneously improving the level of integration and sophistication of capabilities and increasing the amount of data available on the system. The ultimate goal is a centralized system in place where each institution can run the same tests with each other's applications before they release to production.

Specifically:

- A CI/CD tool in place that all institutions have access to and can trigger tests from
- A shared repo with these cross-application spanning tests
- Each of these tests should be quick to run and limited in scope

The Load Testing Tiger Team will identify points of integration likely to contribute to performance degradation, execute test cases and estimate the cost to run test cases.

BioData Catalyst 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 BioData Catalyst ecosystem. The goal of the architecture is to provide a “write once, run many" environment where the app developer can specify the metadata for their application using a format that each of the individual operating environments supports. Version 1.0 of the BioData Catalyst Application Developer Guide provides sufficient guidelines for an "app" developer and an "app" system launcher to build their respective resources in the BioData Catalyst. This guide/document will be updated on an annual basis.

DOCUMENTATION REVIEW

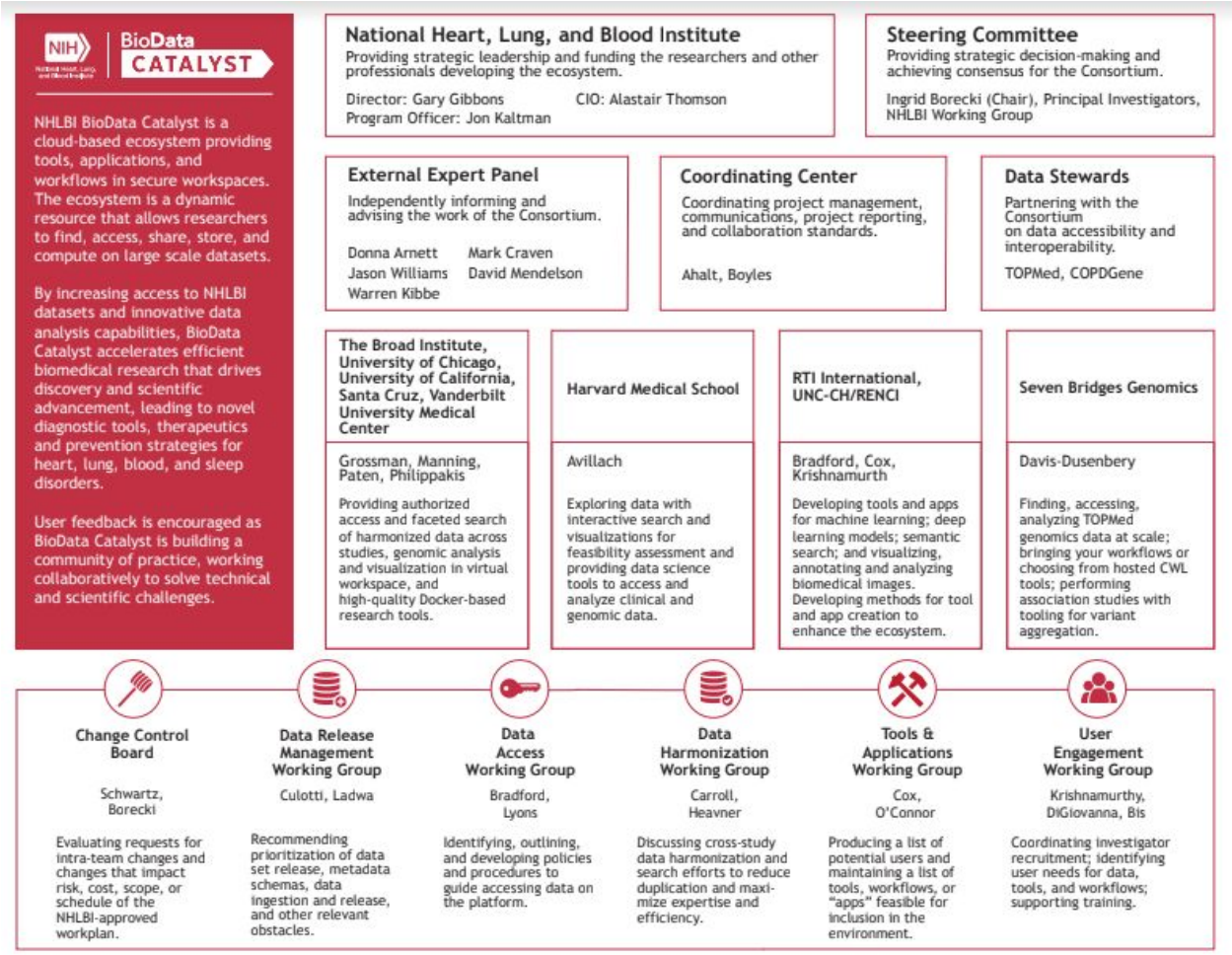
Planning documents, including the Strategic Framework Plan, Implementation Plan, and [this] Project Management Plan, will undergo periodic reviews by NHLBI, the Steering Committee,

and the Coordinating Center to assess any high-level changes in scope and/or direction of the project. Formal reviews will occur on an annual basis. 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 routed to the Change Management process (identified earlier in this document).

RESOURCE MANAGEMENT

ORGANIZATIONAL CHART



STAKEHOLDER MANAGEMENT

All members of the BioData Catalyst 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 BioData Catalyst Consortium Directory serves as the **Stakeholder Registry;** it 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. This

password-protected website engages consortium members on a multitude of pages designed to communicate project goals, initiatives, and events. Consortium members are also engaged through Working Groups and Tiger Teams and regular virtual meetings. Communication to consortium members is initiated through Slack and email. Engagement amongst consortium members and NHLBI is also conducted through Jira by commenting/tagging other users and linking Issues as appropriate.

User Services

Established during WP3.0 and led by the Coordinating Center to help coordinate the various operational elements of User Services with the User Engagement Working Group [soon to be replaced by the User Experience Coordination WG] across the BDCatalyst Ecosystem; user onboarding, training, documentation, metrics, help environment (including the forum and FAQs), and feedback solicitation/collection.

- Consortium Guidance
- Documentation
- Tutorials
- BDCatalyst Draft Help Desk SOP (workflows)
- BioData Catalyst Public Cohort Users
- User Feedback SOP

User Services will be managed and coordinated in WP3.1 under the following Features and supporting planning documents :

- (User Support: Phase II) Feature A: Fellows and User Support
  - Planning Doc: FEATURE-57_Fellows and User Support_User Support: Phase II
- (User Support: Phase II) Feature B: User Support Documentation
  - Planning Doc: FEATURE-58_Documentation_User Support: Phase II
- (User Support: Phase II) Feature C: Outreach to researchers
  - Planning Doc: FEATURE-59-Outreach-to-researchers
- (User-requested features, tools, and improvements) Feature A: UI/UX

User Services Contacts: Chris Erdmann, Chris Lenhardt and Paul Kerr

Fellows Program

The NHLBI BioData Catalyst Fellows Program provides early-career researchers (graduate students, postdocs, junior faculty, and others) the opportunity to receive funding to help support research on novel and innovative data science and data-focused research problems. The program is open to academic disciplines conducting biomedical research in heart, lung, blood, and sleep (HLBS) domains and may include focus on the impact of SARS-COV-2 in those domains.

The BDCatalyst Coordinating Center, under the guidance of NHLBI, helps facilitate and coordinate the Fellows Program. Fellows Cohorts I and II have been onboarded and are actively using BDCatalyst, while Cohort III applicants are being solicited.

Fellows Program Contacts: Sarah Davis, Patrick Patton, Chris Erdmann and Ingrid Borecki

User Experience Coordination Group (UECG)

The User Experience Coordination (UECG) was introduced at the December 2020 Virtual Quarterly Meeting with the following primary goals/objectives:

- Provide coordination for cross-team decision-making related to user-focused work plan activities.
- UECM complements and mirrors the TCM's focus on technical issues
- Develop strategy, specific advice, guidance and coordination of decision-making related to identified focus areas
- Spin off smaller working groups to handle specific areas with specific deliverables. Focus on what the groups produce, rather than what they talk about

The scope and framework of UECG is continuing to evolve. Any updates will be made to the Project Management Plan as well as the drafted Terms of Reference linked below:

Framing for an Updated User Experience Coordination Group (UECG) Terms of Reference (ToR)

BIODATA CATALYST OPERATIONAL RESPONSIBILITY MATRIX

The BioData Catalyst Operational Responsibility Matrix is the RACI, developed by NHLBI, for the key BDCatalyst 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 BDC3 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://www.biodatacatalyst.org/stage-resources/.

CLOUD RESOURCE MANAGEMENT

The Cloud Credits Tiger Team, chaired by Chip Schwartz, will guide the development of Cloud Resource Management for BioData Catalyst:

- Cloud Credits Tiger Team Charter
- Cloud Credit Tiger Team Agenda/Notes
- Cloud Credits Webpage & Form

Additionally, Cloud Credits are identified and managed in WP3.1 via the (Ecosystem Hardening and Scalability) Feature I: Cloud costs and billing systems Feature:

Users on BioData Catalyst will be able to see information about their incurred costs from computation and storage on AWS and Google Cloud. Users will be able to take advantage of STRIDES discounts and get billing groups set up with cloud credits. A user will encounter automated restriction of new analysis due to low cloud credits if the billing account is not sufficient to cover the cost.

- Planning Doc: FEATURE-119_Cloud_Costs_and_Billing_Systems

COMMUNICATIONS MANAGEMENT

INTERNAL COMMUNICATIONS

Communications Management on BioData Catalyst 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, our primary mode, is conducted across the Consortium by means of video conference and face-to-face meetings. The BDCatalyst Consortium Website Calendar is the source of truth for all consortium meetings, both regularly occurring and one-off.

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.

EXTERNAL COMMUNICATIONS

Coordinated external communications, like coordinated branding, assures aligned messaging with a unified voice and clarity of purpose, well-timed distribution of information, and amplification of joint and individual successes. The Coordinating Center, along with guidance and review from the NHLBI, coordinates external communications for BioData Catalyst to achieve these goals.

- BioData Catalyst Branding Guidance
  - BioData Catalyst Style Guide V2.2
- v3 Approved BioData Catalyst Publications Guidelines
- Publications Guidance -- GitBook
- Introduction to BioData Catalyst Slide Library

External Communications and Outreach to users is managed in WP3.1 under the following Feature and supporting planning doc:

- (User Support: Phase II) Feature C: Outreach to researchers
- Planning Doc: FEATURE-59-Outreach-to-researchers

Communication Contacts: Stephanie Suber and Lauren Hochman

BDCATALYST CALENDAR

Please reference the BDCatalyst Consortium Calendar as the source of truth for all consortium meetings (recurring and one-off).

[Updated Statement on Consortium Transparency]

The NHLBI BioData Catalyst Consortium is dedicated to providing an inclusive environment where all voices and opinions can be heard. We believe that fostering this inclusive environment leads to better technical development, stronger governance models, and improved relationships within the Consortium.

Accordingly, the BioData Catalyst Coordinating Center would like to reiterate that all Consortium meetings with more than one team attending are open to all Consortium members, including Fellows. These Consortium meetings are listed on the BDCatalyst Calendar, which includes web conference information and agendas when available.

We recognize that many meetings are more efficient when only key personnel attend. Consortium members should therefore use their discretion not to attend smaller working meetings where they are not likely to be key, active participants. As agendas and meeting notes are generally available, and important decisions and other relevant information are shared regularly with the Consortium at the Steering Committee and quarterly meetings, Consortium members should not feel that meeting attendance is necessary to stay involved in Consortium issues.

Additionally, documents shared among more than one team should be viewable by all Consortium members to ensure transparency among the teams.

WORKING GROUPS & TIGER TEAMS

Formal Working Groups foster cross-collaborative community projects. In consultation with the community, the Steering Committee forms Working Groups. Working Groups serve as an important community-engagement mechanism for providing ongoing synthesis of requirements, identification of community-defined needs, and changes in scientific direction. The current BDCatalyst Working Groups are as follows:

- Data Access WG - **Retired as of 12/2020*
- Data Harmonization WG
- Data Release Management WG
- Tools & Apps WG
- User Engagement WG -- **Soon to be updated by the User Experience Coordination Group*

Tiger Teams, which arise from community-led shared interests, are nominated and approved by the Steering Committee. Alternatively, Tiger Teams may be created by the Steering Committee to address a specific task or problem.

Tiger Teams are specifically designed to incur low overhead and move quickly. Tiger Teams will continue as long as there is interest or need, and they will be provided with access to core BDC3 resources in order to facilitate their interaction. The current BDCatalyst Tiger Teams are as follows:

- Cloud Credits TT

- Integration Testing TT
- Load Testing TT
- Manuscript TT
- Monitoring Dashboard TT

PUSH, PULL, AND INTERACTIVE WRITTEN COMMUNICATIONS TABLE

Communication Type	Communication Vehicle	Examples	Additional notes
Push	Email	• Meeting/due date reminders • Call for document reviews (Governance docs, matrices, etc.)	• Email distributions are determined by the audience. • Lists are formed using the Consortium Directory
Pull	BioData Catalyst Consortium website https://biodatacatalyst.nhlbi.nih.gov/ https://bdcatalyst.gitbook.io/biodata-catalyst-documentation/	• Calendar (hosts all relevant project- and Consortium-level meetings and events) • Collaboration pages • Resources page • Meeting materials • Governance documents and SOPs	• Meeting notes • Consortium members are added to the private website during the onboarding process
Interactive	• Slack • Google Drive • RFCs	• Slack: channels and direct messaging • Google Drive: working documents	Consortium members are added to these shared platforms during the onboarding process

COMMUNICATION GUIDELINES

The Coordinating Center has collected a set of BioData Catalyst Communication Guidelines that overviews meeting best practices and etiquette as well as details on using the BioData Catalyst 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 Jira Monitoring Project for Consortium transparency and to ensure that the assigned Risk Owner takes the necessary

steps to implement the mitigation response in a timely manner.

RISK PROCESS

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 following the process identified below.

Risks

1. Consortium members complete the Risk submittal google form (found on the consortium website Resources page)
2. The submittal is populated in the risk register google sheet
3. The risk register will be previewed by the CC bi-weekly, where it will be decided if the risk can be addressed by CC coordination or needs to be presented to the SC for decision.
4. The CC will create the risk in the BDCatalyst Monitoring Project in "to do" status.
5. The SC will review "to do" risks regularly to decide if the response proposed is acceptable or needs to be amended (document in Comment field).
6. The CC will work with risk assignees to make any SC requested changes and change the status of the risks to "in progress" in the Monitoring Project.
7. The assigned owner will begin to implement the response strategy.
8. The risks will be reviewed in PM, TCM and SC meetings as appropriate.
9. If a risk comes to fruition, the CC will change the issue type from a Risk to an Issue.

Issues

- Any Jira user with Developer access is empowered to create issues in the BDCatalyst Monitoring Project.
- Working Groups and Tiger Teams are encouraged to use the Kanban Board to log and track issues that arise during group meetings, including issues that cross groups.
- If an issue becomes a blocker, the CC will change the issue type from an Issue to an Epic, link the blocker to the Epic(s) it blocks, and ensure it is linked up to the appropriate Feature.

Action Items

- Any Jira user with Developer access is empowered to create Action Items in the BDCatalyst Monitoring Project.
- Working Groups & Tiger Teams as well as note takers in the TCM, F2F, and other Consortium meetings are encouraged to use this Monitoring Project to log and track action items.

Nice-to-haves

- Any Jira user with Developer access is empowered to create Nice-to-haves in the BDCatalyst Monitoring Project.

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.

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

1. BioData Catalyst Risk Submittal Form
2. BioData Catalyst Risk Register
3. BioData Catalyst Monitoring Project (Jira)

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
- 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 BioData Catalyst ecosystem 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 BioData Catalyst ecosystem 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:	AC-17/ 18/ 19

	Encryption Limit remote access points and privileged operations Authenticate wireless	
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		
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) https://csrc.nist.gov/projects/cryptographic-module-validation-program	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	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 BioData Catalyst 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 BioData Catalyst Consortium will incrementally progress towards the vision of the program described in the BioData Catalyst Strategic Framework Plan.	Scope Management
BioData Catalyst Consortium Charter	The charge for the entire BioData Catalyst consortium aligning the visions and goals for multiple OTAs.	Introduction - Executive Summary of Project Charter; Consortium Governance
BioData Catalyst Code of Conduct	Summary of member expectations for a safe environment.	Consortium Governance
BioData Catalyst Publications Guidelines	Set of requirements for publishing content related to BioData Catalyst.	Consortium Governance
BioData Catalyst Working Group Quick Start Guide	A set of parameters for initiating a BioData Catalyst Working Group or Tiger Team.	Consortium Governance
BDCatalyst Change Control Board (CCB) Charter	The scope of the CCB is to make decisions for proposed changes to the BDCatalyst Ecosystem.	Change Management
Procedure for BioData Catalyst Ecosystem Configuration and Change Management SOP	• Identify the types of requests for changes • Identify the workflow process for Change Requests through the CCB • Identify a streamlined process for emergency Change Requests • Guidance on completing a Change Request in Jira	Change Management

BDCatalyst RFC Implementation Process	Outlines the proposed steps for cross-consortium agreement documentation.	Request For Comment (RFC) Process
BDCatalyst Jira SOP	SOP for consortium wide management of Jira	Schedule/Workplan Management
BDCatalyst Jira Projects	Link to all of the BDCatalyst Projects in Jira	Schedule/Workplan Management
BioData Catalyst Data Management Strategy	Consortium approved Data Management Strategy/Governance Doc	Data Management
DRMWG Charter	Data Release Management Working Group Charter	Data Management
nhlbidatastage GitHub	Organization for code sharing, with specific, existing open source projects also using GitHub repositories for source control	Quality Management/Deployment and Testing
BDCatalyst Working Groups	Webpage on consortium website for most up-to-date WG & TT information	Communications Management/Working Groups and Tiger Teams
Budget Template	Template for providing NHLBI with budget per Task/Month/Staff	Financial Management
Governance and Policy Documentation [Folder]	Folder containing BioData Catalyst Governance & Policy Documents.	Quality Management
SOPs [Folder]	Folder containing a collection of internal SOPs utilized by the Coordinating Center.	Quality Management
BioData Catalyst Operational Responsibility Matrix	BioData Catalyst RACI matrix.	Stakeholder Management
Consortium Directory	BioData Catalyst website hosted directory containing consortium contact information.	Stakeholder Management
Coordinating Center Calendar	BioData Catalyst 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

Risk submittal google form	Google form for submitting risks	Risk 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 BDCatalyst Glossary provides definitions and examples for terms relevant to this document.