The Connect.Cyberinfrastructure Portal

Julie Ma
Massachusetts Green High Performance Computer Center, Holyoke, MA
jma@mghpcc.org

Sarah Akbar
Worcester Polytechnic Institute, Worcester, MA
sakbar@wpi.edu

Torey Battelle
Colorado School of Mines, Golden, CO
battelle@mines.edu

Kevin Brandt
South Dakota State Univ., Brookings, SD
kevin.brandt@sdstate.edu

Eric Brown
Univ. of New Hampshire, Durham, NH
eric.brown@unh.edu

Dana Brunson
Internet2, Stillwater, OK
dbrunson@internet2.edu

Dhruva K. Chakravorty
Texas A&M Univ., College Station, TX
chakravorty@tamu.edu

Thomas Cheatham Iii
Univ. of Utah, Salt Lake City, UT
tec3@utah.edu

Bhushan Chitre
Univ. of Kentucky, Lexington, KY
bch229@uky.edu

Adrian Del Maestro
Univ. of Tennessee, Knoxville, TN
Adrian.delmaestro@utk.edu

Andrea Elledge
Univ. of Vermont, Burlington, VT
Andrea.Elledge@uvm.edu

Vikram Gazula
Univ. of Kentucky, Lexington, KY
gazula@uky.edu

John Goodhue
Massachusetts Green High Performance Computer Center, Holyoke, MA
jtgoodhue@mghpcc.org

James Griffioen
Univ. of Kentucky, Lexington, KY
griff@uky.edu

Abigail Hyde
Worcester Polytechnic Institute, Worcester, MA
aghyde@wpi.edu

Douglas Jennewein
Arizona State Univ., Tempe, AZ
douglas.jennewein@asu.edu

Shelley Knuth
Univ. of Colorado, Boulder, CO
shelley.knuth@colorado.edu

Bj Lougee
Federal Reserve Bank of Kansas City, Kansas City, KS
bj.lougee@kc.frb.org

Timothy Middelkoop
Internet2, Colombia, MO
tmiddelkoop@internet2.edu

Sasmita Mohapatra
Univ. of Texas at Dallas, Dallas, TX
sasmita@utdallas.edu

Sia Najafi
Worcester Polytechnic Institute, Worcester, MA
snajafi@wpi.edu

Kaylea Nelson
Yale Univ., New Haven, CT
kaylea.nelson@yale.edu

Lisa Perez
Texas A&M Univ., College Station, TX
perez@tamu.edu

Bruce Segee
Univ. of Maine, Orono, ME
segee@maine.edu

Julia Sheats
Worcester Polytechnic Institute, Worcester, MA
jsheats@wpi.edu

Andrew Sherman
Yale Univ., New Haven, CT
andrew.sherman@yale.edu

Christopher Simmons
Univ. of Texas at Dallas, Dallas, TX
csim@utdallas.edu

Ermal Toto
Worcester Polytechnic Institute, Worcester, MA
toto@wpi.edu

Scott Valcourt
Roux Institute at Northeastern Univ., Portland, ME
s.valcourt@northeastern.edu

Lucas Varella
Worcester Polytechnic Institute, Worcester, MA
lnvarella@wpi.edu

Ralph Zottola
Univ. of Alabama at Birmingham, Birmingham, AL
rzottola@uab.edu

ABSTRACT

The Connect.Cyberinfrastructure.org Portal, originally known as the Cyberteam Portal, was developed to support the management of project workflows and to capture project results for the Northeast Cyberteam (NECT) [3, 5, 6]. Recently, the Portal has expanded to provide support for other programs in the Research Computing ecosystem, creating opportunities for collaboration, and leveraging a consistent, cohesive approach to common challenges. As reported at SC20 [6], a pilot was launched in July 2020 to enable six additional Cyberteam programs to explore the use of the Portal as a management tool for their related programs. In addition, in January of 2021, the Extreme Science and Engineering Discovery Environment (XSEDE) Campus Champions leadership decided to use the Portal to modernize participant management and onboarding functions. Portal details, preliminary results, and future plans are discussed.

PEARC '21, July 18–22, 2021, Boston, MA, USA

ACM ISBN 978-1-4503-8292-2/21/07…$15.00
https://doi.org/10.1145/3437359.3465606

CCS CONCEPTS

• **Social and professional topics** → Professional topics; Computing education; Informal education; Professional topics; Computing profession; Computing occupations.

KEYWORDS

workforce development, research computing facilitator, Ask.CI, Cnct.CI, Northeast Cyberteam, SWEETER Cyberteam, CAREERS Cyberteam, KY Cyberteam, GPN Cyberteam, Campus Champions, RMACC

ACM Reference Format:
Julie Ma, Sarah Akbar, Torey Battelle, Kevin Brandt, Eric Brown, Dana Brunson, Dhruva K. Chakravorty, Thomas Cheatham Iii, Bhushan Chitre, Adrian Del Maestro, Andrea Elledge, Vikram Gazula, John Goodhue, James Griffioen, Abigail Hyde, Douglas Jennewein, Shelley Knuth, Bj Lougee, Timothy Middelkoop, Sasmita Mohapatra, Sia Najafi, Kaylea Nelson, Lisa Perez, Bruce Segee, Julia Sheats, Andrew Sherman, Christopher Simmons, Ermal Toto, Scott Valcourt, Lucas Varella, and Ralph Zottola. 2021. The Connect.Cyberinfrastructure Portal. In *Practice and Experience in Advanced Research Computing (PEARC '21), July 18–22, 2021, Boston, MA, USA.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3437359.3465606

1 ORIGINS

The Connect.Cyberinfrastructure.org Portal, originally known as the Cyberteam Portal, was developed to support the management of project workflows and to capture project results for a series of more than 50 projects launched by the Northeast Cyberteam (NECT). The NECT is a NSF-sponsored program that aims to make regional and national cyberinfrastructure more readily accessible to researchers in small and mid-sized institutions in northern New England [3, 5, 6] by providing research computing facilitation support and aggregating access to knowledge resources.

The Portal performs four principal functions: (1) manage project workflows and capture project results; (2) manage Cyberteam participants; (3) match students and mentors with projects; and (4) provide a front door to the growing research computing knowledge base which has been developed for the community as an outcome of the Northeast Cyberteam program. Most of these functions benefit from broad participation because deep knowledge is widely dispersed throughout the research computing community. Recognizing the potential advantages of a central access point, the Portal was developed with an eye toward making it possible for other programs to participate without incurring significant cost and while maintaining their own branding and project workflows. [1]

As reported in [6], a pilot was launched in July 2020, where custom Portal homepages, referred as "views", were created to enable the CAREERS [7], Great Plains Network [8], RMACC [9], SWEETER [10], Kentucky [11] and, TRECIS Cyberteams [12]; as well as Colorado School of Mines to explore the use of the Portal as a management tool for their related programs. In January 2021, the leadership team of the XSEDE Campus Champions decided to use the Portal to modernize participant management and onboarding functions, and to support a new enhancement to the program. Current status, an example of the efficiency resulting from leveraging Portal functions, and future plans are described below.

2 PORTAL OVERVIEW

2.1 Tags

Central to the Portal is an extensive and ever-evolving tagging infrastructure which informs every aspect of the Portal. The tag taxonomy was initially developed by the Northeast Cyberteam to categorize subject matter relevant to Research Computing Facilitators (RCFs) and is ever changing due to the frequent introduction of new technology in domains that characterize the field of research computing. The Portal makes extensive use of the database and tagging functions that are provided by the underlying Drupal platform. There are several types of information that are available from the Portal: people profiles, project profiles, and three self-service learning resources that comprise the knowledge base. Each element can be assigned any number of relevant tags. Subsequently, tags are used to search the database at a granular, yet curated level to find pertinent content for the task at hand. This infrastructure allows a user to click on a tag from any part of the Portal and obtain a listing of all Portal content, including mentor profiles, project profiles, frequently asked questions, and training resources.

2.2 Portal Development

The development team for the Portal is comprised primarily of undergraduate students from Worcester Polytechnic Institute (WPI) with technical leadership provided by the WPI Advanced Research Computing group and by staff at the University of New Hampshire. The Portal Team uses an agile software development methodology and meets four times per week for brief standups. New features are created in sprints on a development site which mirrors the production site. Once development is complete, the features are pushed to a staging site and subjected to an extensive regression test. After passing the regression test, the features are packaged into a new release that is pushed to the production site. Releases occur every 2-6 weeks depending upon the nature of the sprint.

2.3 Portal Elements

2.3.1 Views. Each participating program has a custom *view* into the Portal which displays core information about the program, as well as featured projects. At present, the Cyberteams adhere to a common page structure for their views; a unique home page view was developed for the Campus Champions.

2.3.2 People/Accounts. Much of the knowledge base information available on the Portal is in the public domain. To access program resources, participants register by creating user accounts. All new accounts are automatically instantiated, although site administrators reserve the right to remove accounts in the event of disruptive behavior. The Portal supports login via username and password and leverages CiLogon support for identities from the InCommon Federation, Google, and other providers. Once an account is created, participants update their user profiles to identify their role(s) in the program (mentor, researcher, student-facilitator, etc.) and add tags representing skills and interests. "Skills" tags identify topics on which one is willing to answer questions and "interest" tags identify topics about which one is interested in learning.

2.3.3 Projects. Projects are used to communicate key activities of the program. On the Cyberteam pages, projects are displayed in a

carousel at the bottom of the page, providing a convenient snapshot of the Cyberteam's activities. The Northeast and CAREERS Cyberteams provide direct support to faculty at small and mid-sized institutions by launching projects that pair a student-facilitator with an experienced mentor to address an immediate faculty research need. [3, 5, 6] Projects draw mentors and students from across the region. Mentors are usually professional RCFs from research computing groups at larger institutions in the region who have subject matter expertise relevant to the project. When a project is created on the Portal, relevant tags are assigned to the project. The program managers then search the database for mentors and students with these tags to extract a list of potential candidates for the project.

Simultaneously, students and mentors can asynchronously check the Portal for new project listings and click on an "I'm Interested" button for projects that are recruiting. "I'm Interested" clicks are reported to the program managers and facilitate the matching of participants to projects. Once a match has occurred, interviews are conducted, and the project is launched when a suitable student and mentor are found for the project. These mechanisms have enabled successful creation of teams to support more than 50 Northeast Cyberteam projects since program inception in May 2017 and over a dozen CAREERS Cyberteam projects since the inception of that program in July 2020.

2.3.4 Knowledge Base. The Portal is used to access and aggregate self-service learning resources into a knowledge base that provides just-in-time information delivery to participants. As described in [3, 5, and 6], one of the most fundamental skills of a successful RCF is the ability to learn quickly enough about new domains and applications to aid and solve problems drawing on parallels with their existing knowledge wherever possible. Tools that enable self-service and just in time learning are, therefore, a core element of the Portal. The knowledge base accommodates three types of information commonly needed by RCFs: (1) Frequently asked questions whose answers evolve over time as technology advances, serviced by Ask.CI, a Q&A Platform for Research Computing [4]; (2) Relatively static information such as introductory training modules for Linux cluster access, programming languages and schedulers, serviced by a Learning Resources Wiki; and (3) Dynamic, situation-specific information needed to solve an immediate problem. Typically handled by a help desk at larger institutions, this need is serviced by a view-specific Help Desk which can be staffed by participants in the program. All of the knowledge base content is tagged and shows up in tag search results.

2.4 Ask.CI Integration

Ask.CI, [2] developed by the Northeast Cyberteam in partnership with the Campus Champions, is implemented on a Discourse.org platform. The Portal Team was able to make use of the Discourse Application Programming Interface (API) to extract relevant content from Ask.CI, based on tags, and display it on the Portal as the result of a tag search. This technique gives the Portal seamless access to Ask.CI content and opens the possibility of future enhancements that leverage other independent resources. Periodic reconciliation of tags between Ask.CI and the Portal are performed to ensure that content is classified consistently between the platforms.

3 EXPANSION

As noted in the introduction, the value of the mentor matching and targeted learning resource identification services provided by the Portal grows as the number and diversity of participants grows. [6] Currently, there are seven different programs piloting use of the Portal. Each one is exploring aspects of the Portal that meet their program needs. For example, the Kentucky Cyberteam is exploring use of the Regional Help Desk and its associated ticketing capability to track engagements with its constituents. TRECIS is using an Ask.CI locale as the entry point into their Cyberteam resources. In perhaps the most comprehensive effort, the CAREERS Cyberteam has adopted the NECT experiential learning model and engaged the NECT program manager to serve as a co-program manager to ensure efficient transfer of knowledge. Feedback from these programs has inspired enhancements that are of value to all. Leadership from each of the programs meet monthly to exchange status and ideas, creating an informal and convivial vehicle to find opportunities to collaborate and leverage best practices. Some time is also allocated to prioritize feature requests and to curate the tag taxonomy. The hope is that standardizing on the taxonomy will create a level of cohesion and consistency across the programs to the benefit of the community.

4 CASE STUDY: CAREERS CYBERTEAM

The CAREERS Cyberteam launched in July 2020 with the goal of making cyberinfrastructure more available to researchers at small/medium-sized institutions in six Northeastern states, taking a two-pronged approach to this challenge: (1) Provide direct support to researchers while building the workforce pipeline by creating experiential learning opportunities for student facilitators working with experienced mentors to help a researcher/educator make use of cyberinfrastructure (2) Create a distributed pool of professional RCFs in the region that can provide direct support in specific areas of expertise to researchers/educators at small/medium-sized institutions.

Recognizing the value of leveraging existing tools and experience, CAREERS adopted Northeast Cyberteam methods and made full use of the project workflows embedded in the Portal to support the experiential learning opportunities as described in section 1. Along the way, the Steering Committee identified opportunities to improve the Portal and associated processes with particular emphasis on scaling to the larger number of states covered by CAREERS. As a result, the CAREERS Steering Committee has been able to focus most of its energy on finding students, mentors, and project opportunities rather than building infrastructure and process. This made it possible to launch 16 projects during the first nine months of the program, each one providing a 3–6-month experiential learning opportunity for a student facilitator. CAREERS and Northeast Cyberteam students are encouraged to participate in projects from both Programs.

The CAREERS Steering Committee is currently working on the second prong of its program, a RCF pool, and is evaluating the affinity group features developed for Campus Champions (see below) in combination with the Regional Help Desk for use in this part of the program.

5 NEXT STEPS

Through the expansion pilot, the Portal team is working with several programs that leverage the Portal in ways that go beyond Cyberteam project management and knowledge sharing.

Campus Champions: The leadership team recognized that the Regional program, which encourages members to make connections with nearby Champions, was having varying degrees of success depending upon accessibility and geography, and decided to replace it with one that is organized around the notion of "affinity groups", which encourage Champions to gather based on common interests - discipline, technology interests or needs, or proximity. Shortly after the Portal expansion pilot was launched, it was observed that the tag infrastructure could be used to identify Champions with similar affinities. A recently introduced enhancement to the Portal now allows participants to dynamically create groups based on common interests, with each group defined by a collection of tags. Affinity Groups are being introduced to the community at the time of this writing, and results will be reported in subsequent publications.

Other Cyberteams: As noted earlier, most of the Cyberteam programs participating in the expansion pilot are making use of or planning to make use of portal features in some capacity. Case studies of successes and challenges resulting from these efforts will be developed as they reach maturity.

6 CONCLUSION

Connect.Cyberinfrastructure.org started as a tool with the specific purpose of supporting Northeast Cyberteam project workflows and aggregating knowledge resources for its constituents. Through the expansion pilot, it has found use in programs that are solving similar problems and appears to be useful in several related areas. Opportunities to partner with other communities of practice are constantly sought after and welcomed as the strength of Portal features such as mentor matching and aggregating learning resources grows with the size of the population participating, and the resulting knowledge base and collective expertise deliver benefits that are greater than the sum of individual parts. We will continue to update the community on how this adventure is playing out in future publications.

ACKNOWLEDGMENTS

This material is based upon work supported by the National Science Foundation under Grants No. OAC-1659377, "CC* Cyber Team: Improving Access to Regional and National Cyberinfrastructure for Small and Mid-Sized Institutions", OAC-1925764, "CC* Team: SWEETER: South West Expertise in Expanding Training Education and Research", OAC-2018873, "CC* Team: CAREERS: Cyberteam to Advance Research and Education in Eastern Regional Schools", OAC-1925687, "CC* Team: KyRC - A Kentucky Research Computing Team", OAC-1548562, "XSEDE 2.0: Integrating, Enabling and Enhancing National Cyberinfrastructure with Expanding Community Involvement", and OAC_1925681, "CC* Team: Great Plains Regional CyberTeam".

REFERENCES

[1] Goodhue, J., Ma, J., Del Maestro, A., Najafi, S., Segee, B., Valcourt, S., Zottola, R. "Northeast Cyberteam - Building an Environment for Sharing Best Practices and Solutions for Research Computing," 2020 IEEE High Performance Extreme Computing Conference (HPEC), 2020, pp. 1-5, doi: 10.1109/HPEC43674.2020.9286254.

[2] Ma, J., Culich, A., Sochat, V., Najafi, S., Segee, B., Braiterman, Z., Battelle, T., Goodhue, J., Brunson, D., Hill, C., Zottola, R., Singh, R., Bulekova, K., Pessin, J., Cheatham, T., DelMaestro, A., Valcourt, S., Thoelen, R., Smith, J. "Ask.Cyberinfrastructure.org: Creating a Platform for Self-Service Learning and Collaboration in the Rapidly Changing Environment of Research Computing." SEHET20, Portland, OR, July 26-30, 2020. Unpublished.

[3] Goodhue, J., Ma, J., Del Maestro, A., Najafi, S., Segee, B., Valcourt, S., Zottola, R. "Northeast Cyberteam:Workforce Development for Research Computing at Small and Mid-sized Institutions." Proceedings of the Conference on Practice and Experience in Advanced Research Computing (PEARC20), Portland, OR, July 26-30, 2020, doi:10.1145/3311790.3396662.

[4] Ma, J., Culich, A., Sochat, V., Najafi, S., Segee, B., Braiterman, Z., Battelle, T., Goodhue, J., Brunson, D., Hill, C., Zottola, R., Singh, R., Bulekova, K., Pessin, J., Cheatham, T., DelMaestro, A., Valcourt, S., Thoelen, R., Smith, J. "Ask.Cyberinfrastructure.org: Creating a Platform for Self-Service Learning and Collaboration in the Rapidly Changing Environment of Research Computing." The Journal of Computational Science Education, vol. 12, no. 2, 2021, pp. 37–40., doi:10.22369/issn.2153-4136/12/2/9.

[5] Goodhue, J., Ma, J., Del Maestro, A., Najafi, S., Segee, B., Valcourt, S., Zottola, R. "Northeast Cyberteam Program – A Workforce Development Strategy for Research Computing." The Journal of Computational Science Education, vol. 11, no. 1, 2020, pp. 8–11., doi:10.22369/issn.2153-4136/11/1/2.

[6] Goodhue, J., Ma, J., Del Maestro, A., Najafi, S., Segee, B., Valcourt, S., Zottola, R. "Extended Abstract: Northeast Cyberteam - Methods, Results and Expansion via the Cyberteam Portal." 2020. Proceedings of Supercomputing (SC20). Accepted for Publication.

[7] Home. CAREERS Cyberteam. (2021, May 25). https://careers-ct.cyberinfrastructure.org/.

[8] CyberTeam. Great Plains Network. (2020, July 10). https://www.greatplains.net/community-initiatives/cyberteam/.

[9] RMACC Cyberteam. (2021, May 25). https://rmacc.cyberinfrastructure.org/contact/rmacc_cyberteam.

[10] Texas A&M High Performance Research Computing. CC* SWEETER | High Performance Research Computing. (2021, May 25). https://hprc.tamu.edu/sweeter.

[11] KyCyberteam. (2021, May 25). https://kycyberteam.ccs.uky.edu/.

[12] Home. TRECIS Cyberteam. (2021, May 25). https://trecis.cyberinfrastructure.org/.