

MEETING SUMMARY

National Artificial Intelligence Research Resource Task Force Inaugural Meeting

July 28, 2021

Meeting Summary

The first meeting of the National Artificial Intelligence Research Resource Task Force (NAIRR TF) was held online via Zoom on July 28, 2021, 1:40–5:00 PM EDT.

Welcome and NAIRR TF Overview

The meeting started at 1:40 PM EDT.

Dr. Erwin Gianchandani (National Science Foundation), NAIRR TF Co-Chair, opened the public meeting and introduced the agenda. Dr. Lynne Parker (White House Office of Science and Technology Policy), NAIRR TF Co-Chair, presented an overview of the TF in terms of the legislative mandate, goals, and timeline. The goal of the NAIRR is to expand access to the cyberinfrastructure that fuels AI research and development, enabling all of America's diverse AI researchers to participate in exploring innovative ideas for advancing AI, including communities, institutions, and regions that have been traditionally underserved. The TF is required to deliver an interim report with recommendations in May 2022 and a final report with an implementation roadmap in November 2022.

A Request for Information was posted to the Federal Register on July 23, 2021, and is open for public comment until September 1, 2021.

TF members discussed the importance of understanding the needs of potential users in specific roles and inquired about the scope of the ideal user base. The need to determine the desired outcomes and measures of success, the theory of change, and reasons for public investment were also discussed. It was noted that desired outcomes and metrics of success will be a focus of the August meeting.

The session ended at 2:10 PM EDT.

Discussion: Desired Characteristics for the NAIRR and Key Challenges

The session started at 2:10 PM EDT.

Dr. Gianchandani led a discussion on desired characteristics of the NAIRR and expected challenges. The TF discussed the diversity of users whom the NAIRR would support, spanning different types of institutions, sectors, and regions, along with the standards and metrics for assessing success. Members also discussed challenges such as allocation of finite resources, cost relative to impact, and the potential for misuse of data. They noted the importance of measuring research impact, including potential negative impacts. In summary, Dr. Gianchandani discussed the need to determine the value proposition

and the use cases in order to be able to design a resource appropriate for the intended users. He also noted that, as the next session would illustrate, there is much activity already underway that could form the basis for a NAIRR.

The session ended at 2:34 PM EDT.

Presentation: Lessons Learned from Federal Cloud Pilots Supporting AI R&D

The session started at 2:34 PM EDT.

As an example of the various efforts already underway to enhance access to advanced computing infrastructure, Dr. Gianchandani summarized ongoing pilots by Federal agencies to leverage commercial cloud resources to support intramural and extramural AI research and development (R&D). He highlighted promising practices and lessons learned from agencies' efforts. Key lessons included the importance of user authentication, having dedicated cloud administration teams, building expertise among the user base, and the benefit of collaboration between government and non-governmental researchers. Key challenges included helping users locate data resources, ensuring appropriate privacy, and overcoming problems associated with billing.

After the presentation, TF members discussed the consideration of how to scale infrastructure needs as projects mature and as the number of users grows. Members also discussed how to design resources to align with users' needs, so that researchers would be incentivized to use them. TF members noted the prospect of connecting the NAIRR to Internet-of-Things (IoT) and embedded-systems testbeds and living laboratories to enhance access to such capabilities.

The session ended at 3:00 PM EDT and moved into a break for 15 minutes.

Presentation: Cloud Pilot Case Study – STRIDES

The session started at 3:15 PM EDT.

Andrea Norris (National Institutes of Health) presented an overview of the NIH Science and Technology Research Infrastructure for Discovery, Experimentation, and Sustainability (STRIDES) Initiative. STRIDES arose to help transition resource-intensive NIH research to the cloud while enabling collaboration and access to training and professional-services consultations. NIH-funded researchers can sign up to participate in STRIDES, then purchase cloud services from AWS, Google Cloud, or Microsoft Azure at reduced rates. Work to improve interoperability across these platforms is ongoing.

After the presentation, TF members asked about user outreach and training through STRIDES. Ms. Norris indicated that NIH manages training delivery but does not deliver the training. Members also inquired about success metrics. Ms. Norris responded that metrics include the number of people enrolled, participation in training sessions, quantity of data stored and compute cycles used, dollars saved compared to costs at normal market rates, and user feedback. Ms. Norris noted that STRIDES is also working with users to improve understanding of estimating budgets and billing for cloud services.

The session ended at 4:08 PM EDT.

Discussion: Task Force Proposed Workplan

The session started at 4:08 PM EDT.

Dr. Parker discussed the deliverables, phases of work, meeting schedule, role of working groups, and other key considerations. The interim report will include recommendations for the NAIRR, and the final report will include a detailed implementation plan. Researchers from the IDA Science and Technology Policy Institute (STPI) will be providing research and analytical support. The TF's work will be organized into three phases:

- July 2021 – February 2022: Assessment Phase;
- March 2022 – May 2022: Interim Report Development (due in May); and
- June 2022 – November 2022: Final Report Development (due in November).

Several working groups of 3-6 TF members each will be established, meeting between full TF meetings. The tentative schedule for meetings with proposed topics is:

- August 2021: Full TF meeting to discuss goals and evaluation metrics, ownership and administration options, governance models, and compute capabilities;
- September 2021: Working Groups on governance models and compute resources;
- October 2021: Full TF meeting to discuss desired capabilities, and barriers and solutions to dissemination of government data;
- November 2021: Working Groups on data resources, user interfaces, educational tools, and testing resources;
- December 2021: Full TF meeting to discuss security requirements and access controls, and privacy, civil rights, and civil liberties requirements.
- January 2022: Working Groups on user access; security; privacy, civil rights, and civil liberties; and technical integration; and
- February 2022: Full TF meeting to discuss sustainment and public-private partnerships opportunities as well as the interim report outline.

TF members will be asked to volunteer for three of the Working Groups throughout the Assessment Phase.

After Dr. Parker's presentation, TF members highlighted the importance of understanding the challenges the NAIRR would address, intended users, and expected outcomes. Example outcomes could include increasing AI research, educating more students, improving government partnerships with the private sector including large and small businesses, and advancing the way federal agencies handle data through a thoughtful consideration of privacy, civil rights, and civil liberties. Members also noted that the governance model should be driven by the specific goals.

Dr. Gianchandani noted that the NAIRR TF's focus is on advanced computing systems and services, but that this can include resources such as testbeds and living laboratories. Dr. Parker concluded by asking TF members for exemplar resource models and potential stakeholders whom the TF should engage as part of its work, and to send their preferences for Working Groups by the end of the day, August 3, 2021.

The session ended at 4:50 PM EDT.

Discussion: Questions from Meeting Attendees

The session started at 4:50 PM EDT.

Dr. Gianchandani moderated a discussion among TF members of questions submitted by attendees via Zoom's Q&A portal. Attendees asked about how they could further engage in the TF's work going forward, the impacts of the ongoing supply-chain challenges on the eventual acquisition or construction of the NAIRR, and the viability of including testbeds and living laboratories as part of the NAIRR.

Dr. Gianchandani concluded the session, thanking members of the TF, STPI, NSF and OSTP staff, and the public. He noted that meeting minutes and slides would be posted on the NAIRR TF website.

The next meeting is scheduled for Monday, August 30, with details to be posted shortly to the Federal Register.

The meeting adjourned at 5:00 PM EDT.

Appendix I: Attendance for NAIRR TF Meeting 1

TF Members Present:

Erwin Gianchandani, National Science Foundation (Co-Chair)

Lynne Parker, White House Office of Science and Technology Policy (Co-Chair)

Daniela Braga, DefinedCrowd

Mark Dean, retired (formerly IBM and University of Tennessee, Knoxville)

Oren Etzioni, Allen Institute for AI

Julia Lane, New York University

Fei-Fei Li, Stanford University

Andrew Moore, Google

Michael Norman, University of California, San Diego

Dan Stanzione, University of Texas, Austin

Frederick Streitz, Department of Energy

Elham Tabassi, National Institute of Standards and Technology

TF Members Absent:

None.