

MEETING SUMMARY

National Artificial Intelligence Research Resource Task Force

Meeting #6

April 8, 2022

MEETING SUMMARY

Meeting Summary

The sixth meeting of the National Artificial Intelligence Research Resource (NAIRR) Task Force (TF) was held online via Zoom on April 8, 2022, 11:00 AM–5:00 PM EDT.

Welcome and Administrative Remarks

The meeting started at 11:05 AM EDT.

Dr. Manish Parashar, NAIRR TF Co-Chair, opened the meeting and provided background and logistical information, then called for a motion to approve the summary from the fifth NAIRR TF meeting held on February 16, 2022; the motion passed. Dr. Parashar then introduced the agenda.

The meeting had four objectives:

1. Discuss the interim report draft;
2. Consider the 2022 workplan and public engagement plans;
3. Develop key questions for implementation of the NAIRR; and
4. Consider models for implementation of responsible and ethical AI research processes.

In addition, time was reserved at the end of the meeting to address questions from the public, with any unanswered questions reserved for the next meeting. Dr. Parashar noted that no unanswered questions had been carried over from the prior meeting of the TF.

The session ended at 11:11 AM EDT.

Discussion of the Interim Report Draft

The session started at 11:11 AM EDT.

Dr. Lynne Parker, NAIRR TF Co-Chair, started the discussion of the interim report. The TF had previously reviewed the first draft of the report and provided written feedback, so the discussion during the meeting provided the opportunity for the group to discuss the second version of the interim report. Dr. Parker noted that based on the first round of feedback the Co-Chairs identified some specific topics where the TF may not have reached a consensus or where there may be differences in understanding or terminology. The discussion focused on these issues, which included terminology, the NAIRR user access portal, and evaluation methods for NAIRR investments, as detailed further below.

Dr. Parashar led the discussion on NAIRR terminology. He began by asking TF members whether “NAIRR” should refer to the management entity contemplated by the interim report, the cyberinfrastructure it manages, or both. Dr. Parashar presented slides with proposed definitions for the various components that the TF has envisioned as making up the NAIRR. He asked TF members if the definitions on the slide reflected their understanding. TF members spoke about the need to clearly differentiate between the roles of NAIRR management, governance, and advisory bodies. Members agreed that “NAIRR” should refer to the full collection of entities (the management and governance bodies, and the cyberinfrastructure, including the resources); however, when only one part of the NAIRR is being discussed, it can be named explicitly (e.g., “NAIRR cyberinfrastructure” or “NAIRR governance body”). Members discussed the possibility of highlighting education, research, and industry partnerships as core functions of the NAIRR. They also discussed other NAIRR management functions, including updating the NAIRR as technologies advance and user needs evolve, and outreach and partnership development.

Dr. Parker guided discussions around the user access portal and evaluation methods. TF members had previously raised questions about whether requiring a single access portal could be overly constraining to users of the NAIRR. To address this issue, Dr. Parker noted that the draft language for the second draft was adjusted to avoid this interpretation. TF members agreed that the revisions were clarifying and noted the access portal should be a starting point for new and returning users to discover and access NAIRR resources, but not the only way to access those resources. When discussing evaluation methods, Dr. Parker asked TF members about how to integrate evaluation methods into the NAIRR and at what level it would be appropriate to incorporate counterfactual analysis. Some TF members suggested specific evaluation methods while others said that a prescriptive approach might be too restrictive. The group agreed that recommending the use of “established best practices” for portfolio-based evaluation would be appropriate for the interim report.

The TF discussed a question from a member of the public about whether the NAIRR would have centralized policies to support user access to resources, especially data. TF members responded that the portal would have core policies for contributed data to support ease of use. TF members noted the NAIRR will in general aggregate and vet data cleaned and processed by others, and aggregated data might have limitations on how it could be shared. Another member suggested the NAIRR should work to integrate heterogeneous data sets drawn from multiple sources in order to provide insights into topics that are poorly represented in data science today because available data are of varying quality.

The session ended at 12:49 PM EDT.

Discussion of the Post-Interim Report Workplan and Public Engagement

The session started at 12:50 PM EDT.

Dr. Parker reviewed TF progress to date, the timeline governing the TF’s work, and a proposed breakdown of next steps to allow the TF to develop a more detailed roadmap and implementation plan for the NAIRR as its final report. Dr. Parker noted that a public request for information (RFI) will be issued to solicit public feedback on the findings and recommendations made in the interim report as well as input on their implementation. In addition, the TF will hold a public listening session, an interagency roundtable, and an

international roundtable to discuss how plans for the NAIRR can leverage or complement other activities that are ongoing and planned. The TF will continue to receive briefings from experts on topics of importance to the NAIRR in future public meetings.

The session ended at 1:02 PM EDT.

Break: 1:02–1:40 PM EDT

Discussion of Implementation Questions for the Final Report

The session started at 1:40 PM EDT.

Drs. Fei-Fei Li and Dan Stanzione, NAIRR TF members, led a discussion to identify open topics that should be addressed in the TF work plan, as well as topics taken up in the interim report which require further refinement. One TF member suggested that a working group (WG) undertake a gap analysis to identify resources currently lacking in the target user community, and other members agreed. One suggested that each WG incorporate some sort of gap analysis into its work. Another member suggested that members conduct an “anti-gap” analysis to name what NAIRR should not do, based on its vision and mission statement. Another member said that some level of detail about user needs could be necessary for a gap analysis.

TF members also discussed the potential scale of the NAIRR and whether the roadmap should lay out detailed actions and timelines for establishing the NAIRR or be more notional. One member suggested that the NAIRR roadmap be framed ambitiously and in detail, and that the TF avoid underestimating resource requirements. One member commented that the TF, as an assembled panel of experts, is in a position to provide in-depth input and its recommendations would likely be welcome; other members agreed that specific recommendations are more actionable.

Several TF members noted that there is a need to define the scale and scope of the resources, including the size of the user community—which will be important in the logic model for evaluating the NAIRR’s impact. Other areas raised by the TF included better defining the NAIRR management entity and governance structure, as well as agency roles; understanding the next generation of AI and data researchers and research topics; conducting a deeper user/customer analysis; and ensuring that the NAIRR be application-agnostic.

The TF also discussed open questions related to the NAIRR’s business/sustainability model and the extent and duration of Federal funding for the NAIRR. One TF member suggested that NAIRR could be a long-term investment, analogous to the Large Hadron Collider; another wondered whether long-term Federal funding could be assured. A third suggested that too much overhead could make the program too expensive, as a major motivator for the NAIRR is to provide cheaper or easier access (for example, through a standard contracting process) to compute cycles for academic researchers—and not only those at prestigious institutions. Similarly, if the NAIRR could streamline the approval process for data access and sharing—for example, by creating a standard data sharing agreement that covers all users—negotiation on an institution-by-institution basis could be avoided.

A TF member commented that the NAIRR should fill a gap that will not be met by the current marketplace, and that needs of the user community could be different in 10 or 15 years. Engagement with industry could help to ensure that the NAIRR follows the emerging cyberinfrastructure landscape. The NAIRR could also help to broker arrangements and understanding between academia and industry. TF members discussed types of partnerships that the group might consider in the second phase of their work, including the potential for international engagement—a topic already on the agenda for an upcoming meeting. TF members also briefly discussed the kinds of administrative and regulatory burdens that the sponsoring Federal agencies might face, such as difficulty in transferring funds from one agency to another.

In response to a question from the public, TF members explained that, in terms of intellectual property, the intent is for the NAIRR to align with applicable laws and regulations with respect to federally-funded research, such as the Bayh-Dole Act. Another question from the public was whether the NAIRR would own and manage assets beyond the user platform. A TF member said that the NAIRR would not in general own assets, but would provide users access to a variety of state-of-the-art resources.

The session ended at 3:09 PM EDT.

Break: 3:10–3:40 PM EDT

Panel: Building Responsible AI Review Processes for the NAIRR

The session started at 3:40 PM EDT.

The panel comprised the following speakers:

- Beena Ammanath, Author, Trustworthy AI and Head of Global Deloitte AI Institute
- Michael Bernstein, Associate Professor of Computer Science, Stanford University
- Arvind Narayanan, Associate Professor of Computer Science, Princeton University
- Beth Plale, Michael A. and Laurie Burns McRobbie Bicentennial Professor of Computer Engineering and Executive Director, Pervasive Technology Institute, Indiana University Bloomington
- Christo Wilson, Associate Professor of Computer Science, Northeastern University

Dr. Oren Etzioni, NAIRR TF member, introduced the panelists, who each spoke for approximately five minutes. Panelists provided ideas and illustrative examples of concrete methods the TF could consider for promoting responsible AI research within the NAIRR, including:

- Define principles or an ethics statement that all researchers are expected to uphold;
- Embed checks throughout the process of system use that prompt users to pause and consider potential harms as well as the guardrails to prevent them;
- Implement an ethics review process for individual projects or uses that helps to identify risks and potential mitigations, as Institutional Review Boards do not address most of the relevant risks of AI research;

- Identify common issues and mitigations as a guide for researchers;
- Provide a mechanism by which projects could be rejected should the risks of the research be deemed too high, along with a mechanism to adjudicate appeals;
- Provide governance and infrastructure that supports rigorous data use, for example, communication between users and data providers; documentation of data sets and standardized data citations and licenses;
- Establish a reporting method for potentially unethical activity within NAIRR;
- Partner with other entities that are considered “good regulators” of data;
- Translate the notion of bug bounties (where individuals receive compensation for identifying cybersecurity vulnerabilities that can then be fixed) to AI systems;
- Leverage university tech transfer offices more effectively to obtain not only market assessments but also intended use assessments;
- Leverage data management plans that are part of research proposals in making NAIRR allocation decision.;
- Conduct impact assessments of AI models rather than data sets alone; and
- Create a team to audit a sample of near-complete projects to serve as a deterrent and promote adherence to ethical principles.

Following panelist remarks, Dr. Etzioni moderated a discussion of the panelists and TF members. Panelists commented on possible downsides to the approaches recommended, including the risk that steps instituted to protect privacy and civil liberties could turn into a compliance exercise without much payoff unless controls are evaluated and adjusted to optimize their value. It was also noted that the controls instituted at early stages of research might not be protective of downstream harms that might nonetheless be attributed to the NAIRR. One panelist suggested that the cost of democratizing access to AI research could be the need to set a high bar for ethical practices. Another commented that trust is necessary for an institution to have legitimacy in the public eye, and this should inform NAIRR design and policies.

A TF member asked whether NAIRR review processes should be more stringent than those already instituted as part of Federal agency proposal reviews. A panelist suggested that ethics review requires expertise from across a range of disciplines beyond AI, such as anthropology and psychology—and that standard review processes are not necessarily equipped to do this. Another panelist commented that approaches to ethical AI are nuanced; even if agency practices are followed, there is still a need for the NAIRR to serve as an anchor point for helping AI researchers to think about AI ethically.

A TF member asked to what extent it is possible for ethics assessments to be objective and reproducible. A panelist noted that, while such assessments have some inherent subjectivity, they can be structured for reproducibility—for example, using a checklist approach. Participants discussed the possibility of heuristics of ethics or harms, as well as risks of potential censorship of research that lies outside of the

mainstream. Other topics discussed included the benefits of an ethics review process for increasing researchers' awareness of ethical issues, designing controls into the NAIRR that both provide protections and build trust in the system, and the relevance of OSTP efforts to develop an AI Bill of Rights.

The session ended at 4:58 PM EDT.

Questions from Public and Meeting Close

The session started at 4:58 PM EDT.

Dr. Parashar concluded the session, thanking members of the TF, NSF, OSTP, STPI, and the public for their participation. Meeting summaries, slide presentations, and details about upcoming meetings can be found at <https://www.ai.gov/nairrtf/>. One question from the Q&A chat remained at the end of the meeting.

The next meeting is scheduled for May 20, 2022, from 2:00 PM–3:00 PM EDT. Details are posted at [Federal Register Notice 87 FR 16034](#).

The meeting adjourned at 5:00 PM EDT.

Appendix I: Attendance for NAIRR TF Meeting #6

TF Members Present:

Manish Parashar, National Science Foundation (Co-Chair)

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

Daniela Braga, Defined.ai

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 Streit, Department of Energy

Elham Tabassi, National Institute of Standards and Technology

TF Members Absent:

None