Support for Interoperability Projects of the NIH Cloud Platform Interoperability Program

Key Dates

- **Publication date: March 3, 2023**
- **Letters of Intent (LOI) Due Date: Monday, April 3, 2023** by 5:00 PM local time of applicant organization.
- **Proposal Questions Due Date: Monday, May 1, 2023** by 5:00 PM local time of applicant organization.
- **Proposal Due Date: Monday, May 15, 2023** by 5:00 PM local time of applicant organization. Late submissions to this ROA will not be accepted.
- **Earliest Start Date**: **July 31, 2023**

Purpose

To fund *NCPI interoperability projects*, which are technical development projects that test or implement tools, approaches, or capabilities with the aim of facilitating researcher-driven analyses of datasets across multiple NIH cloud-based platforms and repositories to foster research and development ultimately helping patients and families.

Background

Connecting NIH's many data systems is a critical step toward improving researchers' access to all types of data. NIH has invested significant resources into the generation of large-scale datasets, such as human genomic, clinical, and imaging data. Many of these data are stored in different cloud based repositories that are stewarded by multiple ICs across the NIH. This can cause researchers to struggle to find, access, aggregate, and co-analyze datasets across different data repositories. For example, when a researcher is interested in studying a specific disease, pertinent data may reside on multiple cloud repositories. It is difficult to discover where all the relevant data reside, and there is often no efficient mechanism to combine data that are not all hosted by the same repository or to perform an integrated analysis. Doing so on a case-by-case basis is costly, time-consuming, and requires expertise in cloud computing and computer programming.

The goal of the NIH Cloud Platform Interoperability (NCPI) program is to enable a federated data ecosystem that will facilitate researcher-driven analyses of datasets across multiple cloud-based platforms and repositories. This will be accomplished through testing and implementing standards and approaches for systems interoperability and universal authentication. To achieve this, technical cohesion among the participating platforms will need to be maintained while solving cutting-edge interoperability and cloud engineering problems. Current work is driven by key scientific data interoperability projects that cut across the Analysis Visualization and Informatics Lab-space (AnVIL), supported by the National Human Genome Research Institute (NHGRI), BioData Catalyst (BDC) supported by the National Heart, Lung, and Blood Institute (NHLBI), the Cancer Research Data Commons (CRDC) supported by the National Cancer Institute (NCI), the Gabriella Miller Kids First Data Resource (Kids First) supported by the NIH Common Fund, and the dbGaP/SRA supported by the NIH National Center for Biotechnology Information (NCBI). Empowering researchers to co-analyze data from these systems will accelerate biomedical research in pursuit of improving public health.

Research Objectives

NCPI interoperability projects are technical development projects that test or implement tools, approaches, or capabilities with the aim of facilitating researcher-driven analyses of datasets across multiple NIH cloud-based platforms and repositories. Interoperability projects should result in the development and/or adoption of foundational, extensible technical tools, approaches, or capabilities for a connected NIH data ecosystem, and ultimately lead to improved Findability, Accessibility, Interoperability, and Reproducibility (FAIR) of data.

Potential areas of interest underlying the *interoperability projects* may include but are not limited to: the ability to search across the NCPI partner systems to different NIH repositories; executing the same analysis workflows on different platforms and validating the workflow equivalency; testing new IT standards to foster cross-platform analysis; creating resources for estimating cloud computing costs for popular workflows in various platforms; and improving semantic and syntactic interoperability to support a specific combined analysis.

Personnel from partner systems who participate in approved *interoperability projects* are expected to participate in and contribute to relevant NCPI activities including but not limited to NCPI-wide workshops, training sessions, and working groups relevant to their *interoperability project* and use case. PIs of partner systems represented in active interoperability projects are also expected to serve on an NCPI Leadership Group that will coordinate closely with the Administrative Coordination Center (ACC) and NIH staff.

Letter of Intent

Interested applicants should submit a Letter of Intent (LOI) outlining the following:

- A cover page that includes:
    - the name and email addresses for the Contact PI and the Authorized Organizational Representative.
    - the names and email addresses of any co-PI and their institutions
    - the list of the NCPI partners involved in the project (e.g., BDC), the specific platform (e.g., PIC-SURE) and any non-NCPI collaborating site.
- An overview of the planned activities and approach. (This section should be no more than 3 pages.)
- Estimated overall annual budgets, for all 2 years of the award, broken out by institution.

LOIs will be reviewed by NIH staff for informational and planning purposes only.

Letters of intent must be submitted by email as a PDF attachment to ncpi-info@od.nih.gov.

Proposal Submission Information

Proposals must:
- Be submitted as a PDF document and Excel budget request using the template provided to ncpi-info@od.nih.gov.
- Follow the page limits described below.
- Align to NCPI's mission to facilitate researcher-driven analyses of datasets across multiple cloud-based platforms and repositories.

- Involve at least two current NCPI partner systems (or "partner system"). Additional systems that are not currently part of NCPI may be involved in the proposal as well.
  - o Current NCPI partner systems include: AnVIL, CRDC, Kids First, BioData Catalyst, and dbGaP/SRA.
  - o All system partners participating in a given interoperability project must work together to compose a proposal. This includes NCBI staff for proposals that involve dbGaP/SRA. Each partner system may participate in multiple interoperability projects. NIH encourages and expects systems and platforms to prioritize the proposals in which they participate based on the potential for substantive involvement and the potential impact on the performance and mission of the system.
- Propose feasible, justifiable, and focused projects to ensure outcomes that meet NCPI goals.

Content

Proposals should explicitly include:

A Title page (1 page max)

An abstract / summary (1 page max)

A proposal narrative of no more than 10 pages containing the following:

A clear **scientific use case** that serves as the motivation for the interoperability project and addresses researcher needs. The use case should include:

- A detailed description of the use case, the motivating **scientific question**(s), and potential scientific impact.
- The name of the **researcher** who proposed and would benefit from successful implementation of the use case. (This person should also be named as personnel on the proposal).
- The **datasets** involved, including the systems in which they reside and the dataset level identifiers (e.g., PHS accession numbers) if feasible and appropriate. Note that interoperability projects must ultimately improve the use of publicly released data, shared with the larger research community or broader public, including data that may be controlled access.
- The **repositories, platforms, or other components** (e.g., portals, data services) of the NCPI partner systems that need to be tested/improved/updated or implemented to address the use case.

A detailed description of the **technical approach** for the interoperability project that will bring the use case to fruition, including the following:

- The **technical goals** of the interoperability project and how this supports the use case and aligns with mission to facilitate researcher-driven analyses of datasets across multiple cloud-based platforms and repositories
- A description of how the **datasets** involved will be used by developers to achieve the technical goals, including maintaining the security of the dataset and adhering to all NIH and/or system requirements. Developers requesting access to controlled access data from other repositories will use that repository's data access request processes used by researchers and their institutions, which may include and are not limited to those used by dbGaP and DUOS.

- **System diagrams** that reflect both the current state and, if applicable, planned changes to systems. A description of standards that will be tested and/or adopted to facilitate the use case and a rationale for using these standards.
- **Plans for testing** the use case that aligns with all relevant security requirements (e.g., RAS ISA, requirements of each system).
- A **timeline** for completing interoperability projects and at least **quarterly milestones and deliverables** including a description of dependencies, foreseeable risks, alternative plans, and risk mitigation strategies for achieving milestones.

Plans for developing public **documentation**. All interoperability projects must be well documented to capture lessons learned and to enable other systems to test or implement.

A brief description of the **teaming arrangements and personnel** involved in the project including developers and researchers and the role they will play. At minimum, there should be two technical leads (one from each system) and one research lead for each project. (Additional information about personnel should be provided as biographical sketches in the appendix.)

A plan for **communicating** and **coordinating** with the NCPI Administrative Coordinating Center (ACC) and across NCPI and the public, including a proposal for how the project plans to engage current NCPI working groups, build off existing use cases, and share lessons learnt to reduce the potential for duplication of efforts. Describe any plans for creating and disseminating relevant training materials in collaboration with the ACC. Approved *NCPI interoperability projects* are expected to provide semi-annual progress updates at NCPI-wide meetings to be coordinated by the ACC, and recipients may be asked to present in broader NIH-wide or public settings.

An appendix with biographical sketches for key personnel

Budget
- Proposals are expected but not required to have total direct costs of $500K or less, not including requests for cloud credits. Requested budgets above $500K should be commensurate and justified by the number of participating platforms and the complexity of the proposed scope.
- Budget request for NCBI should be included in the overall request as a performer institution. NCBI will be supported directly by ODSS through the co-funding process rather than as a sub-award from the ACC.
- Each project should submit separate Excel documents for each participating institution. The file name of the Excel document should reflect the institution name for each separate budget proposed.
- Key personnel are expected to attend one in-person meeting for the NCPI per year. The budget request should include funds for travel, as appropriate.

Eligibility Information
- Proposals must involve at least two current NCPI partner systems. Additional partner systems who are not currently part of NCPI may be involved in the proposal as well.
  - Current NCPI partner systems include: AnVIL, CRDC, Kids First, BioData Catalyst, and dbGaP/SRA.
  - All system partners participating in a given interoperability project must work together to compose a proposal. This includes NCBI staff for proposals that involve dbGaP/SRA.

- o Each partner system may participate in multiple interoperability projects.

Evaluation Criteria and Review Questions

- To what extent does the use case aim to address a real scientific need identified by the research community?
- To what extent are the technical goals aligned to the NCPI's mission to facilitate researcher-driven analyses of datasets across multiple cloud-based platforms and repositories?
- Is the technical approach feasible and adequate to fulfil the proposed use case? Why or why not?
- Are the milestones and deliverables likely to result in completion of interoperability projects within the proposed timelines? Are the plans for addressing dependencies and potential risks appropriate?
- Are the technical personnel appropriate for the project? Does the proposal plan for sufficient time and effort from each person to achieve the milestones and deliverables?
- To what extent are the proposed standards appropriate for meeting the scientific and technical goals?
- To what extent do the plans for testing adequately address the technical goals and security requirements?
- To what extent is the plan for documenting, communicating, and coordinating across NCPI sufficient to share lessons learned and reduce duplication of efforts?

Other Information

- This is a one-time solicitation, non-recurring solicitation.
- Funding for all partner systems except NCBI will be provided as sub-awards from the NCPI ACC Other Transaction award. Sub-awards are supported on a cost-reimbursable basis.
- Project period/duration: Up to 2 years
- NIH will make all funding decisions.

The NCPI ACC will:

- Monitor progress, project timelines, and milestones and deliverables within and across approved projects.
- Provide support for resolving technical and project management challenges. Raise any significant challenges or risks to the projects to the NIH in a timely manner.
- Coordinate the allocation of cloud credits to project members.
- Coordinate presentations and updates on interoperability projects progress.
- Coordinate the timely sharing and posting of technical documentation provided by interoperability projects and ensure information is up-to-date.
- Collate and share training materials generated as products of interoperability projects
- Advise on potential technical solutions, as needed.
- Establish subawards and report on projects' spending
- Provide a synthesis of interoperability projects' progress reports, timeline and milestone updates and other information to the NIH on a regular, and an as-needed basis. This information will be used by NIH staff when making decisions regarding the approval, continuation, and termination of work associated with interoperability projects.

- Facilitate the onboarding of new NCPI partner systems who are part of approved interoperability projects.

Inquiries

Inquiries regarding the purpose and scope of this announcement should be directed to: ncpi-info@od.nih.gov

Inquiries regarding the role and support from the NCPI Administrative Coordinating Center (ACC) should be directed to: ncpi-acc-info@rti.org