Scaling HPC Education

SUSAN MEHRINGER, Center for Advanced Computing, Cornell University, USA
KATE CAHILL, Ohio Supercomputer Center, USA
CHARLIE DEY, Texas Advanced Computing Center, USA
RICH KNEPPER, Center for Advanced Computing, Cornell University, USA
JOHN-PAUL NAVARRO, University of Chicago, Argonne National Lab, USA
JEAIME H. POWELL, Texas Advanced Computing Center, USA
MARY THOMAS, San Diego Supercomputing Center, University of California San Diego, USA

Throughout the cyberinfrastructure community there is a large range of resources available to train faculty and young scholars about successful utilization of computational resources for research. The challenge that the community faces is that training materials abound, but they can be difficult to find, and often have little information about the quality or relevance of offerings. Building on existing software technology, we propose to build a way for the community to better share and find training and education materials, through a federated training repository. In this scenario, organizations and authors retain physical and legal ownership of their materials by sharing only catalog information, organizations can refine local portals to use the best and most appropriate materials from both local and remote sources, and learners can take advantage of materials that are reviewed and described more clearly.

CCS Concepts: • **Social and professional topics** → **Informal education**; • **Applied computing** → **Digital libraries and archives**.

Additional Key Words and Phrases: education, training, community engagement

ACM Reference Format:
Susan Mehringer, Kate Cahill, Charlie Dey, Rich Knepper, John-Paul Navarro, Jeaime H. Powell, and Mary Thomas. 2018. Scaling HPC Education. In . ACM, New York, NY, USA, 2 pages. https://doi.org/XXXXXXX.XXXXXXX

1 EXTENDED ABSTRACT

A federated training repository of vetted and tested training will allow resource providers, campus portals, schools, and other institutions to incorporate training from multiple sources into their own familiar interfaces. The results of a survey [1] conducted October-November 2022 showed that most respondents were both interested in, and able to, work toward community efforts to share and discover materials. The federated repository will identify and vet training resources from this broad set of offerings and provide metadata and characterization that support successful discovery and utilization of training resources, and their incorporation into portals for research computing groups at universities, HPC centers, schools, domain-centered institutions and elsewhere. Local centers will be able to include content identified in the repository and offer it to their local faculty members with local branding and identification side-by-side with local offerings, and in turn their training offerings will be vetted and shared with the broader computational science community.

Manuscript submitted to ACM

The base technology for the repository has been tested and is complete, under the https://software.xsede.org site and the https://research.illinois.edu sites, which provide information about HPC software and research resources respectively. The project team will establish a similar repository for training materials that collects information about the materials: location information, title, and metadata about content and topics. Training materials will remain in their original location and be discovered from the repository itself or from within an HPC Center website that uses the HPC-ED repository via API. In this way, centers can present complete training material sets without the difficulty of creating and maintaining them consistently over time. By leveraging a federated model, the training community can highlight the best possible training resources and emphasize competencies developed individually.

However, building a repository is not enough; we plan to implement quality assurance measures, work with the community to determine a set of metadata tags, and extend the base technology to be accessible for general usage. efforts for the first area of content. The next step in ensuring the further use of the project is to provide outreach that recruits further training material providers and users to the project.

The project will extend and enhance the ability of universities, departments, research computing groups, HPC centers, and domain-specific collaborations to discover and incorporate relevant and proven training materials within their own websites, portals, and science gateways. Leveraging the federated training repository will allow communities access to advanced CI-related training materials without the burden of creating and maintaining large sets of materials, and facilitate the professional development of individuals served by those institutions.

A community-wide project can only thrive when it has input, feedback, and use by the community. A working group has been formed within the ACM SIGHPC Education Chapter to discuss metadata standards for sharing materials across all interested organizations. An affinity group within ACCESS was created as a result of discussion at a PEARC23 BoF. We have a number of partners who confirmed their interest and intention in integrating their training resources as early adopters, who will help us to grow the repository and to collect feedback on the product and procedure.

REFERENCES

[1] K. Cahill, D. Joiner, S. Lathrop, S. Mehringer, J. Navarro, and A. Weeden, "Final Results: National Survey on Educational and Training Materials Repositories." [Online]. Available: https://www.cac.cornell.edu/about/pubs/Survey2022.pdf