

**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) Module Microlab3
June 18, 2021**

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

I. Welcome

Jocelyn Tejada, facilitator, Knowinnovation (KI), welcomed the participants and provided an overview of the Microlab agenda.

James Gao, Ph.D., Division of Extramural Science Programs, National Eye Institute, welcomed participants and provided an overview of Microlab 3.

- This Microlab 3 meeting covered the Data Acquisition module and Skills and Workforce Development module in separate breakout discussions. Each module breakout session began with an introduction and overview. A third integration breakout session explored the intersection of the modules. Participants had an opportunity to meet other potential applicants and to form teams during the three breakout sessions.
- Participants were provided links to the Slack channels and KISstorm and Wonder platforms and were encouraged to participate using these platforms. The Zoom platform was used for the plenary and breakout sessions.
- Participants were able to access technical support from the KI staff through Zoom and Slack messaging, as well as email.

II. Data Acquisition Module Breakout

Dr. Gao explained that the purpose of the Data Acquisition module is to implement a data acquisition plan needed to generate new data sets that are ready for artificial intelligence (AI) and machine learning (ML). Data acquisition processes should incorporate appropriate tools, standards, and ethical and trustworthy AI and ML principles. The Data Acquisition module will require domain experts in biomedical and behavioral research; data scientists with experience in big data collection and analytics; and individuals with AI and ML expertise in modeling and data analysis. The module should identify data sources and modalities; justify the sample sizes for AI and ML data analyses; and demonstrate that the data collected will be reliably used for training, testing, and validating the AI and ML models. The module also should contain well-defined data collection, processing, and curation workflows that follow ethical and trustworthy principles to ensure the data collected are equitable and protect individual's privacy. The proposed data acquisition process should adhere to adopted or developed data standards to facilitate data harmonization and interoperability. Tools should be developed to support the data collection process and ensure that data are reliable and reproducible. In addition, the Data Acquisition module should include plans for data analysis and quality control to ensure data completeness and the quality of the metadata.

Finally, the module should include a data management and sharing plan for making project data readily available to the broader research community in a timely manner.

- NIH expects the Data Acquisition module to work closely with other Data Generation Project modules, especially the Ethics, Standards, and Tools modules. The Teaming module will be important for cross-module interactions and the efficiency of the entire project team. The Skills and Workforce Development module can be used to develop workforce capabilities for employing the Data Acquisition module and to disseminate best data acquisition practices across all Bridge2AI Data Generation Projects.
- The first discussion prompt for the Data Acquisition module breakout session was, How do we incorporate active learning into the Bridge2AI data acquisition process? Dr. Gao raised the possibility of using AI and ML analysis of data collected during the Bridge2AI Data Generation Project to refine the data acquisition process as the project progresses. AI/ML analysis of Bridge2AI data would occur through yearly jamborees and hackathons with the community.
- The second discussion prompt for the Data Acquisition module breakout session was: How do we generalize the Bridge2AI data generation processes to other data collection efforts, including future efforts to generate AI and ML-ready data sets?

III. Skills and Workforce Development Module Breakout

Lanay M. Mudd, Ph.D., Program Director, Division of Extramural Research, National Center for Complementary and Integrative Health, introduced the Skills and Workforce Development module breakout session and emphasized the role of this module in supporting all the other Data Generation Project modules. The twofold purpose of the Skills and Workforce Development module is to enhance skill development while attracting and developing an AI and ML-trained biomedical and behavioral research workforce. To accomplish these goals, this module will require expertise in developing training materials and activities, training and mentoring, promoting inclusive and supportive environments, and delivering instruction in scientific rigor. The components of the Skills and Workforce Development module will include descriptions of the skills development processes used to develop, implement, and assess these activities, which also will promote the products from the other five modules. The workforce development component of the module will describe the support of career development and mentoring activities across all modules to develop transdisciplinary expertise and encourage cross-training in different areas of science and technology while developing “soft skills.” Dr. Mudd emphasized that the Skills and Workforce Development module must address how each component supports the implementation of the Plan for Enhancing Diverse Perspectives.

- The Skills and Workforce Development module will interact with all of the other modules to develop skills in the areas relevant to Teaming, Ethics, Standards, Data Acquisition, and Tools.
- The first discussion prompt for the Skills and Workforce Development breakout session was, What are some of the challenges of bridging all of the different fields to build a new AI/ML biomedical and behavioral science workforce? Skills and Workforce Development breakout participants also should consider the challenges of building the new, interdisciplinary workforce required to implement Data Generation Projects successfully.
- The second discussion prompt for the Skills and Workforce Development breakout session was, What has been your experience with developing new skills in others? What aspects should this module

consider when creating skills development activities that would be useful for team members with a broad range of backgrounds, technological capabilities, and education?

IV. Integration Breakout

Ms. Tejada introduced the integration of the Data Acquisition and Skills and Workforce Development modules.

- Discussion prompts for the Integration breakout session were as follow: How can skills and workforce development help the data acquisition processes in the data generation projects? Which module do you see yourself working on primarily? How might the other modules contribute to the work you would perform in your selected module?
- Ms. Tejada asked that participants capture insights on the KISform Integration page.

V. Upcoming Bridge2AI Sessions and Next Steps

Ms. Tejada introduced the NIH representatives and asked them to discuss Grand Challenges.

- Grace C. Y. Peng, Ph.D., Program Director, Division of Discovery Science and Technology (Bioengineering), National Institute of Biomedical Imaging and Bioengineering, reiterated that Bridge2AI Grand Challenges are motivators for creating new, ethically sourced, multimodal, multiscale, multiscreen, multidomain, and multi-disease area data sets that can be used to predict health outcomes. Grand Challenges should focus on current data gaps and problems that currently are not addressed. She asked applicants to consider current data gaps and what types of data need to be collected to respond to major biomedical and behavioral research challenges around the world. Applicants should consider the types of data that need to be collected and synchronized with other data. Dr. Peng also encouraged the collection of data for exploratory research that are not restricted by specific research questions. Applicants should consider the types of data that would need to be available 5 years from now to respond to major research challenges.
- Dr. Peng shared the KISform page, showing the Data Generation Grand Challenge criteria and instructions for how to post ideas. She explained that NIH might combine modules from different Grand Challenge proposals. Each Grand Challenge proposal should be generalizable enough to produce the data needed across the various Grand Challenges. When posting Grand Challenge ideas, participants should provide a title, the types of data that will be required, the ethical issues that might be encountered in carrying out the proposed Grand Challenge, and the types of perspectives needed on the Grand Challenge team.
- Shurjo K. Sen, Ph.D., Program Director, Division of Genome Sciences, National Human Genome Research Institute explained that participants should focus on the word “Grand” when developing their ideas and applications. Successful Grand Challenge project proposals should extend beyond the scope of any R01-funded project, regardless of the size of that project or the quality of the science. Dr. Sen encouraged participants to consider projects that would not be doable with existing data sets or tools and that would require the use of AI. He added that NIH representatives do not wish to unduly influence application ideas, and for this reason, participants should not submit ideas to the NIH representatives participating in this meeting to determine whether their ideas are appropriate for a Data Generation Project Grand Challenge.
- Dr. Sen suggested that participants imagine the Grand Challenge ideas as “Hackathon” challenges. He encouraged individuals and small groups that lack the resources and expertise to support a project involving all six modules to use the Expo to connect with other teams proposing ideas of interest to

them. All participants should be able to use the Expo to form or join a team that can support a successful Grand Challenge application.

- Dr. Gao also encouraged participants to begin developing project ideas by focusing on the field they know best. He remarked that Grand Challenges need not focus solely on areas where data are missing, but also on topics for which data are incomplete or inconclusive and could be augmented with the use of AI. Dr. Gao further suggested that participants post Grand Challenge proposal ideas for other participants to review.
- Dr. Mudd supported her NIH colleagues' remarks regarding the need for Grand Challenge projects to push the limits of existing science. She remarked that six Grand Challenge ideas presented in the OT2 Research Opportunity Announcement (ROA) should be viewed as examples of the scope of the projects that NIH is seeking for the Bridge2AI Data Generation Projects. Applicants also could build upon those previously funded projects or develop new Grand Challenge ideas within the spirit and culture intended for Bridge2AI. Dr. Mudd encouraged participants to post all ideas on the Grand Challenge page to see if the idea interests' others in community.
- Dr. Mudd explained that the Expo on June 23, 2021, will focus on team building. Some participants might wish to use the Expo to identify individuals with the expertise needed for their planned project teams. Other participants might not have an idea for a proposal but be looking for team to join. Dr. Mudd added that participants should post project ideas that might seem too small for a Grand Challenge because other participants might be able to provide input on how to expand the scope of those ideas.
- Andy Burnett, Managing Director, KI, clarified several logistical and technical issues. He explained that sticky notes simply show what participants want to discuss during the Expo. Mr. Burnett recommended that those who are submitting project ideas go to the Participants Group page on KISform to seek people with particular skills needed for their project. The Participants Group page contains profiles for everyone who has identified expertise around skills and workforce development.
- Conversations will be structured around the Grand Challenge ideas rather than modules during the Expo. KI staff, however, will establish breakout rooms where participants can engage in "Birds of a Feather" conversations focusing on more specific topics upon request. Mr. Burnett clarified that three separate sessions will be held to allow different participants to join conversations about the same proposed topics. Slack channels and Birds of a Feather Zoom breakout rooms can both be used to discuss topics separately.
- Sticky notes should be used to submit project ideas and recruit team members for those projects. By the morning of June 21, 2021, KI staff will create a page where participants can sign up for sticky note ideas that interest them.
- A participant remarked that many potential partners are available to join a relatively small number of teams. He is not interested in joining a project team but is interested in providing laboratory equipment for Grand Challenge teams. Dr. Mudd clarified that Bridge2AI Expo meetings will not center around services. Grand Challenges should focus on a specific idea rather than a service or technical capability.
- In response to a question about the size of teams, Mr. Burnett referred participants to the Frequently Asked Questions (FAQ) page on KISform. Participants also could pose this question to NIH staff on

Slack. Mr. Burnett referred participants to Dr. Sen for more information about the requirements of the U54 Bridge Center funding opportunity.

- The Bridge2AI Resources page includes links to the Bridge2AI Funding Opportunities, which are U54 BRIDGE Center FOA and the Data Generation Projects OT2 ROA, which provides guidelines on how applicants should format their proposals.

Action Items

None