# HPC-ED: Building a Sustainable Community Driven CyberTraining Catalog

Susan Mehringer
Center for Advanced Computing, Cornell University
shm7@cornell.edu

Katharine Cahill
New Jersey Institute of Technology
katharine.cahill@njit.edu

Charlie Dey
Texas Advanced Computing Center
charlie@tacc.utexas.edu

Brian Guilfoos
Ohio Supercomputer Center
guilfoos@osc.edu

David Joiner
Kean University
djoiner@kean.edu

Richard Knepper
Center for Advanced Computing, Cornell University
rich.knepper@cornell.edu

John-Paul Navarro
Argonne National Lab
navarro@anl.gov

Jeaime H. Powell
Omnibond Systems
jeaime@omnibond.com

Mary P. Thomas
University of California, San Diego
mpthomas@ucsd.edu

Zilu Wang
Center for Advanced Computing, Cornell University
zw427@cornell.edu

Jiesen Zhang
University of California, San Diego
jiz147@ucsd.edu

## ABSTRACT

HPC-ED is working to improve discovery and sharing of CyberTraining resources through the combination of the HPC-ED CyberTraining Catalog, an effective and flexible interface, thoughtful metadata design, and active community participation. HPC-ED encourages authors to share training resource information while retaining ownership and allows organizations to enrich their local portals with shared materials. By basing the architecture on an established, flexible framework, HPC-ED can provide a range of solutions people and organizations can employ for sharing and discovering materials.

In this paper we describe the initial pilot phase of the project, where we prototyped the HPC-ED catalog, established an initial metadata set, provided documentation, and began using the system to share and discover materials. We gathered community feedback through a variety of means, and are now planning an implementation phase based on evolving our architecture and tools to meet community needs and feedback through improved interfaces and tools designed to address a range of preferences.

## KEYWORDS

Education, Training, Community Engagement, HPC, Cyberinfrastructure, Metadata, Globus


https://doi.org/10.22369/issn.2153-4136/16/1/2

## 1 INTRODUCTION

A wealth of CyberTraining material exists, but locating appropriate resources can be challenging. HPC-ED's primary goal is to improve discovery and sharing of CyberTraining resources through the HPC-ED CyberTraining Catalog, an effective and flexible interface, and active community participation. HPC-ED enables training material owners to share metadata describing their resources, allows organizations to enrich their local portals with shared materials, and provides individuals with the ability to share and discover training resources. The catalog is based on Globus Search [6], an established, flexible framework; this allows HPC-ED to provide a range of solutions people and organizations can employ.

During the initial stage of the project, we prototyped the HPC-ED catalog, developed a set of basic metadata [15], provided instructions for sharing and discovering materials, and began populating the catalog using different sources. We gathered community feedback through surveys, presentations, tutorials, and hackathons, explored collaborations and engaged with projects interested in utilizing the catalog. Currently, the project is evolving to address community needs and feedback; we are adapting interfaces and building tools to address a range of preferences, including downloading filtered search results and using Jupyter Notebooks to share resources. Details of the pilot project can be found in [30].

Building on recent improvements, our immediate goals are to implement an improved approach for hosting the HPC-ED catalog in production mode, expand its user base, and continue to improve tools and interfaces, such as integrating Large Language Models (LLMs) to generate relevant and contextually appropriate metadata and content and supporting a new approach to share material metadata (see Section 3). By implementing a federated catalog with a robust interface, HPC-ED can broaden the reach of training materials, increase their usage, and expand local learning resources.

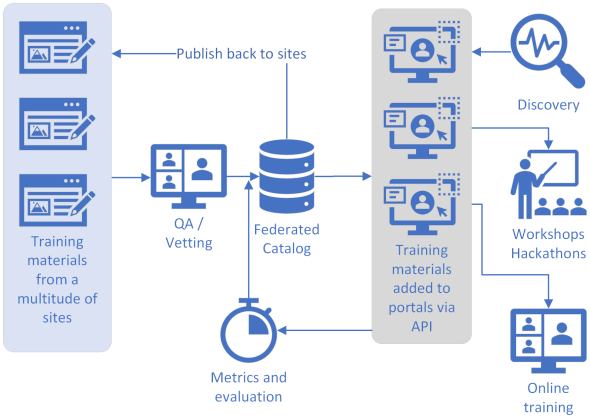


**Figure 1: Pilot project architecture: Key features include: materials remain in original location; only metadata describing the materials are stored; individuals or smaller organizations can share their materials more broadly.**

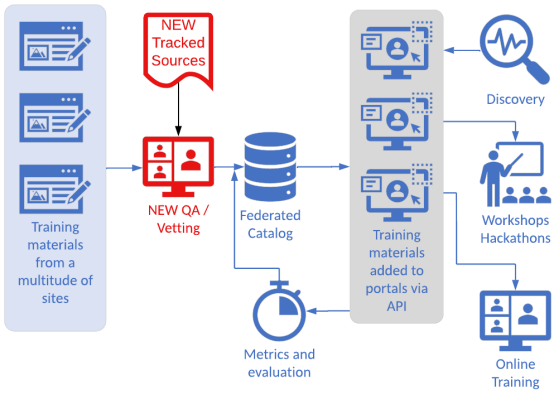


**Figure 2: Implementation project architecture: key change is a decentralized *pull model* with indirect interaction with the HPC-ED catalog, providing improved security, quality checks, and enabling tool development.**

## 2 PILOT: PROOF OF CONCEPT

The pilot project and results are highlighted in this section. Details about the project can be found on the project web page [14] and additional progress and results can be found in a recent paper [30].

### 2.1 Architecture

The pilot project architecture is designed to enable CyberTraining material sharing and discovery. The community currently relies on finding resources on local training catalogs and discovery portals hosted at larger research or commercial organizations, or by using internet search engines. The HPC-ED Pilot architecture (see Figure 1) leverages and builds on the strengths and flexibility of organization-specific training catalogs and portals while enabling individuals to share their materials and discover training materials from all HPC-ED catalog participants.

An overview of the process as implemented in the pilot project is presented in Figure 1. Key features of this architecture are that materials remain in their original location; only metadata describing the materials are stored in the HPC-ED catalog; individuals or smaller organizations can share their materials more broadly; thoughtful metadata design results in more appropriate search results; and the robust software base used is stable and provides excellent flexibility for mid-level tools.

Using Globus Search commands, materials can be shared or discovered. Sharing materials is done by compiling HPC-ED specified metadata in JSON (JavaScript Object Notation [17]) format and sharing it with the HPC-ED catalog. Sharing to the pilot catalog requires authorized credentials, which are issued by the HPC-ED project. Searching the catalog is done by issuing a Globus Search command or by using any web portal that is populated with catalog materials. Authorization credentials are not needed to search public HPC-ED catalogs. Sharing and discovery is described in detail on the project wiki page. [16]

### 2.2 HPC-ED Pilot Metadata

The HPC-ED project is committed to ensuring that the digital objects stored in the repository are findable and reusable. This is accomplished by following the Findable, Accessible, Interoperable and Reusable (FAIR) principles, [5]. The HPC-ED Pilot project website describes the metadata that is being used for an initial demonstration of capability. [15] We discuss our implementation plans in Section 5.2. This includes custom metadata and existing sets including: Metadata required by the Globus search, the Research Data Alliance (RDA) ontology for recommended minimal metadata set for training materials [12], the set of HPC-ED metadata, and the Dublin Core Learning Resource Type. [3]

HPC-ED uses Globus search tools, which require a JSON formatted document that is used to share information on resources and events to the HPC-ED catalog. [6] Details on the fields in the JSON document can be found in the HPC-ED Metadata Description and the Globus Search Overview. Currently there are 9 required metadata fields and 12 optional fields, including Title, Description, Authors, Publisher, Type, Language, Cost, Format, License, Target Group, Expertise Level, Certification details, and very importantly, Persistent Identifiers, Tags, or Keywords. One optional field is "keywords," which can be used to add additional search metadata.

A challenge for the repository lies in the area of searching and discovery, especially when the number of objects is large. The metadata and keywords used by the HPC-ED project will be used to create an HPC training materials taxonomy, knowledge graph, and eventually an ontology. This data will be used to support a variety of clients to share and discover materials. These can be used to create an LLM for discovering training materials.

### 2.3 Communication and Community

To facilitate effective communication and participation in the community, a dedicated website and wiki were established as part of the project. The website provides a comprehensive overview of the project, including details of recent and upcoming events. The

wiki serves as a repository for procedural documentation, which is regularly updated to ensure clarity and accessibility, featuring resources such as a quick start guide.

In addition to these resources, a Slack channel and an ACCESS affinity group were created to offer support to users testing the new tools. Regular emails have been distributed throughout the project year to keep mailing list subscribers informed of new developments, features, and future events. An informational session was held in April for people interested in utilizing the publication and sharing tools. During this session, processes were reviewed and feedback was sought. For those unable to attend, the session was recorded and made available on the project website to ensure broad accessibility.

Project documentation is provided on the HPC-ED web site [14] which includes a project overview, community activities, events, and contact information. Technical documentation for sharing and discovering HPC-ED catalog materials are available on the HPC-ED wiki pages [16]. The wiki includes a process overview, a full description of the required and optional metadata, along with detailed and quick start documents on how to share and discover materials. The recording of an early tester online event is available online [13].

HPC-ED is seeing growing interest in engagement from people and organizations, and coordination with projects as a result of meetings, presentations, and hackathons 5.1. Recent presentations include *Practice and Experience in Advanced Research Computing 2024: Human Powered Computing* [30], *Assessing Shared Material Usage in the High Performance Computing (HPC) Education and Training Community* [19], *The HPC Federated Learning Catalog* [18], and *Scaling HPC Education* [20]. In addition, a tutorial *Publishing and Discovering CyberTraining Materials Across the HPC and CI Research Communities* [21] was held at PEARC24 [24].

With our objectives for the pilot project successfully testing the concept met, effort has begun on initiating an action plan for implementation.

# 3 IMPLEMENTATION: IMPROVING USABILITY AND SCALING

The HPC-ED *implementation plan* includes a decentralized architecture that retains flexibility and adds security with the integration of several new features: implementing a distributed model for sharing of data by contributing partners, adding security by controlling which partner data is added to the catalog, validating and vetting partner and shared metadata, developing different sharing and discovery approaches for different types of partner capabilities, and engaging in collaborations to develop new tools and build community. These new features are described the sections below.

## 3.1 Architecture

We are shifting to a new architecture, as shown in Figure 2, that will be more flexible, secure, and easier to use. The primary change is that rather than giving all community members direct access to share resource metadata with the HPC-ED catalog, we are shifting to a decentralized *pull model* where contributors will not directly share metadata with the HPC-ED catalog. Instead, the relevant data will be pulled from local catalogs to the HPC-ED catalog. When the data is pulled, identity information will be added, and quality checks will be conducted. In addition, only the data that is necessary for the HPC-ED catalog will be pulled. This change has many advantages: the community member does not need to obtain a provider ID or create unique subject tags as they will be automatically created by the pull software, the community member can choose the automatic sharing interval (e.g. daily, weekly, or monthly), and each person or site can choose the method that best suits their needs. By removing the ability to share directly with the catalog, the project benefits from an interim step that allows security and quality checks to be applied.

This architectural shift enables a variety of sharing and discovery options, such as those described in the following sections.

## 3.2 Sharing Training Material from Local Catalogs

The new HPC-ED architecture supports a decentralized data collection system (see Figure 3). The HPC-ED architecture supports the creation of local catalogs created by different organizations and projects. When registered, the system will identify relevant contributions, and then will pull, inspect and validate the material. Examples of different contribution mechanisms are described below.

For organizations with multiple resources to share, but choose not to maintain a local catalog, the system can accept shared materials from a variety of document formats. A few examples are provided below.

*3.2.1 Spreadsheets:* A spreadsheet template is provided that includes dropdown selections, help tabs, and common values, all intended to simplify choosing appropriate metadata to describe the shared materials. The spreadsheet can either be shared once by email to the project email address, or, if shared with the project via a google spreadsheet, it would be used to update the shared material at the selected periodicity. This method also simplifies the process by removing the requirement to provide a provider ID and unique subject fields.

*3.2.2 Simple Text Files.* The system architecture can accept multiple file formats, including a text file. In this example, we use "key=value" pairs based on the HPC-ED required metadata. A snippet of the code is shown in Figure 4.

## 3.3 Discovery Clients

The federated catalog has a search API that is client agnostic. [15] Currently the search query returns the JSON formatted metadata information. Other query mechanisms will be created as part of the implementation phase of the project.

*3.3.1 Download search results to a web page.* A simple method to incorporate select materials from the catalog into a local web site is to download those results directly to an html file. This method has been implemented on the development portal [23]. After using the filters available on the portal to find the desired materials, a button on the page shares a downloaded web page with the results that can be incorporated into a local web site.

*3.3.2 A Web Application to Discover Materials from a Local Portal.* A collaborative effort is underway to provide a web application

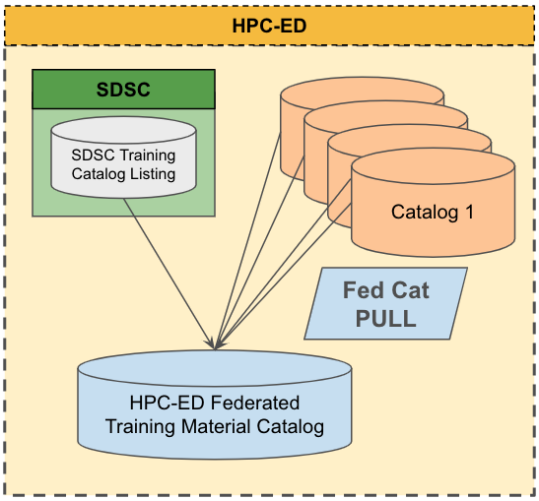


**Figure 3: Diagram demonstrates how resource material is pulled from remote catalogs based on validation criteria and permissions.**

```
"Title": "Expanse Webinar: Introduction to Neural Networks, Convolution Neural Networks and Deep Learning on Expanse",
"Abstract": "This webinar will be a quick introduction and overview of neural networks, convolution networks...",
"Authors": "Paul Rodriguez, Ph.D. ",
"Keywords": [ "Expanse", "HPC Training", "Industry"],
"Resource_URL_Type": "URL",
"Learning_Resource_Type": "recorded lesson",
"Start_Datetime": "2022-05-19T18:00:00.000000",
"URL": "https://education.sdsc.edu/training/interactive/202205_ExpanseWebinar-Intro-to-Neural-Networks-and-Deep-Learning-on-Expanse"
```

**Figure 4: Example showing the SDSC Training Materials text file used to share materials, consisting of KEY=VALUE pairs, where KEY is based on the HPC-ED metadata. [15]**

that can be used on any local site to let visitors search the HPC-ED catalog without leaving their local community portal. This would be available to download from GitHub in the form of an iframe plug-in.

*3.3.3 Training Materials Web Page.* The discovery example in Figure 5 shows the SDSC Training Materials client (Webhost) discovering relevant metadata stored in the catalog (via a python script titled "intvid.py") to display training material to the different clients. The diagram shows the specific case of retrieving videos from YouTube [25] based on video url information contained in the federated catalog.

## 4 COLLABORATIONS

The collaborative efforts of multiple organizations have contributed significantly to the advancement of this project. ACCESS Support and ACCESS Operations have played a crucial role in metadata development and hosting a pilot catalog search portal. Future plans include the integration of ACCESS training materials into the repository and the provision of our search tools to ACCESS users, including a proposed NAIRR (nairr.org) training catalog collection. This aligns with the project's goal of enhancing metadata interoperability and improving user discovery.

The ACM SIGHPC Education chapter has emerged as a valuable partner for dissemination through its Education committee. This collaboration facilitates the broader distribution of project resources and findings within the high-performance computing education community. Additionally, the Outreach committee is developing a web application to integrate the repository into the chapter's website, which will enhance accessibility and user engagement.

These partnerships have been instrumental in addressing key challenges in metadata management, including analyzing metadata requirements, adopting appropriate schemas, and creating metadata content. The collaborations also support the project's focus on improving search functionality and providing web services, which are essential for enhancing the discoverability and utility of digital resources. We will continue to reach out through our website and email updates and outreach opportunities to connect with the HPC Education and Training Community and identify new collaborations.

## 5 QUALITY ASSURANCE

Our quality assurance plan is focused on allowing for both human collection and automated collection of quality assurance and review information on catalog items. Defining quality variables in digital libraries can contain a large number of dimensions [7, 31]. After a review of different potential quality variables, we chose 4 items that would help us to maintain a process easy enough for the end user to encourage participation and enough data to help with the overall goal of connecting end users to materials hosted by catalog maintainers. These 4 items include 2 human contributed review items, and 2 automated items.

For human contributed review metadata, we will focus on whether the end user found materials that were accurate and well crafted (content correctness), and whether the search process helped the user to find what they were looking for (search relevance). As our goal is to provide a federated search across catalogs, our focus is at the item level and not the collection level. Reviews will ask end users two questions, with a Likert scale response (displayed as star ratings) and an optional text exposition of the user's answer. First, "Did you find what you were looking for?" Second, "How would you rate the quality of the materials you accessed?" Likert scale ratings will allow for a no response option in cases where a response is not applicable (A user who did not find what they were looking for, for example, may not choose to follow up with a statement on material quality.)

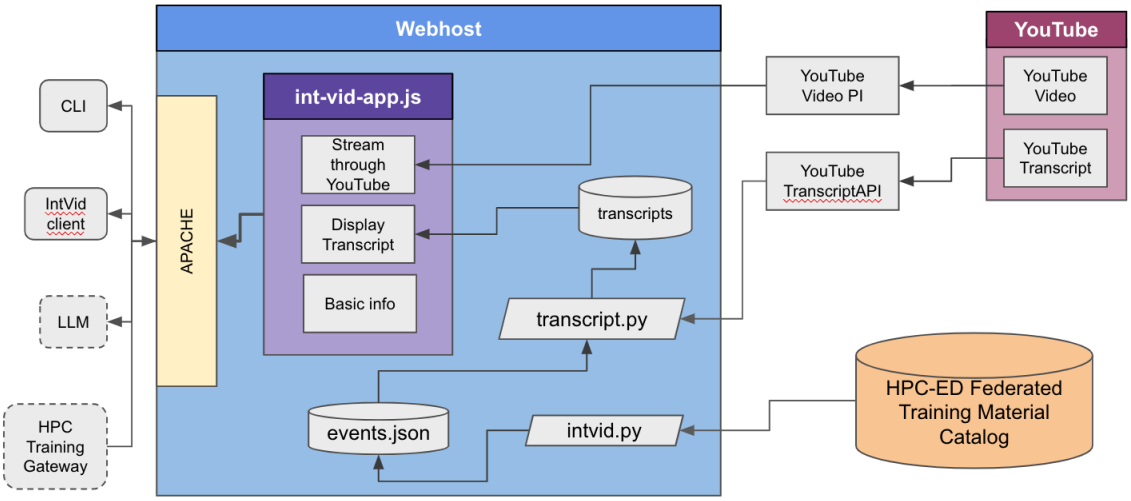


**Figure 5: Discovery example showing the SDSC Training Materials client discovering relevant metadata stored in the catalog to display training material to the different clients. Diagram shows specific case of retrieving videos from YouTube [25]**

For automated review metadata, we will focus on availability and accessibility, as measured by standard uptime measures which can be checked periodically and ADA compliance, for which there are multiple online tools available which have been used for studies in educational settings (See for example [29]).

Review metadata will be stored in a single Globus collection managed by the project, to ensure consistency of metadata format, and to allow for one to many relations between items and reviews. Reviewer identity will be managed and collected through the user's CI login, and stored by ORCID ID. For automated review metadata, a list of datetime values for which the URI was checked along with response will be stored, and ADA compliance will include metadata on the datetime of each check, the compliance tool used, and the result.

We will develop JavaScript based web applications that can be embedded into both item pages and into search results.

## 5.1 Workshops/Hackathons

Hackathons have become valuable tools in high-performance computing (HPC) education and training, providing an intensive, hands-on environment for participants to engage with real-world computational challenges. These events foster collaboration and innovation by bringing together students, professionals, and researchers to solve complex problems using HPC resources. Through hackathons, participants not only deepen their technical skills in parallel computing, data analysis, and algorithm optimization, but also develop critical soft skills such as teamwork, problem-solving, and communication. This immersive learning experience accelerates the development of expertise in HPC, making it a highly effective model for educational and training programs in the field.

A central repository of high-performance computing (HPC) training materials can significantly enhance the effectiveness and impact of HPC hackathons. By providing participants with ready access to comprehensive resources, such as tutorials, code libraries, case studies, and best practices, the repository ensures that both novice and experienced participants can quickly upskill and troubleshoot during the event. This centralized resource hub allows teams to focus on applying knowledge to real-world HPC challenges, rather than spending time searching for learning materials. Furthermore, the repository promotes consistency in training, helping hackathon organizers standardize the educational experience and ensure that participants have the foundational knowledge needed to maximize their productivity and success during the hackathon.

Hackathons can play a pivotal role in improving the development of high-performance computing (HPC) training materials by providing a real-world testing ground for educational content. During these events, participants encounter various challenges and gaps in their understanding, which can reveal shortcomings in existing materials. This feedback from participants, especially in terms of the resources they found most helpful or lacking, offers valuable insights for refining and expanding training content. Additionally, the collaborative and fast-paced nature of hackathons often leads to the creation of new tools, workflows, and techniques, which can be incorporated into future HPC training modules. The dynamic environment of a hackathon thus fosters continuous improvement of training resources, ensuring that they remain relevant and aligned with the evolving needs of the HPC community.

As a means to grow the data contained in and the community around the HPC-ED catalog, workshops, tutorials, and hackathons have been held at multiple cyberinfrastructure-related events. The events began with mini-workshops to engage with the community in a fashion that both spurred conversation and gathered terminology data from the audience. Later workshops and tutorials provided overviews for alpha and beta users. Additionally, a student-centered

hackathon saw the creation of example web applications built with the HPC-ED catalog API.

*5.1.1 Mini-Workshops.* Initial community engagement with the HPC community began through the use of mini-workshops in which terminology linked to categorization was gathered. Two such events included the SGX3 Faculty Hackathon [9] and later Gateways Conference [26]. The workshops prompted the audience members to define overloaded terms common to cyberinfrastructure in the community. The results were collected with the interactive whiteboard application Google Jamboard [8]. The results were displayed during the sessions with commentary to induce conversation and dissemination further. The audience members produced 211 definitions across six common terms. The results will aid in future HPC-ED categorization methodologies.

*5.1.2 Hackathon.* In June of 2024 the HPC-ED API was selected by three undergraduate student groups as the focus of their hackathon projects. The event, SGX3's HackHPC@ADMI24 [11] hosted by the The Association of Computer Science Departments at Minority Institutions (ADMI) [22] and funded by Science Gateways Community Institute [28] through SGX3 [1] for workforce development. The three teams were comprised of four students and a team mentor. In summary, the teams created web applications that implemented HPC-ED catalog search and publication [10]. Deliverables included GitHub repositories, presentations, and posters of which two were submitted to the Gateways24 Conference [27].

*5.1.3 Workshops and Tutorials.* Several workshops and tutorials have been conducted focusing on community usage and education. From workshops with concept dissemination to alpha user tutorials, HPC-ED has used both virtual and in-person events for directed community engagement. One such event took place at the Practice & Experience in Advanced Research Computing PEARC24 [24] annual conference as a tutorial titled *Publishing and Discovering Cybertraining Materials Across the HPC and CI Research Communities* [21]. The session was attended by over 30 educators/trainers in the HPC field. The tutorial included the use of a developmental catalog that was populated by the audience members. As true alpha users, the engagement both highlighted future enhancements and features through conversations and discussions fostered during the event. The audience members provided real-time feedback as a very diverse sample of clients.

## 5.2 Metadata and Metrics

As described in Section 2.2, the federated catalog metadata is designed to enable training material owners to share materials by publishing metadata associated with those materials to the federated catalog. In the pilot phase, we adopted a "minimal" metadata set approach, details of which can be found on the HPC-ED metadata description web page [15]. These metadata terms were used to conduct simple searches.

In the implementation phase, we will extend the existing metadata fields to include the following categories: metadata that describes the training material, its access methods, and educational characteristics; metadata that identifies the publisher and source of the training material so that when an individual selects a specific training item they can be directed to the source catalog that published that material in order to browse all available information and to access that training material; metrics metadata about educational material reviews, ratings, and access metrics. Information about the metrics will be drawn from the quality assurance part of the project, as described earlier in Section 5.

Defining, categorizing, and standardizing the metadata will be a significant effort. Where possible, we will identify and use existing metadata sets, taxonomies, and ontologies. [2, 4, 32] Where needed, we will add new terms to these existing ontologies and work with existing communities to update them or to extend the HPC-ED training materials ontology.

# 6 FUTURE WORK

Looking to future directions for the HPC-ED project, we will extend interface functionality, expand the types and quantity of catalog offerings, and pursue additional outreach opportunities to raise awareness of the project. The primary value of any directory of training materials is in the breadth and quality of the offerings available, and our immediate plan is to make adding entries to the catalog an easier task. There are still a number of organizations who have expressed interest sharing their materials but are waiting until the process is easier. Improving the process to add materials is a critical element to supporting the growth of the catalog and bringing additional partners to the project.

As the process for listing more entries in the catalog improves, we will continue to work to build awareness of the project and the potential for discovering and sharing more materials. This includes meetings within the national cyberinfrastructure community that bring together trainers and learners (PEARC, SCxy, and other meetings), as well as reaching out through the CyberTraining program in order to support these projects in making their training offerings easier to discover and share.

With additional items discoverable through HPC-ED, providing fully-functional quality assurance for catalog listings will be a necessity to ensure that users are directed to items that are relevant, salient, and effective. HPC-ED will incorporate a familiar content ratings system ("one to five stars") that allows users to contribute scores for materials appearing in the catalog, and will experiment with leveraging AI tools to provide summaries and quality ratings.

Looking further to extending the capabilities of the project, we will make use of the idea-generating potential of Hackathons and Workshops where the HPC-ED team can participate and help identify refinements and improvements to the catalog interface options, search, quality assurance, and integrations. Our team will also look to the capabilities of LLMs for providing better training results and summaries of content. Throughout the process of supporting sharing and discovery of training materials by a wider group of trainers and learners, we will continue to note the contributions of users of the catalog in order to identify capability improvements and extensions.

# ACKNOWLEDGMENTS

Our thanks for collaborative efforts go to the SIGHPC Education Chapter [1]. This project is funded through the NSF CyberTraining:

[1] SGX3 was funded by the National Science Foundation under award number 2231406

Pilot: HPC ED: Building a Federated Repository and Increasing Access through CyberTraining (#2320977); and supported by multiple NSF grants (#2230127, #2017767, #2320934, #2112606). We appreciate being granted access to NSF resources and services including: the NSF ACCESS Track 3 Award: COre National Ecosystem for CyberinfrasTructure (CONECT) (#2138307); and several HPC resources (SDSC Expanse #1928224 and TACC Stampede # 1663578).

## REFERENCES

[1] ACM. 2023. Association for Computing Machinery: SIGHPC Education chapter. https://sighpceducation.acm.org/

[2] Ayesha Ameen, Khaleel Ur Rahman Khan, and B. Padmaja Rani. 2012. Creation of Ontology in Education Domain. In *2012 IEEE Fourth International Conference on Technology for Education*. 237–238. https://doi.org/10.1109/T4E.2012.50

[3] Dublin Core. 2043. LRMI Learning Resource Type vocabulary. Retrieved September 8, 2024 from https://www.dublincore.org/specifications/lrmi/concept_schemes/learningResourceType

[4] Dong Dai, Yong Chen, Philip Carns, John Jenkins, Wei Zhang, and Robert Ross. 2019. Managing Rich Metadata in High-Performance Computing Systems Using a Graph Model. *IEEE Transactions on Parallel and Distributed Systems* 30, 7 (2019), 1613–1627. https://doi.org/10.1109/TPDS.2018.2887380

[5] Leyla Garcia, Bérénice Batut, Melissa L. Burke, Mateusz Kuzak, Fotis Psomopoulos, Ricardo Arcila, Teresa K. Attwood, Niall Beard, Denise Carvalho-Silva, Alexandros C. Dimopoulos, Victoria Dominguez Del Angel, Michel Dumontier, Kim T. Gurwitz, Roland Krause, Peter McQuilton, Loredana Le Pera, Sarah L. Morgan, Päivi Rauste, Allegra Via, Pascal Kahlem, Gabriella Rustici, Celia W.G. Van Gelder, and Patricia M. Palagi. 2020. Ten simple rules for making training materials FAIR. *PLoS Computational Biology* 16, 5 (2020), 1–9. https://doi.org/10.1371/journal.pcbi.1007854

[6] Globus. 2024. Globus Search Service. https://docs.globus.org/api/search/.

[7] Marcos André Gonçalves, Bárbara L. Moreira, Edward A. Fox, and Layne T. Watson. 2007. "What is a good digital library?" – A quality model for digital libraries. *Information Processing & Management* 43, 5 (2007), 1416–1437. https://doi.org/10.1016/j.ipm.2006.11.010 Patent Processing.

[8] Google. 2024. Google Jamboard. https://jamboard.google.com

[9] HackHPC. 2023. CiberInfrastructure. https://hackhpc.github.io/facultyhack-gateways23/assets/slides/pre-event_training/FacultyHack-HPC-ED.pdf

[10] HackHPC. 2024. Hackathon Teams. https://hackhpc.github.io/sgx3admi24/teams.html

[11] HackHPC. 2024. SGX3's Association of Computer Science Departments at Minority Institutions (ADMI) Hackathon. https://hackhpc.github.io/sgx3admi24/

[12] N. J. Hoebelheinrich, K. Biernacka, M. Brazas, L. J. Castro, N. Fiore, M. Hellström, E. Lazzeri, E. Leenarts, P. M. Martinez Lavanchy, E. Newbold, A. Nurnberger, E. Plomp, L. Vaira, and A. van Gelder, C. W. G. & Whyte. 2022. Recommendations for a minimal metadata set to aid harmonised discovery of learning resources. https://doi.org/10.15497/RDA00073

[13] HPC-ED. 2024. Early Tester Training. https://www.youtube.com/watch?v=fD9g5z5qZFQ

[14] HPC-ED. 2024. HPC-ED Documentation. https://hpc-ed.github.io/

[15] HPC-ED. 2024. HPC-ED Metadata. https://github.com/HPC-ED/HPC-ED.github.io/wiki/Metadata-Description

[16] HPC-ED. 2024. HPC-ED Wiki Documentation. https://hpc-ed.github.io/wiki

[17] JSON. 2024. JavaScript Object Notation. https://www.json.org/

[18] Richard Knepper. 2024. The HPC Federated Learning Catalog. In *Thirty-First International Conference on Learning*.

[19] Susan Mehringer, Kate Cahill, Scott Lathrop, Charlie Dey, Mary Thomas, and Jeaime H Powell. 2023. Assessing Shared Material Usage in the High Performance Computing ( HPC ) Education and Training Community. In *Sixth Workshop on Strategies for Enhancing HPC Education and Training (SEHET23)*. https://doi.org/10.22369/issn.2153-4136/14/2/4

[20] Susan Mehringer, Mary P Thomas, Charlie Dey, Kate Cahill, David Joiner, Richard Knepper, and Jeaime H Powell. 2023. Scaling HPC Education. In *Tenth SC Workshop on Best Practices for HPC Training and Education BPHTE23*. Denver, 41–46. https://doi.org/10.22369/issn.2153-4136/15/1/8

[21] John-Paul Navarro, Katharine Cahill, Susan Mehringer, Je'Aime Powell, and Richard Knepper. 2024. Publishing and Discovering Cybertraining Materials Across the HPC and CI Research Communities. https://github.com/HPC-ED/PEARC-Tutorial-2024/tree/main

[22] Association of Computer Science Departments at Minority Institutions. 2024. What is ADMI? https://admiusa.org

[23] ACCESS Operations. 2024. HPC-ED Developers Portal. https://search-pilot.operations.access-ci.org/.

[24] PEARC. 2024. PEARC24. https://pearc.acm.org/pearc24

[25] SDSC. 2024. Interactive Video Tutorials. https://education.sdsc.edu/training/interactive.

[26] SGX3. 2023. Science Gateways 23 Annual Conference. https://sciencegateways.org/gateways2023

[27] SGX3. 2024. Gateways24. https://sciencegateways.org/gateways2024

[28] SGX3. 2024. SGX3 - A Center of Excellence to Extend Access, Expand the Community, and Exemplify Good Practices for CI Through Science Gateways. https://sciencegateways.org/

[29] Tatiana I Solovieva and Jeremy M Bock. 2014. Monitoring for Accessibility and University Websites: Meeting the Needs of People with Disabilities. *Journal of Postsecondary Education and Disability* 27, 2 (2014), 113–127.

[30] Mary P Thomas, Susan Mehringer, Katharine Cahill, Charlie Dey, Brian Guilfoos, David Joiner, John-Paul Navarro, Jeaime H Powell, and Richard Knepper. 2024. Building a Federated Catalog for CyberTraining Materials: The HPC-ED Pilot Project. In *Practice and Experience in Advanced Research Computing 2024: Human Powered Computing* (New York, NY, USA), Vol. 1. ACM, 1–5. Issue 1. https://doi.org/10.1145/3626203.3670586

[31] Hong Xie. 2006. Evaluation of digital libraries: Criteria and problems from users' perspectives. *Library and Information Science Research* 28, 3 (2006), 433–452.

[32] Aolong Zhou, Kaijun Ren, Xiaoyong Li, Wen Zhang, Xiaoli Ren, and Kefeng Deng. 2021. Semantic-based discovery method for high-performance computing resources in cyber-physical systems. *Microprocessors and Microsystems* 80 (2021), 103328. https://doi.org/10.1016/j.micpro.2020.103328