Training as Infrastructure
Workshop Report

November 21, 2025
New York University
New York, NY

Executive Summary....1
Key Takeaways....2
Workshop Overview....3
Plenary Presentations and Discussion....3
Working Session 1: Understanding Training as Infrastructure....4
Working Session 2: Empire AI Priorities and New York Needs....5
Working Session 3: Drafting Priorities for Training as Infrastructure....6
Conclusion....7
Appendix 1: Worksheet....8
Appendix 2: Workshop Participants....8

Executive Summary

With significant public investments in AI infrastructure, including the National Artificial Intelligence Research Resource (NAIRR) at the federal level and Empire AI at the state level, a shared understanding has emerged: **AI infrastructure requires building human expertise as a resource**, alongside compute, data, and software.

From July 29-31, 2024, an in-person workshop was convened at New York University, bringing together leading experts in the cyberinfrastructure, AI policy, governance, and responsible AI communities to inform the development of priorities and policies for the NAIRR, with a particular focus on advancing trustworthy AI.[1] The workshop was convened by Dr. Julia Stoyanovich, Institute Associate Professor of Computer Science and Engineering, Associate Professor of Data Science, and Director of the Center for Responsible AI at New York University, and co-hosted by the NYU Tandon Center for Responsible AI and the NYU Center for Data Science. The workshop was supported by NSF Award No. 2432040 and by Omidyar Network, a non-governmental partner of the NAIRR Pilot.

On November 21, 2025, Stoyanovich convened a follow-up in-person meeting at New York University. The workshop was designed to connect lessons from the NAIRR Pilot with New York State's developing **Empire AI initiative**, and to examine how training, mentoring, facilitation, and user support can function as integral components of AI infrastructure. Supported by NSF Award No. 2432040, the workshop took place at a pivotal moment, as states are beginning to invest in AI-ready supercomputing capacity to ensure that research, education, and workforce

[1] See workshop report at https://r-ai.co/nairr_report

development are not limited to a small number of institutions or private-sector actors. Empire AI is among the most ambitious of these initiatives.

The workshop brought together researchers, educators, community organizations, and philanthropic foundations. The agenda included short presentations followed by structured discussions and collaborative working sessions. Participants actively engaged in identifying gaps, emerging needs, and opportunities to strengthen the training and human-capacity dimension of AI infrastructure, with insights intended to inform both the ongoing development of the NAIRR and New York's Empire AI initiative.

Key Takeaways

Through the presentations and working sessions, participants distilled several key takeaways about how large-scale public investments in AI infrastructure, including the **National Artificial Intelligence Research Resource (NAIRR)** and **New York's Empire AI initiative**, can support broad participation in AI research, education, and workforce development. While access to advanced computing resources is essential, participants emphasized that infrastructure investments must also address the human expertise required to use these resources effectively across institutions with widely varying levels of capacity.

First, training and human support functions are themselves infrastructure. Compute capability alone is insufficient. Effective participation requires facilitation, mentoring, onboarding support, project scoping guidance, and ongoing technical assistance embedded within infrastructure operations. Participants emphasized the need for clear and differentiated entry points for users with different levels of experience and institutional readiness.

Second, Empire AI and NAIRR should be designed to be complementary. Participants emphasized the importance of aligning state and federal AI infrastructure efforts. Empire AI can play an important role in strengthening institutional readiness, expanding training capacity, and supporting onboarding and facilitation structures that enable researchers, educators, and students to effectively engage with national AI infrastructure resources such as the NAIRR.

Third, shared resources and regional networks can accelerate capacity building. Participants highlighted the value of leveraging local collaborations and institutional networks to support the effective use of AI infrastructure. Shared resources, including curated repositories, curriculum models, documentation templates, and communities of practice, can reduce duplication across institutions and make it easier for researchers, educators, and students to build on existing expertise. By connecting local contexts and networks to national infrastructure efforts, these shared resources can help accelerate AI research, development, and education while expanding participation across institutions.

Workshop Overview

The workshop agenda featured opening remarks and short presentations from leaders of several initiatives focused on expanding access to AI infrastructure and training, including programs connected to the NAIRR Pilot and related capacity-building efforts. These presentations were followed by active discussions among participants. The majority of the event was devoted to structured working sessions designed to foster interactive discussion and to identify priorities for integrating training, mentoring, and facilitation into emerging AI infrastructure efforts, including both the NAIRR and New York's Empire AI.

Plenary Presentations and Discussion

The workshop opened with remarks by Dr. Julia Stoyanovich, who situated the discussion within the trajectory of the NAIRR Pilot and the earlier NAIRR Community-Informed Policies and Best Practices Workshop. She emphasized a central finding from that earlier workshop: human expertise must be treated as a resource alongside compute, data, and software. Framing the day's discussion around this premise, she posed the workshop's guiding question: **What would it mean to treat training and mentoring as infrastructure, rather than as an afterthought or a compliance requirement?** She also underscored the workshop's dual purpose: to strengthen the training and education dimension of the NAIRR and to help articulate how New York's Empire AI initiative could complement NAIRR by investing in human expertise at the state level.

Following the opening presentation, participants heard three lightning talks describing existing models for training, facilitation, and capacity-building. Dr. Forough Ghahramani presented the Campus Champions program, an NSF-supported effort within the NAIRR ecosystem that expands participation in advanced computing by embedding local champions who help researchers, educators, and students navigate resources, training opportunities, and mentorship networks. Dr. Gabrielle Allen described work to support research-emerging institutions as they build capacity to engage with NAIRR resources, highlighting challenges faculty encounter in identifying appropriate data, understanding resource requirements, and navigating large repositories. Dr. Jamie Payton presented NSF Level Up AI, a Computing Research Association initiative designed to expand access to AI education through structured community engagement, roundtables, and pilots aimed at surfacing barriers and opportunities in pedagogy, infrastructure, and access. Together, these presentations grounded the workshop in concrete examples of how training and support are already being operationalized across institutions and initiatives, and helped set the stage for the working sessions that followed.

In the discussion that followed, participants reflected on common themes emerging across the presentations. They noted that, while significant investments are being made in AI infrastructure, many institutions continue to face barriers in accessing and effectively using these resources. These barriers include uncertainty about how to scope projects, limited local technical support, and gaps in training pathways for researchers, educators, and students with varying levels of experience. Participants also emphasized that training needs extend beyond introductory instruction to include ongoing mentorship, facilitation,

and domain-specific guidance. Across the discussion, participants highlighted the importance of treating these support structures as integral components of AI infrastructure in order to ensure that new investments translate into meaningful research, education, and workforce development outcomes.

Participants also highlighted the opportunity to leverage existing institutional networks and communities of practice in the New York region to share expertise, develop common resources, and expand participation in AI infrastructure.

Working Session 1: Understanding Training as Infrastructure

The first working session invited participants to reflect on patterns, lessons, and emerging needs revealed by the four initiatives presented during the morning lightning talks. The session aimed to surface common experiences across these programs, examine how the broader AI landscape has evolved since the NAIRR vision was first articulated, and identify foundational elements of what it would mean to treat training and mentoring as infrastructure.

In the first part of the discussion, participants reflected on how researchers, educators, and students currently learn to engage with AI systems, datasets, and infrastructure. Across the initiatives discussed, participants noted that effective learning often occurs through social and institutional structures rather than through formal instruction alone. Mentorship networks, local champions, consulting-style support, and communities of practice were repeatedly cited as critical mechanisms through which individuals gain the confidence and skills needed to use complex AI resources. Participants observed that successful programs tend to combine technical instruction with hands-on engagement around concrete research problems, enabling users to connect computational tools to domain-specific questions.

Participants also discussed where existing initiatives have been successful in meeting community needs and where gaps persist. Programs such as Campus Champions and related training initiatives were seen as valuable entry points that help researchers navigate available resources and build initial expertise. At the same time, participants noted persistent unmet needs in areas such as project scoping, domain-specific guidance, and sustained mentorship. Many researchers and educators struggle to determine whether AI approaches are appropriate for their problems, what resources are required to pursue them, and how to translate research questions into feasible computational workflows. Participants also highlighted disparities across institutions in terms of staffing, technical support, and access to educational computing resources, which can limit the ability of some communities to take full advantage of new infrastructure investments.

These reflections led to a broader discussion of what it means to conceptualize training as infrastructure. Participants suggested that training-as-infrastructure should include not only instructional materials and courses, but also facilitation roles, consulting services, shared repositories of training resources, and mechanisms for knowledge exchange across institutions. In this framing, training becomes an ongoing ecosystem that supports researchers and educators throughout the lifecycle of a project rather than a one-time educational intervention.

In the second part of the session, participants considered how the landscape of AI research and infrastructure has evolved since the NAIRR Task Force first proposed the concept of a national AI research resource. Participants observed that rapid advances in AI tools, growing demand for computational resources, and expanding adoption of AI methods across scientific disciplines have significantly increased the need for training and support. At the same time, expectations around responsible and trustworthy AI have grown, introducing new requirements related to data governance, ethical considerations, and interdisciplinary collaboration. These changes have expanded the scope of skills and guidance required for effective participation in AI-enabled research.

Participants also discussed how NAIRR leadership has begun responding to these shifts through pilot programs that emphasize capacity building, access to shared resources, and support for emerging institutions. These efforts were viewed as important steps toward operationalizing the "people layer" of AI infrastructure. However, participants noted that several training and facilitation needs have become more visible as the ecosystem evolves. These include the need for sustained mentoring networks, support staff dedicated to facilitating infrastructure use, clearer training pathways for users with different levels of experience, and mechanisms for connecting local communities of practice to national infrastructure initiatives.

Taken together, the discussion in this session helped clarify both the current gaps in training ecosystems and the structural elements required to address them. These insights informed the subsequent working sessions, which focused on identifying concrete priorities and drafting proposals for integrating training and mentoring into AI infrastructure efforts at both the national and state levels.

Working Session 2: Empire AI Priorities and New York Needs

The second working session focused on the emerging role of Empire AI within the evolving AI infrastructure landscape and explored how state-level investments in computing and capacity-building could interact with national efforts such as the NAIRR. The session sought to surface perspectives from New York–based institutions and partners regarding opportunities, gaps, and priorities for strengthening the state's AI ecosystem.

The discussion began with **an overview of Empire AI's emerging priorities**, including plans to build shared AI computing infrastructure and to support research, education, and workforce development across New York State. Participants noted that Empire AI has the potential to expand access to advanced computing resources beyond a small number of institutions, helping to ensure that researchers, educators, and students across the state can participate in AI-driven discovery. Participants also emphasized that the success of such investments will depend on accompanying efforts to build institutional capacity, including training pathways, facilitation structures, and support for interdisciplinary collaboration.

New York–based stakeholders reflected on how these priorities resonate with the needs they observe within their own institutions and networks. Participants highlighted the diversity of institutions across the state, including large research universities, public university

systems, community colleges, and workforce development programs, and noted that each faces distinct challenges in adopting AI technologies. Several participants emphasized the importance of workforce development and educational programming that can prepare students and professionals for a rapidly evolving technological landscape. Others highlighted the need to support faculty and researchers in fields that have traditionally relied on scientific computing but are now beginning to integrate AI methods into their work.

Participants also discussed the importance of connecting Empire AI to existing institutional networks and community partnerships across the state. New York was seen as having a strong foundation of universities, research centers, workforce initiatives, and nonprofit organizations that could collaborate to expand training opportunities and build communities of practice around AI-enabled research and education. Leveraging these networks, participants suggested, could help accelerate the development of shared resources and knowledge while reducing duplication of effort across institutions.

In the second part of the discussion, participants considered how Empire AI and NAIRR might operate as complementary components of a broader AI infrastructure ecosystem. Participants noted that while NAIRR is designed to provide national access to AI resources and to support a diverse community of researchers and educators, state-level initiatives such as Empire AI may be particularly well positioned to build institutional readiness and to support localized training and facilitation efforts. In this view, Empire AI could help institutions develop the expertise, partnerships, and project ideas necessary to engage effectively with national infrastructure resources.

At the same time, participants identified **several open questions** about how Empire AI's programs and services will evolve and how institutions across New York can best engage with the initiative. These questions included how training and workforce programs will be structured, how support for teaching and research will be coordinated across institutions, and how state-level infrastructure can be integrated with national initiatives such as NAIRR. Participants emphasized that continued dialogue among state agencies, universities, and community partners will be important for ensuring that Empire AI's investments translate into broad participation in AI research, education, and workforce development.

Working Session 3: Drafting Priorities for Training as Infrastructure

The final portion of the workshop was devoted to structured working sessions aimed at **drafting priorities for treating training and mentoring as infrastructure** within emerging AI infrastructure initiatives. Participants were divided into two groups that worked in parallel using a shared worksheet and a common set of prompts, see worksheet in Appendix 1 for details.

The groups were asked to identify cross-cutting needs and gaps related to training and mentoring, explain why these challenges should be considered infrastructure issues rather than one-off programming needs, propose ideas for complementarity between the NAIRR and Empire AI, and sketch early elements of a training-as-infrastructure model.

Across both groups, participants identified a range of needs related to training pathways, research support, and institutional capacity. Participants noted that many researchers and educators lack guidance on how to scope AI-enabled research projects, identify appropriate datasets and methods, or estimate the compute requirements needed to carry out their work. Participants also emphasized the need for differentiated training pathways tailored to different goals and levels of readiness, ranging from AI literacy and computing fundamentals to domain-specific training for researchers seeking to integrate AI into scientific workflows. Several participants highlighted the importance of teaching support, community building, and workforce training aligned with regional and state-level needs.

Participants also discussed why these challenges should be understood as infrastructure issues rather than isolated programmatic gaps. Many of the needs identified, including training materials, consulting support, curated learning resources, and computing access for education, are difficult for individual institutions to develop and maintain independently. Participants emphasized that addressing these needs requires:

- durable structures,
- shared processes, and
- roles dedicated to facilitation and
- knowledge sharing.

Statewide infrastructure could enable institutions to share curriculum models, case studies, and training resources, reducing duplication and lowering barriers to participation.

Finally, participants discussed potential complementarities between the NAIRR and Empire AI. Participants suggested that Empire AI could serve as a regional hub and on-ramp to national infrastructure, helping institutions build capacity and develop projects that could later leverage NAIRR resources. At the same time, state-level initiatives may be well positioned to support localized communities of practice, provide consulting and facilitation support, and develop shared training resources tailored to regional needs. These discussions reinforced the view that training, mentoring, and facilitation should be treated as integral components of AI infrastructure, enabling institutions with diverse capacities to participate more fully in emerging AI research and education ecosystems.

Conclusion

This workshop explored how training, mentoring, facilitation, and user support can function as integral components of emerging AI infrastructure initiatives. Participants emphasized that investments in compute and data must be accompanied by sustained investments in human expertise, including clear training pathways, mentoring networks, and support structures that enable researchers, educators, and students to use these resources effectively.

Participants also highlighted the opportunity to align national and state efforts. While the NAIRR provides a national framework for expanding access to AI resources, initiatives such as Empire AI can help build institutional readiness and strengthen regional training ecosystems. Beyond

Empire AI, participants noted that the lessons and **recommendations from this workshop have broader relevance for regional AI hubs and related infrastructure efforts across the United States**. Taken together, the discussions reinforced a central conclusion: training and human support functions are foundational elements of AI infrastructure, necessary for ensuring broad participation in AI-enabled research and education and **for accelerating scientific discovery and workforce readiness**.

Appendix 1: Worksheet

Section A: Needs/Gaps (bulleted)

A list of 4–6 cross-cutting needs/gaps (from NAIRR, Empire AI, institutions, faculty, students, workforce, etc.)

Section B: Why these are infrastructure issues

A short statement of why these are "infrastructure needs": avoid one-off programming, focus on structures, roles, processes

Section C: NAIRR / Empire AI complementarity

Three draft ideas for federal–state complementarity (state initiative strengthens, extends, or operationalizes NAIRR elements) Example forms:

- *"NAIRR does X; Empire AI could extend by Y"*
- *"State-level infrastructure should fill gap Z that NAIRR structurally cannot"*
- *"Empire AI could serve as the on-ramp to NAIRR for under-resourced NY institutions"*

Section D: Early sketch of a training-as-infrastructure model

One early idea for a structured training pipeline model, something like:

- *entry-level onboarding support*
- *domain-specific training kits*
- *mentorship / fellows networks*
- *facilitator support layer*
- *community engagement and reflection cycles*

Appendix 2: Workshop Participants

Aankit Patel	City University of New York
Brendan Collins	NYC Tech Talent Pipeline
Deborah Najjar	Alfred P. Sloan Foundation

Forough Ghahramani	NJ Edge
Gabrielle Allen	University of Wyoming
Jamie Payton	CRA Level Up AI; New Jersey Institute of Technology
Joshua Brumberg	City University of New York / Empire AI
Julia Stoyanovich (**convener**)	New York University
Rene Parmar	SUNY New Paltz
Sonali Chadha	Alfred P. Sloan Foundation
Thorsten B Joachims	Cornell University
Anya Kozorez (**scribe**)	New York University
Caterina Fuligni (**logistics**)	New York University