

Integrated Network Data Coordinating Center & Centralized Repository Request for Applications (RFA)

September 22, 2022

KEY DATES	
RFA Released	September 22, 2022
Submission Deadline	November 17, 2022
Review & Presentations*	January 16-31, 2023
Projected Award Notification Date	February 10, 2023
Project Start	March 2023

*If deemed necessary, applicants may be requested to clarify and/or present their proposal orally. BD² will notify applicants in the event such a presentation is necessary.

About BD²

[BD²: Breakthrough Discoveries for thriving with Bipolar Disorder](#) is a collective force to transform what we know about and how we treat bipolar disorder. It's a commitment to the 40 million people living with bipolar disorder, those not yet diagnosed, and their loved ones.

The Baszucki, Brin, and Dauten families united with the Milken Institute to create BD² to advance discoveries for families like theirs. For too long, there have been limited advances in the study and treatment of bipolar disorder due to lack of collaboration and funding. It's time for a new approach.

BD² is establishing a network of interdisciplinary investigators to apply cutting-edge biotechnology, big data analytics, and an unprecedented data ecosystem to address bipolar disorder in an innovative, equitable, and rigorous way.

About Bipolar Disorder

Bipolar disorder is a highly complex and heterogeneous disorder that is often debilitating. Even though it is prevalent in about 3% of individuals worldwide, and is recognized as a leading cause of disability, little is known about its biology. Advancements in our understanding and treatment of bipolar disorder to date remain far from ensuring that everyone living with it can manage their condition and lead independent, fulfilling lives.

Mission

BD² was launched to realize the vision of a world where the etiology and biological mechanisms of bipolar disorder are well-understood, allowing the development of effective interventions and optimized outcomes for all individuals with bipolar disorder.

Opportunity Snapshot

BD² is launching a multidisciplinary initiative to increase our understanding of the natural history and underlying biology of bipolar disorder and ultimately, to identify novel strategies for improved care and intervention. In partnership with people living with bipolar disorder, clinicians, and researchers, BD² is facilitating a longitudinal study and a learning health network, together forming the BD² Integrated Network. The Integrated Network will generate clinical, imaging, digital tracking, and multi-omics biological data from a total of 4,000 study participants and is the focus of this request for applications (RFA). BD² intends to provide up to \$5,000,000 over five years to the selected Data Coordinating Center (DCC).

Purpose

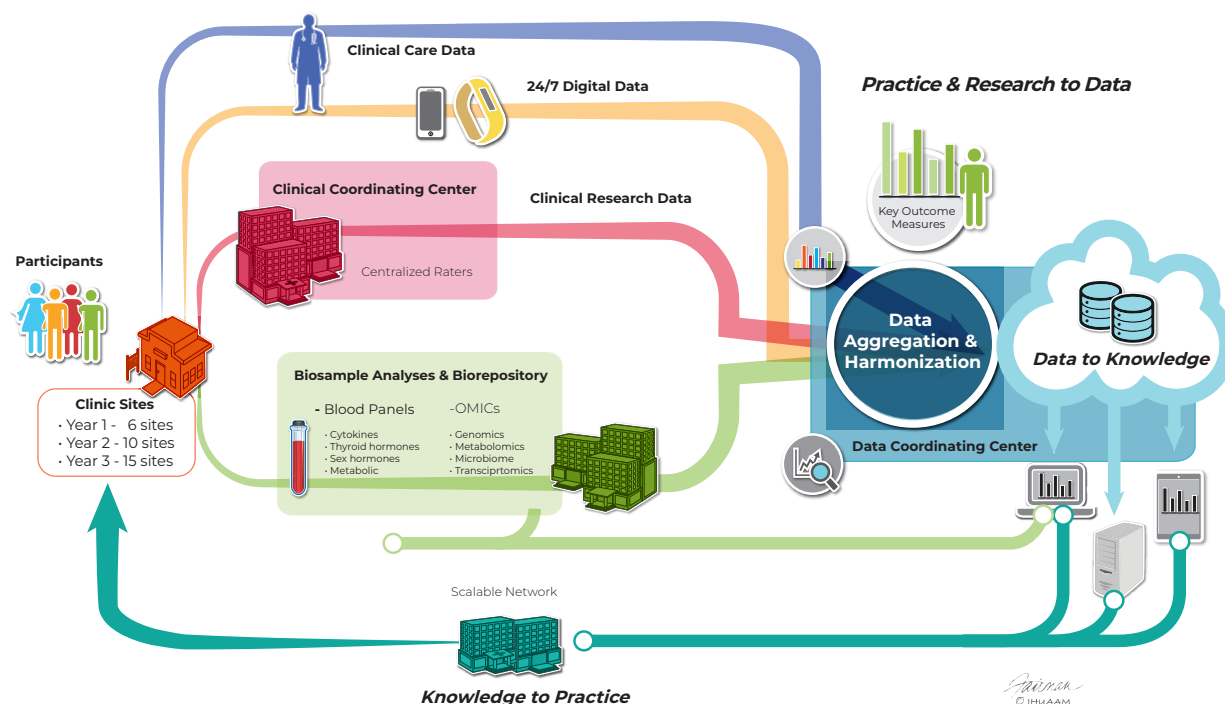
BD² seeks to identify a DCC to provide technical and logistical support to sites participating in the Integrated Network, and to build and maintain a centralized data repository (CDR). The selected organization will be responsible for coordinating and executing activities across three main areas:

- Data aggregation, harmonization, and storage;
- Data governance and access; and
- Data analysis and reporting.

The DCC will play a critical role in providing resources, governance, and technical expertise to support iterative cycles of data aggregation, analyses, and implementation of evidence-based recommendations – all with the goal to improve care for those living with bipolar disorder as rapidly and effectively as possible.

The CDR will form the backbone of a bipolar disorder data ecosystem that will serve up to 15 sites and 4,000 bipolar disorder participants (recruited and enrolled over five years) within the Integrated Network. It will support a learning health network as well as a platform to facilitate the open sharing and analysis of program-generated datasets.

Figure 1. Overview of the BD² Integrated Network



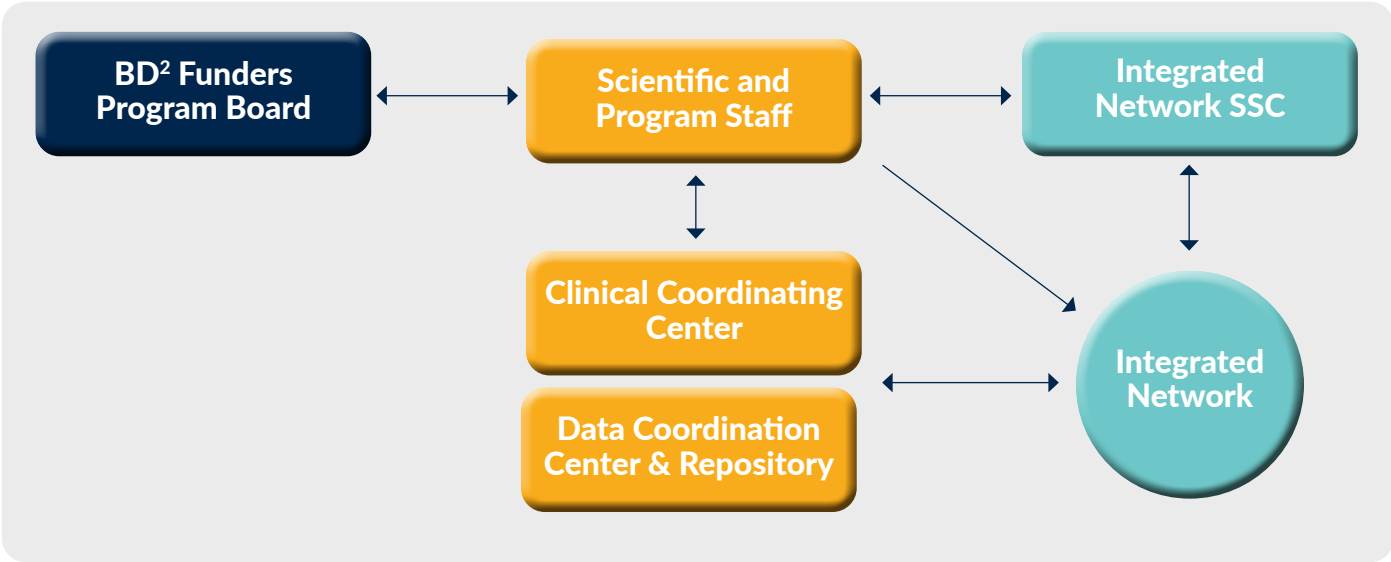
As outlined in Figure 1, participant data will be aggregated from several sources to build in-depth clinical and biological profiles of bipolar disorder. The overarching goal of the Integrated Network is to use this data to identify key outcome measures for bipolar disorder as well as interventions and strategies to improve clinical care and practice.

Over a five-year period, BD² will enroll a cohort of 4,000 patients with a diagnosis of bipolar disorder type I. In-person visits will occur at baseline and every 12 months for five years (constituting a total of 24,000 study visits), supplemented by more frequent remote tracking of mood and other related outcomes. The program will launch with six sites across the USA and expand globally to 15 sites over three years. Additional sites will be added in subsequent years to achieve a global presence and enhanced diversity of participants.

Governance

BD² is managed by a Program Staff, in consultation with a funder's Program Board (PB) (Figure 2). The PB receives input and strategic advice from a scientific steering committee (SSC) who in turn guide and direct the operations of the Integrated Network, along with its centralized services, including a Clinical Coordinating Center (CCC) and the DCC.

Figure 2. BD² Integrated Network Governance Structure



The overall direction of the awardee’s work will be governed by the Integrated Network SSC. The SSC is comprised of leading experts in the fields of psychiatry, clinical psychology, neuroimaging, genomics, community outreach, and mental health informatics. The SSC will meet regularly with the awardee organization to guide program development. DCC integration with the SSC and Program Staff will be critical to the successful design and execution of the Integrated Network.

A primary goal of the program is to aggregate consistent and complete data from sites and core facilities. To this end, the Integrated Network SSC will oversee data standardization and interoperability between sites and cores. The CCC and team of centralized raters will support and enable the collection of standardized data. Data from other sources, including different health systems, electronic health records (EHRs), and claims, will require engagement with the leadership of participating healthcare systems to establish and optimize collection. While the principal investigator at each participating site will be responsible for leading these engagements, they will need input and support from the DCC.

Funding

BD² intends to make a single award for the scope of work (SOW) described in this RFA. Proposals should include a competitive budget that captures all direct and indirect costs associated with the SOW over a period of five years. All funding totals are inclusive of indirect costs up to a maximum of 15% of the total award budget.

BD² intends to provide up to \$5,000,000 over five years to complete the SOW. The initial contract will be awarded for 24 months and will be eligible for a non-compete renewal every 12-24 months thereafter until the completion of the SOW. The anticipated award date is approximately February 10, 2023.

Organizational Eligibility

Proposals will be accepted from any private sector organization including nonprofit and for-profit organizations, universities, hospitals, laboratories, healthcare systems, and research organizations.

Organizations with prior experience in health informatics, outcomes research, and the development of large, multi-center programs are encouraged to apply. An understanding of and proven ability to comply

with applicable data protection regulations, including the privacy and security regulations associated with implementing the Health Insurance Portability and Accountability Act (HIPAA), as well as the General Data Protection Regulation (GDPR), is essential.

Project Goals and Scope of Work

The Integrated Network will generate clinical, imaging, digital tracking, and biological data. As we work to better understand and treat bipolar disorder, these data will be accessed and analyzed by people with bipolar disorder, clinicians, health system administrators, and researchers.

Overview of Scope of Work

The selected organization will be responsible for establishing a DCC and CDR to facilitate the work of BD² and the Integrated Network.

Data Coordinating Center (DCC): Under the direction of the SSC and the Program Staff, the DCC will organize and support interactions between all parties within the Integrated Network. The DCC will provide regular reports to the Integrated Network Program Director (PD) and will provide support in the following general categories:

- **Program Infrastructure:** The DCC will be responsible for establishing program infrastructure including, but not limited to, a data portal, data workbench (for in-platform data manipulation and analysis), and multi-faceted data storage.
- **Program Management:** The DCC will provide technical and logistical support to Integrated Network sites and partners to ensure timely development of the network.
- **Technical Assistance:** The DCC will offer technical assistance and guidance to accommodate the needs of the Integrated Network sites and partners in accordance with SSC direction. This will include but is not limited to data aggregation from all sources (e.g., EHRs, REDCap, wearables, claims data), establishment of data capture and quality control workflows, data harmonization, and data de-identification.
- **Data Governance:** The DCC will solicit and disseminate recommendations for policies and practices to support data integrity, usability, sharing, and security across all operations and partners of the program.
- **Leadership Support:** The DCC will offer support as needed to the SSC and BD² Program Staff for leadership meetings, including the PB and SSC.

Centralized Data Repository (CDR): In addition to the responsibilities stated above, the selected organization will also design, build, and maintain a CDR. The CDR will:

- Build and maintain data dashboards for multiple stakeholders (including patients, physicians, and researchers);
- Support the aggregation, harmonization, and storage of Integrated Network clinical and biological data;
- Provide the necessary security for both identifiable and de-identified data; and
- Facilitate efficient and effective data access, queries, and analysis.

Central to the success of BD² is the formation of a bipolar disorder data ecosystem. This ecosystem will support open sharing of the data aggregated within the CDR alongside analysis and visualization tools to facilitate engagement with the global research, clinical, and lived experience communities. As outlined in Figure 1, clinicians will be able to access their own patient's data, as well as aggregated data from across the Integrated Network. The latter will be used to identify new strategies and clinical practices to improve care.

Researchers at the sites will be able to access the various data domains to formulate and test hypotheses, analyze trends, and produce new insights into bipolar disorder to develop novel and improved interventions. Participants in the study will also be able to access their own data to track changes of various relevant metrics. A more detailed list of feature requirements is included in Appendix A.

Proposal Requirements

Applicants should submit a five (5) page proposal, with at least 11-point font and 1-inch margins on all sides of the document, which adequately address the points below. All applications will be submitted via the [Submittable online grant portal](#).

Organization Profile

- Please include organization name, age, mission statement, overview of offered services and products, and a brief history.
- Diversity and inclusion strategies and practices within the organization.

Scope of Work

Solution approach to requirements and services listed within this RFA. Specifically, the proposal should address:

- Approach to data collection from multiple sources (e.g., different health systems, EHRs, REDCap, other data sources such as claims).
- Proposed workflows for data capture into the CDR.
- Approach to data harmonization and potential use of common data models (e.g., FHIR, OMOP).
- Processes and workflows specific to identifiable vs. de-identified research data.
- Data management including plans to:
 - Implement technical, physical, and administrative data security controls;
 - Conduct data privacy training for all individuals with access to the CDR;
 - Establish and ensure compliance with all governmental regulatory requirements at sites;
 - Document and maintain study protocols and standard operating protocols (SOPs);
 - Establish and complete procedures for routine quality control (QC) checks of the data;
 - Generate routine QC reports; and
 - Process requests for data access.
- Key project milestones, deliverables, and associated timelines.

Relevant Experience

The applicant should outline their previous experience and technical and subject-matter expertise relevant to planning and conducting the required activities.

- Qualifications and experience leading a data coordinating center.
- Relevant practice and experience developing CDRs for multi-faceted data types, along with multiple stakeholders such as researchers, clinicians, people with lived experience, and hospital and research institution administrators.
- Applicant's demonstrated experience with and proven understanding of compliance with HIPAA privacy and security regulations and all other data protection requirements applicable to the DCC and CDR.

Budget and Supplemental Information (not included in 5-page limit; [please use provided templates](#))

Applicants can apply for up to \$5,000,000 USD of funding over five years to cover the cost of carrying out the SOW described in this RFA.

- As noted above, all funding totals are inclusive of indirect costs up to a maximum of 15% of the budget.
- The budget should reflect costs across five years (up to a total of \$5,000,000 USD). The initial contract will support years one and two, with the option for a non-compete renewal every 12-24 months thereafter through to completion of the SOW.

Please also include:

- Project team roles and responsibilities;
- Proposed team members, their qualifications, and biographies; and
- A letter of commitment from the CEO, President, or lead of the organization ([template](#)).

Review Process

Written proposals will be reviewed by the Integrated Network leadership, including members of the SSC, mental health and informatics experts, the PD, and Program Staff. Proposals will be evaluated based on the responses to all requirements in this RFA. The evaluation of an applicant's ability to provide the required services will be based on the written material submitted, interviews, and, if requested, presentations.

Each proposal will be competitively evaluated on its strengths and weaknesses. Below is an example of the criteria that may be used as part of the evaluation process:

- History of building successful DCCs for multi-institutional biomedical or clinical efforts;
- Demonstrated expertise in provisioning and supporting services described in this RFA;
- Global reach and the ability to scale operations up or down as required;
- Applicability, creativity, and innovation (i.e., better process, lower cost, synergies) of applicant's solutions;
- Timely responsiveness to this RFA and the schedule contained herein, as well as follow-up to additional requests for clarification and information;
- Business and staffing continuity plan to ensure continuous delivery levels; and
- Continuous improvement and transformation initiatives.

Finalist applicants may be asked for additional interviews, site visits, and additional materials.

Final Selection

Following proposal review, the BD² Integrated Network PD, Program Staff, and SSC will convene and select the successful applicant. Final decisions will be confirmed by the PB. Once notified, the organization will work with the PD and Program Staff to begin work.

Funding Awarded in BD²'s Discretion

Responding to this RFA and/or submitting an application does not entitle any individual or organization to receive funding from BD². Funding, if any, would be provided in BD²'s sole discretion pursuant to the terms of a written agreement executed by BD² and the selected organization, the terms of which BD² may require to be acknowledged by the awardee.

Contact Information

An automated email confirmation is generated upon application submission. If you do not receive confirmation within 24 hours of submitting your application, please check spam filters then contact integratednetwork@bipolardiscoveries.org.

For inquiries about scientific priorities, eligibility requirements, and application submission please contact integratednetwork@bipolardiscoveries.org. For all other questions, including general and media inquiries related to BD², please contact: info@bipolardiscoveries.org.

Appendix A

Description of Services

The BD² Integrated Network requires the CDR to function with the following features:

1. Accounts for Users, Administrators, and Groups Must Be Supported
 - A. Some CDR features do not require a User Account.
 - B. Individuals are granted access via a User Account.
 - C. User Accounts may be added to Groups and Groups may be assigned permissions.
 - D. User Accounts and Groups may only be assigned permissions by an Administrator.
2. Manage Access to Datasets, Metadata, and Features
 - A. All Public website features and metadata do not require a User Account.
 - B. All Protected website features, metadata, and data require a User Account.
 - C. All Protected data downloads log the user, time(s), date, and any errors.
3. Data Validation
 - A. The system will log any validation errors during system initialization. (The portal will log any metadata records that it does not understand.)
 - B. Data validation rules are configurable (e.g., via dictionary, schema, possibly custom business rules).
4. Webpage Interfaces
 - A. The set of webpages should be editable by an approved administrator. Administrators should be able to define and update a small set of content pages:
 - i. A splash page to provide an overview of the initiative's data and release roadmap.
 - ii. An experimental metadata page or section provides an overview of the experiment design and protocols for data collection.
 - iii. A summary data page to provide interactive summary charts and high-level descriptions of the latest data releases.
 - iv. A release notes page to describe the latest release, previous releases, and release plans.
 - v. A request access page with a Data Access Request form.
 - vi. An interactive metadata search page to provide search and download features.
 - vii. A downloads page providing links to metadata features and tools.
 - B. The summary data page must provide summary charts derived from live queries or cached nightly updates of current data summaries in SQL tables.

5. Metadata Search Capabilities

- A. The metadata search page must provide (initially) faceted search by:
 - i. Clinical metadata and case/control/diagnosis fields;
 - ii. Variants and known mutations fields; and
 - iii. Datatype and biosample inventory fields.
- B. The metadata search page must provide interactive feedback based on user selections including:
 - i. Participant count;
 - ii. Sample count;
 - iii. Biosample count; and
 - iv. File count (number and size).
- C. The metadata search page must be configurable by the administrator (not a developer or code change) to support new facets, new fields, and new data field ranges.
- D. The metadata search page must support “http” download directly from the browser for filtered results for (initially) the following metadata types:
 - i. Participant clinical data;
 - ii. Sample metadata (e.g., genomics, metabolomics, proteomics);
 - iii. Assay metadata;
 - iv. Assay files metadata (e.g., file sizes, file names, directories); and
 - v. Biosample metadata.
- E. The metadata search page must support the ability to search for data based on user specified metadata features (e.g., datasets marked “genomics” & “level 3” & “patient age > 50”).
- F. The metadata search page must support free-text metadata search.
- G. The metadata search page must support search combinations. (Search criteria from each user selection is combined.)
 - i. An example of a similar additive search utility can be found at <https://amp-pd.org/>

6. Architecture and User Workflow Preference

- A. Support a user’s “round trip” workflow through tools, materials, and services. The search features would enable users to:
 - i. Start your search selections with a list of subject ID or other relevant metadata that the user may upload. This could appear as a facet the user could turn on or off.
 - ii. Start your search selections with a list of subject IDs or other relevant metadata that are provided by another tool.
 - iii. Save your selections and come back to it later.
 - iv. Extensibility: Launch or link to other tools, providing the results of your selections as inputs to the next tool.
 - a. The data download tool or visualization tool.
 - b. External applications (biorepository for samples).

7. Download Tools

- A. The system will provide a simple Graphical User Interface standalone app or web interface that provides:
 - i. A thin user interface to the download functions;
 - ii. Authentication of an individual's User Account;
 - iii. Specification of a destination directory;
 - iv. A message showing total download size to be transferred;
 - v. A progress bar depicting bytes transferred/remaining and text display of number of files transferred/remaining;
 - vi. A download recovery option for partial/failed downloads; and
 - vii. A message pane to display statuses and error outputs.
- B. The system will provide a simple-to-use command line executable that provides the same capability as the GUI to support computational biologists, systems, and scripts.
- C. The download tool must:
 - i. Support cloud to local download and cloud to cloud download;
 - ii. Inspect the destination disk for adequate disk space and inform the user;
 - iii. Provide controls for recovering failed downloads at the file level; and
 - iv. Provide instructions to the user and links for support.

8. Visualization Tools

Researchers would benefit from the ability to visualize complex data. Specifically, the ability to review metadata for each dataset and inspect data in a tabular, graphical, and mapping view. Below are examples of common visualizations that would be of interest.

- A. Omics heat map for all genes or a gene list/symbol, and view all available data in one comprehensive overview, likely including proteomics and metabolomics.
- B. Compare a selected feature (gene/protein/geneset/etc.) for a selected data type across multiple groups of samples.
- C. View one or many genes/proteins/gene sets for a selected data type, with or without patient subgroups:
 - i. Quickly check data for expression of particular genes; and
 - ii. Determine pathway activation (e.g., data from GO enrichment)

9. Governance

- A. Auditing of downloads and data queries. All interface activities are logged.
- B. Must be able to track all data downloads by user account.
- C. Only an administrator should be able to delete portal data; this activity requires some fail-safes and confirmations.

10. Security and Availability

- A. All interface operations must be secured (e.g., SSL/TLS, MFA/2SV).
- B. The solution should have a disaster recovery plan.
- C. The solution will pass vulnerability scans (e.g., Nessus).

11. System and Performance Testing

- A. All portal user interfaces, APIs, and tools must be tested with representative metadata and download datasets:
 - i. A representative metadata dataset includes metadata for each of the following files:
 - a. Participant clinical data;
 - b. Omics sample metadata;
 - c. Biosample metadata; and
 - d. File metadata.
 - ii. A representative download dataset includes the full download of all files.
- B. Local download tests must include at least one 200 GB concurrent download test to US West and US East locales.
- C. Website content pages must load in 2 seconds or less.
- D. Website interactive content pages must load in 5 seconds or less.
 - i. Click responses in a loaded page must complete in complete in under 1500 milliseconds.
 - ii. Click response in a loaded page should average 750 milliseconds or better.
- E. All tests must log test name, command, start and end time metrics, and be packaged in GitHub with release notes.

12. Source Control, Documentation, and Reports

- A. All provided interfaces and tools are clearly documented.
- B. Any system interface APIs include test clients for all operations.
- C. Data access reports, system uptime reports, usage (by user, site).
- D. All installation, configuration, and deployment steps are clearly documented.
- E. All system requirements and input file requirements are clearly documented.
- F. All software source code, configuration files, devops tools, test clients, test data, test files, and supporting documentation must be version controlled and stored in a managed GitHub repository.
- G. High level architecture, component architecture, and software documentation must be documented, updated, and packaged with each release.
- H. Release notes denoting changes to software, devops procedures, configuration files, and dependencies must be documented, updated, and packaged with each release.

13. Support Services

- A. Software Maintenance: Would like the software, platforms, and tools that are used in the solution to be mature, robust, and/or have an active support community to fix defects and develop enhancements.
- B. Solution Maintenance: Would like the overall solution to be maintained by a solution vendor or if not fully supported, have adequate documentation/on call support so that an administrator can maintain the solution.

14. Licensing and Cost Management

- A. While open-source technologies are not required, would like the solution and its components to be free of long-term licensing costs.
- B. Would like to limit the cost incurred by user activities like large file downloads and big queries.