



High-Performance Computing Skills Platform
and European Collaboration for Training

D1.1 – Architecture for the EuroHPC Training Platform

WP2: Design and Evolution of the EuroHPC Training
Platform

Grant Agreement Number: 101136267



Funded by
the European Union



EuroHPC
Joint Undertaking

This project has received funding from the European High-Performance Computing Joint Undertaking under grant agreement No. 101136267.

Document Information

Deliverable Number	D2.1
Deliverable Name	Architecture for the EuroHPC Training Platform
Due Date	30/09/2024 (M9)
Deliverable Lead	USTUTT
Authors	Miriam Koch, USTUTT
Responsible Author	Miriam Koch, USTUTT
Keywords	HPC training platform, skills framework, platform design
WP	WP2
Nature	Report
Dissemination Level	Public
Final Version Date	30/09/2024
Reviewed by	Simon Wong, ICHEC
PMT Approval	30/09/2024

Funded by the European Union. Views and opinions expressed are however those of the author only and do not necessarily reflect those of the European Union or the European High-Performance Computing Joint Undertaking. Neither the European Union nor the granting authority can be held responsible for them.

Document History

Partner	Date	Comments	Version
USTUTT	15/08/2024	Initial version	0.1
USTUTT	05/09/2024	First draft for internal review	1.0
USTUTT	19/09/2024	Revisions made according to feedback	1.1
USTUTT	30/09/2024	Final version for submission	1.2

Executive Summary

The HPC SPECTRA project aims to develop a comprehensive EuroHPC Training Platform which forms a key part of the EuroHPC JU's training strategy. This platform is designed to foster high-performance computing (HPC) skills across Europe, providing a centralised hub for accessing and promoting HPC training resources. The project supports European training providers by facilitating the promotion of events and materials, while offering users a centralised access point for finding and discovering relevant training opportunities.

The training platform will align with existing European efforts by integration with the C2ISS (CASTIEL 2 Information Sharing System), which will become a central hub in dissemination of European HPC activities. It has been developed to support the establishment of a competency framework that standardises HPC skills development to address the heterogeneity of European HPC training. By leveraging tools like Indico for event management and Zenodo for training materials, the HPC SPECTRA project aims to build upon prior efforts and experiences rather than duplicating them.

The architecture of the training platform has been developed with user experience in mind so that training information can be provided, displayed and made accessible in a convenient manner for its target audiences, from training providers to potential trainees. Additional features such as event management and personalised training recommendations are also incorporated to enhance the user experience. The architecture is also designed to evolve with stakeholder input, to ensure the long-term sustainability of HPC training efforts across Europe.

In the next phase, the HPC SPECTRA project will focus on deploying the pilot platform by 2025, with ongoing collaboration with CASTIEL 2 and other stakeholders to meet the evolving needs of the HPC training community.

List of Abbreviations

C2ISS	CASTIEL 2 Information Sharing System
CASTIEL 2	Project to coordinate the activities of the NCCs and CoEs at the European level
CERN	European Council for Nuclear Research
CoE	Centre of Excellence
DOI	Digital Object Identifier
ELIXIR	European life sciences infrastructure
FocusCoE	European project to coordinate CoE activities (2018-2022)
ICHEC	Irish Centre for High-End Computing
JU	(EuroHPC) Joint Undertaking
NCC	National Competence Centre
PRACE	Partnership for Advanced Computing in Europe
TeSS	Training eSupport System (developed by ELIXIR)
URL	Uniform Resource Locator
USTUTT	University of Stuttgart
WP	Work Package

List of Tables

Table 1. Training event metadata	12
Table 2. Training material metadata	13
Table 3. Skills metadata	14
Table 4. Tracks metadata	15
Table 5. Modules metadata.....	15
Table 6. Key architectural concepts of the platform components.....	26

List of Figures

Figure 1. High-level structure of the C2ISS platform	17
Figure 2. Indico event management interface	20
Figure 3. Example of HPC training material on Zenodo.....	22
Figure 4. Overview of datasets and metadata locations.....	26



Table of Contents

Document Information	2
Document History	2
Executive Summary	3
List of Abbreviations.....	4
List of Tables	5
List of Figures.....	5
Table of Contents	6
1 Introduction	8
2 Background to the Training Platform.....	9
3 Architectural Requirements.....	10
3.1 Target Audience / Platform Users.....	10
3.2 Relevant Datasets and Metadata.....	12
3.2.1 Training Events	12
3.2.2 Training Materials	13
3.2.3 Skills.....	14
3.2.4 Modules & Tracks.....	14
4 Architecture Platforms and Solutions.....	16
4.1 C2ISS.....	16
4.2 Training events management.....	18
4.2.1 Indico	18
4.2.2 TeSS.....	20
4.3 Training Materials Management.....	22
4.3.1 Zenodo	23
5 Proposed Architecture Overview.....	25



5.1	Location of datasets and metadata	25
5.2	Key architectural concepts	26
5.3	User scenarios	27
5.3.1	User looking for an upcoming training event	27
5.3.2	Registered user looking for relevant training events/materials	27
5.3.3	Training provider promoting a new training event	28
5.3.4	Training provider uploading a new training material	28
6	Conclusions and Next Steps	29
7	References	30

1 Introduction

The HPC SPECTRA project addresses two main areas of priority in the EuroHPC JU's training strategy. Firstly, it aims to develop an EuroHPC Training Platform that promotes and advances European HPC training. Secondly, it collaborates with international partners to organise the 2024 and 2025 editions of the International HPC Summer School. This deliverable represents a key step of the process to develop the training platform for the first objective of the project.

The training platform is intended to serve as a central hub for HPC related training information and access point to training initiatives across Europe. It caters for European training providers to promote training offers (e.g. courses, workshops) and facilitates platform users to identify and find relevant courses and opportunities. The platform will also support the development of a competency framework to better elucidate the competencies being covered by the training offerings in a standardised manner and to help establish accreditation and certification mechanisms. The platform, centred on development of HPC competencies, will be a key pillar in the European HPC ecosystem.

This deliverable describes some of the initial work and outcomes from the HPC SPECTRA project to define the architecture for the training platform. Section 2 provides some background developments, including evolution of the European training landscape, that are relevant to the design and architecture of the platform. Section 3 covers the requirements engineering process and data that underpin the functions of the platform. Section 4 includes a review and evaluation of some of the relevant technologies that addresses the requirements and desired functionalities. Section 5 is a description of the platform architecture that are to be implemented in the first phase of the project. The deliverable concludes with an overview of the next steps in the realisation of the training platform. The architecture may evolve over time with input from key stakeholders and users such as the NCCs and the CoEs. This work was done and will be continued by a working group consisting of members from all work packages led by Task 2.1.

2 Background to the Training Platform

It is important to note that HPC training in Europe has grown and evolved over many years prior to this project, with actors mainly from academia and HPC centres. In the past, several projects and initiatives had fostered and structured the heterogeneous landscape, both on national and European levels. First efforts in harmonising that landscape were made, one of the most relevant for this project was the pilot implementation of the HPC in Europe portal [1], which was developed by PRACE [2] not only for its own extensive training offerings but also used by the NCCs and CoEs. On the training aspect there had been an active collaboration between PRACE (who coordinated over 100 HPC training courses annually), the FocusCoE project (who coordinated the CoEs from 2018 to 2022) and CASTIEL (who coordinated the NCCs) to ensure that the pilot portal catered for heterogeneous training providers and users. PRACE had also developed an event management platform, based on adapting the Indico software from CERN [3], which had been used to manage over 1,000 HPC courses. HPC SPECTRA is committed to not duplicate efforts but build upon some of these prior activities and experience in developing the EuroHPC training platform.

Subsequent to the launch of the HPC SPECTRA project, CASTIEL 2 [4] has begun the development of the C2ISS platform which aims to cater for NCC activities including training but also other services (e.g. industry) and success stories. Importantly, the C2ISS platform aims to serve pan-European HPC providers and stakeholders (i.e. beyond the NCCs and CoEs), similar to the original intentions of the HPC in Europe portal. Hence rather than building an independent, standalone EuroHPC Training Platform, the HPC SPECTRA project will build its training platform that integrates seamlessly with C2ISS. While C2ISS will necessarily dictate some of the visual design elements (i.e. the principle look and feel), the technical development and functionalities of the training component will be managed by the HPC SPECTRA project. In essence, HPC SPECTRA will be responsible for rolling out the training section of the C2ISS platform.

3 Architectural Requirements

The goal of the first phase of the architecture definition was to establish the requirements without specific solutions in mind to ensure the final product is well aligned within the ecosystem. The first step for this is to define, what the EuroHPC Training Platform is and what functionalities it should include:

- HPC SPECTRA project aims to develop an open, interactive, user-oriented and inclusive **central database of HPC training services and material** that will serve both to standardise and promote training initiatives from across the EuroHPC ecosystem.
- The ambitions of the EuroHPC Training Platform extend beyond an aggregation of training services to also serve as a foundation to implement a common standardised EuroHPC skills framework for the development of a **pan-European HPC training and certification scheme** for existing and new courses offered within the European HPC ecosystem.
- Additionally, the platform will include **event management functionalities** for the training providers, e.g. registration capacities for the training events, provision of attendance certificates or contacting the training attendees. An important feature of the former HPC in Europe portal will also be the capacity to **produce tailored extracts and exports** for countries (NCCs), training institutions and research communities.
- Finally, the platform will include a **trainee corner** that will collect past trainings attended and access to their related learning materials as well as suggestions of follow-up or new training events based on the past career and personal interests of each trainee.

3.1 Target Audience / Platform Users

Deriving from the aforementioned requirements, it becomes apparent that there are several types of user roles that needed on the EuroHPC Training Platform, which are:

1. Training Providers: Upload and manage events and material

- Provide information (metadata) about a new training event to the platform (e.g. an upcoming HPC course)
 - Update event information on the platform
 - Upload course materials to be accessible from the platform, these materials are typically associated with an event being promoted on the training platform
 - Manage training event registrations, attendee selection (optional), timetabling, attendee feedback, etc.
 - Promote the full or a subset of the courses promoted on the training platform on their own websites, e.g. courses in own country from any provider.
2. Attendees: Track past training and material as well as get guidance for the next trainings
- Find all upcoming HPC courses in Europe, as well as having the ability to filter this list according to different criteria, e.g. geographic region, online vs face-to-face/hybrid events, language, topic of interest.
 - Access important course information such as the content of the course, the provider, ways to register and costs (if any).
 - Have the option of registering for an account that will provide extra functionalities such as, easier registration, viewing of past courses attended, guidance on relevant courses for the user.
 - Explore learning pathways/tracks that provides guidance of the skills/competences that are typically required in different career profiles/trajectories.
3. Platform Admin: Administration of page structure, users and content
- Have overall responsibility for the infrastructure that runs the platform (e.g. security updates on the underlying operating system).
 - Maintain and update the platform software (e.g. installing and maintaining webserver software along with any customised codes from the HPC SPECTRA project).
 - Manage the platform's overall page/content structures, templates and associated databases.

- Oversee user accounts management on the platform, from training providers to attendees.

3.2 Relevant Datasets and Metadata

In order to address the functionalities and requirements stated above, the training platform needs to collect, store, process and manage a variety of different datasets and associated metadata. Subsections 3.2.1 to 3.2.4 outlines some of the key datasets and metadata that are envisaged to be used for the training platform.

3.2.1 Training Events

The metadata on training events have mainly been derived from that from the existing HPC in Europe portal [1]. This metadata have been deliberated and agreed upon from previous European projects that oversaw hundreds of HPC courses annually (e.g. FocusCoE, CASTIEL and PRACE). Much of the metadata deals with training event logistics (e.g. what, when, where) but also encompasses some indication of content type (e.g. domain area, level, target audience). In a new development, training events can now be attributed to specific skills (part of a “skills tree” framework) as defined by the HPC Certification Forum [5]. Training providers can also specify skills that are pre-requisites for certain events.

Table 1. Training event metadata

Field Title	Data Type
Event ID	Number
Title of Training Course/Event	Text Field
Training Description	Text Area
Training Image	Image
Start Date & Time	Date
End Date & Time	Date
Registration Link	URL
Event Type	Term (Taxonomy)
Address	Text Area
City	Term (Taxonomy)
Country	Term (Taxonomy)
Language	Term (Taxonomy)

Level	Predefined list
Fees (if applicable)	Free Text
Organiser	Term (Taxonomy)
Co-organisers	Term (Taxonomy)
Existing User	Ref User
Non-existing User	Term (Taxonomy)
Skills	Ref Link
Scientific Domain	Predefined list
Pre-requisites (Logistics)	Text Area
Pre-requisites (Skills)	Ref Link
Target Sector	Predefined list
Target Profile	Predefined list

3.2.2 Training Materials

The metadata on training materials encompass basic description about the material, including author and licensing information. Materials also contain key references to the skills framework, indicating the topical area(s) of relevance that enhances its discoverability. In addition, references to the training event where the material had been used also allows it to be findable via its referenced training event.

Table 2. Training material metadata

Field Title	Data Type
Material Title	Default Title
Material Description	Text Area
Author	Text Area
Training Material Type	Select List
Material Upload	Upload File
Material Link	URL
Licensing	Text Area
Skills	Ref Link
Training Event Ref	Ref Link

3.2.3 Skills

The metadata on skills are mostly derived from previous and ongoing efforts to formulate and categorise different HPC skills into a common framework, via the the HPC Certification Forum [5] whose founding member is a part of the HPC SPECTRA project consortium. The main concept is a HPC skills tree which can serve as a comprehensive encapsulation of all the key skills that are relevant to HPC that caters for different profiles from application users, developers and platform developers. Both training events and materials metadata contain references to this skills framework, which can then be used to group events/materials by skills (so users can find all relevant courses that covers particular areas, or stakeholders can review the availability, or lack of, courses in particular areas). This further supports the creation of specific skill “tracks” to indicate training pathways for different user groups, e.g. application developers.

Table 3. Skills metadata

Field Title	Data Type
Skill ID	Text
Skill Title	Default Title
Background	Text Area
Aim	Text Area
Outcomes	Text Area
Parent Skill(s)	Reference(s) to other skill ID(s)
Child Skill(s)	Reference(s) to other skill ID(s)

3.2.4 Modules & Tracks

The key objective of “tracks” is to facilitate the definition of training or learning pathways, i.e. a progressive list of skills for learning, that can be conceived for different HPC profiles, e.g. application users, developers. The tracks metadata contains an ordered list of skills or groups of skills (called Modules) that can be specified for different profiles. An example usage of tracks would be to guide portal users in identifying relevant courses according to their own chosen profiles or ‘profession

tracks’. Apart from a Tracks table, an associated Module table is shown where each module encapsulates one or more skills that represents the “stage” of a track.

Table 4. Tracks metadata

Field Title	Data Type
Track ID	Number
Track Title	Default Title
Track Description	Text
Track Image	Image
Modules	Ordered list of ref link to Modules

Table 5. Modules metadata

Field Title	Data Type
Module ID	Number
Module Title	Default Title
Module Description	Text
Skills	Ref Link to Skill(s)

4 Architecture Platforms and Solutions

Having the requirements for functionalities, users and data set, several candidate platforms and components were evaluated for their suitability in an operational framework for the EuroHPC Training Portal. It is important to note that the HPC SPECTRA project had always anticipated the adoption of pre-existing platforms/tools and inter-operability with other EuroHPC platforms from the outset, i.e. it did not set out to develop an entirely new training platform from scratch, nor it will be developed in isolation with other EuroHPC platform developments.

As described in Section 2, the training platform will be developed with close integration with the C2ISS platform that will become a central web presence of the European HPC ecosystem (subsection 4.1). Hence the architecture of the training portal will necessarily take into account some of its design decisions in collaboration with CASTIEL 2. The other key platforms/tools that have been considered concerns the user interface and management of training events and materials, which are covered in subsections subsections 4.2 and 4.3.

4.1 C2ISS

The C2ISS (CASTIEL 2 Information Sharing System, working title) is developed in CASTIEL 2. The ultimate goal for the platform is to become the central point of access to information on EU-funded HPC actors, which will greatly improve visibility and accessibility of the projects. CASTIEL 2 is using the HPC in Europe portal [1] as a foundation and build the following structure:

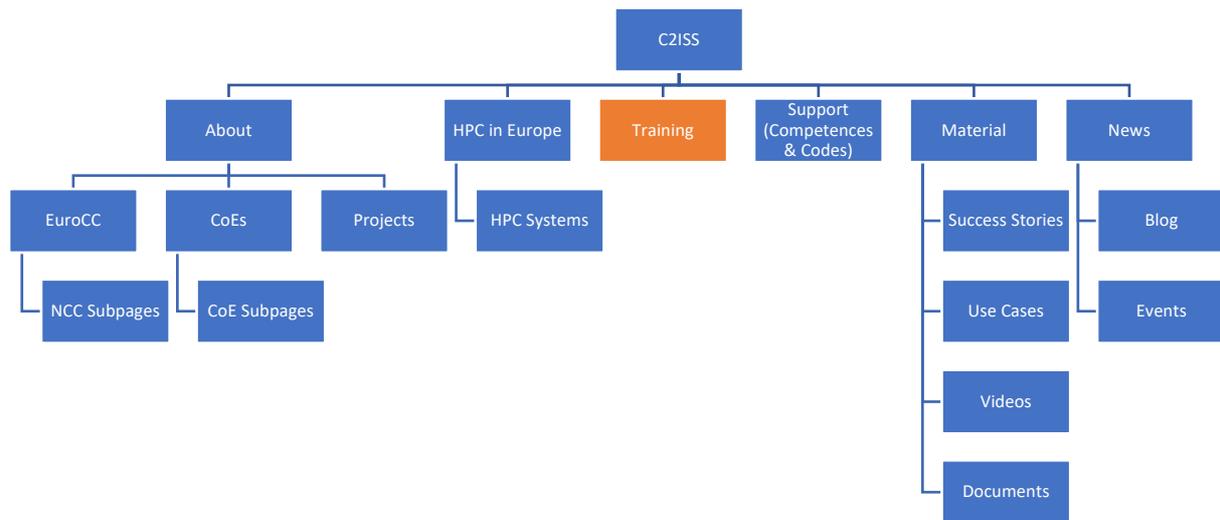


Figure 1. High-level structure of the C2ISS platform

The C2ISS is based on NCCs, CoEs and other projects providing content about them. This means, every project or centre can have a user profile that can contribute to different sections:

- the NCCs, CoEs and other EuroHPC JU funded projects,
- HPC in Europe (ecosystem overview, HPC systems)
- training
- support offered by the actors on the platform
- resources and materials
- news

As described in Section 2, the responsibility of developing the training section of the C2ISS platform will be delegated to the HPC SPECTRA project. This entails close collaboration in ensuring a consistent look and feel, i.e. visual elements, which is not a significant challenge as both CASTIEL 2 and HPC SPECTRA are building on top of the HPC in Europe portal.

In terms of inter-operability with C2ISS, some technical considerations need to be taken into account. The C2ISS platform is being developed based on the commonly-used Drupal content management platform [6] for large websites and portals. While it could potentially handle nearly all the functionalities that are needed for the training

portal, e.g. management of training events and materials, the Drupal platform would act as a gateway to other solutions (see subsections 4.2 and 4.3). The Drupal platform has plenty of scope and flexibility to be inter-operable and share data with other tools and solutions.

To ensure CASTIEL 2, HPC SPECTRA and possible future projects have a legal foundation for working on the portal and executing their DoAs, a memorandum of understanding is being developed that will serve as a terms of reference for collaboration.

4.2 Training events management

The solution that would facilitate the management of training events for the EuroHPC training portal has requirements that goes beyond those offered by basic “event management” tools, e.g. built-in Drupal function that offer registration forms and attendee management. These additional requirements include the ability to incorporate skills metadata, agenda/scheduling management, (optional) review processes and analytics/reporting tools. One solution in particular, called Indico, has been successfully adopted for many years in the European HPC training ecosystem (subsection 4.2.1). Another solution, called TeSS, has also been evaluated as a potential candidate (subsection 4.2.2).

4.2.1 Indico

Indico [3] is an event management tool developed by CERN to handle the organisation, scheduling, and management of events such as conferences, meetings, lectures, and workshops. It was created to meet the specific needs of the scientific community but has since grown into a versatile platform used by various organisations worldwide (see Figure 2 for a sample screenshot of its event management interface).

Key Features are:

- **Event Creation and Management:** Indico allows users to create and manage events of various types, including conferences, workshops, meetings, and seminars. It provides tools to manage event agendas, session scheduling, speaker management, and more.

- **Registration and Payment Handling:** The platform includes features for managing participant registration, including custom registration forms, ticketing, and payment processing. It also supports managing multiple registration tiers and handling discounts or special offers.
- **Abstract Submission and Reviewing:** Indico facilitates the submission of abstracts, papers, and presentations for events. It includes a reviewing system that allows organisers to manage the review process, assign reviewers, and handle feedback.
- **User and Access Management:** The tool provides robust user management features, allowing organisers to define roles and permissions for different users, such as participants, speakers, and reviewers.
- **Collaboration Tools:** Indico supports collaboration by allowing participants to share files, take notes, and engage in discussions. It also integrates with other tools like videoconferencing systems to support virtual or hybrid events.
- **Customisation and Branding:** Organisers can customise the appearance and branding of their event pages to match their organisation's identity.
- **Analytics and Reporting:** The platform offers reporting tools to analyse attendance, engagement, and other key metrics, helping organisers assess the performance of their events.
- **Multi-language Support:** Indico is designed to support multiple languages, making it accessible to a global audience.

Indico is widely used within the academic and research communities, especially in the European HPC ecosystem. It had been used successfully by PRACE to for a decade to host over 120+ training events per annum. Some HPC centres have since adopted the same platform for organising their national programmes of training events, as well as some CoEs. Furthermore, Indico is an open-source platform and offers APIs to connect the platform to others.

Conference 1 19 July

Created by Anamaria Costache

Event actions

General settings

Timetable

Material

Programme

Registration

Call for Abstracts

Contributions

Paper Reviewing

Chat Rooms

Video Services

Evaluation

Advanced options

Lists

Protection

Tools

Layout

Logs

General Settings

Title

Conference 1 [\(edit\)](#)

Description

No description

[\(edit\)](#)

Place

Location: CERN
 Room:
 Address:

[\(edit\)](#)

Start/End date

Starts : 19/07/2011 08:00
 Ends : 19/07/2011 18:00

[\(edit\)](#)

Timezone

Europe/Zurich

Additional info

No description

[\(edit\)](#)

Support

Caption : Support
 Email :

[\(edit\)](#)

Default style

[\(edit\)](#)

Visibility

Everywhere [\(edit\)](#)

Event type

conference [\(edit\)](#)

Keywords

Short display URL

There is not any short url yet. Click "Modify" to setup.

Chairpersons

Screen dates

Tuesday 19 July 2011 08:00 (*normal*) -> Tuesday 19 July 2011 18:00 (*normal*)

Types of contributions

Figure 2. Indico event management interface.

In summary, Indico can be deemed a sustainable, well-established solution, which is already familiar to many in the EuroHPC training community (e.g. many NCCs and the CoEs). Crucially, another advantage of Indico over other solutions is the ability to de-couple the event management interface hosted on Indico (for training providers) from the event display/browsing interface hosted on C2ISS within Drupal (for training portal users). Hence it is a tool that had been indicated in the project proposal to be deployed as part of the training portal.

4.2.2 TeSS

TeSS (Training eSupport System) [7] is a platform designed to serve as a central hub for trainers and trainees to find online resources, including training materials, events, and interactive tutorials. Training providers can use TeSS to promote their events and news while contributing to a growing repository of materials. For trainers, the platform offers a space to share materials and event details. For trainees, it serves as a

convenient gateway to discover relevant training opportunities and resources, as well as to perform guided analysis tasks through customised training workflows.

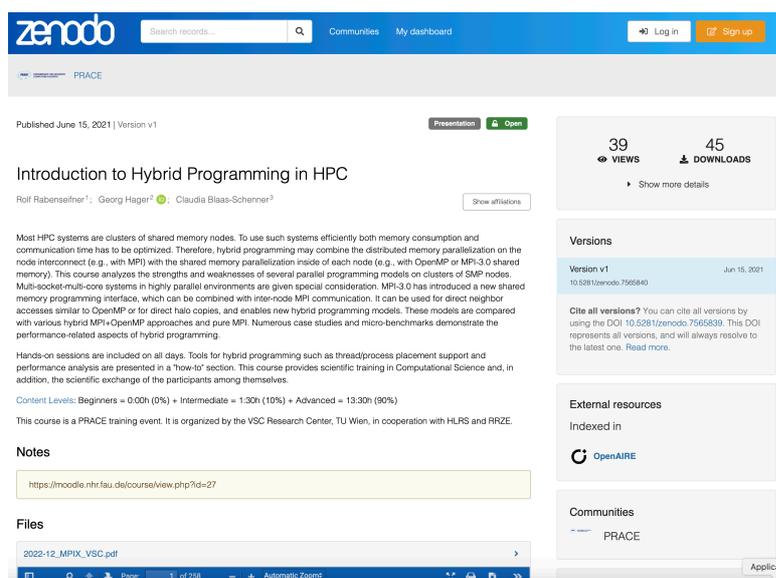
TeSS was initially created as part of ELIXIR [8], Europe's distributed infrastructure for life-science data. A key objective of ELIXIR is to train research scientists in effectively utilising computational infrastructures to tackle research challenges. The system facilitates access to both in-person training and a wide range of training materials and resources that had been and continues to be scattered across Europe.

Significant considerations were given to examine the potential to adopt TeSS as a solution for HPC training in Europe. Its objectives and use case are very well aligned with the objectives of the EuroHPC training portal, e.g. collecting training event information from disparate providers, management of training materials and pathways. However, TeSS is developed very much as a stand-alone platform and challenges remain in terms of how it could be integrated with C2ISS, e.g. how the interface to browse/search training events can be done within C2ISS (based on Drupal) without significant changes and forked maintenance of the TeSS codebase. Furthermore, there is less familiarity with TeSS among the European HPC ecosystem compared with Indico.

4.3 Training Materials Management

Another aspect of the EuroHPC training portal is to provide a system to manage HPC training materials that allows for its use among the community. This includes materials that have been produced/developed as part of training events, as well as standalone training resources that have been developed for remote learning. Therefore it encompasses a wide range of files (e.g. presentation slides, documents with text and graphics, audiovisual files such as videos, to compressed packages of example codes).

While there are many types of both open source and commercial “document archival” solutions for such files, a key idea of the HPC SPECTRA project is to ensure that the solution does not simply offer a naïve file repository but that these resources are annotated with the relevant metadata that can be widely accessible and discoverable. Therefore, the solution being proposed by HPC SPECTRA project would be to leverage the Zenodo digital repository [9] as the platform to store materials and their metadata while C2ISS would serve as a portal for users to navigate and find these materials. The bonus is that such materials remains discoverable via means beyond the training platform, e.g. via Zenodo itself or search engines. It also obviates the need for C2ISS to act as a file repository, i.e. that the materials will remain accessible beyond the HPC SPECTRA project (see Figure 3 as an example). An overview of the features offered by Zenodo is provided below.



The screenshot shows a Zenodo record for a presentation titled "Introduction to Hybrid Programming in HPC". The record includes the following information:

- Metadata:** Published June 15, 2021 | Version v1. It is a presentation and is open access.
- Statistics:** 39 views and 45 downloads.
- Authors:** Rolf Rabenseifner¹, Georg Hager², and Claudia Blas-Scherner³.
- Abstract:** Most HPC systems are clusters of shared memory nodes. To use such systems efficiently both memory consumption and communication time has to be optimized. Therefore, hybrid programming may combine the distributed memory parallelization on the node interconnect (e.g., with MPI) with the shared memory parallelization inside of each node (e.g., with OpenMP or MPI-3.0 shared memory). This course analyzes the strengths and weaknesses of several parallel programming models on clusters of SMP nodes. Multi-socket/multi-core systems in tightly parallel environments are given special consideration. MPI-3.0 has introduced a new shared memory programming interface, which can be combined with inter-node MPI communication. It can be used for direct neighbor accesses similar to OpenMP or for direct halo copies, and enables new hybrid programming models. These models are compared with various hybrid MPI+OpenMP approaches and pure MPI. Numerous case studies and micro-benchmarks demonstrate the performance-related aspects of hybrid programming.
- Content Levels:** Beginners = 0.00h (0%) + Intermediate = 1.30h (10%) + Advanced = 13.30h (90%).
- External Resources:** Indexed in OpenAIRE.
- Files:** A PDF file named "2022-12_MPIX_VSC.pdf" is available for download.

Figure 3. Example of HPC training material on Zenodo.

4.3.1 Zenodo

Zenodo [9] is an open-access digital repository developed by CERN under the European OpenAIRE project. It was launched in 2013 to provide researchers across various disciplines with a platform to share and preserve their research outputs. Zenodo supports the open science movement by enabling the sharing of research data, publications, software, and other scientific materials in a freely accessible manner.

Key Features are:

- **Open Access Repository:** Zenodo is a free and open repository, allowing researchers to upload and share their work with the global community. All uploaded content is assigned a Digital Object Identifier (DOI), ensuring that it is easily citable and permanently accessible.
- **Diverse Content Types:** The platform supports a wide range of research outputs, including datasets, publications, software, reports, posters, presentations, and even multimedia content. This flexibility makes Zenodo suitable for researchers from all fields.
- **Versioning and Updates:** Zenodo allows users to upload new versions of their work while maintaining links to previous versions. This feature is particularly useful for software and datasets that may be updated over time.
- **Metadata and Searchability:** Users can add detailed metadata to their uploads, making it easier for others to find and cite their work. The platform also integrates with search engines and other indexing services, increasing the visibility of the content.
- **Licensing Options:** Zenodo offers a variety of licensing options, enabling researchers to choose how their work can be used by others. This includes open licenses like Creative Commons, which promote sharing and reuse.
- **Community Collections:** Researchers and institutions can create their own communities on Zenodo, where they can group related research outputs under a single collection. This is useful for organising and promoting specific projects or thematic areas of research.



- **Integration with GitHub:** Zenodo integrates with GitHub, allowing developers to easily archive their software projects and code repositories. When a release is made on GitHub, Zenodo can automatically create a DOI for that release, ensuring proper citation.
- **Long-Term Preservation:** As a service hosted by CERN, Zenodo benefits from the robust infrastructure and long-term preservation policies of one of the world's leading research institutions. This ensures that the content remains accessible over time.

5 Proposed Architecture Overview

The evaluation of the platforms and solutions covered in Section 4 was carried out while keeping in mind the evolving operational framework of the first proposed platform architecture of the training portal. The following outlines how the chosen platforms/solutions are brought together into coherent, functional training portal for various users.

5.1 Location of datasets and metadata

It may be useful first to visualise where the various datasets reside on the various platforms that form the training portal. Figure 4 shows the following:

- Training event metadata is entered by training providers on Indico. These metadata are then synchronised with C2ISS for display and navigation by portal users, i.e. Indico serves as a back-end for displaying training events on C2ISS. Most users browse/search upcoming or relevant training events via C2ISS.
- Training materials and associated metadata are deposited (pushed) onto Zenodo by training providers (via an interface on C2ISS). C2ISS again acts as the hub for displaying the materials that are available; portal users can browse/search for relevant training materials on C2ISS but download materials via links to Zenodo.
- The other relevant datasets/metadata regarding skills, tracks and modules will be stored and managed on C2ISS, initially by HPC SPECTRA partners involved in development of the skills framework. These are then synchronised with Indico (so that training providers can annotate events).

Apart from the datasets and metadata shown, another architectural consideration is that a single sign-on solution would be supported for users of C2ISS and Indico. This enables training providers, such as NCCs/CoEs, to easily enter information on both platforms. For users, this allows for customised training recommendations (based on user input of preferred tracks and/or training courses attended) on C2ISS as well as a more convenient registration process into courses managed on Indico.

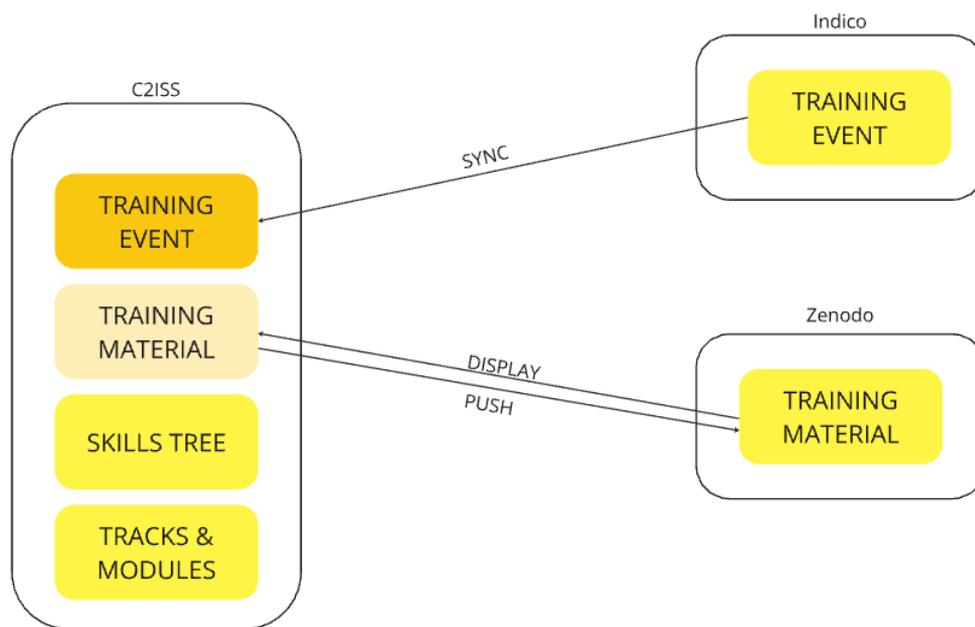


Figure 4. Overview of datasets and metadata locations

5.2 Key architectural concepts

Having conceptualised the flow of data between the key platform components (C2ISS, Indico and Zenodo), Table 6 outlines the key architectural concepts and functions of each component. It highlights that the Training section of C2ISS remains to be the focus of user interaction, i.e. individuals looking for training events/materials, while Indico and Zenodo provides supporting functionalities that supplements C2ISS, mainly for training providers.

Table 6. Key architectural concepts of the platform components

C2ISS	Indico	Zenodo
Central point of interaction with trainees (browsing, searching relevant training events/materials) – the landing page for these users.	Main portal for training providers to provide training events information, synchronised with	Mostly a file repository for training materials with metadata
Provides user (trainee) registration functionality for training recommendations based on input	C2ISS	Alternative route to discover materials

Displays courses relevant to specific parts of the skills framework, or within tracks (learning pathways or professional profiles)	Optionally training providers may make use of event management capabilities such as user registration, feedback, selection, agenda
Provides interface for training providers to upload materials with metadata	

5.3 User scenarios

Within the proposed architecture, some of the key user scenarios are foreseen in their interaction with the portal.

5.3.1 User looking for an upcoming training event

1. User navigates to the Training section of C2ISS.
2. User finds an upcoming training course (via browsing or filtering using skills, geographic location, etc.).
3. User clicks on the link for more information on a course, he/she may be taken to:
 - a. Indico course information page with all details of the course including agenda, etc., where training providers utilise the EuroHPC training portal to manage the event. User can register to the course directly on the Indico instance.
 - b. External course information page with all details of the course including agenda, etc. User registers on an external site as the training provider does not make use of the event management aspects of the EuroHPC Indico instance. However, key course metadata are captured on Indico nevertheless.

5.3.2 Registered user looking for relevant training events/materials

1. User registers for a C2ISS account.
2. User provides additional information regarding training, e.g. preferred training pathways/tracks, participation in courses, etc.

3. User navigates to the Training section of C2ISS.
4. C2ISS provides recommended training events/materials in the Training Section, based on user input, e.g. upcoming courses in relevant track.

5.3.3 Training provider promoting a new training event

1. Training provider signs into Indico using credentials (same as that for C2ISS).
2. Creates a new event in Indico and provides the relevant metadata, e.g. description, location, relevant skills.
3. The information is automatically synchronised with C2ISS, i.e. the event will be displayed and discoverable on C2ISS without any intervention by training provider.
4. Optionally the training provider can make use of the registration, scheduling functions offered by Indico to manage registrations, participant surveys, etc. If the training provider wishes to use an external site for registration, he/she simply provides the URL for external registration link.

5.3.4 Training provider uploading a new training material

1. Training provides signs into C2ISS using credentials (same as that for Indico).
2. Goes to a materials uploads page to upload training material and enter relevant metadata.
3. The materials and metadata are uploaded onto Zenodo automatically by C2ISS.
4. C2ISS retains the materials metadata (so it can be found on C2ISS), however, any download of the material will be via a link to Zenodo.

6 Conclusions and Next Steps

This deliverable outlines the key architectural requirements, components, decisions and concepts of the EuroHPC Training Portal. While the European HPC training landscape evolves, the training platform provides an important tool to support various audiences from HPC users, developers, training providers and HPC stakeholders. The design of the architecture was carried out with considerable experiences from the consortium partners in developing similar solutions in the past, while also catering for likely future developments, e.g. further elaboration of the HPC skills framework. The architecture was also conceived in a manner as not to “reinvent the wheel” and with an eye on sustainability, i.e. ensuring that the platform can be usefully maintained beyond the life of the project within reasonable resources.

The next steps will be for the HPC SPECTRA consortium to deploy a functional training portal for pilot users in the next few months. Preliminary technical implementation has already begun and close collaboration with CASTIEL 2 in the C2ISS roll-out, as well as the wider HPC training community, e.g. the NCCs and CoEs, will ensure that the pilot portal begin serving its target audiences in 2025.

7 References

- [1] HPC in Europe Portal: <https://hpc-portal.eu/>
- [2] Partnership for Advanced Computing in Europe website: <https://prace-ri.eu/>
- [3] Indico software by CERN: <https://indico.cern.ch/>
- [4] Information about the CASTIEL 2 project: <https://www.eurocc-access.eu/about-us/the-projects/>
- [5] HPC Certification Forum: <https://www.hpc-certification.org/>
- [6] Drupal Content Management System: <https://www.drupal.org/>
- [7] Training eSupport System by ELIXIR: <https://tess.elixir-europe.org/about>
- [8] ELIXIR infrastructure for life science: <https://elixir-europe.org/>
- [9] Zenodo digital repository: <https://zenodo.org/>