

Communities of Practice and Cyberinfrastructure Adoption in Big Science Organizations:

Decision-Making, Adaptation, and AI Readiness

Mahedi Hasan, PhD Candidate, Texas Tech University

This dissertation investigates how Communities of Practice (CoPs) shape cyberinfrastructure (CI) adoption and AI-related transitions within National Science Foundation–funded Big Science Organizations, including Major Facilities (MFs) and Mid-Scale Research Infrastructures (MSRIs). Although NSF invests billions of dollars in CI to sustain data-intensive and collaborative science, adoption remains uneven. Formal implementation plans often coexist with underused systems, informal workarounds, and persistent coordination challenges.

Drawing on Everett Rogers’ Diffusion of Innovations (DOI) theory and Etienne Wenger’s Communities of Practice framework, the study reconceptualizes CI adoption as a multilevel communicative and community-driven process. Rather than treating adoption as a discrete organizational decision, it examines how informal knowledge-sharing networks, governance expectations, and everyday practices interact to shape decision trajectories, risk perceptions, and institutionalization outcomes.

The dissertation addresses four research questions:

RQ1. How do Communities of Practice influence the adoption and ongoing adaptation of CI, including emerging AI enabled services, within RIs?

RQ2. a) What organizational, technical, and cultural challenges do CoPs address as RIs adopt and adapt cyberinfrastructure, including AI related tools, and

b) how do these communities help overcome barriers related to complexity, interdisciplinary coordination, and stakeholder engagement?

RQ3. How do RIs define “AI readiness” for data and CI, and how do AI related decisions reorganize workflows, governance arrangements, trust, and risk management?

RQ4. How do manual qualitative analysis and computational topic modeling compare in identifying patterns in discourse about CI and AI related decision making across RIs?

The study employs a mixed-methods, multi-site design. Ethnographic data collection includes participant observation and semi-structured interviews at two NSF-funded facilities: a Major Facility and a Mid-Scale Research Infrastructure. To date, 14 days of participant observation aboard the NSF-supported research vessel R/V Sikuliaq have generated approximately 80 pages of fieldnotes documenting CI coordination, troubleshooting, and training practices. The project also includes 24 in-depth interviews with scientists, engineers, data managers, and administrators engaged in CI decision-making. Additional fieldwork is underway at a Mid-Scale facility. For the final data collection, this project aims for multisite ethnography and around 50-70 interviews.

Data analysis integrates constructivist grounded theory with computational topic modeling using Latent Dirichlet Allocation (LDA). This methodological integration allows for systematic comparison between interpretive coding and computational pattern detection, contributing both substantive findings on CI adoption and methodological insights into qualitative data analysis in organizational research.

Preliminary findings indicate that Communities of Practice function as critical yet often invisible infrastructures for CI adoption. CoPs reduce uncertainty, enable tacit knowledge exchange, and build relational trust across disciplinary and organizational boundaries. Effective adoption depends on alignment between formal governance directives and community-developed practices. In the context of AI-related initiatives, CoPs also shape how “AI readiness” is defined, how risks are interpreted, and how emerging tools are incorporated into scientific workflows.

The dissertation advances theory by integrating Communities of Practice into diffusion scholarship, demonstrating how informal social structures mediate innovation trajectories in complex, federally funded scientific systems. Methodologically, it bridges ethnographic and computational approaches to qualitative analysis. Practically, it offers actionable insights for NSF-funded facilities, intermediary organizations, and policymakers seeking to design more community-informed CI governance and AI transition strategies.

Beyond Big Science, the findings speak to other high-stakes sectors navigating technological transformation, including healthcare, energy, and finance. By centering communicative and community processes in infrastructure adoption, this research highlights the social foundations of technological resilience and organizational innovation.