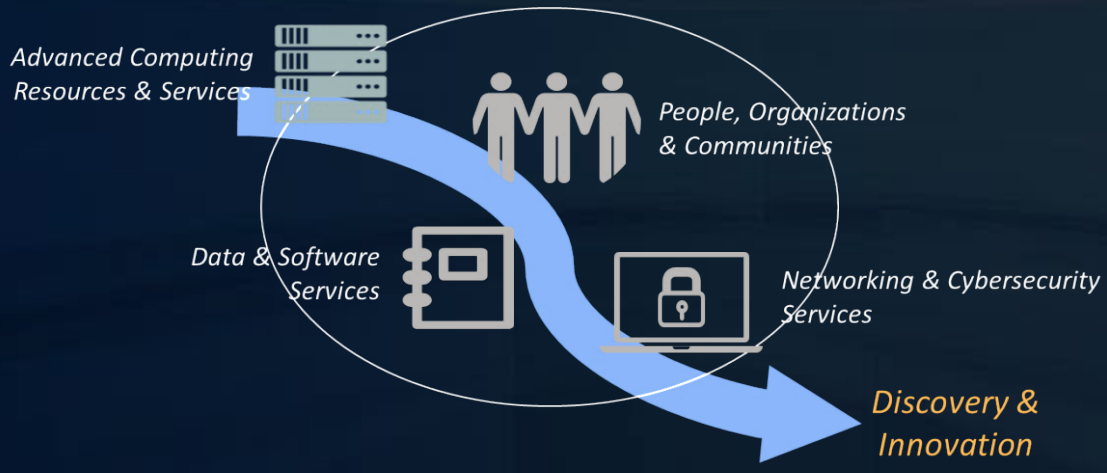
Transforming Science Through Cyberinfrastructure

NSF's Blueprint for a National Cyberinfrastructure Ecosystem for Science and Engineering in the 21st Century



Blueprint for Cyberinfrastructure Learning and Workforce Development

Office of Advanced Cyberinfrastructure
Directorate for Computer & Information Science & Engineering
National Science Foundation

August 2021

Transforming Science Through Cyberinfrastructure: Cyberinfrastructure Learning and Workforce Development

NSF's Blueprint for Cyberinfrastructure Learning and Workforce Development for Accelerating Science and Engineering in the 21st Century

Executive Summary

The national research cyberinfrastructure (CI) has become critical to computational and data intensive research across all of science and engineering (S&E) in the 21st century. The National Science Foundation (NSF) recently shared a vision, developed by its Office of Advanced Cyberinfrastructure (OAC), for ***an agile, integrated, robust, trustworthy, and sustainable CI ecosystem that drives new thinking and transformative discoveries in all areas of science and engineering (S&E) research and education***. The envisioned CI ecosystem deploys advanced CI resources, services, and expertise towards collectively enabling new, transformative discoveries across all of S&E. This document is the fifth in a series of blueprint documents published by OAC that outline NSF's plan for realizing this vision. It presents a holistic and forward-looking blueprint for CI learning and workforce development (LWD) that broadens participation in all relevant communities of concern including CI contributors, CI users, and CI Professionals, along with their institutions. This blueprint is motivated by insights from community workshops, Requests for Information, national initiatives, S&E researchers, high-performance computing (HPC) stakeholders, and CI Professionals (i.e., practitioners). This LWD blueprint aims to: (1) support the creation of structures, processes, and materials to address the learning and training needs of diverse researchers and CI developers; (2) promote commitments to mentoring plans and professional development of CI Professionals and the establishment of career pathways at research institutions; and (3) foster and nurture the self-organization of sustainable CI communities, such as regional or domain-centered collaborative networks. Blueprint actions will incentivize continual advancement of skills and ensure well-connected capacity to achieve the goals of S&E researchers across the NSF CI ecosystem. The overarching goal is to foster and nurture a diverse, recognized, and skilled CI workforce community that can accelerate and amplify the transformative impact of CI across all S&E research and education.

Table of Contents

Table of Figures

1 Introduction

The national research cyberinfrastructure (CI) has become critical to computational and data intensive research across all of science and engineering (S&E) in the 21st century. It is a key catalyst for discovery and innovation and plays a critical role in ensuring US leadership in S&E, economic competitiveness, and national security, consistent with the National Science Foundation's (NSF's) mission. Through the Office of Advanced Cyberinfrastructure (OAC), NSF has shared a vision[1] that calls for the broader availability and innovative use of ***an agile, integrated, robust, trustworthy, and sustainable CI ecosystem*** that can ***drive new thinking and transformative discoveries in all areas of S&E research and education***.

NSF's vision for an agile, integrated, robust, trustworthy and sustainable CI ecosystem that drives new thinking and transformative discoveries in all areas of S&E research and education

- *View CI more holistically: CI continuum seamlessly integrating a spectrum of resources, tools, services, and expertise to enable transformative discoveries.*
- *Support translational research: Core innovations → development of community tools and frameworks → deployment and operation of sustainable production CI.*
- *Balance innovation with stability: Ensure continuity in production computational capacity while fostering innovation and transition to production.*
- *Couple discovery and CI innovation cycles : Rapidly address new challenges and opportunities in an era of disruptive technologies and evolving science needs.*
- *Improve usability: Ease pathways for discovering, accessing, understanding and using powerful CI capabilities and services to enhance researcher productivity and scientific impact.*

This document is the fifth in a series of blueprint documents that outline NSF's plan for realizing this vision. This forward-looking document outlines NSF's plan, a holistic and forward-looking blueprint for CI learning and workforce development (LWD) that broadens participation in all relevant communities of concern including CI contributors, CI users, and CI Professionals, as well as their institutions. This blueprint has been informed by the community through advisory bodies, requests for information (RFIs), community workshops and conferences, and national initiatives (including those listed in the vision document[1]). The overarching goal of this blueprint is to foster and nurture a diverse, recognized, and skilled CI workforce that can accelerate and amplify the transformative impact of CI across all S&E research and education.

[1] "Transforming Science Through Cyberinfrastructure: NSF's Blueprint for a National Cyberinfrastructure Ecosystem for Science and Engineering in the 21st Century," https://www.nsf.gov/cise/oac/vision/blueprint-2019/

2 CI Workforce: Vital Importance of Communities

The powerful spectrum of skills the CI workforce brings to research computing includes inventing novel methods, developing and deploying tools and services, and applying scalable capabilities to transformative research and education programs. The CI LWD communities of interest are illustrated in Figure 1, and include CI contributors, CI Professionals, and CI Users.

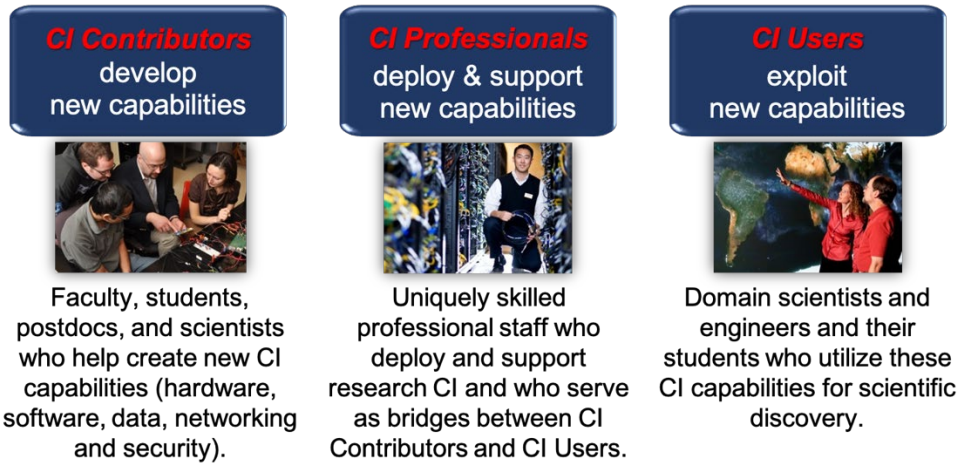


Figure 1. CI LWD communities of interest include CI Contributors who help create new CI capabilities, CI Professionals who deploy and support research CI, and CI Users who use CI capabilities. These communities collectively accelerate and amplify the transformative impact of CI across all S&E research and education.

CI Contributors: This is the community of computational, data, and domain scientists and engineers who research and develop new CI capabilities, approaches, and methods. The CI Contributors community includes faculty and research scientists in computing as well as other fields, including their undergraduate and graduate students and postdoctoral fellows, who develop new knowledge in the design, development, and utilization of robust research CI. CI Contributors explore all aspects of advanced CI including new scalable models and simulations as well as architecture and middleware for extreme-scale systems, scalable algorithms and applications, software at various levels of the scientific software stack, and a robust advanced CI ecosystem to enable major advances in potentially transformative fundamental research.

CI Professionals: This is the community of research CI and professional staff who deploy, manage, and support the effective use of research CI. The broad CI Professionals community includes the information technology professionals, scientists, and engineers who work closely with computational and data-enabled scientific and engineering researchers at colleges and universities, supercomputing and other centers, and research laboratories. Examples of CI Professionals include CI system administrators, CI research support staff, CI research software engineers, and CI facilitators, and may also include computational research scientists and engineers and non-tenure-track faculty. CI Professionals provide a broad spectrum of skills and expertise to the scientific and engineering research enterprise, whether for a single research project or a set of research projects, for a single department or campus or across multiple departments and campuses, or for a single scientific discipline or across multiple disciplines.

CI Users: This is the community of domain scientists and engineers who effectively exploit – or seek to exploit more effectively – advanced CI capabilities and methods for S&E research. The CI Users community includes the much larger, diverse groups of science and engineering faculty and researchers, including their undergraduate and graduate students and postdoctoral fellows, who are the current and potential users of new advanced CI capabilities for their S&E research activities.

In addition to addressing the preparation, mentoring, professional development, and career growth across all aspects of the CI workforce, a key focus of this blueprint is on further nurturing and broadening these communities of CI Professionals, Contributors and Users. For example, CI Professionals can build bridges to advance each element of the CI ecosystem (Figure 2) throughout the scientific discovery and innovation life cycle. Increased community among the CI workforce can also broaden the pool of agile talent able to pursue emerging technologies across the ecosystem. Collaborations among S&E faculty, students, postdocs, and CI Professionals can heighten awareness of opportunities, for example, to better support research projects that have been allocated advanced computing resources.

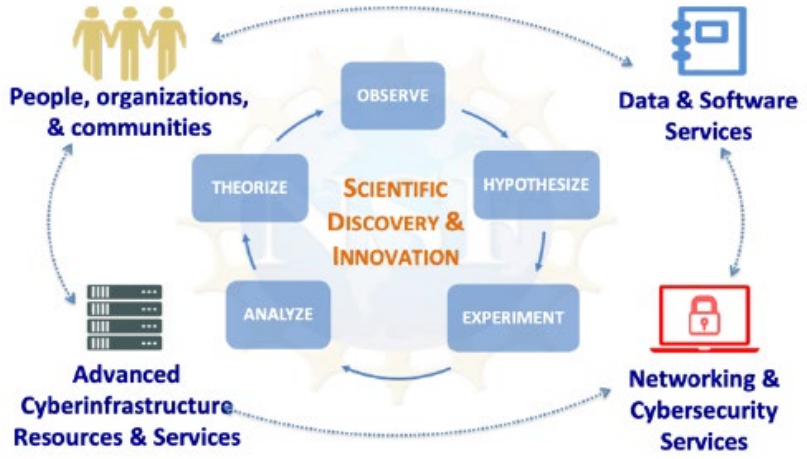


Figure 2. The CI workforce is embedded in the life cycle to heighten efficacy and awareness.

The interdisciplinary, sometimes geographically distributed, and highly diverse nature of CI workforce contributions can also present challenges, particularly related to their recognition, career paths, and sustainability. For example, these roles may not align with traditional structures and processes at academic institutions and as a result, may lack appropriate professional career paths. To be referred to as a "CI unicorn" can be both a compliment and a source of frustration for such an individual. Increasing community among these experts can improve both recognition of their vital roles in the research enterprise and the long-term sustainability of the overall CI workforce. CI workforce technical excellence has been encouraged successfully through self-organized communities, e.g., Cyber Teams, fueled by shared regional or discipline/field-specific research interests. Their sustainability is strengthened as communities focus on the triad of technical excellence, awareness to match research needs with services, and recognition of mutual interests.

This blueprint especially recognizes the multifaceted nature of increasing the sustainability of CI Professionals communities. This sustainability can be viewed analogously to the definition of sustainable software: it is essential for capacity to endure, adapt and grow over time, across

technical platforms, and to meet new needs[2]. CI Professionals community capacity encompasses the complete lifecycle from organization and recruitment, to training and research, communications, and deployments, to skill refinement and reorganization. CI Professional community endurance is achieved when CI ecosystem stakeholders realize a favorable balance of benefits and costs. Each type of CI ecosystem stakeholder may perceive the value of these transactions differently[3]. Researchers can gain valuable scientific insights that are more efficiently achieved through funding team members who are CI Professionals. CI Professionals' investment of training time and effort may be compensated monetarily, as well as via job satisfaction, advancement along a career path, resume building, and community involvement. Although their levels of recognition may differ, institutions can see returns on their investment in CI Professionals' development in the form of new research awards, publications, and an ability to recruit faculty, students, and staff. Industry partners' support of CI Professionals can yield benefits such as product revenue, brand equity, awareness of customer needs, talent recruitment, and corporate citizenship.

Advancement of CI LWD is essential to NSF's mission to support research in all fields of S&E. These advances are pursued holistically through a portfolio of investments across all aspects of the interdisciplinary CI ecosystem, i.e., resources, tools, services, and expertise. NSF's support to institutions of higher education nationwide creates opportunities to heighten awareness and formalization of burgeoning career pathways for CI Professionals. This blueprint highlights numerous successful programs that will be leveraged to implement actions, such as promoting mentorship and professional-development opportunities among CI Professionals and codifying job titles across institutions. NSF's ongoing investments in LWD will be fully aligned with relevant recommendations for CI workforce development from the national strategic plan produced by the Subcommittee on Future Advanced Computing Ecosystem (SC-FACE) of the National Science and Technology Council[4].

This blueprint outlines how NSF, through OAC, CISE, and other directorates and divisions, will continue to strengthen and scale up community development across the CI Professionals, Contributors and Users communities with ready capacity, heightened awareness, greater diversity, agile capability, and long-term sustainability.

3 NSF's Learning and Workforce Development (LWD) Investments

NSF's CI Learning and workforce development portfolio, through programs led by OAC, supports research and education to prepare, nurture, grow and coordinate the national scientific workforce for creating, employing, and supporting advanced CI. This includes cross-cutting OAC **programs** in advanced cyberinfrastructure research; in education and training; and in research workforce development and career advancement. Projects funded as part of this portfolio support development of new knowledge in the innovative design, development, and utilization

[2] "Working Towards Sustainable Software for Science – Practice and Experiences," http://wssspe.researchcomputing.org.uk/
[3] Berente, N., Howison, J., King, J.L., Cutcher-Gershenfeld, J., Pennington, R., Leading Cyberinfrastructure Enterprise: Value Propositions, Stakeholders, and Measurement (March 26, 2014). Available at SSRN: http://ssrn.com/abstract=2416247
[4] "Pioneering the Future Advanced Computing Ecosystem: A Strategic Plan," https://www.nitrd.gov/pubs/Future-Advanced-Computing-Ecosystem-Strategic-Plan-Nov-2020.pdf

of robust research CI, sustainable and scalable models of education and training, and research institutions' workforce development or career advancement of current and future generations of talented experts. Stakeholders interacting in the CI ecosystem include the computing, computational and data scientists and engineers who are researchers and developers of new capabilities, the research computing and professional staff who support the productive use, sustainable maintenance, and effective governance of hardware and software systems, and the scientists and engineers who effectively exploit advanced CI capabilities to advance their discovery and innovation efforts. CISE is committed to broadening participation in computing (BPC[5]) and strongly encourages meaningful actions that address the longstanding underrepresentation in computing and closely related disciplines of various populations including women, minorities (African Americans/Blacks, Hispanic Americans, American Indians, Alaska Natives, Native Hawaiians, Native Pacific Islanders and persons from economically disadvantaged backgrounds), and persons with disabilities.

3.1 Training-based Workforce Development for Advanced Cyberinfrastructure

While aspects of LWD are integrated in all OAC programs[6], a flagship program in NSF's LWD portfolio is **Training-based Workforce Development for Advanced Cyberinfrastructure (CyberTraining)[7]**. The CyberTraining program seeks to prepare, nurture, and continue to grow the national scientific research workforce. The goals of CyberTraining are to: (i) ensure broad adoption of CI tools, methods, and resources by the research community in order to catalyze major research advances and to enhance researchers' abilities to lead the development of new CI; and (ii) integrate core literacy and discipline-appropriate advanced skills in advanced CI as well as computational and data-driven science and engineering into the Nation's educational curriculum/instructional material fabric spanning undergraduate and graduate courses for advancing fundamental research. The CyberTraining program calls for innovative, scalable training, education, and curriculum/instructional materials—targeting one or both solicitation goals—to address the emerging needs and unresolved bottlenecks in scientific and engineering research workforce development, from the postsecondary level to active researchers. The funded activities, spanning targeted, multidisciplinary communities, are expected to lead to transformative changes in the state of research workforce preparedness for advanced CI-enabled research in the short- and long-terms.

[5] "Broadening Participation in Computing (BPC)" https://www.nsf.gov/funding/pgm_summ.jsp?pims_id=505891
[6] For example, a prior CC* solicitation (NSF 20-507, https://www.nsf.gov/publications/pub_summ.jsp?ods_key=nsf20507) explicitly targeted facilitation of the use of campus computing resources to fund Cyber Team expertise.
[7] "Training-based Workforce Development for Advanced Cyberinfrastructure (CyberTraining)," https://www.nsf.gov/funding/pgm_summ.jsp?pims_id=505342

3.2 Early-Career Development

The NSF **Faculty Early Career Development (CAREER)**[8], the **CISE Research Initiation Initiative (CRII)**[9], and the **Engineering Research Initiation (ERI)**[10] programs are all intended to support faculty in the early stages of their research and teaching careers. These programs typically support for mentoring graduate students and postdocs as well. The **Research Experiences for Undergraduates Sites and Supplements (REU)** program[11] supports undergraduates to engage in meaningful research across all S&E fields supported by NSF, to reach even earlier in the research workforce pipeline. OAC's involvement in these NSF-wide program ensures that all aspects of CI and the broad range of CI communities' interests are served.

3.3 Additional LWD Professional Development Programs and Activities

Multiple programs have built momentum for development of the cyberinfrastructure workforce. Cyber Teams formed with support from the **Campus Cyberinfrastructure (CC*)**[12] program have demonstrated novel networking methods to form communities and support researchers. The OAC Core Research program[13] supports all aspects of advanced CI research that will significantly impact the future capabilities of advanced research CI, as well as the research career paths of computer as well as computational and data-driven scientists and engineers. The network of Advanced Cyberinfrastructure Research and Education Facilitators (ACI-Ref)[14] assisted researchers to better leverage shared HPC resources. ACI-REF also highlighted the value of facilitation to advance S&E research methods. Novel learning pathways for CI workforce development are being produced in the **Data Science Corps**[15] program, part of the NSF-wide Harnessing the Data Revolution (HDR) Big Idea[16]. The ongoing **Cyberinfrastructure for Sustained Scientific Innovation (CSSI)** [17] program supports many projects where CI Professionals develop and sustain advanced CI software and data tools and services.

4 Building on Community Inputs

While issues related to LWD permeate across all advanced CI related events, discussions, and reports, the following recent events have significantly informed this blueprint.

[8] "Faculty Early Career Development Program (CAREER)," https://www.nsf.gov/funding/pgm_summ.jsp?pims_id=503214
[9] "Computer and Information Science and Engineering (CISE) Research Initiation Initiative (CRII)," https://www.nsf.gov/funding/pgm_summ.jsp?pims_id=504952
[10] "Engineering Research Initiation (ERI)," https://www.nsf.gov/funding/pgm_summ.jsp?pims_id=505888
[11] "Research Experience for Undergraduates (REU)," https://www.nsf.gov/funding/pgm_summ.jsp?pims_id=5517
[12] "Campus Cyberinfrastructure (CC*)," https://www.nsf.gov/funding/pgm_summ.jsp?pims_id=504748
[13] "Computer and Information Science and Engineering (CISE): Core Research Programs," https://www.nsf.gov/funding/pgm_summ.jsp?pims_id=505667
[14] "Advanced Cyberinfrastructure Research and Education Facilitators (ACI-Ref)," https://aciref.org
[15] "Harnessing the Data Revolution (HDR): Data Science Corps (DSC)," https://www.nsf.gov/funding/pgm_summ.jsp?pims_id=505536
[16] "Harnessing the Data Revolution" https://www.nsf.gov/cise/harnessingdata/
[17] "Cyberinfrastructure for Sustained Scientific Innovation (CSSI)," https://www.nsf.gov/funding/pgm_summ.jsp?pims_id=505505

4.1 Key Sources of Community Feedback

Responses were received from the *NSF Request for Information (RFI) on Future Needs for Advanced Cyberinfrastructure to Support Science and Engineering Research (NSF CI 2030)*[18]. That RFI, initiated on behalf of the NSF Advisory Committee on Cyberinfrastructure (ACCI), explored the advanced CI needs of the S&E community over the next decade. NSF received over 130 responses comprising over 300 individual contributors from across the research community – from research institutions, science-related organizations, industry, as well as many of NSF's major multi-user research facilities. The responses spanned an extraordinary range of research fields and were uniformly thoughtful and urgent in highlighting the many technical, process, and practice innovations needed to address present and anticipated trends in the scientific discovery enterprise. The results have been discussed widely across the NSF research directorates and the resultant ACCI report is available online[19].

The Subcommittee on Future Advanced Computing Ecosystem (SC-FACE) of the National Science and Technology Council strategic planning activities, co-chaired by NSF, developed a strategic plan, *Pioneering the Future Advanced Computational Ecosystem: A Strategic Plan*, informed in part by a convening among multiple public and private stakeholder groups held in August 2020[20]. SC-FACE identified workforce development as a key concern and major area of investment across all federal agencies. The FACE strategic plan specifically identified expanding a diverse, flexible workforce as critical to building and sustaining the future advanced computing ecosystem needed to further key national priorities and maintain U.S. leadership in science and engineering.

NSF supported the *Cyberinfrastructure Development Workshop 2020*: *Building the research innovation workforce: A workshop to identify new insights and directions to advance the research computing community*, that was held virtually in August-September 2020[21]. The primary goal of the workshop was to explore challenges, strategies, and solutions to address the need for an advanced CI research innovation workforce, specifically focusing on issues related to the CI Professional workforce. The workshop produced recommendations intended for a broad set of stakeholders, including the CI research computing and data community itself, institutions of higher education and other research organizations, and government agencies.

Over the past several years NSF has invested, and will continue to invest, in a number of workshops with LWD components. The CyberBridges series of workshops[22] focused on some of the same issues related to CI Professionals as the 2020 Cyberinfrastructure Development Workshop, while others have focused on specific programs, such as principal-investigator (PI)

[18] "Request for Information on Future Needs for Advanced Cyberinfrastructure to Support Science and Engineering Research (NSF CI 2030)," https://www.nsf.gov/cise/oac/ci2030/ .
[19] "CI2030: Future Advanced Cyberinfrastructure: A report of the NSF Advisory Committee for Cyberinfrastructure," https://www.nsf.gov/cise/oac/ci2030/ACCI_CI2030Report_Approved_Pub.pdf
[20] "OSTP Convening: Pioneering the Future Advanced Computing Ecosystem," https://www.nitrd.gov/ostp-convening-pioneering-the-future-advanced-computing-ecosystem/
[21] "CI Workforce Development Workshop 2020," https://www.rcac.purdue.edu/ciworkforce2020/
[22] "2016 Cyberbridges workshop," http://www.cyberbridges.org/2016/ ; "2015 Cyberbridges workshop," http://cyberbridges.org/2015/

meetings for the CSSI[23] [and its predecessors, Software Infrastructure for Sustained Innovation (SI)[24] and Data Infrastructure Building Blocks (DIBBs)[25]], CC*[26], and Big Data Regional Innovation Hubs (BD Hubs)[27] programs.

NSF's overall LWD strategy and program portfolio receives guidance and input from the NSF ACCI; the NSF cross-directorate Assistant Directors (AD) Council that includes ADs and Office Heads from all the NSF research directorates and offices; the Cyberinfrastructure Strategy Group, which includes senior leadership from the NSF research directorates and offices; and directly from the research community through multiple sources including PI meetings, workshops, sessions at professional conferences, community blue-ribbon studies, and public RFIs.

4.2 Integrating Community Feedback into an Architecture for Actions

Feedback from the convenings and LWD investments leads to a CI community-building architecture with three pillars of actions, where community sustainability is maintained and strengthened through the organization of the CI workforce and a measurement of the impacts (see Figure 3). Each pillar depicts needs for action expressed by stakeholders as necessary to increase communities' shared sustainability.

The first pillar speaks to actions that promote community visions and organization. CI experts who support S&E research sometimes experience isolation because their roles may not align with departmental/institutional organization charts, or their roles continually require new funding sources. NSF-funded Cyber Teams (funded by the NSF CC* program) have demonstrated how visions common to a region, such as the Massachusetts Green High Performance Computing Center[28], the Great Plains Network[29], and SouthWest Expertise in Expanding, Training, Education and Research (SWEETER)[30] team, can enable partnerships with S&E researchers to recruit across institutions and projects, and to enable collaboration with CI Professionals who might otherwise be hard to reach. NSF will promote creation and recognition of CI career pathways at research institutions where they are not well-established. Clearer recognition of CI roles can enable broader participation, recruitment opportunities, and partnerships needed to expand and diversify the workforce.

The central pillar of CI Community Building focuses on sustainability. Sustainability is the capacity to endure, adapt and grow over time, across technical platforms, and to meet new needs. It is challenging for any given institution's staff to maintain full awareness and capacity to grow with the rapidly evolving CI platforms, tools, methods, and requisite skills. CI experts express concern with the level of professional-development support available within individual projects and

[23] "2020 NSF Cyberinfrastructure for Sustained Scientific Innovation (CSSI) Principal Investigator Meeting", https://cssi-pi-community.github.io/2020-meeting/
[24] "Software Infrastructure for Sustained Innovation (SSE, SSI, S2I2)," https://www.nsf.gov/pubs/2017/nsf17526/nsf17526.htm
[25] "Data Infrastructure Building Blocks (DIBBs) ," https://www.nsf.gov/pubs/2017/nsf17500/nsf17500.htm
[26] "2019 NSF Campus Cyberinfrastructure and Cybersecurity Innovation for Cyberinfrastructure PI Workshop", https://www.thequilt.net/public-event/2019-nsf-campus-cyberinfrastructure-and-cybersecurity-innovation-for-cyberinfrastructure-pi-workshop/
[27] "Big Data Regional Innovation Hubs (BD Hubs)," https://www.nsf.gov/funding/pgm_summ.jsp?pims_id=505185
[28] "Massachusetts Green High Performance Computing Center" https://www.mghpcc.org/
[29] "Great Plains Network" https://www.greatplains.net/
[30] "SouthWest Expertise in Expanding, Training, Education and Research" https://sweeter.cyberinfrastructure.org/

institutions. NSF will continue to build ecosystem-wide socio-technical connections to foster greater appreciation for the value of investing in professional development of CI experts. For example, NSF will continue to support national CI Centers of Excellence (CI CoE)[31] to identify and address technical, educational, and organizational challenges, for example by creating novel training materials and sharing best practices.

Organized CI communities can produce compelling rationale for sustainability support from their institutions by measuring the impact of CI expertise, which is the third pillar of the Community Building architecture. Fulton et.al. (2017)[32] describe the intrinsic difficulties faced by research institutions' attempts to quantify the return on investment for research CI, either from a conventional financial basis or from an intellectual basis. Each type of CI ecosystem stakeholder may perceive the value of CI investments differently. The Campus Research Computing Consortium (CaRCC)[33] has identified how appreciation for the research computing and data workforce can differ significantly based on the value proposition metrics of interest to the different stakeholders in the CI ecosystem. The US Research Software Engineer Association (US-RSE)[34] connects RSEs to share knowledge, professional connections, and resources; promotes RSEs' impact on research; creates technical and career development resources; and seeks to increase diversity throughout the community.

Clearer measurement of the impacts of these engagements can inform host institutions' return on investments in CI workforce development, for example, through breakthrough S&E capabilities, research awards and insights, faculty and student recruitment, bolstered endowments, and retention of institution-wide interdisciplinary CI experts. Of institutions' willingness to increase investment in CI career development on campuses, one Cyber Team PI stated, "They [institution leadership] just need to hear about it and see it." NSF anticipates that coordinated networks of CI-based research communities can strengthen collective marketing and promotion of the value propositions created by the deep technical talents of each CI practitioner. NSF can incentivize such human capital investments by promoting institutional plans for CI Professional development and mentoring that can occur within one campus or via collaborative networks.

[31] "CI Centers of Excellence" https://nsf.gov/funding/pgm_summ.jsp?pims_id=505744

[32] "XDMoD Value Analytics: A Tool for Measuring the Financial and Intellectual ROI of Your Campus Cyberinfrastructure Facilities" in Proceedings of PEARC17, New Orleans, LA, USA, July 09-13, 2017. https://dl.acm.org/doi/pdf/10.1145/3093338.3093358

[33] "Campus Research Computing Stakeholders and Value Propositions (draft)," https://CaRCC.org/wp-content/uploads/2018/12/Stakeholders.pdf

[34] "U.S. Research Software Engineer Association," https://us-rse.org

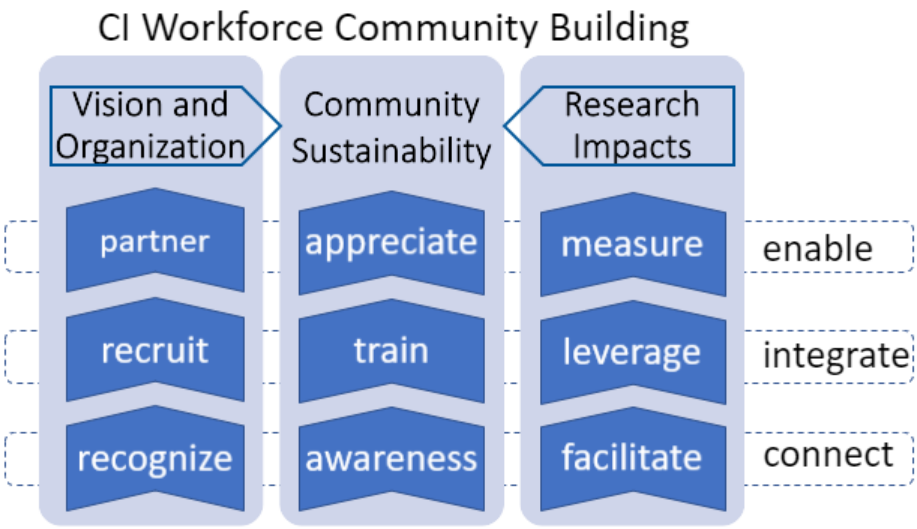


Figure 3. The architecture of a sustainable CI workforce community relies on organization and demonstrable impacts. The architecture builds upon social connectivity of well-trained experts who can integrate technologies to enable valued outcomes.

The middle integration layer across all three pillars of the architecture reflects elements of the national strategic plan for the future advanced computing ecosystem, which calls for workforce diversification, training and upskilling required to build tools and leverage large-scale systems.

5 A Blueprint for Cyberinfrastructure Learning and Workforce Development for Accelerating Science and Engineering in the 21st Century

This section presents a forward-looking blueprint for NSF's investments in CI LWD. NSF will continue and expand programs dedicated to developing instructional material for training, broadening participation, and upskilling all parts of the CI workforce. In addition to continuing NSF's portfolio of career development programs, the blueprint describes additional means to recognize CI Professionals' roles and career paths and foster and nurture associated communities.

This LWD blueprint focuses on: (1) supporting the creation of structures, processes and materials for learning and training; (2) promoting the mentoring and professional development of CI Professionals and the establishment of career pathways within research institutions; and (3) fostering and nurturing the organization of demonstrably sustainable communities. Blueprint actions will incentivize lifelong learning and advancement of skills to ensure well-connected capacity to achieve the goals of S&E researchers across the NSF CI ecosystem.

Transformative S&E research will continue to expand in speed and scale by leveraging the entire talent of the nation and requires fostering sustained collaboration among S&E researchers, CI Professionals, and other CI workforce members, in addition to establishing clear career pathways and suitable on-ramps for CI workforce members.

5.1 Curriculum Development, Training

Existing investments in CI resources and services, such as Extreme Science and Engineering Discovery Environment (XSEDE),[35] provide powerful insights into the continuously evolving CI needs of S&E researchers. These findings guide instructional and training design for both novice and more experienced CI workforce members. NSF investments will continue to enhance programs that produce tools, learning modules, and services for advanced CI training, for example, through the CyberTraining program. Training materials are also produced as part of NSF investments in CI research as well as in experimental systems and services, such as those funded by the Advanced Computing Systems and Services (ACSS) program[36].

Instructional-material development and training activities for advanced CI will also continue to be supported through the demonstration pilots and concept definition studies for CI CoE, which have as one of their focus areas learning and workforce development activities. NSF views CI CoEs as *service-oriented* hubs of expertise and innovation targeting specific areas, aspects, or stakeholder communities, with the aims of driving advancements, expanding utilization, and improving efficiency of the national research CI ecosystem through structured, strongly community-engaging and community-serving approaches. LWD-focused CI CoE activities will target the full range of CI research workforce members, from undergraduate and graduate students to postdoctoral scientists, early-career researchers, and CI Professionals. Those workforce members span all S&E areas supported by NSF, including computer scientists, computational scientists, domain scientists, and data scientists.

5.2 Career Development

Programs such as CAREER, CRII, and REU will continue to lower the cost of entry for new recruits into the CI ecosystem. Similarly, programs such as the OAC Core Research and CSSI programs will continue to support training and career development for research-workforce individuals in developing, creating, and supporting secure, advanced and scalable research CI.

As a complement to developing individuals' technical skills, this blueprint also recognizes institutions' roles in recruitment, mentorship, career development, and retention of sustainable communities of CI Professionals. Self-organized communities are best suited to harmonize definitions of CI Professionals' positions (and job titles) and types of services they offer, and to increase their accessibility. Institutions are expected to adapt so these definitions can be recognized by human-resource development plans. This blueprint encourages novel approaches to organizational structures that embrace and empower CI Professional services, such as service centers that span research departments, or co-funded regional centers with participation by multiple institutions.

NSF anticipates that attention to CI Professional mentoring will enable PIs and institutions to propose LWD-related efforts best suited to their respective organizations' sizes and disciplines. One strategy might be to consider CI Professional mentoring or professional-development plans

[35] "Extreme Science and Engineering Discovery Environment (XSEDE)," http://xsede.org

[36] "Advanced Computing Systems & Services: Adapting to the Rapid Evolution of Science and Engineering Research," https://www.nsf.gov/funding/pgm_summ.jsp?pims_id=503148

as a programmatic element. Plans could be tailorable based on team size, maturity level, number and types of contributing organizations, and opportunities to leverage applicable NSF programs. Likewise, organizations would be able to propose collaborative mentoring, institutional mentor-protégé agreements, or perhaps industry joint appointments or externships, within whichever business models can best advance their CI workforce community's sustainability goals.

The CI CoE program has already funded a demonstration pilot project that is investigating models to provide institutions and individuals the tools, support, and community to build and sustain successful CI operations, expand the development of current and new CI Professionals, and provide a connector for the broad CI Professional community[37].

5.3 Develop Capacity and Nurture Coordination

Scalable and timely provisioning of agile, responsive, reliable CI Professional services is of utmost importance to NSF. NSF appreciates the continuing growth in demand for CI Professional services to support transformative S&E research, as well as the increasing diversity across thus demand. Ongoing achievements facilitated by XSEDE demonstrate how coordinated communications with CI Professionals enables S&E research productivity. The Advanced Cyberinfrastructure Coordination Ecosystem: Services & Support (ACCESS)[38] program is expected to provide a framework for continuing this support in an agile and scalable manner by coordinating a community of CI Professionals who are appropriately located to maximize their impact on a diverse set of NSF S&E researchers. NSF aims to further broaden such coordination as additional networked communities emerge regionally and within research domains. NSF recognizes growth in demand requires increases in capacity, as well as continuous increases in capability across a spectrum of CI services, platforms, and techniques.

For example, capacity will be nurtured by better understanding the breadth and depth of the CI workforce, including CI Professionals, and developing that workforce into a set of related communities. CaRCC has facilitated institutional awareness of capacity via their Research Computing and Data Capabilities Model[39]. Furthermore, a census would enable quantification of the national landscape of the CI workforce. Finally, a CI CoE will enable understanding of requisite skills, provide a means of coordinating best practices across communities, and investigate case studies to promote greater awareness of the value and market demand for CI workforce individuals.

5.4 Plan into Action: Programs, Projects, & Opportunities to Foster Community

NSF intends to implement the different aspects of this strategic plan through a network of interrelated and coordinated programs as illustrated in Figure 4. This includes expanding existing programs as well as exploring new ones. These programs, and others, focus on CI LWD to varying

[37] "CI CoE: Demo Pilot: Advancing Research Computing and Data: Strategic Tools, Practices, and Professional Development," https://www.nsf.gov/awardsearch/showAward?AWD_ID=2100003

[38] "Advanced Cyberinfrastructure Coordination Ecosystem: Services & Support (ACCESS)," https://www.nsf.gov/funding/pgm_summ.jsp?pims_id=505866

[39] "Research Computing and Data Capabilities Model," https://carcc.org/rcdcm/

degrees. Some focus primarily on CI, with inclusion of some LWD (e.g. PATh), and others are LWD-focused with elements of CI (e.g. CAREER, INCLUDES).

A key focus in the implementation will be on growing and strengthening the CI workforce community. Fostering community between S&E researchers and CI Professionals occurs by reinforcing each of the fundamentals of sustainable ecosystems (Figure 4). Improved definitions of CI workforce members' unique roles will enable clearer recognition of those roles and their associated skills, which will in turn lead to clear communication structures. A related focus areas is promoting (and incentivizing) the mentoring, professional development, and recognition of CI Professionals. NSF can incentivize such human capital investments by promoting institutional plans for CI professional development and mentoring that can occur within a campus or via collaborative networks and partnerships.

The community's interdependence, mutual interests, and shared value propositions can potentially be realized through existing programs such as Research Coordination Networks (RCN)[40], new coordination hubs, and further coalition building. Performance monitoring and community feedback must continue to be gathered to ensure that CI workforce communities have requisite skills and ample capacity to serve programs and projects such as ACCESS, ACSS, and the Partnership to Advance Throughput Computing (PATh)[41]. NSF programs will promote diversity by broadening participation of institutions serving underrepresented communities, institutions of all types and sizes, and from across the county. Enabling inclusion of a diverse set of communities will scale up the pool of CI expertise available to support S&E research.

[40] "Research Coordination Networks," https://www.nsf.gov/funding/pgm_summ.jsp?pims_id=11691
[41] "Partnership to Advance Throughput Computing," https://path-cc.io/

NSF plans to move ahead aggressively to put this blueprint for CI LWD into action in part by leveraging and expanding existing programs, such as NSF INCLUDES[42] and CISE-MSI[43], and specific projects responsive to current and emerging S&E needs, and also in part by exploring new programs, all with input from the broad NSF S&E community.

Balance and timing are essential among investments in the advanced computing platforms and the programs that provide services and capacity to leverage those resources (see Figure 5). The ubiquitous nature of CI workforce contributions in each component of the CI ecosystem reinforces the holistic, interdependent strategy for LWD investments. Ongoing input from community-driven workshops, as well as lessons learned during each program's annual reviews, will inform investment priorities. The staggered periodicity across programs enables tactical adjustments, while ensuring uninterrupted provision of CI services at speed and scale.

Figure 5. Notional timeline of programs dedicated to CI LWD.

6 Ongoing Strategic Planning and Community Engagement

Fully consistent and aligned with the federal strategic plan for a future advanced computing ecosystem, NSF will continue its strategic planning and community engagement activities as it explores other elements of the advanced computing ecosystem, including software and data systems and services, networking, cybersecurity, and LWD. NSF notes that this blueprint is the fifth in a series of blueprints focused on different elements of the CI ecosystem to be developed in partnership with the community.

Looking to the future, the road ahead for research CI promises to be exciting with many new opportunities for growth in technology integration and workforce development. In the near term, the NSF Big Ideas[44] are moving into full gear with multiple new solicitations and are complemented by recent relevant national initiatives in areas such as Quantum Information Science (including a National Quantum Initiative Act signed into law in December 2018[45]) and a

[42] "Inclusion across the Nation of Communities of Learners of Underrepresented Discoverers in Engineering and Science," https://www.nsf.gov/funding/pgm_summ.jsp?pims_id=505289
[43] "Computer and Information Science and Engineering Minority-Serving Institutions Research Expansion Program (CISE-MSI)," https://www.nsf.gov/funding/pgm_summ.jsp?pims_id=505854
[44] "NSF Big Ideas" https://www.nsf.gov/news/special_reports/big_ideas/index.jsp
[45] "National Quantum Initiative," https://www.congress.gov/115/bills/hr6227/BILLS-115hr6227enr.pdf

National Artificial Intelligence Initiative Act[46] signed into law in January 2021 (including an Executive Order on *Maintaining American Leadership in Artificial Intelligence*[47]). For example, NSF's investments through the Quantum Leap Big Idea[48], including the Quantum Computing & Information Science Faculty Fellows (QCIS-FF) program (NSF 19-507)[49], as well as foundational and use-inspired Artificial Intelligence[50] research, including the National Artificial Intelligence (AI) Research Institutes program[51], will help define the nature and structure of the CI ecosystem over the longer term. NSF looks forward to continuing to work with the community to define the future of cyberinfrastructure research and research cyberinfrastructure, with the overarching goal of realizing an integrated, expansive CI ecosystem that transforms all of S&E research and education.

7 Conclusion

The NSF-funded CI ecosystem, comprised of a rich foundation of resources, a fabric of services, and a continually evolving expert workforce, is playing an increasingly critical role across all of S&E research and education, enabling discoveries and driving innovation. It is an important part of the national CI ecosystem that is critical for ensuring US leadership in S&E, economic competitiveness, and national security. As a result, it is essential that NSF strategically rethink, evolve, and incent sustainability of this CI ecosystem in response to the changing nature of needs of S&E, driven by the dynamic technology landscape, evolving career pathways, and informed by community inputs. This document builds on NSF's recently articulated vision for a national CI ecosystem. The envisioned CI ecosystem integrates computational, data, software, networking, and security resources, tool and services, and computational and data skills and capacity of expertise towards collectively enabling new, transformative discoveries across all of S&E. This document provided NSF's blueprint for cyberinfrastructure learning and workforce development to support science and engineering research and education in the 21st century. It also outlined a plan to implement this blueprint by producing LWD materials, enhancing recognition of CI experts' career pathways, nurturing and broadening community development, and thereby expanding CI ecosystem capacity. The vision and blueprint presented in this document have been informed by the community through advisory bodies, public RFIs, community workshops and conferences, and national initiatives. NSF intends to continue to work with community members, institutions, and industry to evolve and implement the vision and blueprint presented in this document, as well as to develop complementary blueprints for other CI elements.

[46] "National Artificial Intelligence Initiative Act" https://www.congress.gov/116/crpt/hrpt617/CRPT-116hrpt617.pdf#page=1210
[47] "Executive Order on Maintaining American Leadership in Artificial Intelligence," https://www.federalregister.gov/documents/2019/02/14/2019-02544/maintaining-american-leadership-in-artificial-intelligence
[48] "NSF Quantum Leap Big Idea," https://www.nsf.gov/news/special_reports/big_ideas/quantum.jsp
[49] "Quantum Computing & Information Science Faculty Fellows (QCIS-FF)," https://www.nsf.gov/funding/pgm_summ.jsp?pims_id=505535
[50] "Artificial Intelligence (AI) at NSF," https://nsf.gov/cise/ai.jsp
[51] "National Artificial Intelligence (AI) Research Institutes," https://www.nsf.gov/funding/pgm_summ.jsp?pims_id=505686