Dear reader,
please note that this document represents the submitted deliverable to the European Commission (EC). The review process of the EC is ongoing and once the reviewed document is available, it will be uploaded here.

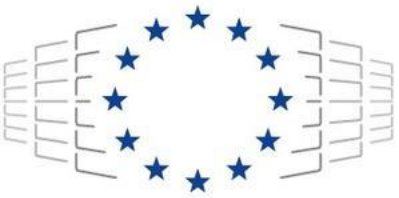
DIGITAL-EUROHPC-JU-2022-NCC-01

Coordination & Support
for National Competence Centres on a European Level Phase 2

Project Number: 101102047

D2.2

Second year report on NCC/CoE Networking, Mapping of Competences, Codes and Services



This project has received funding from the European High-Performance Computing Joint Undertaking (JU) under grant agreement No 101102047. The JU receives support from the Digital Europe Programme and Germany, Italy, Spain, France, Belgium, Austria.

Work package:	2	NCC/CoE Networking and Mapping of Competences, Codes and Services	
Author(s):	Laura Morselli		CINECA
Approved by	Project Management Team		2024-12-13
Reviewer 1	Aline Melinette		BSC
Reviewer 2	Jisika Yono		HLRS
Dissemination Level	PU - Public		

Date	Author	Comments	Version	Status
2024-11-24	Laura Morselli	Initial draft	V0.0	
2024-12-02	Laura Morselli	Comments from WP2 partners included	V0.1	
2024-12-10	Laura Morselli	Comments from internal reviewers included	V0.2	
2024-12-11	Laura Morselli	Version sent to PMT for final checks	V0.3	
2024-12-12	Laura Morselli	Comments of PMT implemented	V0.4	
2024-12-12	Laura Morselli	Final version to PMT	V0.5	

List of abbreviations

AI	*Artificial Intelligence*
BSCW	*HLRS Social Workspace Server*
C2ISS	*CASTIEL 2 Information Sharing System*
CI/CD	*Continuous Integration / Continuous Deployment*
CM	*Competence Map*
CoE	*Centre of Excellence*
CSA	*Coordination and Support Action*
DL	*Deep Learning*
EuroHPC JU	*European High Performance Computing Joint Undertaking*
GA	*Grant Agreement*
HPC	*High Performance Computing*
HPC3	*HPC CoE Council*
HPDA	*High-Performance Data Analytics*
KPI	*Key Performance Indicator*
ML	*Machine Learning*
MLops	*Machine Learning Operations*
NCC	*National Competence Centre*
PMT	*Project Management Team*
QC	*Quantum Computing*
TPR	*Technical Progress Report*
WP	*Work Package*
WPL	*Work Package Leader*

Executive Summary

This document outlines the objectives, activities, and outcomes of Work Package (WP) 2 for CASTIEL 2 during the reporting period of January 1, 2024, to December 31, 2024. CASTIEL 2 builds on the mission of the CASTIEL Horizon (H2020) project by coordinating and supporting the National Competence Centres (NCCs) while expanding its scope to include the Centres of Excellence (CoEs) in High-Performance Computing (HPC) in the second phase. WP2, titled "NCC/CoE Networking and Mapping of Competences, Codes, and Services," aims to enhance networking and collaboration between NCCs and CoEs as well as among the NCCs and CoEs themselves. This effort focuses on mapping the competences and codes within these entities to make them accessible to the broader European HPC community.

The primary objectives of WP2 are: *i*) to facilitate regular discussions between NCCs and CoEs on shared interests, with support from the HPC CoE Council (HPC3) and EuroCC2; *ii*) to collaborate with HPC3 to maintain a high-level dialogue with the CoEs on strategic aspects related to large-scale HPC application; *iii*) to contribute to the design of the CASTIEL 2 Information Sharing System (C2ISS), specifically regarding specifications and technical requirements for storing, indexing, and referencing HPC competences, codes, and libraries; *iv*) to organize workshops on specific topics and skills, tailored to the needs of NCCs and CoEs.; *v*) to steer WP2 activities based on observed improvements in competences.

One of the major outcomes of the reporting year was the finalisation of the **upgraded Competence Map**, which categorizes expertise from NCCs into five macro-areas: *Cluster Infrastructure and Management, Building and Deploying Parallel Programs, Big Data and Artificial Intelligence, Quantum Computing,* and *Domain-Specific Software.* The mapping exercise aggregated individual-level expertise to derive NCC-level competences and also resulted in the development of the new **Experts Database**. The database, shared internally via BSCW, is a user-friendly, HTML-based tool that allows users to search for experts based on macro-categories and sub-categories of competences. This effort provides an accessible resource for connecting stakeholders with relevant expertise.

WP2 also focused on strengthening collaboration between NCCs and CoEs through activities such as the **"2nd NCCs-CoEs Online Meeting"**, which introduced new CoEs (EoCoE, POP-3, dealii-X, and MICROCARD-2) and projects like EPICURE. Additionally, WP2 developed the **"CoEs' Codes White Book,"** a comprehensive catalogue of 72 tools and codes managed by CoEs. This resource provides valuable insights into the software landscape and supports knowledge-sharing initiatives. Additionally, WP2 organized targeted webinars and workshops, including a six-part series on **Energy-Efficient Computing**. Complementing this effort, WP2 continued to support the organization of the **"Code of the Month" webinars**, providing technical insights on the selected codes managed by the CoEs and encouraging collaboration with NCCs and end users.

Table of Contents

List of Figures

1 Introduction

This document serves as the second periodic report for WP2 of CASTIEL 2 [1], providing an overview of the WP objectives and results for the reporting period from January 1, 2024, to December 31, 2024. The CASTIEL 2 Coordination and Support Action (CSA) builds upon the mission of the CASTIEL Horizon (H2020) project by coordinating and supporting National Competence Centres (NCCs) and extending its scope to include Centres of Excellence (CoEs) in High-Performance Computing (HPC). WP2, titled "NCC/CoE Networking and Mapping of Competences, Codes, and Services," aims to enhance networking and collaboration both within and between NCCs and CoEs and to map the competencies and codes of these entities for the benefit of the entire European HPC ecosystem.

Specifically, WP2 has the following objectives:

1. Facilitate regular discussions among NCCs and CoEs on topics of shared interest, with support from the HPC CoE Council [2] (HPC3) and EuroCC2 [3].
2. Engage with HPC3 to promote collaboration among CoEs.
3. Contribute to the design of the Information Sharing System (C2ISS), focusing on technical requirements and specifications for storing, indexing, and referencing HPC competencies, codes, libraries, and services—tailored to the needs of both information providers (NCCs, CoEs) and academic/government end-users.
4. Develop and conduct workshops on topics outside of HPC, such as soft skills and complementary competencies, based on needs expressed by NCCs and CoEs, and make the materials available.
5. Guide WP2 activities based on observed competence improvements, types of information accessible through the EuroCC Access portal, and provide KPI input for WP1.

WP2 activities are organized into two main tasks:

- **T2.1 “Contribution to the C2ISS and the EuroCC Access Portal”:** Focuses primarily on mapping competencies and codes for integration into the C2ISS and the EuroCC Access Portal.
- **T2.2 “CoE/NCC Networking and Collaboration Support”:** Aims to enhance communication and collaboration among NCCs and CoEs and organizes relevant events (e.g., webinars, workshops).

WP2 (as all CASTIEL 2 WPs) collaborates with NCCs and CoEs through designated WP2 “champions” and their deputies—individuals appointed by each NCC/CoE to act as contact points between CASTIEL 2 and their respective centres and to report on WP2 activities within their organizations. Together, WP2 partners, champions, and deputies make up the WP2 Working Group.

WP2 continued its active collaboration with the CoEs that started in January 2023, fostering ongoing partnerships and facilitating shared activities within the European HPC ecosystem. These are: BioExcel [4], CEEC [5], ChEESE [6], EXCELLERAT-P2 [7], ESiWACE-3 [8], HIDALGO-2 [9], MaX-3 [10], MultiXscale [11], Plasma-PEPSC [12], and SPACE [13]. In addition to these established relationships, WP2 expanded its network in the second year by integrating four new CoEs: EoCoE [14] and POP-3 [15], which joined in January 2024, and dealii-x [16] and MICROCARD-2 [17], which came on board in November 2024.

Section 1 serves as the introduction to the document, providing an overview of its purpose and structure. Section 2 of this deliverable summarises the activities undergone to obtain the upgraded Competence Map of the NCCs. Section 3 focuses on the activities organised to foster the collaboration between NCCs and CoEs. Section 4 presents the activities aimed at enriching the competences of the network of NCCs and CoEs. Finally, Section 5 presents the main achievements of WP2 during the second year of CASTIEL2.

2 The Upgrade of the Competence Map

2.1 What happened in Year 1

Since the start of CASTIEL 2, WP2 recognized that mapping NCCs' competences effectively, as specified by the project's objectives, requires a clear, shared definition of "competences". In addition, a clear purpose of the Competence Map and of its users need to be clear for this task to be successful.

As stated in D2.1, the challenge of a clear definition of competences was raised in various occasions during the first year of CASTIEL2. Given that the Competence Map created in CASTIEL 1 mixed competences and services, WP2 prioritized coordination with other CASTIEL 2 WPs to set boundaries for the Competence Map's content and visualization on EuroCC Access [18] and its inclusion in the future C2ISS database. Following that, WP2 presented an upgrade plan for the Competence Map, outlining the strategy and the roadmap for the activity. At the end of the first year of the project, these main points on the Competence Map were considered as key:

1. **Purpose of the Competence Map**: To provide an ecosystem-wide overview of HPC competences, facilitate collaboration between NCCs, CoEs, and other EU projects, assist end-users in identifying competence providers, and help the Project Management Team (PMT) and Joint Undertaking [19] (JU) monitor competence growth.
2. **Definition of Competences**: Competence involves the knowledge, skills, and experience required to perform a task effectively. It enables service provision but is distinct from the service itself. For example, possessing the competence to port codes to GPUs could lead to consultancy or training services in this area if resources are available.
3. **Strategy for the upgrade of the Competence Map**: Distinguish competences from services and exploit available material to define the new Competence Map categories. WP2 launched a task force to support the Competence Map upgrade.

D2.1 presented the five macro-categories of competences that were the results of the work of the task force dedicated to the upgrade of the Competence Map. Each macro-category is additionally divided into two levels of subcategories. Figure 1 shows the macro-categories and first level sub-categories of the upgraded Competence Map (after the implementation of the feedback from WP2 champions, described in Section 2.2). The second level of sub-categories is reported in the Annex (Section 8).

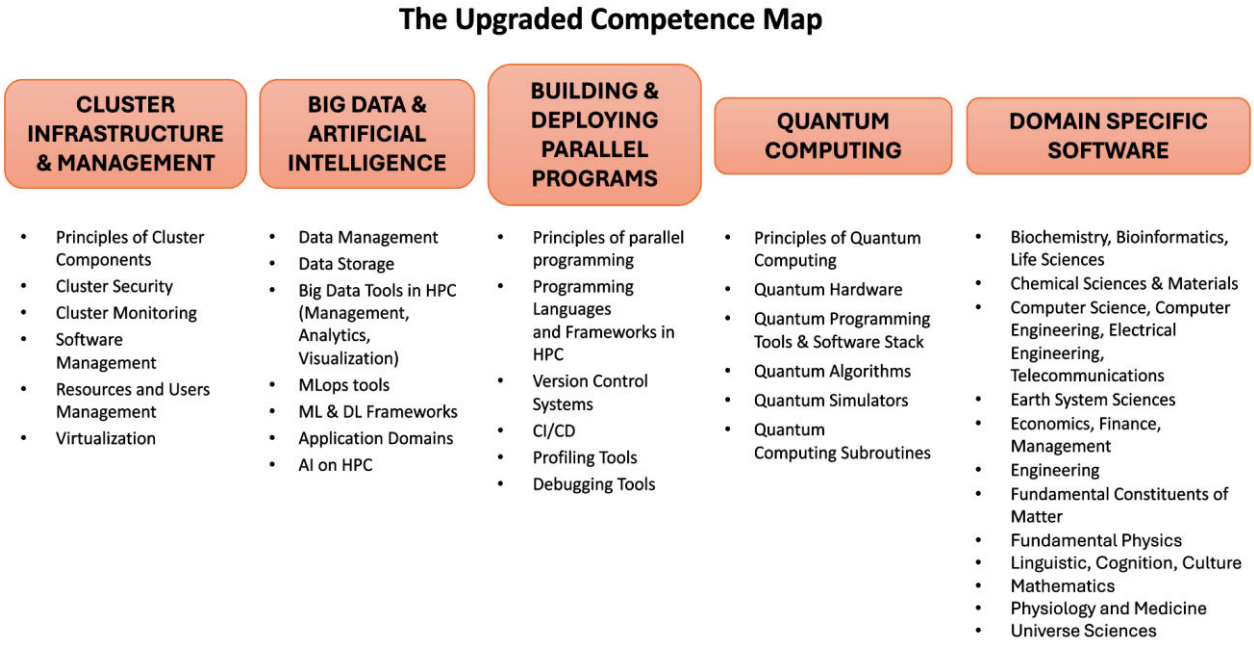


Figure 1: The macro-categories and first level sub-categories of the upgraded Competence Map.

2.2 Feedback from the WP2 champions and first year review of CASTIEL2

The structure of the upgraded Competence Map, fruit of the works of the task force, was presented to the WP2 Working Group at the end of the first year of the project and a questionnaire was launched to collect feedback on the same. The collection of answers to the questionnaire took place in January 2024. The following questions were asked to the WP2 champions:

1. The following 5 macro-categories have been identified: Cluster Infrastructure & Management; Big Data & Artificial Intelligence; Building & Deploying Parallel Programs; Quantum Computing; Domain-specific software. Thinking about your NCC or CoE, do you think these categories encompass most of the knowledge (related to HPC, AI and HPDA) within your NCC/CoE?
2. For the macro-category "Cluster Infrastructure & Management" the following subcategories have been identified: Basics of Cluster Components; Cluster Security; Cluster Monitoring; Software Management; Resources and Users Management; Virtualization; Valorisation. If applicable, do you think these sub-categories well describe the competences within your NCC/CoE in the category "Cluster Infrastructure & Management"?
3. For the macro-category "Big Data and Artificial Intelligence" the following sub-categories have been identified: Data Management; Data Storage; Big Data Tools in HPC; MLops tools; ML & DL Frameworks; Multi GPU AI; Valorisation. If applicable, do you think these sub-categories well describe the competences within your NCC/CoE in the category "Big Data and Artificial Intelligence"?
4. For the macro-category "Building and Deploying Parallel Programs" the following sub-categories have been identified: Basics of parallel programming; Programming Languages and Frameworks in HPC; Version Control Systems; CI/CD; Profiling Tools; Debugging Tools; Valorisation. If applicable, do you think these sub-categories well

describe the competences within your NCC/CoE in the category "Building and Deploying Parallel Programs"?

5. For the macro-category "Quantum Computing" the following sub-categories have been identified: Basics of Quantum Computing; Quantum Hardware; Quantum Programming Tools & Software Stack; Quantum Algorithms; Quantum Simulators; Quantum Computing Subroutines; Valorisation. If applicable, do you think these sub-categories well describe the competences within your NCC/CoE in the category "Quantum Computing"?
6. For the macro-category "Domain Specific Software" the following sub-categories have been identified: Computational Fluid Dynamics; Molecular Dynamics; Computational Materials Science; Computational Chemistry; High Energy Physics; Weather and Climate; Astrophysics. If applicable, do you think these sub-categories well describe the competences within your NCC/CoE in the category "Domain Specific Software"?

Additional questions were added to give the chance to respondents to provide their feedback on each category and on the Competence Map as a whole as free text.

Figure 2 shows the answers provided on the six questions listed above. The following aggregated feedback was drawn from the results:

- The large majority of respondents (78%) considered the new structure of macro-categories complete and sufficient to express the competences of their NCCs/CoEs.
- For all the macro-categories, the respondents that declared that the suggested sub-categories do not fit the knowledge of their NCCs/CoEs (thus answering “No”) are always below 10%.
- For the macro-categories “Building and Deploying Parallel Programs” and “Domain Specific Software” the fraction of respondents that answered “Partially” is large, above 40%. Thus, special attention was given to the feedback provided by the respondents in the form of “open answers” when finalizing the sub-categories of these two macro-categories. As an example, several respondents suggested the use of the classification of the HPC in Europe portal for the classification of the domains in the “Domain Specific Software” macro-category and this feedback was implemented in the final version of the upgraded Competence Map.

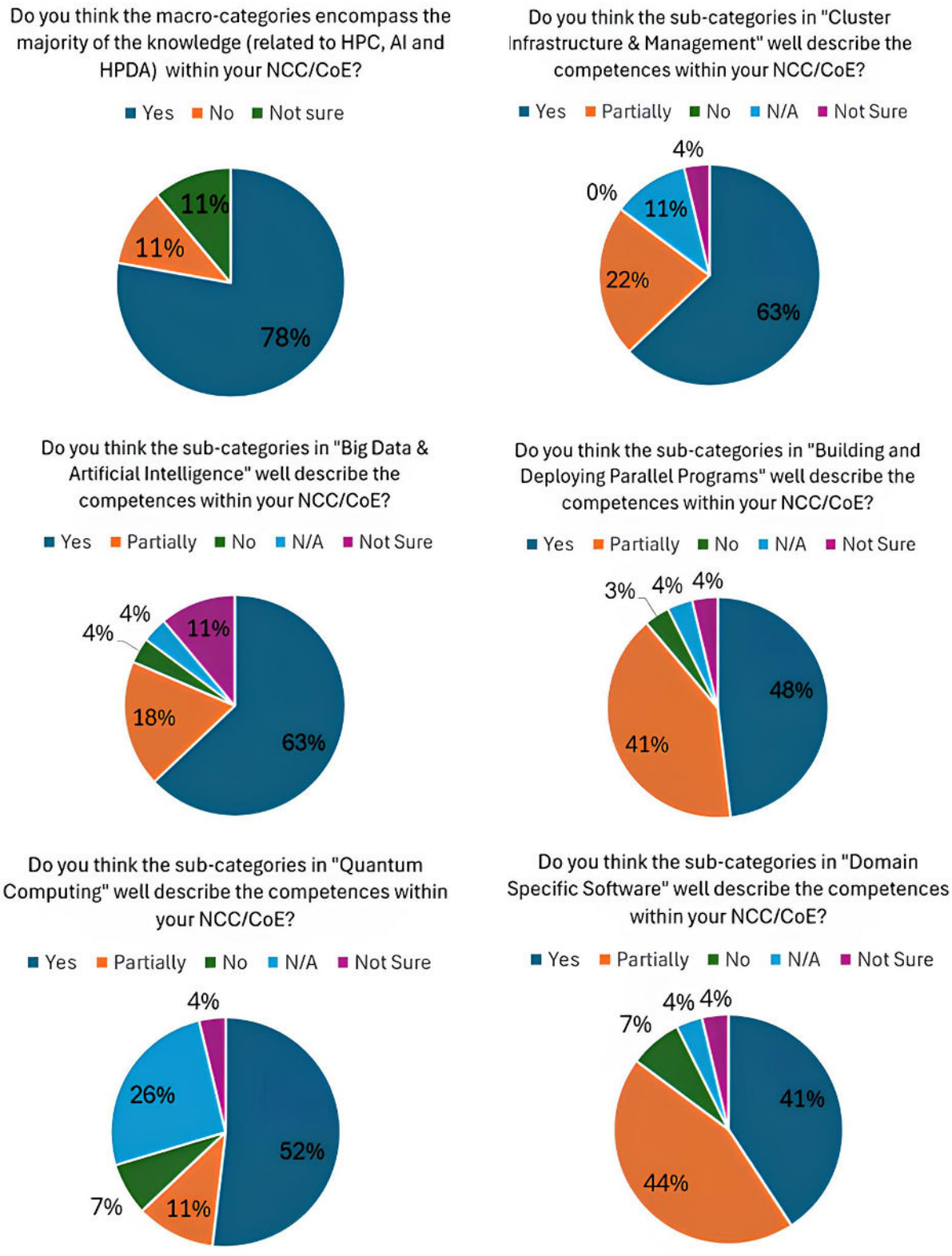


Figure 2: Feedback on the upgraded Competence Map collected via a questionnaire following the WP2 working group meeting in December 2023.

The upgraded Competence Map (including its new categories and sub-categories) was shown at the first-year review of CASTIEL2 (February 2024). The feedback collected from the reviewers underlined that:

- While the new structure is appreciated, the consortium should make sure that the complexity to maintain the Competence Map does not increase;
- The information in the Competence Map is standardized to enable its use;
- The size of the Competence Map and the ontology should remain the focus to ease the use of the Competence Map itself;
- A clear purpose and destination of the collected information needs to be established;
- Challenges such as NCCs and CoEs having an intrinsically different organizational structure and target group should be taken into account.

2.3 Setting the stage for the efficient management of the Competence Map

To address the challenges and points raised by the reviewers, of which the consortium and WP2 in particular was already well aware of, WP2 decided to proceed as follows.

1. **Manage the Complexity of the Map.** It is important to recognize that NCCs are heterogenous entities composed of organizations that might (or might not) have a specific focus within the HPC ecosystem, in terms of application domains as well as in terms of end-users. As a consequence, it has always been challenging to ensure that the information collected from NCCs is complete and standardized, to represent a European-wide perspective. Additionally, NCCs and CoEs are intrinsically different organizations (e.g., national versus pan-European, users-oriented versus developer-oriented, focus on several domains versus specific focus on a specific domain). At the end of the first year, WP2 decided that to avoid unnecessary complexity of the Competence Map, the CoEs would be added to the solely under the "Domain-specific software" category, thus emphasising their domain of expertise. Concerning the collection of competences of the NCCs, WP2 decided to approach this starting from the competences and expertise of the individuals working in NCCs and to use them as a starting point for drawing the competences at the NCC level.

2. **Synchronization with the EuroCC2 PMT.** In Y1, EuroCC2 PMT started the so-called Experts Database: a living document shared internally with NCCs on which individuals could enter information on their expertise, so that the network of NCCs could benefit from a list of experts in given topics. While this is a commendable effort with the ability to significantly impact the network, the first version of the database was not characterized by a standardized structure, meaning that individuals could provide information at a very different level of detail. Also, the number of entries was relatively low with respect to the number of individuals in the EuroCC2 network. Thus, CASTIEL2 WP2 initialized a synchronization with EuroCC2 PMT and proposed to collect the information needed for the "Experts Database" as a starting point for the upgraded Competence Map, to address point 1.

3. **Strategy for Collecting Information**. Taking into consideration points 1 and 2 listed above, WP2 decided to proceed as follows: i) collecting expertise of individuals involved in NCCs following the upgraded Competence Map structure; ii) using the individual-level information to create the new Experts Database to share internally within the EuroCC2 network; iii) drawing the NCC-level competences by aggregating the expertise of individuals working in the same NCC; iv) starting a feedback loop with

the NCCs' WPLs and WP2 champions to adjust the NCC competences if needed (e.g., in the case that not all individuals working in the NCC answered the questionnaire, resulting in missing competences). It is important to notice that the Competence Map of NCCs will not include the second level of details for each macro-category, as that would make the Competence Map difficult to manage and keep up to date. That level of detail will only be included in the Experts Database. Figure 3 shows the overall strategy designed by WP2 and agreed on with the EuroCC2 PMT for the upgrade of the Competence Map.

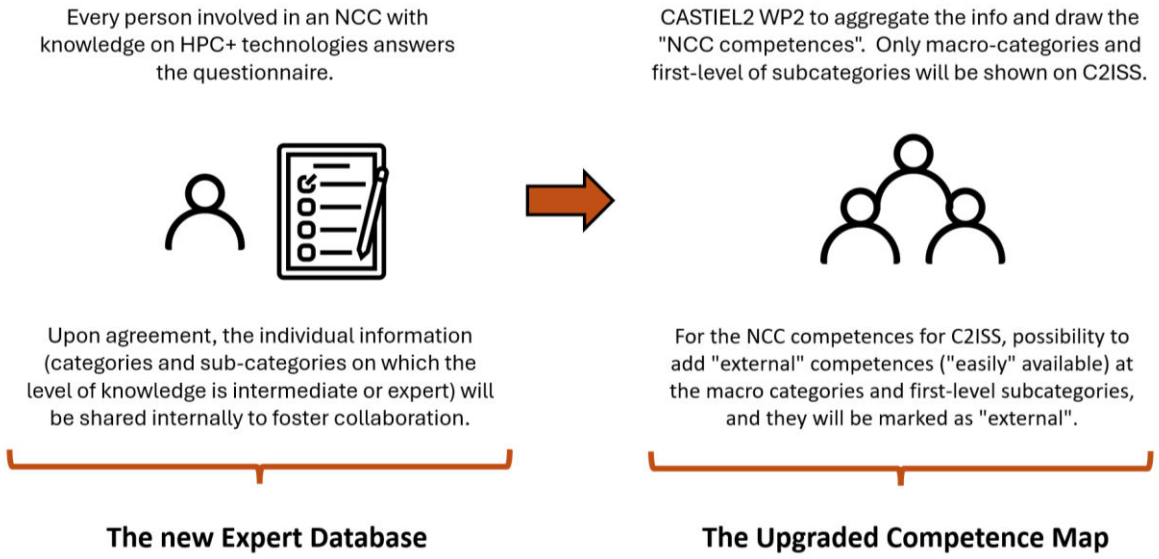


Figure 3: Strategy behind the collection of competences for the upgraded Competence Map.

WP2 had the chance to present the strategy mentioned above as well as the finalised upgraded structure of the Competence Map to NCCs and CoEs during the All-Hands-Meeting that took place in Slovakia in April 2024. At that occasion, the following feedback was collected from NCCs:

- Some NCCs might have "external" competences that are within organisations which are not in the consortium but are tightly connected to the NCCs themselves. Some NCCs requested the possibility to include these competences in the Competence Map.
- The Experts Database is seen as a very useful tool for collaborations but needs to be kept up-to-date and easy to use.
- The way individual competences are collected should allow for easy-to-manage updates, so that people do not need start from anew every time an update is requested.

2.4 The questionnaire on competences

As explained in Section 2.3, WP2 decided to start from the expertise of individuals involved in NCCs to draw the competences at the NCC level while, at the same time, producing the new Experts Database. To this aim, WP2 prepared a questionnaire to share with the NCCs. To ensure that information can be updated easily by individuals, thus addressing a request of the NCCs, as well as to make sure that the provided information remains coherent with time, the questionnaire was created exploiting a platform ensuring that answers can be saved, so that each of the respondent, when asked, can re-open the link, see the previous answers and decide to

modify them if needed (for example, if he/she/they developed new competences on a specific tool) or to leave them unchanged.

For each macro-category and relative sub-categories, the respondents have to assess their level of knowledge in four given categories:

- None
- Basic
- Intermediate
- Expert

The questionnaire is structured so that the questions exploring the level of knowledge on sub-categories appear only when the answer to the related category is “Intermediate” or “Expert”.

A “trial” test of the questionnaire was presented during a session dedicated to competences which took place in the All Hands Meeting in Slovakia in April 2024. Following the meeting, WP2 and EuroCC2 PMT organised a webinar dedicated to the new Experts Database and upgraded Competence Map. The webinar, that took place in May 2024, was organised to explain the strategy behind the new Experts Database and the upgraded Competence Map to all the NCCs, and not only the people that were participating in the All-Hands-Meeting.

Shortly after the webinar, the questionnaire on competences was shared with NCCs with the request to spread it internally and ensure that most people would answer it. Additionally, WP2 participated in the one-to-one feedback sessions that EuroCC PMT organised with NCCs (June 2024). These sessions were beneficial to remind people from NCCs to share the questionnaire internally. At the end of June 2024, the answers to the questionnaire were 302. Of these, 34% declared that they did not wish to include their information on the new Experts Database, thus the information they provided was used only at the aggregated level and for obtaining the NCC competences.

2.5 The new Experts Database

Following the collection of expertise from individuals working in NCCs, WP2 proceeded with the preparation of the new Experts Database to share internally, including only those people who accepted the usage of their information to this aim.

To ensure that the new Experts Database is easy to use, WP2 decided to create an html file that can be open in every browser and that allows the selection of macro-categories and their sub-categories. This file was created exploiting a combination of Python and Javascript. Following this selection, a list of experts with their name, email and NCC appear. Figure 4 shows the structure of the html file. Note that the names of experts appear also when selecting only the macro-category and no further level of detail.

The new Experts Database was shared internally with NCCs via BSCW in September 2024. A one-page guide on using the database, shown in Figure 5, was also provided. The guide highlights the importance of clearly introducing the purpose of the message when contacting an expert. Additionally, it recommends copying both CASTIEL2 WP2 and the expert’s NCC Work Package Leader (WPL) in the email’s CC field.

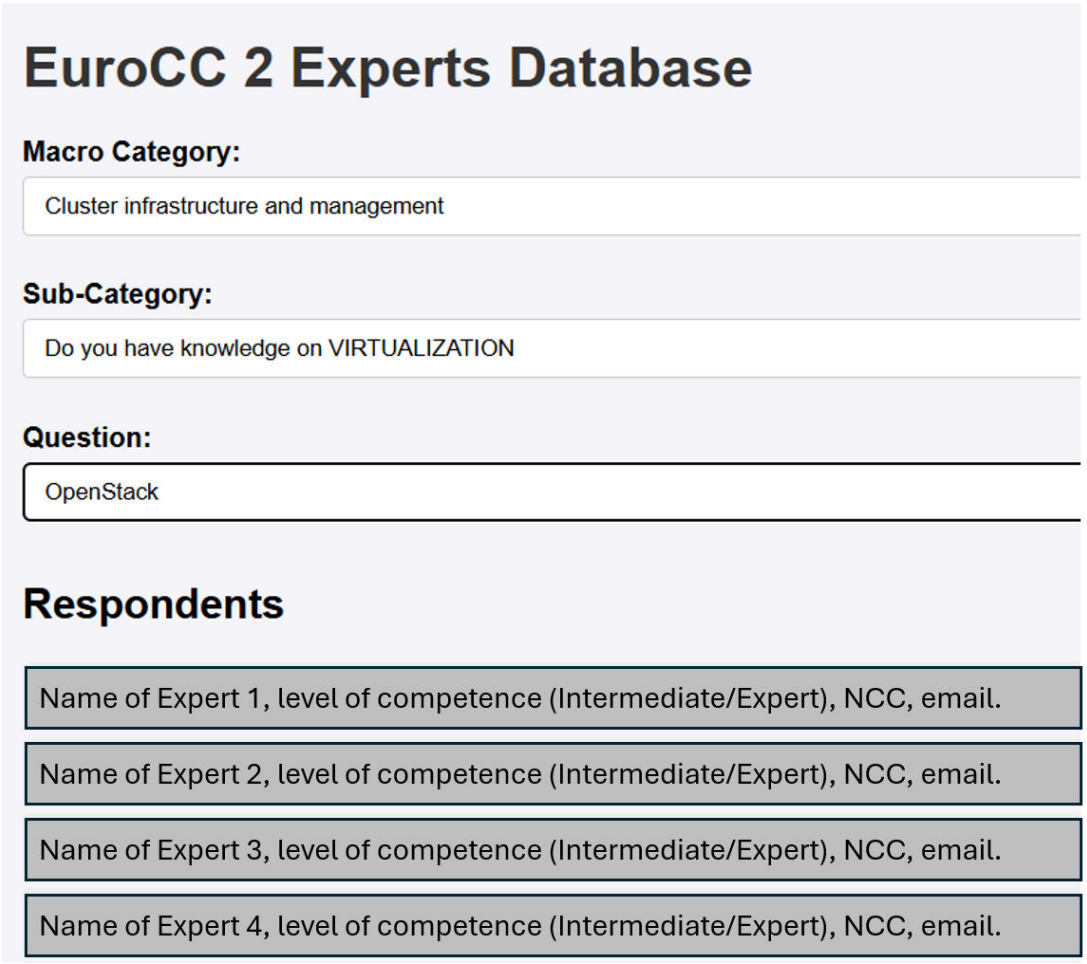


Figure 4: The new Experts Database of EuroCC2.

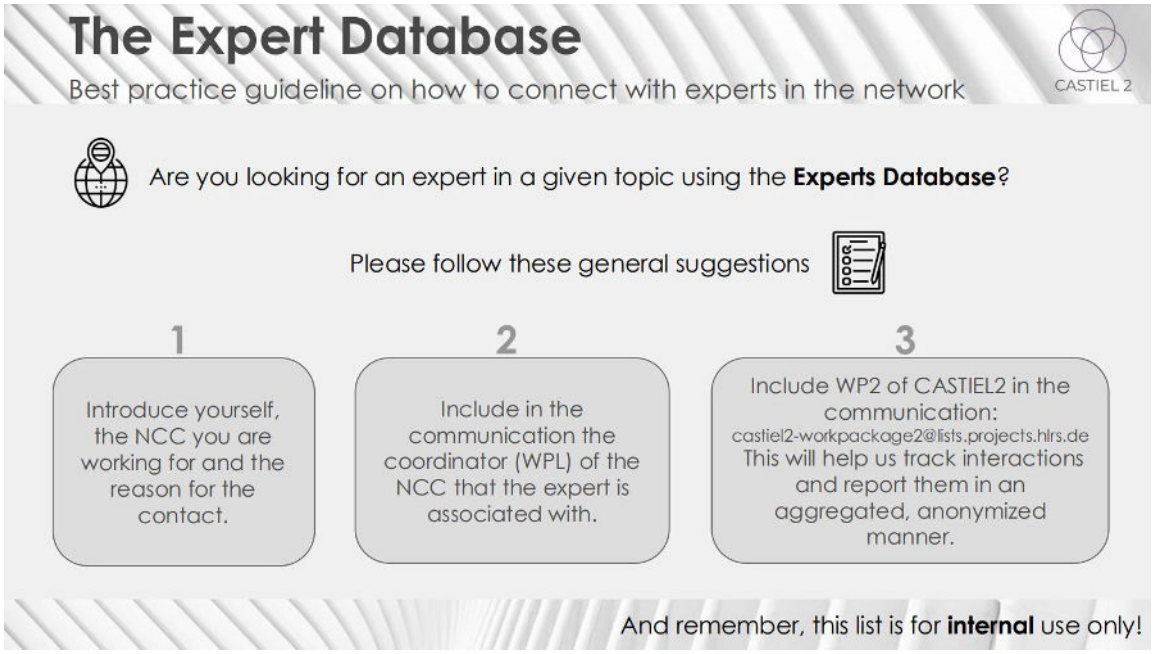


Figure 5: One-page "best practice guide" on how to make best use of the new Experts Database.

2.6 The upgraded NCCs Competence Map

Following the preparation of the Experts Database, WP2 started the aggregation of competences at the NCC level. The following strategy was established:

- Select individuals from the same NCC
- For each sub-category, if at least one individual from an NCC assessed their knowledge as "Expert", then the NCC level on that same competence is also categorized as "Expert".
- In the same way, if at least one individual from an NCC has declared that their expertise in a given sub-category is "Intermediate", and there are no individuals from the same NCC who replied "Expert" on the same category, then the NCC level of that given category is set at "Intermediate".
- To classify the knowledge of an NCC in a macro-category, we follow a structured approach based on the distribution of expertise levels (Advanced, Intermediate, None/Basic) across its sub-categories. The classification prioritizes higher levels of expertise while incorporating thresholds to ensure fairness. Specifically, an NCC is classified as **Advanced** in a macro-category if it demonstrates Advanced-level expertise in at least three sub-categories within that macro-category. If the NCC does not meet this Advanced threshold but has Intermediate expertise in a majority of the sub-categories (e.g., more than 50%), it is classified as **Intermediate**. Finally, if an NCC fails to meet these criteria or has significant None/Basic levels or missing data, it is classified as **None/Basic**. This method ensures a transparent, hierarchical evaluation of expertise, balancing specialization (Advanced classification) and broader competency (Intermediate classification).
- With the aim of facilitating interaction with the NCCs during the feedback process needed to finalise the upgraded NCCs Competence Map, WP2 also kept record of the percentage of respondents and of the provided levels for each category.

The preliminary results of this exercise were shown to the NCCs during the "2nd NCCs-CoEs online meeting" that took place in November 2024. It is important to underline that WP2 is aware that at the time of writing the depiction of NCCs' competences is partial as it is possible that experts from NCCs did not answer the questionnaire. Nevertheless, there are already some interesting trends worth exploring in this deliverable.

Figure 6 shows the overall classification of the expertise of the NCCs within the four macro-categories of the upgraded Competence Map. Based on the information collected by WP2 via the "Questionnaire on Competences", 44% of the NCCs result in advanced knowledge in "Cluster Infrastructure and Monitoring", while 34% are "Intermediate". For the macro-category "Building and Deploying Parallel Programs", the percentages are: 47% "Advanced", 34% "Intermediate" and 19% "None/Basic". For "Big Data and Artificial Intelligence" the fractions are: 44% "Advanced", 37% "Intermediate" and 19% "None/Basic". This is the category with the highest fraction of NCCs having "Advanced" or "Intermediate" expertise. Finally, for Quantum Computing, we observe the smallest fraction of "Advanced" (19%) and "Intermediate" (16%) level of expertise.

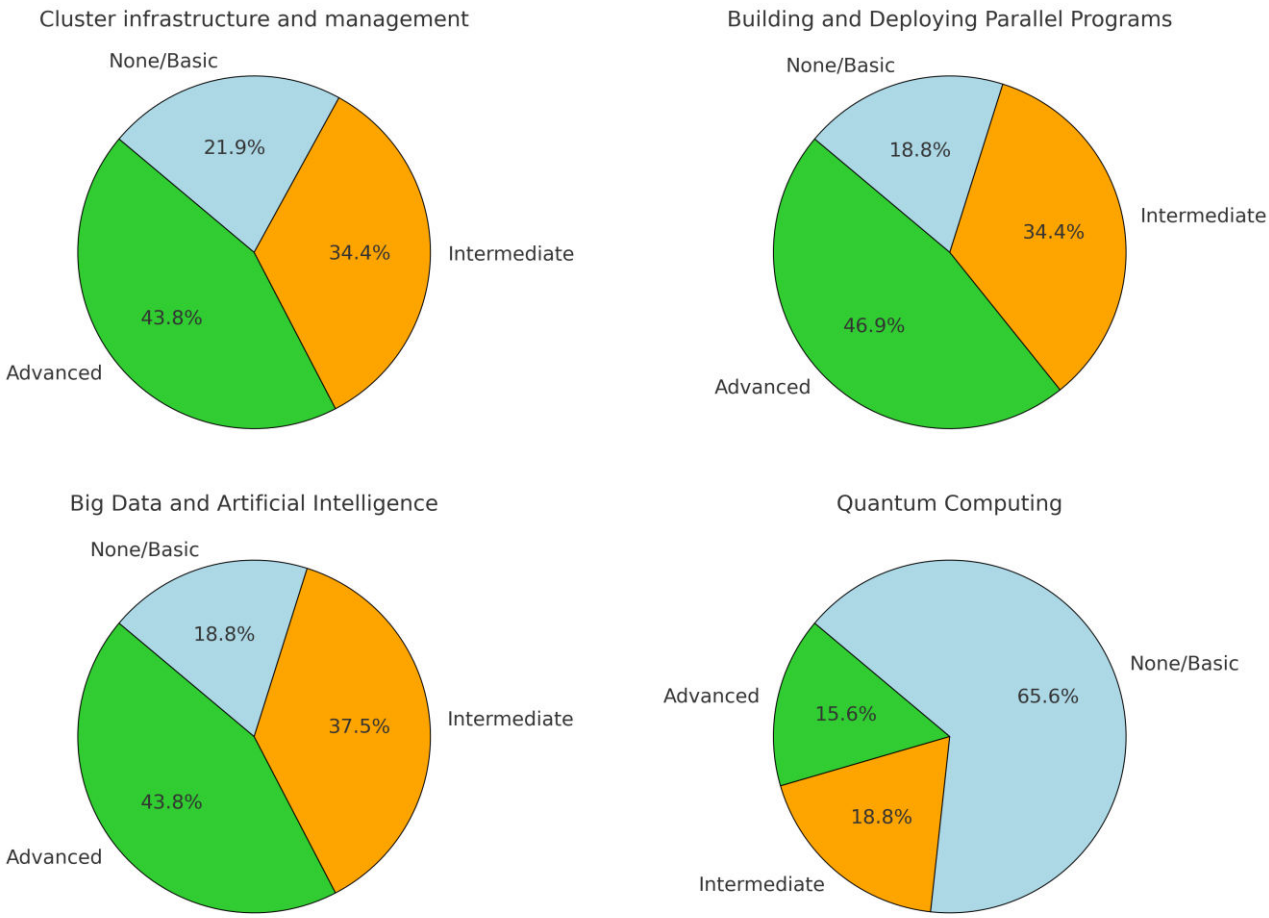


Figure 6: Overall level of knowledge of the NCCs in the four macro-categories of the updated Competence Map.

Figure 7 shows the classification of expertise in the six sub-categories of the "Cluster Infrastructure and Monitoring" macro-category. The sub-category "Principles of cluster components" (the following is the definition provided in the upgraded Competence Map: *defined as the knowledge of cluster infrastructure components refers to the various hardware elements that make up a computing cluster. Effective understanding of cluster components is essential for optimizing cluster performance, reliability, and resource utilization)*, is the category with the largest fraction of "Advanced" knowledge (71%), followed by system and monitoring tools (58%). Overall, the categories for which the sum on "Advanced" and "Intermediate" classes is higher are "Principle of cluster components" and "System and Monitoring Tools".

Figure 8 shows the classification of expertise in the six sub-categories of the "Building and Deploying Parallel Programs" macro-category. The percentages of "Advanced" knowledge are very high for "Principles of Parallel Programming" (defined as *knowledge of the basic concepts of HPC Programming: message passing systems, shared memory systems and general-purpose GPU programming*), for "Programming Languages and Frameworks in HPC" (defined as *knowledge of programming languages and frameworks currently used in HPC),* "Code Optimization" and "Compilers". Interestingly, the fraction of "Advanced" knowledge drops significantly in favour of "Intermediate" for the categories "Debugging Tools" and "Profiling Tools".

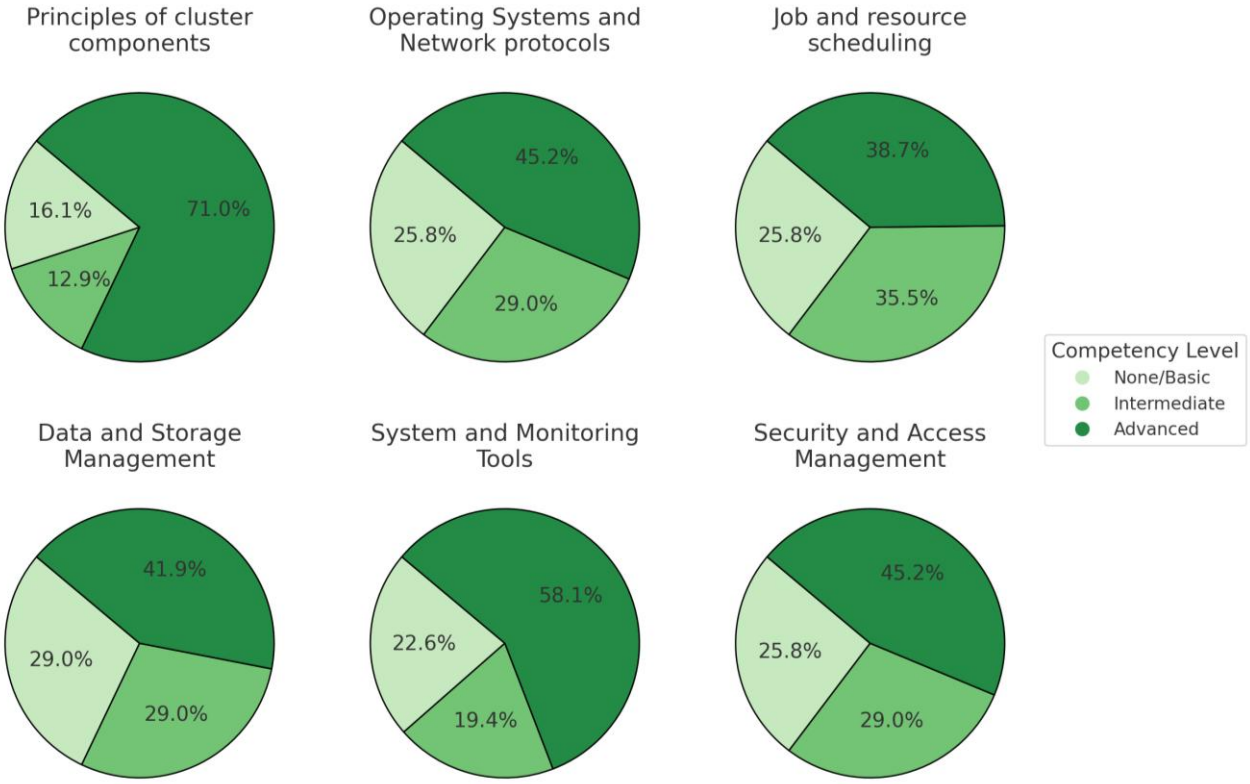


Figure 7: Fraction of NCCs with "Advanced", "Intermediate", "None/Basic" competences in the subcategories of the macro-category "Cluster Infrastructure and Management".

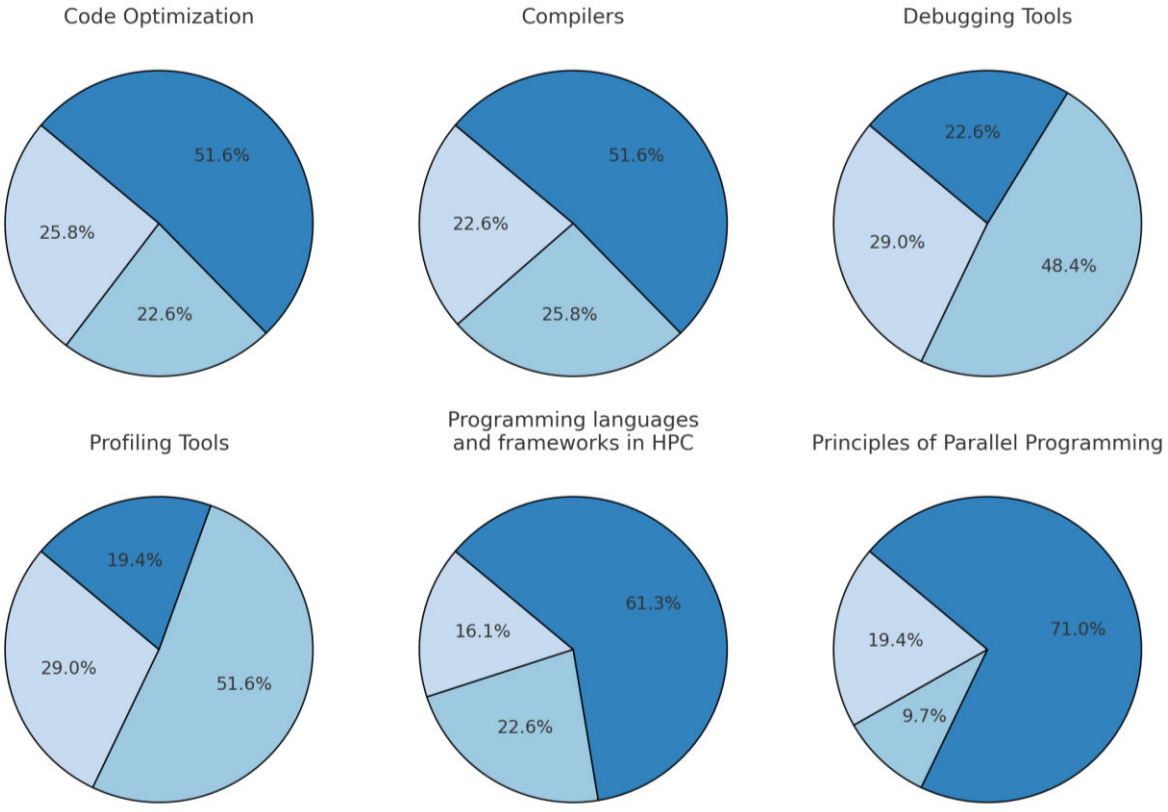


Figure 8: Fraction of NCCs with "Advanced", "Intermediate", "None/Basic" competences in the subcategories of the macro-category "Building and Deploying Parallel Programs".

Figure 9 shows the classification of expertise in the six sub-categories of the "Big Data and Artificial Intelligence" macro-category. The percentages of "Advanced" knowledge are very high for "Machine Learning and Deep Learning applications to solve specific real-life problems" and for Machine Learning and Deep Learning Frameworks". Interestingly, the fraction of "Advanced" knowledge is lower for sub-categories such as "Data Management", "Big Data Tools in HPC (Management, Analytics, Visualization) and "Data Storage" and significantly more so for "MLOps Tools".

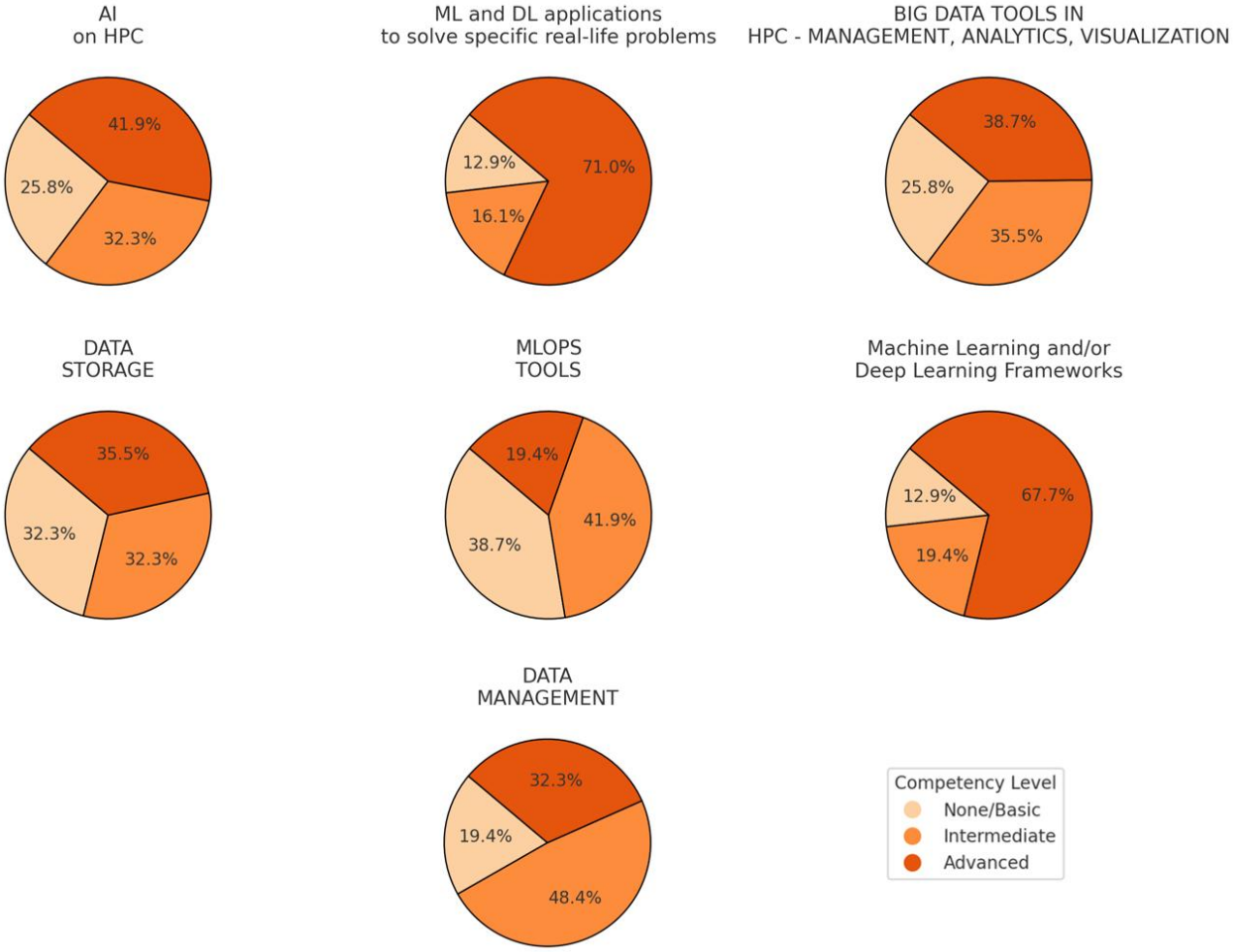


Figure 9: Fraction of NCCs with "Advanced", "Intermediate", "None/Basic" competences in the subcategories of the macro-category "Big Data and Artificial Intelligence".

Finally, Figure 10 shows the classification of expertise in the six sub-categories of the "Quantum Computing" macro-category. As expected, in these sub-categories the fraction of NCCs having "Advanced" knowledge is overall lower with respect to other macro-categories, the highest being the 25% observed for "Principles of Quantum Computing". For the other sub-categories, the fraction of NCCs with "Advanced" knowledge oscillates between 10% and 19%.

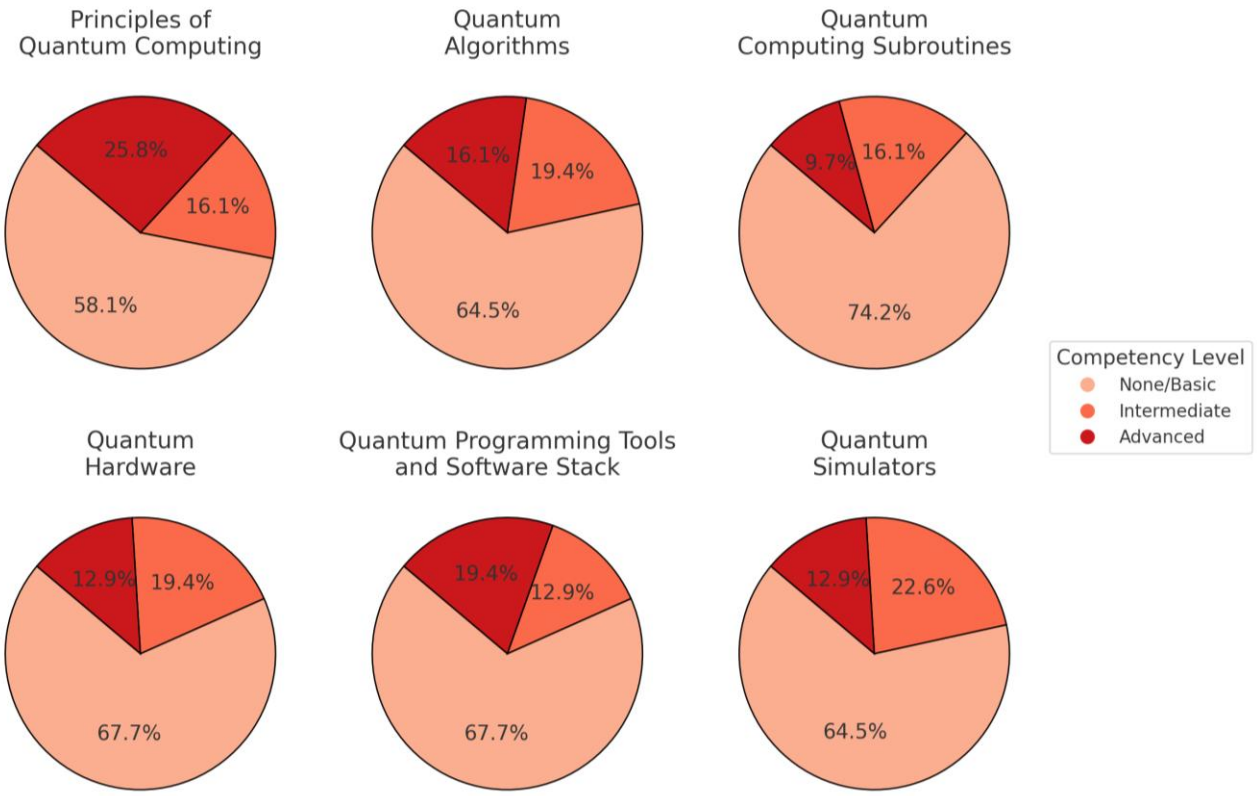


Figure 10: Fraction of NCCs with "Advanced", "Intermediate", "None/Basic" competences in the subcategories of the macro-category "Quantum Computing".

In conclusion, while acknowledging that these results are still preliminary as feedback from each NCC needs to be fully implemented, some important trends are already emerging from this mapping exercise:

- **Embracing AI and Parallel Programming**. NCCs have recognized the pivotal role of Artificial Intelligence in addressing the needs of their national ecosystems and as a powerful tool for attracting users to HPC infrastructure. Notably, 24 NCCs are classified as either Intermediate or Advanced in both "Building and Deploying Parallel Programs" and "Big Data and Artificial Intelligence." However, this represents only a partial picture, and as further feedback is integrated, it is likely that more NCCs will demonstrate deeper expertise in these two areas.
- **Identifying Gaps for Targeted Training**. Some sub-categories, such as "MLOps Tools" or "Profiling Tools," show significantly lower proportions of "Advanced" expertise compared to other sub-categories within the same macro-category. These topics represent opportunities for organizing targeted training activities to address existing gaps and strengthen NCCs' overall competencies.
- **Quantum Computing as an Emerging Area**. Quantum Computing is an emerging field of expertise among the NCCs. While still at an early stage, efforts are underway to build and strengthen capabilities in this area as interest and investment continue to grow.

2.7 Next Steps

At the time of writing this deliverable, the feedback loop with NCC Work Package Leaders (WPLs) and WP2 champions has commenced, and the upgraded Competence Map is expected to be finalized in early 2025. Following its finalization, the next steps are as follows:

- **Integration with C2ISS**: Incorporate the upgraded Competence Map into the new C2ISS platform once it is online.
- **Regular Updates on Individual-Level Competences**: Request updates on individual-level competences from NCC representatives every six months. This interval ensures that the Competence Map reflects any changes in NCCs' capabilities due to personnel transitions, such as individuals joining or leaving the project or their respective organizations.
- **Automatic Feedback Loop**: WP2 aims to develop a dedicated strategy within C2ISS to facilitate the feedback loop with WPLs and WP2 champions. This tool will automate and simplify the process of managing updates on NCC-level competences, ensuring a more streamlined and efficient approach.

In addition to the steps mentioned above, WP2 will also make available to all NCCs the structure of the questionnaire and the scripts developed for the analysis of the answers at the individual and organization level, so that they can reproduce this at their national level.

3 Fostering the collaboration between NCCs and CoEs

This chapter provides an overview of key activities undertaken by WP2 to strengthen collaboration and knowledge-sharing between NCCs and CoEs. Building on the success of the first project year, WP2 has continued its efforts to foster deeper connections, address identified needs and support the HPC ecosystem through targeted initiatives.

One of the central activities has been the **update of the Summary Document on CoEs** (Section 3.1), which now includes detailed information on all CoEs, including those newly funded ones in 2024. This document serves as an important resource, offering insights into the objectives, tools, and codes of each CoE, as well as links to training portals and collaboration opportunities. Additionally, WP2 developed the **"CoEs' Codes White Book"** (Section 3.2) a comprehensive, filterable table summarizing 72 codes and tools from CoEs. This resource not only facilitates collaboration but also lays the groundwork for future integration into the upcoming C2ISS platform.

WP2 also initiated the organization of **developer-oriented technical activities** (Section 3.3) based on feedback from CoEs, such as webinars and workshops on shared challenges like GPU programming. These events are thought to be a platform for CoEs to exchange knowledge and align on innovative approaches to technical challenges.

Additionally, WP2 organized the **2nd NCCs-CoEs online meeting** in November 2024, building on the success of the first meeting. This event allowed new CoEs and collaborative projects like EPICURE [20] to present their objectives and activities. With over 115 participants and focused breakout sessions on collaboration, training, and industry engagement, the meeting reinforced the strategic partnership between NCCs and CoEs while outlining future actions for CASTIEL2.

Finally, during the second year of the project, CASTIEL2 WP2 continued to interact with HPC3, as such interaction is considered a useful complement to the communication structures of CASTIEL2 with the CoEs. HPC3 was established with support from the CSA FocusCoE [21] (December 2018 – March 2022) in May 2019 and was continued by the then active CoEs after the end of FocusCoE. HPC3 is run as an inclusive forum by the CoEs in order to exchange information, coordinate activities and identify common viewpoints. HPC3 meets approximately on a monthly schedule.

3.1 Updates to the summary document on CoEs

In the first year of the project, several NCCs expressed a desire to have easy access to information on CoEs, such as objectives, tools, use cases, contacts, and more, to facilitate collaboration and awareness. WP2 recognized the need for a comprehensive summary on CoEs and decided to produce such a document gathering the following information for each CoE:

- **Coordinator** (name, country, and Principal Investigator)
- **Partners** (names and countries)
- **Project Duration**
- **Webpage Link**
- **Type** (I or II)
- **Funding Phase** (newly funded, second round, or third round)
- **Brief Objectives and Description**
- **Links to Service and Training Portals** (where available)
- **Industry Collaboration Potential**

- **Use Cases**
- **Codes/Workflows**

In the first part of the second year, WP2 updated the document to include the two new CoEs that started in January 2024, EoCoE and POP-3. The new updated document was shared publicly on EuroCC Access [22]. At the time of writing, WP2 is in the process of updating the summary to include dealii-x and MICROCARD-2. **Figure** 11 shows an example of the summary for EoCoE.

3.2 The document on the codes of the CoEs

The **"CoEs' Codes White Book"** is an essential resource developed by WP2 to consolidate and organize a comprehensive overview of the codes, tools, and libraries developed or utilized by the Centres of Excellence. This initiative builds upon the information collected during the preparation of D2.4 due at M12. By leveraging insights from all CoEs, WP2 has created a structured, filterable table designed to foster collaboration and knowledge-sharing across the network. As two new CoEs started in January 2024 while WP2 was already preparing the document, their codes/tools were added to the file before its first release in June 2024. Currently, the file includes, for each code:

- **Name of the code**
- **CoE developing / porting the code**
- **Category (e.g., tool, workflow, application domain)**
- **Short Description**
- **Main use**
- **Use cases**
- **URLs**
- **History / Versions**
- **Owner**
- **OSS**
- **Licensing Details**
- **Term of Use**
- **Contacts**

The total number of entries in the file is 72. Looking ahead, WP2 envisions integrating the White Book into the C2ISS (CASTIEL2 Information Sharing System), transforming it into a dynamic, interactive resource. In the meantime, while updating the file to include the codes of the CoEs that started in November 2024, WP2 is planning to produce an html file that allows to quickly browse through the different categories and filtering options.

Currently, the White Book is intended for internal use within the CoEs and NCCs network, promoting deeper collaboration. Additionally, WP2 is exploring the possibility of making the information included in this document publicly available in the future via the C2ISS, extending its impact and accessibility to the wider scientific and HPC communities.

01.01.2024 - 31.12.2026
Energy - Engineering
https://www.eocoe.eu/
3rd Phase

Coordinator: CEA F (PI: Edouard Audit)	
CEA F	FZJ DE
Barcelona Supercomputing Centre ES	CNRS F
INRIA F	CERFACS F
ENEA IT	E4 IT
MPG DE	CNR IT
FAU Erlangen DE	ULB BE
PSNC PL	DDN F
U. Roma Tor Vergata IT	IFPEN F
UNITN IT	CIEMAT ES

EoCoE III demonstrates the benefit of HPC for the net-zero energy transition for research institutes and also for key industries in the energy sector, channeling its efforts into 5 exascale lighthouse applications covering the key domains of Energy Materials, Water, Wind and Fusion.
The multidisciplinary effort of the consortium partners harnesses innovations in computer science and mathematical algorithms within a tightly integrated co-design approach to overcome performance bottlenecks, to deploy the lighthouse applications on the coming European exascale infrastructure, and to anticipate future HPC hardware developments.

- Service Portal: https://www.eocoe.eu/services/
- Software as a Service portal: https://eocoe.psnc.pl/
- Training events: https://www.eocoe.eu/
- Potential Collaboration with Industry: Yes

CODES

•Gysela-X: 5-D code enabling simulations of the nonlinear turbulent core-edge plasma region at unprecedented spatial and temporal resolution length scales
•libNEGF: Non-equilibrium Green's function library, combined with density functional theory (DFT) models of material properties to enable new advances in photovoltaics
•Parflow: Groundwater modelling framework, combined with the hydropower infrastructure management system HYPERstreamHS to provide high-resolution hydrological forecasting at the continental scale
•ALYA: High Performance Computational Mechanics code focusing on incompressible flow for Wind over complex terrain, including physics required by wind community
•waLBerla-wind : fluid dynamics solver to achieve multiscale offshore and flat terrain onshore wind farms modeling
•AGMG, PSCToolkit, GmgPolar, Maphys++, MUMPS : Exascale-ready Linear Algebra libraries
•Deisa: Library that ensures coupling MPI simulation codes with the Dask Framework

GRAND CHALLENGES / USE CASES

•Large-scale calculations of non-equilibrium charge carrier transport across the heterostructure of two-dimensional layers
•High Fidelity Microscale wind simulations and wind resource assessment, optimal complex terrain scale resolving simulations on GPUs
•Weak and strong scalings on the 4D Gysela axisymmetric version up to 100 GPUs, and up to 5000 GPUs with Gysela-X, i.e. the 5D non-axisymmetric version
•Integrated hydrologic and hydropower infrastructure information at the resolution of a reservoir, stream reach and neighborhood from the continental to the global scale
•Promotion of solutions and codes in the largest European Alliance on Energy research (EERA)

Figure 11: Example of the information collected in the "Summary Document on CoEs" produced by WP2 and shared with all NCCs and CoEs.

As the codes managed by the CoEs were already included in D2.1 and D2.4, here we list the codes/tools that are managed by EoCoE and POP-3 CoEs that started in January 2024 and thus were not part of D2.1 and D2.4. The codes of the two CoEs that started in November 2024 will be added in D2.5 (due at M25) and also in D2.3 (due at M36).

3.2.1 EoCoE

The codes/tools/workflows managed by EoCoE are:

- **GYSELA-X**. GYSELA-X is a global full-f nonlinear gyrokinetic code that simulates electrostatic plasma turbulence and transport in the core of Tokamak devices.
- **libNEGF**. Non-equilibrium Green's function library, combined with density functional theory (DFT) models of material properties to enable new advances in photovoltaics.
- **Parflow**. ParFlow is a parallel, integrated hydrology model that simulates spatially distributed surface and subsurface flow, as well as land surface processes including evapotranspiration and snow.
- **waLBerla-wind**. Fluid dynamics solver to achieve multiscale offshore and flat terrain onshore wind farms modelling.

3.2.2 PoP-3

The codes/tools/workflows managed by PoP-3 are:

- **Extrae**. Tools for performance monitoring, devoted to generate Paraver trace-files for a post-mortem analysis.
- **Paraver**. Tool for performance analysis, was developed to respond to the need to have a qualitative global perception of the application behaviour by visual inspection and then to be able to focus on the detailed quantitative analysis of the problems. Expressive power, flexibility and the capability of efficiently handling large traces are key features addressed in the design of Paraver.
- **Dimemas**. Dimemas is a performance analysis tool for message-passing programs. It enables the user to develop and tune parallel applications on a workstation, while providing an accurate prediction of their performance on the parallel target machine.
- **Scalasca**. Software tool that supports the performance optimization of parallel programs by measuring and analysing their runtime behaviour.
- **Cube**. Performance report explorer for Scalasca and Score-P, a generic tool for displaying a multi-dimensional performance space consisting of the dimensions (i) performance metric, (ii) call path, and (iii) system resource.
- **Score-P**. Tool suite for profiling and event tracing of HPC applications: works with Scalasca, Vampir, and Tau and is open for other tools.
- **MAQAO**. Performance analysis and optimization framework operating at binary level with a focus on core performance.
- **MERIC**. Lightweight C/C++ library with Fortran interface for HPC applications dynamic behaviour detection with a goal in energy consumption reduction - applying READEX approach.
- **MUST**. MPI Runtime Correctness Analysis: MUST detects usage errors of the Message Passing Interface (MPI) and reports them to the user.

3.3 The developers-oriented technical activities

Following the input received by the CoEs during the drafting of D2.4 "Legacy Code Report", WP2 started planning activities related to the organisation of technical, developer-oriented activities like webinars/workshops with the aim of sharing information among CoEs on, e.g., scalability of third-party libraries, porting on different GPUs. The interest in such activities was expressed by several CoEs independently. A technical collaboration table was set up – based on input gathered for D2.4 – and the feedback of CoEs solicited on interest in individual topics, in order to create a solid basis for subsequent collaboration activities. Based on the topics highlighted in Excel file that was used to collect the interest of the CoEs on technical collaborations, a first meeting centred on GPU programming was held in order to prioritise subtopics of interest and determine possible first activities. For each activity, interest was gauged among the CoEs, also to lead related activities or ability to contribute. In parallel to those activities, interactions with the EPICURE project started in order provide them with the gained insights on the technical topics most interesting for the CoE developers, to align technically oriented future activities and to create synergies, taking into account that CASTIEL2 itself is largely lacking the technical competences to drive the necessary activities in full detail and all necessary ramifications. Currently WP2 is organising a workshop or webinar on the Kokkos [23] framework. More details will be given in the upcoming deliverable D2.5, "Update on legacy software report", due at M25.

3.4 The "2nd NCCs-CoEs online meeting"

Following the success of the first NCCs-CoEs online meeting in the previous year, WP2 organized the **2nd NCCs-CoEs online meeting** on the 8th of November 2024 to continue fostering collaboration and strengthening ties between NCCs and CoEs. The timing of the meeting was chosen so that it would not be too close in time to the AHM in Slovakia in (April 2024) and to allow all four CoEs that started in 2024 (two in January and two in November) to attend and present their projects to all the NCCs. At the same time, WP2 decided that the meeting could be a great chance to introduce the EPICURE and MINERVA projects to all NCCs and CoEs, thus they were invited to give a presentation too. Indeed, both projects (that are European support centres for HPC and AI) are set to collaborate with NCCs and/or CoEs, and thus CASTIEL2 already started interacting with both to align objectives and set the strategy for the collaboration.

The afternoon of the meeting was focused on CASTIEL2, with updates from the PMT as well as from the other CASTIEL2's WPs. In the last hour, three parallel breakout sessions dedicate to CASTIEL2 WP2, WP3 and WP4 took place. These discussions enabled participants to share ideas, address challenges, and outline future actions.

The following was the agenda of the meeting:

- 9.30-9.40: WELCOME
- 9.40-10.00: EoCoE CoE
- 10.00-10.20: POP-3 CoE
- 10.20-10.40: dealii-X CoE
- 10.40-11.00: MICROCARD-2 CoE
- 11.00-11.15: BREAK
- 11.15-12.00: EPICURE
- 12.00-12.20: MINERVA
- 12.20-12.30: Q&A
- 12.30-13.45: LUNCH

- 13.45-14.00: CASTIEL2 PMT
- 14.00-15.00: CASTIEL2 WP2-WP3-WP4-WP5
- 15.00-15.15: BREAK
- 15.15-16.15: Break Out Sessions: "NCCs-CoEs Networking and Collaboration", "Training", "Interaction with Industry"
- 16.15-16.30: CLOSING

The meeting proved to be another step forward in fostering strategic partnerships between NCCs and CoEs, paving the way for innovative collaborations in the HPC ecosystem.

Overall, the meeting had a good turn-over, with a peak of more than 115 participants. A short questionnaire to track the participation to the meeting was shared during the meeting and 107 answers were collected. The fraction of respondents divided by project (NCCs, CoEs and CASTIEL2) is shown in Figure 12.

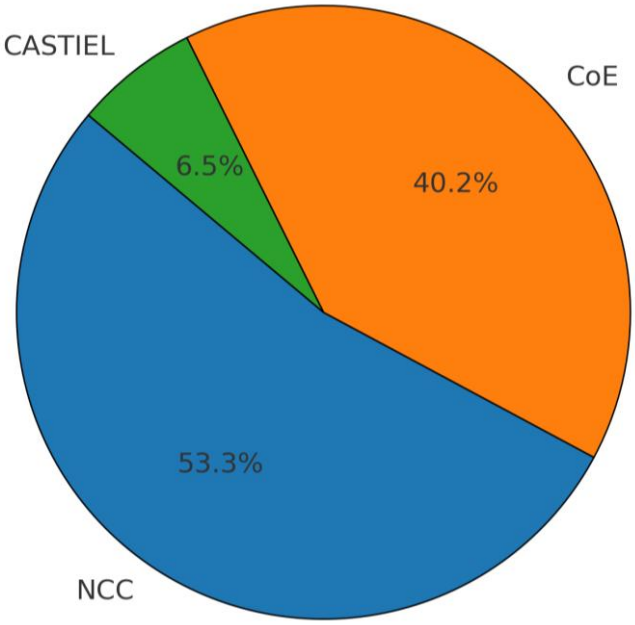


Figure 12: Participants to the "2nd NCCs-CoEs online meeting" divided by type of project.

4 Sharing, fostering and enriching the competences of NCCs and CoEs

One of the objectives of CASTIEL 2 WP2 is to define and run workshops on non-HPC specific relevant topics and competences (soft skills, complementary competences), based on the needs expressed by NCCs and CoEs, and make the materials available (sometimes internally to all NCCs and CoEs, sometimes to the wider ecosystem, depending on the webinar).

D2.1 reported the answers provided by NCCs and CoEs when asked on the topics they would have liked to have webinars organised on. As energy efficient computing was the second most-voted answers, WP2 decided to organise a series of six webinars focused on that topic. These webinars were open to everyone and recorded: slides and recordings are made available on the EuroCC Access website [24].

The following is the list of six webinars with the abstract, information on the speakers, their affiliation and the number of registered participants.

- Title: Introduction to Energy Efficient Computing and Thermal Control in today's HPC systems.

  Abstract: *Reducing power consumption and increasing energy efficiency has nowadays a key role in everyday utilization of HPC systems. Advanced and costly cooling solutions are needed to sustain the high computing densities of high-performance computing equipment. For this reason, detecting and predicting thermal anomalies is vital to assure safe working conditions and avoid performance degradation. In this webinar, the fundamental concepts of power consumption and thermal control in today's high-end supercomputers will be presented.*

  Speakers: Federico Tesser (CINECA) and Andrea Bartolini (University of Bologna).

  Registered participants: 92

  Date: 1st December 2023

- Title: Introduction to Energy and Performance Profiling Techniques in HPC Systems.

  Abstract: *In the first part of this webinar, Federico Tesser will focus on performance monitoring on a referenced architecture x86_64: why it is important, how the user can get information concerning the performances of the running software from the underlying hardware and how a naive Roofline Model can be used to interpret the results. In the second part, Ondřej Vysocký will focus on the energy consumption measurement of HPC applications. A list of available power monitoring systems and their differences will be shown. The participants will be introduced to practical aspects of getting joules consumed and how to make the results comparable from one hardware platform to another.*

  Speakers: Federico Tesser (CINECA) and Ondřej Vysocký (IT4I, NCC Czech Republic).

  Registered participants: 117

  Date: 12th January 2024

- Title: Introduction to low-level Processor Architectural Information on Power Saving Techniques.

Abstract: *To save power and stay within the thermal limits, hardware vendors implement several power-saving mechanisms. In the first part of this webinar, Robert Schöne (TUD) will describe power saving mechanisms and show how they are influenced by user codes on x86 systems. In addition, Robert will talk about developments in the Intel power monitoring interfaces that should be considered. In the second part, Thomas Ilsche (TUD) will explore various factors that affect idle power consumption, along with methodologies and tools for identifying and addressing issues. The presentation will follow a narrative approach, drawing insights gained during the acceptance phase of a 630-node CPU system.*

Speakers: Robert Schöne (TU Dresden) and Thomas Ilsche (TU Dresden).

Registered participants: 69

Date: 23rd February 2024

- Title: Tools for Power Monitoring; ExaMon & COUNTDOWN.

  Abstract: *This webinar will be composed of two parts, each of which related to a tool with a specific role in the power and energy software stack. The first tool is ExaMon (Exascale Monitoring), an operational data analytics framework targeting data-and AI driven data-centre automation. The talk will present the ExaMon key technology aspects and its applications when combined with artificial intelligence technologies. The second covered tool is COUNTDOWN, a run-time library for energy saving in MPI communication primitives. COUNTDOWN identifies and automatically reduces the power consumption of the computing elements during communication and synchronization primitives, filtering out phases which would detriment the time to solution of the application. During the talk, we will see how this is done transparently to the user, without touching his/her code nor needing recompilation of the application itself. And we will see also other characteristics of COUNTDOWN, like its capability to report energy and power information both in a distributed and aggregated fashion, and both in time series and final summary version. Moreover, we will see how COUNTDOWN can extract traces for a performance evaluation of the monitored applications.*

  Speakers: Andrea Bartolini (University of Bologna) and Federico Tesser (CINECA).

  Registered participants: 40

  Date: 15th March 2024

- Title: Energy Analysis and Optimization with EAR.

  Abstract: *This webinar will present the main features of the Energy Aware Runtime (EAR) software, specifically targeted for users of high-performance computing data-centres that use both HPC and AI applications. EAR is the energy management software, which is installed in several European Data Centres such as SuperMUC-NG (LRZ), Snellius (SURF) and now in MN5 (BSC). EAR offers services for system energy monitoring (including system and node power cap), job-level energy and performance monitoring, and energy optimization through a runtime library. We will show how to characterize and optimize applications that represent common CPU intensive, memory intensive, and GPU use cases. During the webinar we will show how to correlate application metrics with energy results, helping users to understand their applications. Some examples of EAR data visualization will be done using Graphana and EAR job analytics tools.*

  Speaker: Julita Corbalan (BSC).

Registered participants: 33

Date: 24th May 2024

- Title: Improving Energy Efficiency of Parallel Applications Using MERIC

  Abstract: *Complex parallel applications show different hardware requirements during their execution. This dynamic behaviour can be exploited for energy savings by changing the hardware power knobs to fit the configuration to the application's needs. The open-source runtime system MERIC is designed to minimize the energy consumption of the HPC infrastructure, executing a parallel application by dynamic tuning of a wide range of hardware power knobs. The idea of dynamic tuning comes from the Horizon 2020 project READEX, under which the development of the MERIC runtime system started. The MERIC and associated tools perform a detailed analysis of complex application behaviour, identify the optimal hardware settings concerning energy consumption and runtime, and provide dynamic tuning during the application runtime.*

  Speaker: Ondřej Vysocký (IT4I, NCC Czech Republic)

  Registered participants: 38

  Date: 15th June 2024

In addition to the Energy Efficient Computing series, WP2 was approached by ENCCS (NCC Sweden) to organise additional webinars that were considered of interest for the whole network:

- **Title: The EBRAINS Project: A Gateway to Collaborative Neuroscience**

  Abstract: *This webinar aims at introducing EBRAINS, a digital research infrastructure that is accelerating collaborative brain research across neuroscience, brain health, and brain-related technologies. The session will commence with an overview of EBRAINS by Prof. Petra Ritter (EBRAINS coordinator), followed by a detailed presentation on the platform's tools and datasets by Anastasiia Andriievska. As a member of the European Open Science Cloud association (EOSC), EBRAINS exemplifies open science, fostering collaboration on new research projects, infrastructure development, and sharing of data and knowledge across disciplines, institutions, and borders. The webinar aims to explore potential collaborations with interested NCCs and CoEs and demonstrate how EBRAINS' commitment to Responsible Research and Innovation practices can shape research in ethical ways that serve the public interest.*

  Speakers: Petra Ritter (Charité Berlin) & Anastasiia Andriievska (ENCCS)

  Registered participants: 104

  Date: 13th February 2024

- Title: **Colony OS: A Meta Operating System for HPC and Cloud**
  Abstract: *ColonyOS is a Meta-OS developed by RISE, designed to simplify and streamline utilization of heterogeneous computing infrastructures, including IoT, edge, cloud, and HPC. It is designed to operate as an overlay to interface with a wide variety of computing environments, enabling seamless integration and operation across different platforms. In ColonyOS, each service runs independently within its designated environment or grid, known as a colony. These colonies are provisioned with precisely the resources they need, and communication is confined within the colony, adhering to a strict zero-trust security protocol. Services in ColonyOS is composed of so-called*

executors that integrate with various underlying platforms based on a distributed microservice architecture. This enables creation of so-called compute continuums. ColonyOS enables development of portable workloads that are compatible both with HPC and cloud-based systems. This flexibility enables workloads to run on a HPC system using Slurm, or a Kubernetes cluster without any modifications. Using a Meta-filesystem (Colony FS) data can be transfer seamlessly between HPC, cloud, or a local development environment. Today, ColonyOS is used commercially by RockSigma AB to analyse seismic events in one of the world's largest underground mines. ColonyOS is also used by ENCCS to make it easier to use EuroHPC systems. During the presentation, Johan will demonstrate how to run a cross-platform workflow on a Kubernetes cluster running on the RISE ICE data centre, combined with two EuroHPC supercomputers (LUMI and Leonardo). Johan will also demonstrate Pollinator, a tool built on ColonyOS designed to simplify ML training. Finally, Johan will demonstrate a web-based ML application built on ColonyOS to analyse satellite images using the Leonardo supercomputer.
Speaker: Johan Kristiansson (RISE - ENCCS)
Register participants: 63
Date: 9th April 2024

Finally, the "Code of the Month" monthly webinars started during the first year of CASTIEL2 continued regularly in the second year of the project, as the codes on which CoEs are working represent an important entry point for collaboration between NCCs and CoEs. As co-organizer of these events with WP4 and WP5, WP2 contributed to the organization of the webinars by pre-selecting potentially interesting codes, by contacting the CoEs and asking for their availability to present their code in the proposed dates. Additionally, WP2 shared the information about the webinars to the WP2 NCCs and CoEs champions and deputies. Every month, for each code two webinars were organised: an internal one for NCCs and CoEs and a public one for anyone interested among the end-users and networks of each NCC/CoE. These are the events that took place in the second year of CASTIEL2:

- Code of the Month Vol. 6. January 2024. *Neko* by CEEC CoE.
- Code of the Month Vol. 7. March 2024. *GROMACS* by BioExcel CoE.
- Code of the Month Vol. 8 April/May 2024. *FALL3D and OpenPDAC* by CHEESE.
- Code of the Month Vol. 9. July 2024. HiDALGO CoE.
- Code of the Month Vol. 10. September 2024. EoCoE CoE.
- Code of the Month Vol. 11. October 2024. *FLEXI/GALÆXI* by CEEC.
- Code of the Month Vol. 12. December 2024. *SIESTA* by MaX.

The public webinars were recorded, and the recordings made publicly available [24]. Detailed information on the "Code of the Month" webinars can be found in Deliverable D4.2 of CASTIEL 2's WP4 "2nd year Report on the Actions performed to support the NCCs and the CoEs' Interactions with Industry".

5 Major Achievements

The second year of CASTIEL 2 WP2 was marked by significant advancements in leveraging collected information, fostering collaboration, and streamlining tools for the NCCs and CoEs. The major achievements are summarized below:

- **Strategy for Expertise and Competence Mapping**. WP2 developed a straightforward and effective strategy to maximize the use of collected information from NCCs. This strategy guided the creation of both the **Experts Database** and the upgraded **Competence Map** of NCCs. These tools aim to enhance collaboration by providing clear, accessible, and actionable insights into the expertise available within the network.
- **Release of the New Experts Database**. Following the collection of expertise information from individuals in NCCs, WP2 prepared the new Experts Database, which includes only those who consented to share their details for internal use. This database was designed for ease of use and shared internally via BSCW in September 2024. To improve accessibility, WP2 created an accompanying HTML interface using Python and JavaScript. This file allows users to filter experts by macro-categories and sub-categories, displaying their names, emails, and affiliated NCCs.
- **Updated Competence Map of NCCs**. WP2 continued its work on upgrading the Competence Map, integrating the newly envisioned strategy for better categorization and usability. Currently in the process of implementing feedback from the NCCs, this upgraded map provides a comprehensive view of NCC-level competences, emphasizing practical examples and links to success stories or training events where specific competences have been applied.
- **"2nd NCCs-CoEs Online Meeting"**. WP2 organized the second iteration of this meeting in November 2024. The event successfully introduced the four new CoEs (EoCoE, POP-3, dealii-X, and MICROCARD-2) to the NCCs and provided updates from CASTIEL2 PMT and other work packages. Additionally, projects like EPICURE and MINERVA, which are set to collaborate with NCCs and CoEs, presented their objectives and planned/ongoing activities.
- **Development of the "CoEs' Codes White Book"**. WP2 finalized the first version of the White Book in June 2024, which now includes 72 codes and tools developed or managed by CoEs. The document was enriched with information from new CoEs launched in January 2024 and is set to be updated with additional codes from CoEs that started in November 2024.
- **Support for Developer-Oriented Activities**. Building on input from the CoEs, WP2 organized developer-focused activities, including a meeting on GPU programming and a planned workshop on Kokkos. These efforts aim to address technical challenges and enhance cross-CoE collaboration on advanced topics such as scalability and porting across architectures.
- **Ongoing Webinar Series**. WP2 maintained its support for the "Code of the Month" webinar series, which highlights codes developed by CoEs. Additionally, WP2 organized webinars on topics such as energy-efficient computing and supported the organization of webinars considered of interest for the whole network.

6 Concluding Remarks

At the time of writing this deliverable, WP2 has achieved several key milestones that establish a strong foundation for continued collaboration and progress. Building on the progress achieved in the second year of the CASTIEL 2 project, the feedback loop with EuroCC2 WPLs and WP2 champions has been initiated, and the first version of the upgraded Competence Map is scheduled for completion in early 2025. The upcoming steps include:

- **Integration with C2ISS**. Once the upgraded Competence Map is finalized, WP2 will incorporate it into the new C2ISS platform, making it a dynamic and accessible resource for all stakeholders.
- **Regular Updates on Competences**. WP2 plans to request updates on individual-level competences from NCC representatives every six months. This regular interval will help ensure the Competence Map remains up to date, reflecting changes in NCC capabilities due to new hires, departures, or evolving expertise.
- **Extension of the Experts Database to include CoEs**. WP2 plans to share the questionnaire that was used to create the Experts Database of EuroCC2 to CoEs as well. On a voluntary basis, we will give the possibility to individuals involved in CoEs to be added to the Experts Database.
- **Automated Feedback Management**. WP2 aims to develop a dedicated tool within C2ISS to automate and streamline the feedback loop with WPLs and WP2 champions. This tool will simplify the process of updating NCC-level competences, enhancing efficiency and consistency.
- **Continue to support technical collaboration of CoEs.** WP2 aims at maximising the synergizes with EPICURE as much as possible to foster technical cooperation between CoEs and experts on topics selected by the CoEs to be of particular interest.
- **Resources for National Implementation**. WP2 will also make available to all NCCs the structure of the questionnaires and analysis scripts used for mapping competences at individual and organizational levels. This initiative empowers NCCs to reproduce similar assessments at the national level, fostering localized knowledge-sharing and strategy development.

In addition to these planned steps, WP2 will continue to focus on fostering collaboration between NCCs, CoEs, and the broader European HPC ecosystem. By leveraging tools like the Experts Database, the CoEs' Codes White Book, and the upgraded Competence Map, WP2 will further strengthen the network, promote knowledge-sharing, and support the overall goals of CASTIEL 2. These efforts ensure that WP2 remains at the forefront of advancing collaboration and competence development across the HPC community.

7 References and Applicable Documents

[1] https://eurohpc-ju.europa.eu/eurocc-2-and-castiel-2-promoting-hpc-boost-digital-skills-jobs-and-industrial-competitiveness-europe-2023-02-03_en.
[2] https://www.hpccoe.eu/hpc-coe-council/.
[3] https://www.eurocc-access.eu/about-us/the-projects/.
[4] https://bioexcel.eu/.
[5] https://ceec-coe.eu/.
[6] https://cheese-coe.eu/.
[7] https://www.excellerat.eu/.
[8] https://www.esiwace.eu/.
[9] http://www.hidalgo2.eu/.
[10] http://www.max-centre.eu/.
[11] https://www.multixscale.eu/.
[12] https://plasma-pepsc.eu/.
[13] https://www.space-coe.eu.
[14] https://www.eocoe.eu/.
[15] https://pop-coe.eu/.
[16] https://eurohpc-ju.europa.eu/research-innovation/our-projects/dealii-x_en.
[17] https://eurohpc-ju.europa.eu/research-innovation/our-projects/microcard-2_en.
[18] https://www.eurocc-access.eu/.
[19] https://eurohpc-ju.europa.eu/index_en.
[20] https://epicure-hpc.eu/.
[21] https://www.hpccoe.eu/focus-coe/.
[22] https://www.eurocc-access.eu/wp-content/uploads/2024/11/CoEs_Summary_Sept24.pdf.
[23] https://kokkos.org/.
[24] https://www.eurocc-access.eu/services/video-library/.

8 Annex: The Upgraded Competence Map

The following are the categories of the Upgraded Competence Map with a description of each macro-category and first level of sub-category.

1. CLUSTER INFRASTRUCTURE & MANAGEMENT

Knowledge in cluster infrastructure & management refers to a deep understanding of the principles and components involved in designing, deploying, and managing high-performance computing clusters. It includes familiarity with hardware components and their functioning, network architecture, storage types and characteristics. Cluster management involves managing users, resources and their allocation, as well as software. It also includes activities for ensuring the safety of the cluster as well as the monitoring of the correct functioning of the system. This knowledge is vital for optimizing resource utilization, scalability, power efficiency, and fault tolerance within parallel and distributed computing environments.

- **Principles of Cluster Components.** *Knowledge of cluster infrastructure components refers to the various hardware elements that make up a computing cluster. Effective understanding of cluster components is essential for optimizing cluster performance, reliability, and resource utilization.*
    - Node Components (RAM, CPU, GPU)
    - Storage Systems (High Performance Storage Hardware, Object Storage, Parallel Filesystem, RAID, Backups, Data Transfer)
    - Power
    - Network (Network Hardware (switches, routers etc.), Network Topologies, Network Protocols, Nftables, NTP, High Speed Interconnect, Bandwidth Management)
    - Interconnect
    - Cooling Systems (air, water, …)

- **Virtualization.** *Virtualization in HPC refers to the use of virtualization technologies to create and manage virtual environments on HPC clusters and knowledge of the hardware components that make it possible.*
    - Networking
    - Image Management
    - Hardware Passthrough
    - QEMU
    - VMWare
    - Libvir
    - OpenStack

- **Cluster Security.** *Security for cluster management in an HPC environment is a critical concern to protect the integrity of the cluster, its resources, and sensitive data.*
    - Best Practices
    - Verifying Software Authenticity and integrity
    - Network Security
    - Attack Monitoring
    - Data Encryption and Secure Data Transfer
    - User Authentication and authorization

- **Cluster Monitoring.** *Cluster monitoring is a key component of cluster management. It involves the continuous tracking and analysis via specific tools, of various aspects of a cluster's performance, resource utilization, and health.*
    - Icinga
    - Prometheus
    - Telegraf
    - Grafana
    - Kubernetes

- **Software Management.** *Software management involves the effective installation, configuration, and maintenance of the software components. Properly managing software on cluster nodes ensures that applications and services run smoothly, vulnerabilities are addressed, and resources are utilized efficiently.*
    - Operating System Package Installation
    - Virtual Environments
    - Environment Modules
    - Building from Source
    - Containerization (e.g., Docker, Docker swarm, Kubernetes, Openshift/OKD, Rancher)
    - SPACK
    - EasyBuild

- **Resources and Users Management.** *Managing resources and users in HPC environments requires a well-defined set of policies, effective job scheduling, user support, and security measures to ensure the efficient and secure operation of the cluster. Proper management practices help maximize resource utilization while maintaining a balance between fair access for users and compliance with cluster policies and regulations.*
    - User Support
    - Job Scheduling Systems (e.g., SLURM, Torque, PBS)
    - Authentication Method
    - Users Management (e.g., LDAP, Kerberos...)

2. BIG DATA AND ARTIFICIAL INTELLIGENCE

HPC is essential for extracting meaningful information from the analysis of large volumes of data through tools and methodologies from Data Science and/or Artificial Intelligence. Competence in this field involves the capability to effectively manage, process, and analyze vast and diverse datasets. This includes knowledge in data collection, storage, and processing, along with the ability to ensure data security and privacy. Proficiency in distributed computing and scalability is crucial for handling growing data volumes and delivering real-time insights. Artificial Intelligence is a powerful tool for extracting insights from data. Competences in AI encompass the expertise to design, train, and deploy machine learning and deep learning models for a wide range of domains, exploiting MLops and frameworks and, when needed, multiple GPUs for training and inference.

- **Data Management.** *Knowledge of data security measures and regulatory compliance requirements to protect sensitive data and ensure legal adherence. Implementing data*

backup and disaster recovery strategies to prevent data loss and ensure business continuity. Awareness of data privacy and ethical considerations, particularly in handling sensitive and personal data.
    - Data Privacy and Ethics
    - Data Security and Compliance
    - Data Backup and Recovery

- **Data Storage.** *Competences in data storage involve the knowledge and skills required to design, implement, and manage efficient and scalable data storage solutions. This includes selecting appropriate storage technologies and configuring storage systems to meet specific data requirements.*
    - SQL Databases
    - NoSQL Databases
    - Object Storage
    - Columnar Databases
    - Data Lakes
    - Data Warehouses

- **Big Data Tools in HPC (Management, Analytics, Visualization).** *Distinguish the benefits and drawbacks of various big data tools in the HPC environment. Apply a data science workflow on existing data using various big data tools. Construct simple data science workflows.*
    - Hadoop
    - Apache Spark
    - Apache Fink
    - Apache Cassandra
    - Dask
    - Apache Kafka
    - Apache Storm
    - Apache Hive

- **MLops tools.** *Competences in MLops tools include proficiency in using them for orchestrating workflows, automating model training and deployment, managing version control, ensuring model scalability, and monitoring model performance and data pipelines.*
    - MLflow
    - KubeFlow
    - Terraform
    - Apache Airflow
    - Jenkins
    - Prometheus
    - Grafana

- **ML & DL Frameworks.** *Competences in ML and DL frameworks involve the knowledge for creating and optimizing models exploiting existing tools for solving real-world problems.*
    - Scikit-learn
    - Tensorflow

  - Pytorch
  - Keras
  - Caffe
  - MXNet
  - Horovod
  - CNTK
  - Theano
  - Chainer
  - PaddlePaddle

- **Application Domain.** *Include the knowledge to solve specific problems exploiting different types of AI models.*
  - Classification
  - Regression
  - Clustering
  - Dimensionality Reduction
  - Anomaly Detection
  - Super resolution
  - Time Series Analysis
  - Natural Language Processing
  - Recommendation systems
  - Computer Vision
  - Autonomous Systems
  - Reinforcement Learning

- **AI on HPC.** *Knowledge of exploiting High Performance Computing environments to train or deploy AI models, e.g., by doing distributed training of a NN.*
  - Data Parallelization
  - Pipeline Parallelization
  - Tensor Parallelization
  - Sequence Parallelization
  - DeepSpeed
  - Accelerate
  - Axolotl
  - Ray
  - Pytorch Distributed
  - Tensorflow Distributed

3. BUILDING AND DEPLOYING PARALLEL PROGRAMS

Competence in building and deploying parallel programs refers to the skill and knowledge required to create and run software applications that can efficiently utilize multiple processing units or cores in a computing environment. This competence encompasses various aspects, including programming languages, parallel programming frameworks and techniques, profiling and debugging tools, and tools for version control and continuous integration / continuous deployment.

- **Principles of parallel programming.** *Knowledge of the basic concepts of HPC Programming.*
  - Message Passing Systems
  - Shared Memory Systems
  - General Purpose GPU Programming

- **Programming Languages and Frameworks in HPC.** *Knowledge of programming languages and frameworks currently used in HPC.*
  - C/C++
  - Python
  - Java
  - Fortran
  - R
  - Matlab
  - Perl
  - Ruby
  - Julia
  - OpenACC
  - Chapel
  - OpenMP
  - OpenCL
  - CUDA
  - MPI

- **Profiling Tools.** *Knowledge of tools for analyzing the performance and behavior of applications running on high-performance clusters. Includes tools for: Profiling and tracing of MPI and multithreaded, tool to trace intel-compiled apps, Detailed profiling for intel applications, General purpose tracing tool, Memory debugging, Profiling and tracing, PAPI, Tracing, I/O, Traces of GPU-accelerated applications*
  - gprof
  - Perf
  - Intel VTUNE profiler
  - ITAC
  - TAU
  - HPC Toolkit
  - Scalasca
  - Arm MAP
  - Score-P
  - Vampir
  - Darshan
  - TAURUS
  - NVIDIA Nsight System
  - Lo2s
  - CUPTI
  - Extrae/Paraver

- **Debugging Tools.** *Knowledge of tools for identifying and resolving issues in parallel and distributed applications running on HPC.*

  - GDB
  - TotalView
  - Arm DDT
  - Parallelware Analyzer
  - PADB
  - ReMPI
  - X86db
  - Compu
  - Cuda-gdb

- **Compilers**
  - GNU (gcc, gfortran,…)
  - Intel (icc (icx), icpc (icpx), ifort (ifx))*
  - IBM (xlf, xlc, xlC)
  - HPC-SDK (Nvidia)
  - LLVM (Clang)

- **Code Optimization**
  - Level1-Level2 optimization**:** local optimization, compromise between compilation speed, optimization, code accuracy and executable size (usually default)
  - Level 3 optimization: high optimization, can alter the semantics of the program
  - Level 4 optimization: Aggressive optimization, depending on hardware
  - Vectorization
  - Register allocation
  - Static vs Dynamic Allocation
  - Dead and redundant code removal
  - Common subexpression elimination (CSE)
  - Strength reduction (e.g., replacing an exponentiation within a loop with a multiplication)
  - Inlining
  - Loop optimizations such as index reordering, loop pipelining, unrolling, fusion, splitting, re-rolling, peeling, versioning, reversal, interchange
  - Cache blocking
  - Optimizing I/O performance

QUANTUM COMPUTING

Competences in quantum computing encompass the knowledge and skills necessary to understand, develop, and harness the power of quantum computers. They include the physical basis of quantum physics, quantum programming tools and software stack, quantum algorithms and simulators and quantum, computing subroutines.

- **Basics of Quantum Computing.** *Knowledge of the foundations of Quantum Computing*
  - Reversibility, DiVincenzo criteria
  - Qubits, Quantum Gates, Universal Gate Set

  - Circuit design, notation, matrix representation
  - Basic quantum programming techniques
  - Complexity Theory, Quantum Complexity Classes, Computational Limitations, Quantum Advantage
  - Universal fault-tolerant quantum computers, NISQ quantum computers

- **Quantum Hardware.** *Knowledge of the basic components of a Quantum Computer and their functioning***.**
  - Superconducting Electronic Circuits
  - Spin Based systems
  - Neutral atoms and ions
  - Photonic systems
  - Quantum State Control
  - Hybrid Quantum Systems

- **Quantum Programming Tools & Software Stack**
  - Graphical platforms
  - Quantum assembler languages and software development kits, Quantum Circuit simulator
  - Quantum Compilers & High-Level programming with pre-defined subroutines
  - Hybrid Quantum-Classical Algorithms and quantum embedding
  - Cloud Platforms
  - Quantum Error correction, quantum error mitigation

- **Quantum Algorithms**
  - Number theory and factorization (e.g., Shor Algorithm)
  - Oracular algorithms and database search (e.g., Grover algorithm)
  - Linear Algebra (e.g., Harrow-Hassidim-Lloyd algorithm)
  - Quantum Optimisation
  - Quantum ML & Quantum Neural Networks
  - Quantum simulation algorithms
  - Noisy intermediate-scale quantum (NISQ) algorithms: Variational Quantum Eigensolver (VQE), Quantum Approximate Optimisation Algorithm (QAOA)

- **Quantum Simulators**
  - Digital Quantum Simulators
  - Analogue Quantum Simulators and (adiabatic) quantum annealers

- **Quantum Computing Subroutines**
  - Quantum Amplitude Amplification
  - Quantum Fourier Transform (QFT), hidden subgroups finding
  - Quantum, phase estimation
  - Quantum linear algebra subroutines, quantum singular value decomposition
  - Other techniques and subroutines for, e.g., quantum, walks, amplitude estimation etc.)

DOMAIN SPECIFIC SOFTWARE

Domain-specific software in HPC is designed to address the specific needs and requirements of particular scientific, engineering, or research domains.

- Abinit
- Alya
- Amber
- Ansys Fluent
- AVBP
- BHAC
- Big-DFT
- BIT1-BIT3
- CHANGA/GASOLINE
- CODA
- EC-Earth4
- ExaHyPE
- ESPResSO
- ELMER/ICE
- FALL3D
- FDO4ESM
- FIL
- FLEXI
- FLEUR
- FLEW
- GENE
- Gysela-X
- GROMACS
- HADDOCK3
- HySEA
- ICON
- IFS
- LaMEM
- LAMMPS
- linNEGF
- m-AIA
- NAMD
- NEMO
- Nek5000
- Neko
- NekRS
- OpenFOAM
- Open GADGET
- openPDAC
- PIConGPU
- PLUTO
- PMX
- pTatin3D
- Quantum Espresso
- RAMSES
- RES – Renewable Energy Sources

- SeisSol
- Siesta
- SPECFEM3D
- Tandem
- VASP
- Vlasiator
- waLBerla
- waLBerla-wind
- xSHELLS
- YAMBO