Participating Organization(s)	National Institutes of Health (NIH)
Components of Participating Organizations	National Center for Advancing Translational Sciences (NCATS)
Funding Opportunity Title	**Biomedical Data Translator: Technical Feasibility Assessment and Architecture Design Projects (OT3)**
Activity Code	OT3 Multi-component Research Project - Other Transaction Award
Announcement Type	New
Related Notices	N/A
Funding Opportunity Announcement (FOA) Number	OT-TR-16-001
Catalog of Federal Domestic Assistance (CFDA) Number	93.350
Number of Applications	Only one application per institution (identified by a DUNS number) is allowed.
Funding Opportunity Purpose	These two-year awards will support data assessment and feasibility analyses for building a data "Translator" that will integrate multiple types of existing data sources, including objective signs and symptoms of disease, drug effects, and intervening types of biological data relevant to understanding pathophysiology. The purpose of the Translator will be to accelerate biomedical translation and to facilitate the generation of new hypotheses for understanding and treating disease.
Objective Review	Objective review will be conducted in two phases. In the first phase, written applications will be reviewed within 2 weeks of the application due date. A subset of those applicants will be invited to present their applications to a review panel after which awardees will be selected. Applicants will not receive written feedback from the reviews.

Application Due Date	June 1, 2016
Earliest Start Date	September 1, 2016
Funding Instrument	Other: An assistance mechanism that is not a grant or cooperative agreement. Other Transactions awards are subject to the requirements of the Other Transaction Award Policy Guide for the Biomedical Data Translator Program: Technical Feasibility Assessment and Architecture Design Projects. Applicants may review this policy guide at the link above.
Eligibility	See Eligible Applicants section of this announcement
Funds Available and Anticipated Number of Awards	NCATS intends to commit $5,000,000 in FY 2016 to fund 2-5 awards. Future year support is contingent upon the availability of funds.
Award Budget	Application budgets are not limited but need to reflect the actual needs of the proposed project. NCATS anticipates a minimum budget of $1,000,000 total costs per year will be requested per application. Because the nature and scope of the proposed research will vary from application to application, it is anticipated that the size of each award will also vary. Performance will be evaluated quarterly and continued support will be based on the outcome of those evaluations.
Award Project Period	Up to 2 years
Funding Opportunity Expiration Date	June 2, 2016

Overview Description

Issue/gap being addressed: Clinicians and biologists think of disease in different ways, and speak different languages. Physicians diagnose and treat disease based on signs and symptoms affecting specific target organs. Clinicians will use whatever tools they have at their disposal to pinpoint the cause of the symptom to ensure the most appropriate treatment is administered. Unfortunately, the molecular aspects of disease are rarely taken into account for diagnosis and treatment, because the tools to use that information are often not available. In contrast, biomedical researchers think of disease in terms of molecular changes in specific proteins, pathways or cell types. In fact, biomedical researchers

do not often study diagnosis and treatment, but rather focus on molecular and cellular processes that are hypothesized to be involved in the pathophysiology.

As a result of recent scientific advances there is a tremendous amount of biomedical research data and data from disease classifications, health records, clinical trials and adverse event reports that are available and could be useful for understanding health and disease and for developing and identifying treatments for diseases. Ideally, one would mine these data collectively to gain insights into the relationship between molecular and cellular processes (the targets of rational drug design) and the signs and symptoms manifested in diseases. Unfortunately, these very rich, but different, data sources exist in different locations, and often in different forms that are not compatible or interoperable with each other. All of these factors limit our ability to get more treatments to more patients more quickly.

Overall goal of the Biomedical Data Translator Program: The purpose of the program is to accelerate biomedical translation by developing a biomedical "data translator" for the research community. NCATS anticipates that this will be a multi-year, iterative effort with the eventual development of a comprehensive, relational, N-dimensional Biomedical Data Translator that integrates multiple types of existing data sources, including objective signs and symptoms of disease, drug effects and intervening types of biological data relevant to understanding pathophysiology. Each data type (see Appendix for examples) must be comprehensive (e.g., all diseases, all pathways, all SNPs). It must be possible for a user to enter the Translator from any data type and identify all cognates/connections in any other data type. In doing so, we will enable a shift from the current symptom-based diagnosis of disease to disease classification that is based on a set of molecular and cellular abnormalities that can be targeted by various preventative and therapeutic interventions. The Translator must be dynamic, i.e. able to incorporate new data and information as it becomes available, and eventually be able to support machine learning to generate prospective hypotheses. The Translator must be open source and completely publicly available for any user, without requirement to subscribe to proprietary resources or tools to be used.

NCATS is aware of the many existing efforts which catalogue or connect individual data types. While these provide proof of principle, the Biomedical Data Translator will be greater in scope, with the intention of revealing potential relationships across the spectrum of data types from signs and symptoms to molecules. Examples of the types of queries the Biomedical Data Translator will allow for the first time could include, but are not limited to:

- Tell me every disease that has symptom X, and/or affects a particular cell type.
- Tell me all molecular pathways which, when perturbed, lead to malfunction of a particular organelle in a particular organ in people with the following genomic characteristics.
- Tell me all the drugs currently approved only for investigational use that perturb any pathway that is dysfunctional in diseases characterized by this clinical sign.

This effort will require unprecedentedly broad teams of experts to work together in a highly collaborative manner with active program management. Input from clinicians during the design and feasibility assessment will be critical to ensuring appropriate inclusion of clinical data.

Intent of this FY2016 funding opportunity, Biomedical Data Translator: Technical Feasibility Assessment and Architecture Design. As the first activity of the Biomedical Data Translator Program, NCATS is using its Other Transactions authority to invite interested applicants to submit innovative proposals for addressing the architecture needs to build a biomedical data "Translator" and to assess its technical feasibility. Successful applicants will not only contribute their expertise and resources but also must be willing to collaborate to revolutionize translational science and propel new discoveries and best practices for practitioners across the translational spectrum, from biologists to chemists to computer scientists, from scientists doing target validation to clinicians seeing patients.

Key Events	Dates	Action needed by applicants
Call for projects posted	April 29, 2016	
Project applications due	June 1, 2016	Email completed application by 5pm local time
Review of written applications completed	June 14, 2016	
Invitations to present in person sent out	June 15, 2016	
Responses to invitations	June 17, 2016	Accept or decline invitation to present
Presentation by invited candidates in Bethesda	June 29-30, 2016	*Candidates and team attend in person
Negotiations begin	July 2016	
SAM and DUNS number submitted	July 15, 2016	**Candidates e-mail their DUNS number and SAM account information
Awards announced	September 2016	
Kick-off meeting in Bethesda	October 2016	Travel to Bethesda

***Presentation in person by at least one team member is required. NCATS may provide limited travel support.

**DUNS and SAM number registration can take 6 weeks or more. Candidates should begin the registration process at least 6 weeks prior to this deadline to ensure completion in time to provide these to NCATS.

Data assessment and infrastructure design

During the 2-year project period, awardees are expected to identify high-value data sources that would be needed for a comprehensive Translator and to design the infrastructure that would be needed to integrate those data.

NCATS expects that this project will be intensely collaborative among research partners and NIH staff, and that the unrestricted exchange of source code and software tools written as part of this program will be essential to a successful outcome. We intend to require participation in a program-directed

source code repository and the publication of that source code, without undue burden on its reuse by others.

NIH believes that data sharing is essential for expedited translation of research results into knowledge, products, and procedures to improve human health. The NIH expects and supports the timely release and sharing of final research data, software and tools that support the publication of these data from NIH-supported programs for use by other researchers. The goal of this programmatic effort is to produce data, software and tools that are open source and completely publicly available for any user, without requirement to subscribe to proprietary resources or tools to be used.

Goals for the 2-year projects

- Identify high-value data sources that should be included in a comprehensive Translator and a means to evaluate validity and utility of those data.
- Develop a plan for integrating across a comprehensive variety of data types.
- Identify integration barriers (e.g., entity recognition, entity resolution, entity stitching, etc.) or data inclusion barriers.
- Develop and test a plan for data quality control and data updates.
- Develop a demonstration project that illustrates innovative integration of as many disparate data types as possible; however, the novelty of the integration of different data types is more important than the number of data types integrated.
- Define the requirements for a comprehensive Translator, including its architecture and development path, that will catalyze getting more treatments to more patients more quickly.

Quarterly face-to-face meetings and collaboration

The initial 2 years should also be seen as a period of team and capacity building through collaborations among the awardees and with NIH staff. There will be quarterly face-to-face events. Awardees will be required to participate in a series of NCATS-organized 2- or 3-day workshops, occurring roughly quarterly, at which participants will critically examine key assumptions and perceived roadblocks of the program. Each meeting will focus on specific needs for developing a Translator (for example, identifying data types, data sources, and quality control of the data) with the goal of sharing ideas for strategies to address the needs and to also identify challenges and barriers to development.

The meetings will help assess the functional capacity of different architectural approaches, the quality and depth of program components and data sources and specific aspects of the data lifecycle such as curation. General topics of each workshop will be announced in advance, and different types of expertise may be especially relevant at different workshops. Specific topic details will be revealed at the workshop. The goal of these workshops is to collaboratively work through different approaches to these topics, arrive at a common understanding of the requirements around a topic, and then continue the development of awardee-specific ideas in the intervening time between workshops. Each work stream will culminate in a collaborative whitepaper on the subject.

Applicants should budget for travel of their teams to these quarterly meetings.

Demonstration projects

While each awardee will be designing a comprehensive Translator that includes a wide variety of data types, the actual development of such a comprehensive system will not be feasible within the constraints of this funding opportunity. However, the development of a demonstration project(s) will be critically important as part of the feasibility assessment of a larger initiative. Therefore, each project plan should include a brief description of a high-value Translator demonstration project and how it addresses an existing challenge, gap or barrier to the development of the Translator project.

Application Content

All pages should be Arial 11pt, single space with 1" margins.

Cover Page

Please use the PHS398 face page form.

Summary Vision Statement

A single page describing:

1) The types of research questions that such a resource will enable (see examples in the overview);

2) How the expertise and resources that the applicant can provide will be used to collaborate with NIH staff and other awardees to address project goals and integrate into the Translator architecture; and

3) Potential data and infrastructure challenges, gaps and barriers to developing this resource and achieving the project goals.

Project Plan

The project plan should clearly describe how the project goals will be achieved and is not to exceed 5 pages. Any graphs, pictures or data tables must be included in the body of the text and will count against the 5-page limit.

Personnel

Include a CV or abbreviated NIH biosketch that is no more than 1-page for each of the key personnel who have committed to participating in the project if it is awarded. In the context of this program, it is especially important to highlight contributions of personnel to existing open source projects, standards, and initiatives as well as evidence of ability to work collaboratively.

Provide a table listing all personnel, including all to be named personnel, role on the project, and percent effort to be committed. Project PIs should commit a minimum of 25% effort. Do not include salaries here.

Resources

Include a 1-page description of resources available for the project.

Data Sharing and Collaboration Plans

Applicants must include a plan for data sharing and a collaboration plan up to 1 page total. If data sharing is not possible, state why. If data sharing is limited, the applicant should explain such limitations in the data sharing plan, including their potential impact on the proposed project. NIH's data sharing policy may be found at the following website:
http://grants.nih.gov/grants/guide/notice-files/NOT-OD-03-032.html

Budget

Provide 1 page stating the overall cost for each of the following: personnel, travel and other. Institutions with an established F&A rate may include the approved rate. Institutions without an established F&A rate may propose a rate for review and negotiation.

Provide a description of semi-annual milestones for the 2-year project period and associated costs (total cost per milestone) in up to 2 pages.

Letters of support

Letters of support for all collaborators/other significant contributors, consultants and sub-award sites should be provided. Letters that are written for the sole purpose of endorsement should not be included and will not be considered during review.

How to submit the application

Complete applications must be emailed to translator-applications@nih.gov by 5:00 p.m. local time, June 1, 2016. Applications must be submitted in text-recognizable PDF (Adobe) format and file size must be no greater than 20MBs. Paper applications will not be accepted. Applications from institutions must be submitted by an authorized organizational representative.

Objective review process

Applications will be evaluated in two stages. The first stage will be based on a written application. A subset of those applicants will be invited to present their concept in person. The in-person presentation, given by at least one member of the team, is required to be eligible to receive an award.

The evaluation will be based on:

- The plan for data gathering of high value data sets across domains from patient data to molecular data, and data assessment to ensure quality of the data to be used in developing a comprehensive system;
- The plan for developing a demonstration project;
- Past performance and expertise of the team members and complementarity with other awardees;

- The potential impact of the team's vision statement if it were successfully implemented; and
- The adequacy and appropriateness of the budget, resources, and data sharing and collaboration plans.

The evaluation will be conducted by an appropriate review group convened by NCATS and will include federal government staff reviewers.

Written feedback about the outcome of the objective reviews will not be provided.

Eligible Applicants

Applications will be accepted from individuals and organizations.

Individuals

Any individual(s) with the skills, knowledge and resources necessary to carry out the proposed project is invited to develop an application for support. Individuals from underrepresented racial and ethnic groups as well as individuals with disabilities are always encouraged to apply for NIH support.

Applicants may be subject to financial analysis and risk assessment conducted by NIH staff.

Organizations

Higher Education Institutions

- Public/State Controlled Institutions of Higher Education
- Private Institutions of Higher Education
- The following types of Higher Education Institutions are always encouraged to apply for NIH support as Public or Private Institutions of Higher Education:
  - Hispanic-serving Institutions
  - Historically Black Colleges and Universities (HBCUs)
  - Tribally Controlled Colleges and Universities (TCCUs)
  - Alaska Native and Native Hawaiian Serving Institutions
  - Asian American Native American Pacific Islander Serving Institutions (AANAPISIs)
- Nonprofits Other Than Institutions of Higher Education
- Nonprofits with 501(c)(3) IRS Status (Other than Institutions of Higher Education)
- Nonprofits without 501(c)(3) IRS Status (Other than Institutions of Higher Education)

For-Profit Organizations

- Small Businesses
- For-Profit Organizations (Other than Small Businesses)

Governments

- State Governments
- County Governments
- City or Township Governments
- Special District Governments
- Indian/Native American Tribal Governments (Federally Recognized)
- Indian/Native American Tribal Governments (Other than Federally Recognized)
- U.S. Territory or Possession
- Eligible Agencies of the Federal Government-NIH Intramural Program

Other

- Independent School Districts
- Public Housing Authorities/Indian Housing Authorities
- Native American Tribal Organizations (other than Federally recognized tribal governments)
- Faith-based or Community-based Organizations
- Regional Organizations

Multiple Principal Investigators

More than one individual may be named as Principal Investigator on a single application.

Foreign Institutions

Non-domestic (non-U.S.) Entities (Foreign Institutions) **are not** eligible to apply.
Non-domestic (non-U.S.) components of U.S. Organizations **are** eligible to apply.
Foreign components, as defined in the Other Transaction Award Policy Guide for the Biomedical Data Translator Program: Technical Feasibility Assessment and Architecture Design Projects, **are** allowed.

Questions about this opportunity should be emailed to translator-questions@nih.gov

Appendix: Example Data Types

The list below gives a sense of the breath of data types to be considered for the Biomedical Data Translator, but is not exhaustive.

Diseases
Signs of disease
Symptoms (from patient registries and natural history studies)
Patient reported outcomes
Electronic health records (EHRs)
Clinical encounters
Prescription data
Health insurance claims data
Diagnostic labs
Biomedical imaging data
Adverse Event Reports
Biomarkers
Organ systems
Sub-anatomy
Tissue types
Cell types
Cell lineages
Cell processes
Organelles
Clinical trial data
Orthologs/Animal models
Microbiome data
Molecular mechanisms
Signaling pathways
Molecular and cellular networks
Environmental factors
Disease etiologies
Proteomes
Proteins
Post translational modifications (PTMs)
Co-factors
Transcriptomes
Epigenomes
Genes
Genetic mutations
Functional polymorphisms
Therapeutic interventions
Intervention exposure
Pharmacokinetics/Pharmacodynamics