

**U.S. Department of Health and Human Services
National Institutes of Health (NIH)
Office of the Director
Division of Program Coordination, Planning, and Strategic Initiatives**

**Bridge to Artificial Intelligence (Bridge2AI) Town Hall
June 9, 2021**

**Draft Summary
Discussion Points, Highlights, and Action Items**

I. Welcome

Knowinnovation (KI) staff welcomed participants to the NIH Bridge2AI Program Town Hall and noted the upcoming team building events: Microlabs and Expo. Today’s agenda included an introduction and overview, question and answer session, with short breakout groups, touching on understanding the spirit of the modules, and next steps.

- Attendees were provided links to the [Slack channel](#) and [KISstorm platform](#), which will be used for breakout sessions and networking during the Bridge2AI meeting series.
- Attendees were reminded that technical support is available through Zoom (by searching for “KI” in the participant list), Slack, and KI’s support page.

II. Bridge2AI Introduction and Overview

Grace C.Y. Peng, Ph.D., Program Director, National Institute of Biomedical Imaging and Bioengineering (NIBIB), also welcomed participants and provided an introduction and overview of the Bridge2AI Program.

- The new Bridge2AI Integration, Dissemination, and Evaluation (BRIDGE) Center [funding opportunity announcement](#) was published on June 8, 2021, and a full BRIDGE Center video will be posted to KISstorm shortly.
- The Bridge2AI Working Group includes more than 100 members from 30 NIH Institutes, Centers, and Offices, as well as programs from other government agencies.
- The program is intended to foster a culture change in the research community to generate new data that can be interpreted without human inferences and human design hypotheses.
- These data would possess attributes that would encompass the context in which the data are collected, as well as the conditions and the environment. The projects should instill a culture of ethical inquiry, rather than compliance, for the data that are being collected.
- The data should meet standards for findability, accessibility, interoperability, and reusability (FAIR), as well as adhere to transparency, responsibility, user focus, sustainability, and technology (TRUST) principles.
- The program goals are to use the biomedical and behavioral Grand Challenges to motivate the generation of data sets, emphasize ethical best practices, prepare the data in an artificial intelligence (AI)/machine learning (ML)–friendly way, and promote diverse teams. The outputs will be standardized data attributes, automated tools and cross-training tools for workforce

development, and dissemination of products and best practices to the community for long-term use.

- The Grand Challenges will motivate the collection of data that would require multiple modes of data types, as well as multiple spatial temporal scales from the behavioral and biomedical domains to predict health outcomes.
- The Bridge2AI program is focused on the data preparation part of the cycle.

A. Welcome from the NIH Institute and Center Directors (Video)

- Patricia Flatley Brennan, Ph.D., RN, Director, National Library of Medicine, stated that the program's unique approach will accelerate the research community's ability to use AI techniques for discovery. For this outcome, a wide range of perspectives and skills is needed. The approach will help improve understanding data on individuals, family and household behavior, and community activities (e.g., for predicting "hot spots" of infection during the COVID-19 pandemic).
- Michael F. Chiang, M.D., Director, National Eye Institute, stated that most existing clinical systems are used to diagnose specialized conditions. In practice, however, clinicians combine images with clinical, demographic, and genetic data to perform diagnoses. Most AI systems cannot answer complex questions by integrating multiple data types. Data sets should be developed that allow researchers to expand from one disease to multiple diseases and to incorporate clinical data from electronic health records, as well as genetic data.
- Eric Green, M.D., Ph.D., Director, National Human Genome Research Institute, reflected that the study of genomics has yielded new insights on human health, but many questions remain (e.g., understanding the roles and relationships of genes, regulatory elements, pathways, networks, and human disease). AI solutions will be crucial for addressing genomics-related challenges and for advancing genomic medicine.
- Helene Langevin, M.D., Director, National Center for Complementary and Integrative Health, noted that the research community's understanding of health has failed to keep pace with advancements in the understanding of disease. New tools and methods are needed to obtain a comprehensive understanding of health in individuals. AI is well suited for this task.
- Bruce J. Tromberg, Ph.D., Director, NIBIB, stated that the bioengineers are committed to the concept that biological processes can be represented symbolically by mathematical equations that have explainable free parameters. Observational data, however, often are limited and imperfect. With new observations and computational approaches, AI will help researchers deduce new laws that describe the emergent properties of complex biological systems.

III. Questions and Answers

A. Brief the Breakouts and Breakouts for Question Generation

Attendees were assigned to random breakout rooms to discuss thought-provoking questions that the NIH should address.

B. Discussion Points

- Several attendees asked whether the program would allow data collection for multiple diseases and data types. Gene Civillico, Ph.D., commented that many existing applications are limited by

the narrowness of scope. The program is intended to allow researchers to consider multiple diseases and to analyze broader data sets.

- Several attendees asked about using existing data versus creating new data. Dr. Peng stated that new data generation can be motivated by the Grand Challenges; standards can be harmonized, and tools and standards can be substantiated and made publicly available. Existing data sets might be more challenging to incorporate and harmonize. AI/ML approaches could be used to address those data sets. Bridge2AI is focused on the generation of new data. Dr. Peng also stated that the creation of synthetic data also could be considered.
- Several attendees expressed concern about retrospective consent and associated ethical and legal issues surrounding the new applications of data. Shurjo K. Sen, Ph.D., stated that the answer to this question is complex. The associated groups must develop solutions for retrospective consent as the research continues to progress.
- An attendee asked about the definition of scope without an underlying hypothesis. Lanay M. Mudd, Ph.D., pointed out that this concept falls within the program's Grand Challenge and will require multilevel data sets across spatiotemporal scales. Dr. Sen stated that the Grand Challenges offer flexibility for applicants to develop questions that are broad and challenging enough that AI approaches should be employed.
- Several attendees asked about the development and dissemination of standards in data generation projects. Dr. Peng reiterated that the program is intended to consider and substantiate standards through an interoperability process. The standards would be considered in the corresponding module and implemented in tools that are in development. The existing standards represent a topic of debate within the research community. Converters could be applied for different data sets, or consensus on one set of standards could be reached. Dr. Mudd also commented that the Slack channel includes module-specific pages for further discussion.
- An attendee asked about the scope of the BRIDGE Center's responsibilities. Dr. Sen emphasized the importance of preserving objectivity to examine the program, rather than individual projects. The evaluation aspect will consider both data generation and tool development. Ethical practices also should be considered.
- Several attendees asked about the number of Grand Challenges being considered by the NIH. Dr. Mudd encouraged the attendees to propose their own Grand Challenges and to develop new ideas. She added that NIH intends for five to eight data generation projects to be awarded funding in the first round of evaluation, depending on merit.
- Several attendees inquired about database ownership and access. Dr. Mudd emphasized that frequent data sharing will be required beginning in the second year of funding. The databases will be publicly available. James Gao, Ph.D., added that NIH policies for data sharing will be implemented, but data repositories will not be specified. Applicants will be expected to identify the appropriate repository for their data sets. Dr. Gao noted that controlled-access phenotypic data is stored in the database of Genotypes and Phenotypes (dbGaP). The NIH Common Fund website includes a list of appropriate repositories.
- Dr. Peng clarified that all six modules are required for the Data Generation OT2 application. She noted that more detailed application-specific questions will be addressed during the next meeting. Dr. Sen added that the U54 mechanism for the BRIDGE Center allows researchers to apply for a subset of the cores; for data generation projects, however, researchers must apply to all six modules. He clarified that the researchers do not need to respond to all six grand challenges because they were provided only as illustrative examples. Dr. Peng added that the required

diversity component is intended to encourage researchers to gain a broader perspective through collaborations to meet the module requirements.

- Dr. Sen emphasized that ethical concerns are as equally important as scientific merit alone. These topics could be addressed through new collaborations between various investigators (e.g., clinicians, trainees, postdoctoral researchers, bench scientists). Workforce and skill development are important components of the projects. Dr. Mudd informed the attendees that networking opportunities are available through KISstorm; users can be filtered and identified by their expertise. She encouraged the attendees to initiate conversations through the Slack channel and noted that additional networking opportunities will be provided in future meetings.

C. Parallel Sessions

The attendees were asked to select one of the six modules for open parallel discussions, which are accessed through KISstorm.

IV. Next Steps

- The upcoming Microlab sessions will occur in pairs. The June 14 session will address Teaming and Ethics; the June 16 session, Standards and Tools; and the June 18 session, Data Acquisition and Skills and Workforce Development. The attendees were reminded that they do not need to register for these meetings separately.
- The attendees were invited to attend an informal networking session through the Wonder platform, which can be accessed through KISstorm. The team clarified that content shared through KISstorm is visible to the event organizers; Slack is encouraged for communication between events.

Action Items

- The team will review the submitted questions and update the Frequently Asked Questions section on the NIH Common Fund website with answers.