EURO Greece

Training Needs for HPC, HPDA, AI, HPAI in Greece

2022 Report

Table of Contents

List of Figures

List of Tables

Official Title: Training Needs for HPC, HPDA, AI, HPAI in Greece: 2022 Report

Authors and Contributors
Xenia Ziouvelou, NCSR Demokritos (GR)
Stelios Karozis, NCSR Demokritos (GR)
Vangelis Karkaletsis, NCSR Demokritos (GR)

Statement of originality
This report contains original unpublished work, except where clearly indicated otherwise. Acknowledgement of previously published material and of the work of others has been made through appropriate citation, quotation, or both.

Disclaimer

Copyright notice

Note
For anyone interested in having more information about the current study, please contact us at: xeniaziouvelou@iit.demokritos.gr
For anyone interested in having more information about the EuroCC@Greece project, please contact us at: contact@eurocc-greece.gr

Reference
Ziouvelou X., Karozis S., and Karkaletsis V., (2022), "Training Needs for HPC, HPDA, AI, HPAI in Greece: 2022 Report". EuroCC@Greece Report 2022.

ABSTRACT

HPC constitutes a strategic resource in the digital decade dominated by an increasing number of data-intensive applications and services and an ever increasing need to address our macroscopic and microscopic challenges for the benefit of citizens, businesses, researchers and public administrations around the world. Thus, increasing the usability of HPC, and bridging the gap in HPC, High Performance Data Analytics (HPDA) and Artificial Intelligence (AI), aligned with the needs of the relevant stakeholders is of key importance for Europe and the aim of the Euro CC project.

As part of this project, the Greek National HPC Competence Center (EuroCC@Greece) aims to create knowledge, skills and provide support for the Greek ecosystem to enable academic and industrial researchers and HPC users to take advantage of HPC and the broader European supercomputing ecosystem. Special emphasis is given to the HPC training needs in Greece. As such the first HPC Training Needs Analysis (TNA) in Greece was conducted, aiming to provide information about the HPC landscape in Greek academia, industry and public sector, and at the same time identify the Greek HPC training needs as those were depicted by the survey respondents during 2021.

The survey results provide us with an initial understanding of the distinct needs of the Greek users (multi-stakeholder and cross sectorial) as well as those that want to become HPC users. Our aim is to assess these training needs on a yearly basis in order to create training portfolios for the coming years, enhancing this way the adoption of High-Performance Computing in Greece.

1 Background

High-Performance Computing (HPC) uses supercomputers and computer clusters with extremely high computational power, to solve advanced computation problems, has grown significantly over the past three decades into a mature technology that acts as a key tool for science and industry.

HPC is a key tool for processing and analysing the constantly growing volume of data, from 64.2 zettabytes in 2020 to an expected 180 zettabytes in 2025 (1 zettabyte is equal to 1 trillion gigabytes) (Statista, 2022). As such, HPC has a large number of application areas that range from climate change, monitoring and mitigating planning to the production of safer and greener vehicles and treating COVID-19 pandemic to the advancement of knowledge in almost every scientific field and industrial domain.

Consequently, HPC constitutes a strategic resource in the digital decade dominated by an increasing number of data-intensive applications and services and an ever increasing need to address our macroscopic and microscopic challenges for the benefit of citizens, businesses, researchers, and public administrations around the world. Thus, HPC is one of the key digital domains for Europe and a strategic investment priority, playing a significant role in Europe's path towards recovery[1], with an EU investment expected to significantly increase in the next Multiannual Financial Framework (2021-2027)[2].

1.1 The EuroCC project

The EuroCC project[3] has established and operates National High Performance Computing (HPC) Competence Centres (NCCs), in the EuroHPC Joint Undertaking (JU) Participating States, that will connect to a network of hubs to provide access to leading-edge HPC technology and knowledge. The aim of EuroCC is to bridge the gaps in HPC, High Performance Data Analytics (HPDA) and Artificial Intelligence (AI), aligned with the needs of the relevant stakeholders and with level of maturity of each country, in order to increase the usability of these technologies in the different countries and thus provide a European excellence baseline.

1.1.1 The EuroCC@Greece: The National HPC Competence Center in Greece

The EuroCC@Greece is one of 33 national nodes of the EuroCC project, which started during September 2020, built in the framework of the European High Performance Computing Joint Undertaking (EuroHPC JU). The aim of EuroCC is to advance competitiveness in research, improve effectiveness of government services and promote innovation in industry.

[1] https://ec.europa.eu/commission/presscorner/detail/en/ip_20_940
[2] https://digital-strategy.ec.europa.eu/en/activities/funding-digital
[3] The EuroCC project has received funding from the European High-Performance Computing Joint Undertaking (JU) under grant agreement No 951732. The JU receives support from the European Union's Horizon 2020 research and innovation programme and Germany, Bulgaria, Austria, Croatia, Cyprus, the Czech Republic, Denmark, Estonia, Finland, Greece, Hungary, Ireland, Italy, Lithuania, Latvia, Poland, Portugal, Romania, Slovenia, Spain, Sweden, the United Kingdom, France, the Netherlands, Belgium, Luxembourg, Slovakia, Norway, Switzerland, Turkey, Republic of North Macedonia, Iceland, Montenegro.

EuroCC@Greece aims to create knowledge, skills and provide support for the Greek ecosystem to enable academic and industrial researchers and HPC users to take advantage of High-Performance Computing (HPC) and the broader European supercomputing ecosystem.

As such, in order to develop and support the creation of a highly competitive and innovative High-Performance Computing (HPC) ecosystem in Greece, EuroCC@Greece will place special emphasis on the skills development and deliver trainings activities, in order to address the different user needs, enabling them to scale up their HPC workloads and enhance their HPC competencies.

1.1 HPC Training Mapping and Training Needs Analysis

Aligned with the EuroCC@Greece aim, it is important to assess the Greek ecosystem of HPC current and future users, and their familiarity with HPC and at the same time to identify, understand and articulate the distinct training HPC needs of the Greek users as well as those that want to become HPC users. As such a survey study (Training Needs Analysis (TNA)) was designed and implemented, during 2021, and its findings of this first study of the Greek HPC ecosystem and training needs mapping, will identify needs, discrepancies, or gaps between current skills and skills required for the effective HPC implementation across various stakeholder segments with specific HPC needs. These training needs will thus facilitate the development of specialized training portfolios for the coming years, aiming to enhance the skills and the use of High-Performance Computing in Greece.

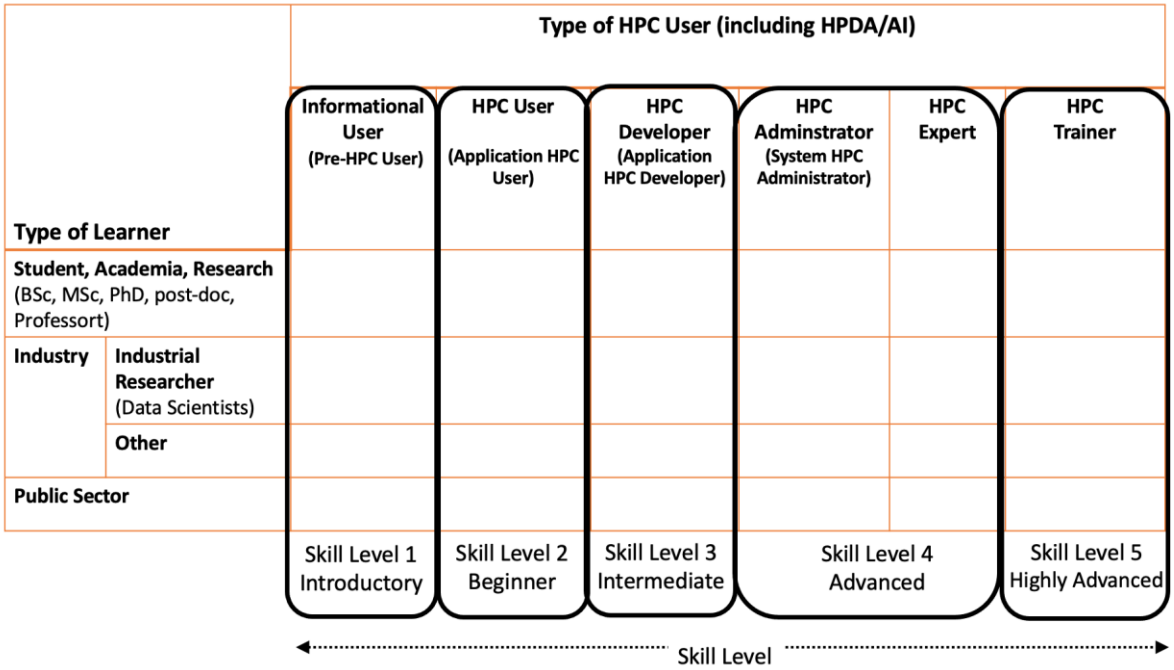

Figure 1: An HPC Training Mapping Framework

In order to capture these needs, we have developed a training mapping framework (see Figure 1) that depicts in the horizontal axes, the different *types of HPC Users and their associated skill level,* into 5 key categories:

- **Level 1**: Informational (pre-HPC) type of HPC user - Introductory skill level
- **Level 2:** HPC User (application user) - Beginner skill level
- **Level 3:** HPC Developer (application developer) - Intermediate skill level

- **Level 4:** HPC Expert type of user and HPC Administrator type of user- Advanced skill level
- **Level 5:** HPC Trainer - Highly Advanced skill level

In the vertical axes, the *types of learner stakeholder segments,* into 4 categories:

- **Academia/Research:** including both students and the researchers (PhD and post-doctoral researchers), professors
- **Industry**: **(a) Industrial Researcher:** Data Scientists and research software engineers from industrial communities (research positions only) and **(b) Other non-research industrial positions** in small and medium sized enterprises (SMEs) and large companies
- **Public Sector:** including all public sector organizations

This framework has facilitated both the creation of the survey and will also enable us to depict our study findings in order to understand the HPC training needs of the Greek users ("demand" side of the HPC training).

This report is structured as follows. Section 2 presents the study methodology. The study findings are provided in Section 3, detailing the HPC user profiles (Section 3.1), the HPC familiarity and usage (Section 3.2), the HPC training needs (Section 3.3) for industry (Section 3.3.1), academia (Section 3.3.2) and public sector (Section 3.3.3). Finally, this section concludes with the user suggestions (Section 3.4). The final section of this report presents the conclusions and key points of this study.

2 Methodology

The research methodology for identifying the HPC training needs in Greece, involved an online questionnaire survey. The questionnaire design involved four thematic sections: (1) generic information (individual information and background), (2) HPC Familiarity and Usage, (3) HPC Training Needs, and (4) Suggestions (see Annex 1).

Overall, nine questions were included, using a Likert-type scale (5 or 7 type scale) and response drop-down boxes or selection of single responses by ticking boxes. Open ended responses were included in specific questions aiming to enable participants to provide additional information and express their views.

To facilitate the dissemination of the survey, an online questionnaire was created utlising as a survey platform Google Forms, respecting all ethical and privacy considerations (informed consent, privacy and confidentiality, right to withdrawal) (Regmi et al., 2016).

2.1 Sample and procedure

The study started in October 2021 and ended in December 2021, covering a period of three months. The sampling process involved the EuroCC@Greece partners networks as well as online dissemination of the survey in order to ensure a random sample of respondents.

The survey asked people from academia, industry, and public sector to submit their answers and their training needs in relation to HPC, HPDA, AI. The participation was adequate, and more than 80 replies were received. In the next sections, the analysis of the profile, HPC usage and familiarity, as well as training needs are presented.

3 Study Findings

3.1 HPC User Profile

The elements that comprise the profile of the Greek HPC user, that participated in our study, are the position the stakeholder currently have, alongside the scientific and/or expertise domain. It is found that the stakeholder segment with the highest number of respondents was from ***academia and research*** (professors, researchers, students)[4] with a total of 74%. In particular, students accounted for 34.5%, post-doctoral researchers for 21% and university professors for 18.5% of the sample (see Figure 2). Private sector represented 16%[5] of the respondents whereas the Public sector accounted for 6.2%.

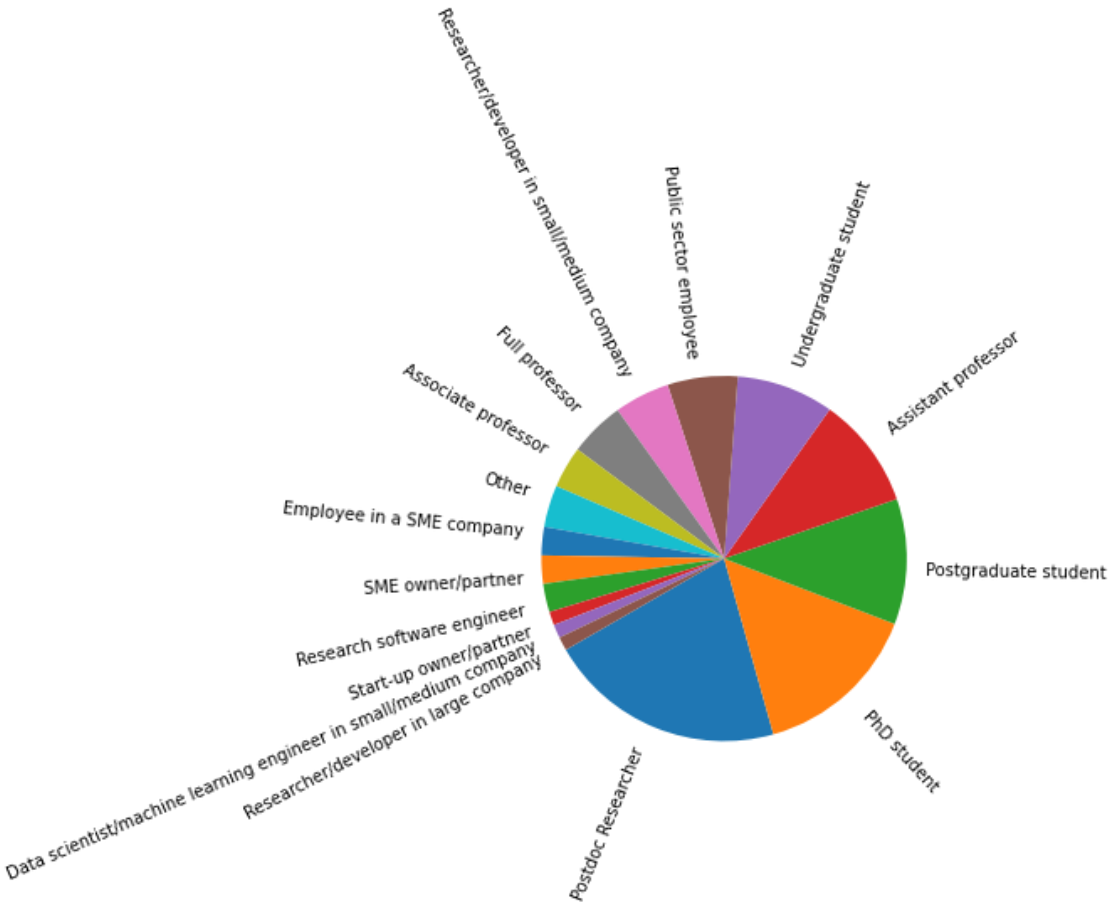

Figure 2: Respondents stakeholder segment [N=81]

[4] Students: PhD students (14.8%), postgraduate student (11.1%), undergraduate student (8.6%)
Researchers: post-doctoral researcher (21%)
University Professors: Assistant professor (9.9%), Associate professor (3.7 %), Full professor (4.9%)

[5] Research software engineers (2.5%), Researcher/developer in an SME (4.9%), Researcher/developer in a large company (1.2%), Data scientist in an SME (1.2%), Employee in an SME (2.5%), Start-up owner/partner (1.2%), SME owner/partner (2.5%)

Table 1*:* Percentages of user profiles that correspond to pie plot of Figure 2

Position	Percentage
Postdoc Researcher	21.0%
PhD student	14.8%
Postgraduate student	11.1%
Assistant professor	9.9%
Undergraduate student	8.6%
Public sector employee	6.2%
Researcher/developer in small/medium company	4.9%
Full professor	4.9%
Associate professor	3.7%
Other	3.7%
Employee in a SME company	2.5%
SME owner/partner	2.5%
Research software engineer	2.5%
Start-up owner/partner	1.2%
Data scientist/machine learning engineer in small/medium company	1.2%
Researcher/developer in large company	1.2%

3.1.1 Scientific Domain

In terms of scientific domain, an aggregation of different answers (see Figure 3 & Table 2) to groups with common expertise was performed. Two new groups were created:

- Natural Sciences: Physical Sciences, Chemical Sciences, Earth and Related Environmental Sciences, Biological Sciences
- Engineer & Tech: Chemical Engineering, Mechanical Engineering, Materials Engineering, Electrical Engineering, Electronic Engineering, Information Engineering

As such, the Engineer & Tech had the highest share accounting for 40.6%**,** followed by Natural sciences (26%) and computer and information sciences (25.9%). Other categories are under 10% each, due to the lack of participants or to the kind of research that does not utilize HPC services, so far.

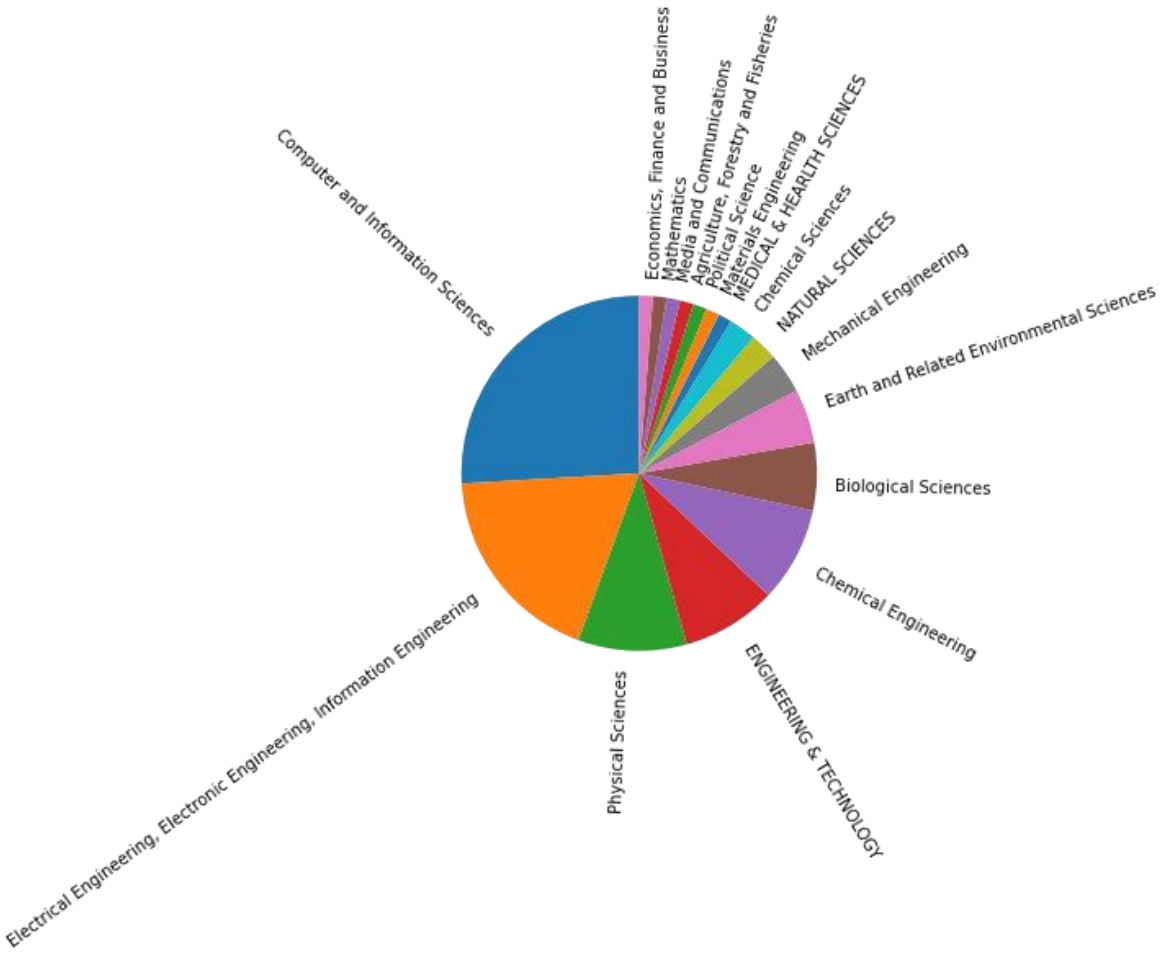

Figure 3: Scientific Domain [N=81]

Table 2*:* Percentages of scientific domain and expertise that correspond to pie plot of Figure 3

Scientific domain	Percentage
Computer and Information Sciences	25.9%
Electrical Engineering, Electronic Engineering, Information Engineering	18.5%
Physical Sciences	9.9%
ENGINEERING & TECHNOLOGY	8.6%
Chemical Engineering	8.6%
Biological Sciences	6.2%
Earth and Related Environmental Sciences	4.9%
Mechanical Engineering	3.7%
NATURAL SCIENCES	2.5%

Chemical Sciences	2.5%
MEDICAL & HEALTH SCIENCES	1.2%
Materials Engineering	1.2%
Political Science	1.2%
Agriculture, Forestry and Fisheries	1.2%
Media and Communications	1.2%
Mathematics	1.2%
Economics, Finance and Business	1.2%

3.2 HPC Familiarity and Usage

Having identified that the profile of our survey participants comes mainly from Academia, in the current section, we assess the respondent's familiarity of HPC infrastructure and/or usage. As it can be seen in Figure 3, below, the vast majority of our respondents appear to have basic information (Level 2: Beginner HPC user level, see the training mapping framework in Figure 1) accounting for 37% of the respondents.

Users with intermediate familiarity with HPC (Level 3: Intermediate HPC user level), represented 21% of our respondents, followed by non-familiar users that accounted in total for 16.1 %, comprising users that are either not familiar (6.2%) or not familiar but interested in learning more (9.9%). Advanced (Level 4: HPC Expert, HPC Administrator) and highly advanced (Level 5: HPC Trainer) user segments account for 8.6% and 7.4 % accordingly.

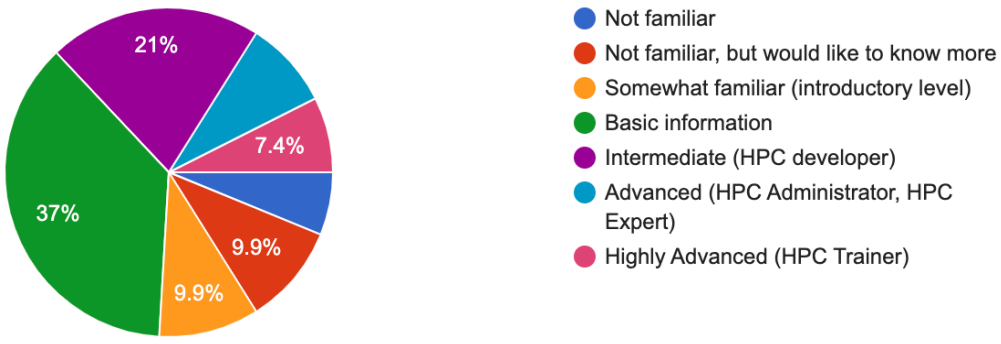

Figure 4: Pie chart of user familiarity with HPC [N=81]

The sections that follow provide an analysis of our findings depicting HPC familiarity based on the type of user (section 3.2.1), the fields that are using HPC infrastructure or services (section 3.2.2), and the participants skill level (section 3.2.3).

3.2.1 HPC Familiarity by Type of User

In order to analyze further the HPC familiarity of users we present a series of bar plots, per type of user (i.e., user's position/occupation), with x-axis describing the level familiarity and y-axis the percentage of sample accordingly.

The results, presented in Figure 5, show a progressive familiarity level for **academia** positions (see plots A-G, Figure 5), from "not familiar, but would like to know" being dominant among Postgraduate students, to "Advanced" and "Highly Advanced" for professors of all levels.

On the other hand, the diffusion of HPC usage in the **industry** has been found to be low, according to the sample, as in the respective plots (see plots I-M, Figure 4) most answers appear to show "Somewhat familiar" with HPC for industrial stakeholders. As it is expected, software engineering companies are an exception in this area.

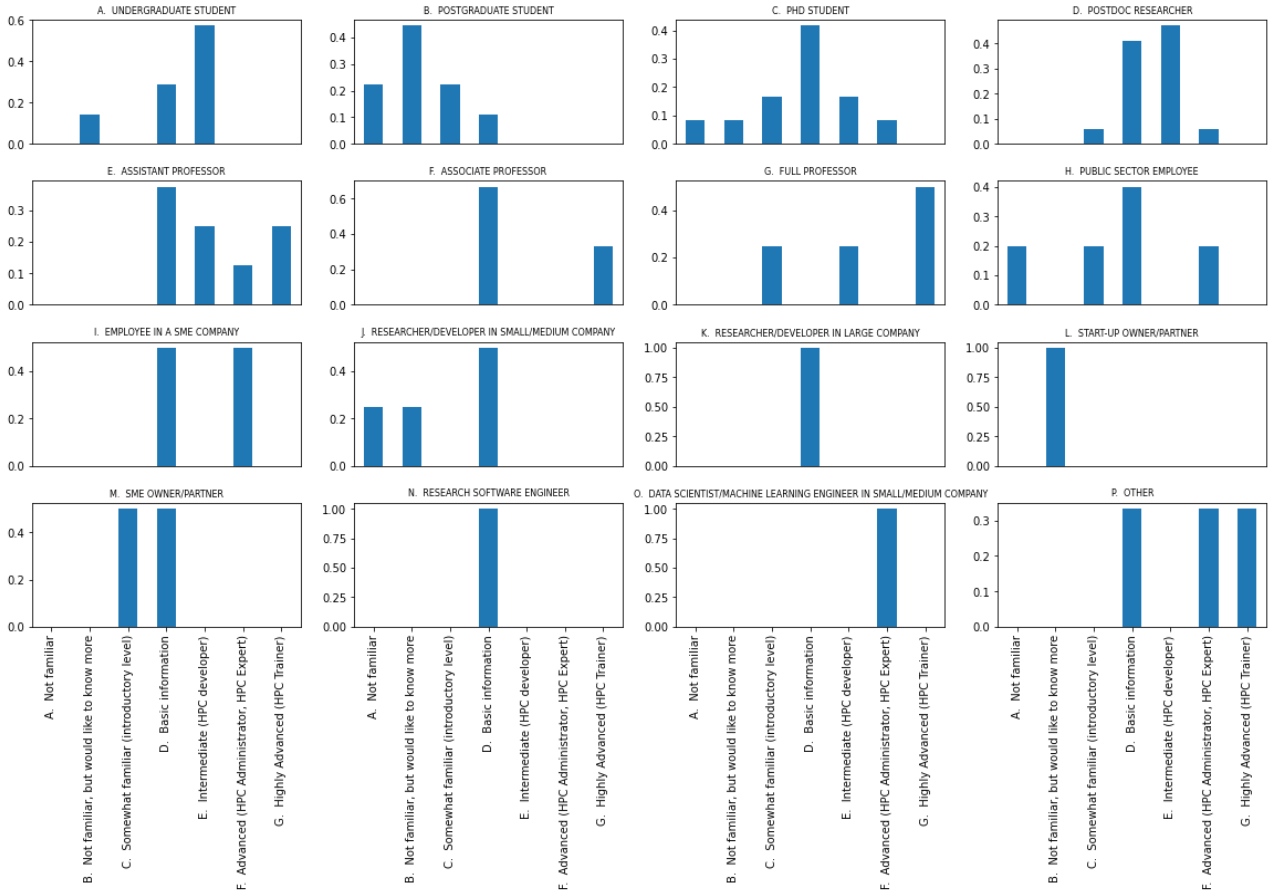

Figure 5: Bar plots familiarity of HPC over type of user [N=81]. The individual x-axis describes the level familiarity and y-axis the percentage of sample per each position.

The HPC familiarity in the **public sector** (see plot H, Figure 4) appears to be high compared to the composition of the public sector[6]. The latter may be considered to be an outlier, if we take into account

[6] Public sector typically includes representatives mainly from administration, ministries, IRS, ministries, etc., where the HPC is not prioritised in the daily activities. As such the basic and intermediate familiarity with HPC that is depicted in our sample appears to be not expected and may be attributed to the fact that the few public

that the sample of survey participants were probably linked with the national HPC infrastructure of Greece or in its network.

The category **Other** (see plot P) is difficult to be assigned or explained, but it is found to be above "Intermediate" familiar with HPC.

In order to assess the familiarity of different sectors of interest, an aggregation of the user types was performed according to the following:

- **A. Academia/Research:** A. Undergraduate student, B. Postgraduate student, C. PhD student, D. Postdoc Researcher, E. Assistant, professor, F. Associate professor G. Full professor
- **B. Industry (non-industrial positions):** I. Employee in a SME company, J. Researcher/developer in small/medium company, K. Researcher/developer in large company, L. Start-up owner/partner, M. SME owner/partner
- **C. S/W Eng, DataScience (Industrial Researchers):** N. Research software engineer, O. Data scientist/machine learning engineer in small/medium company

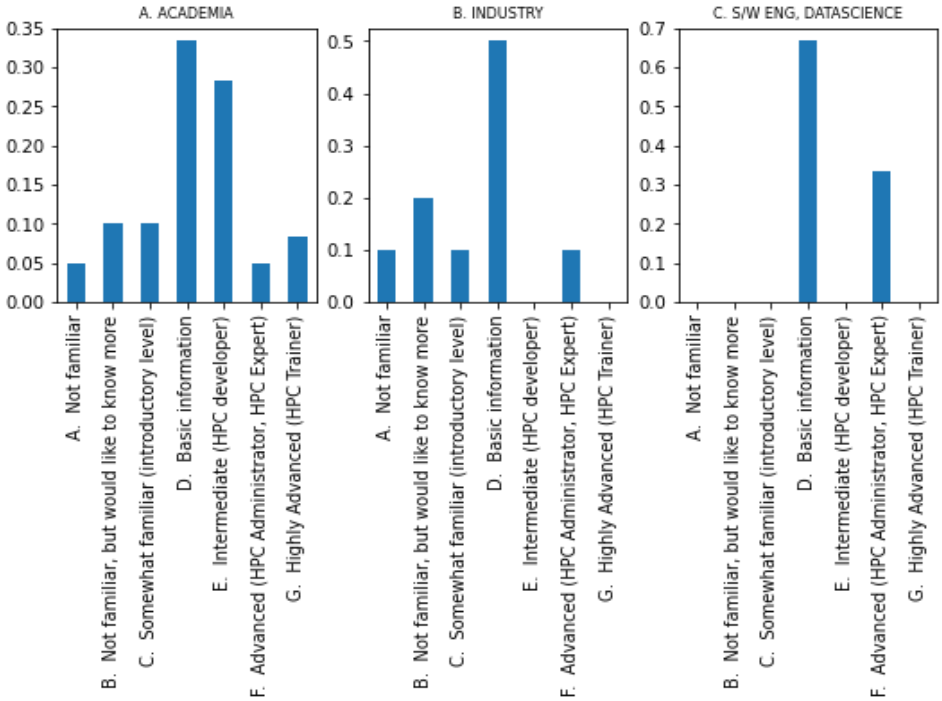

Figure 6: Bar plots familiarity of HPC over type of user [N=81]. The individual x-axis describes the level familiarity and y-axis the percentage of sample per each position.

Figure 6 presents the aggregated bar plots. In the case of **academia/research** a normal distribution of the levels of familiarity exists. Whereas, in **industry (non-industrial positions)** a positive skewed distribution depicts the low diffusion of HPC services in the sector. However, when the **S/W engineers and Data Scientist (industrial researchers)** are considered, which constitutes a category of their own, we can see that they are either advanced users or have basic information for the subject due to their background. These findings validate our classification framework (see section 1.1), since if we considered Industry as a single category, the industrial research community would have been masked inside this category providing us with ambiguous results.

sector representatives in our sample are most probably from people directly or indirectly related with Greek HPC services of ARIS.

3.2.2 HPC Familiarity by Domain/Field

Having analysed our findings based on the type of user, we now present the analysis that was performed to pinpoint the fields that are using HPC infrastructure or services. The fields are grouped as described in Section 3.1 and the result is presented in Figure 7 as heatmap representation.

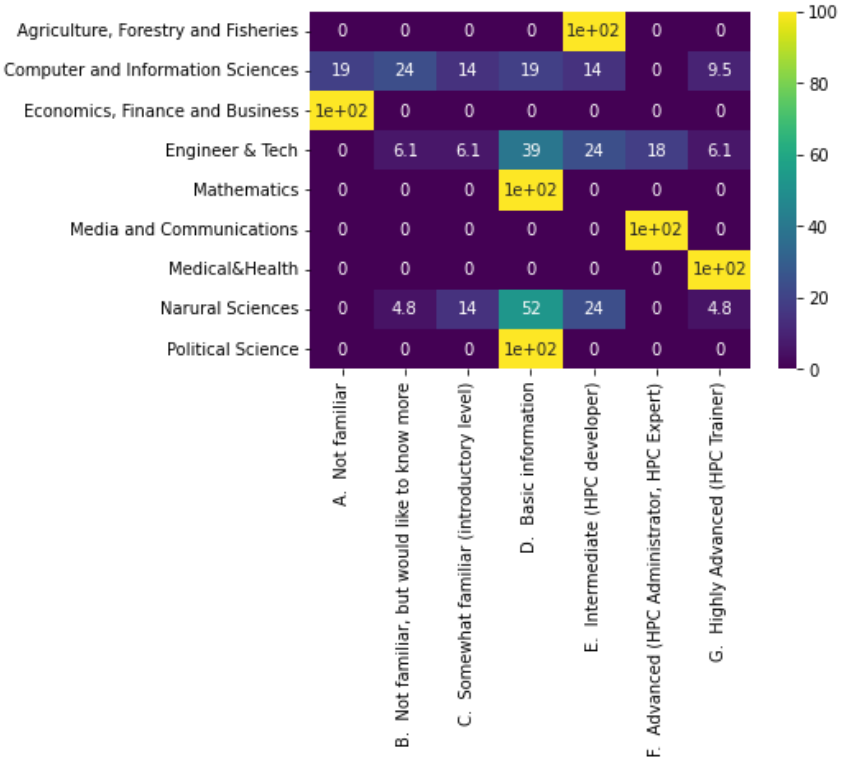

Figure 7: Map of field of grouped expertise[7] per level of familiarity [N=81]. The x-axis is the level of familiarity and y-axis is the field under study. The coloring goes from dark purple (0%) to yellow (100%) depending on the counts in the sample.

The Engineer & Technology (E&T), Natural Sciences (NS) and Computer and information Sciences (CIS) fields present a range of results, with E&T and NS to be more familiar with HPC compared to CIS (~60% of CIS is Somewhat familiar or not familiar with HPC). The latter may be due to the nature of problems under study of engineers and natural scientists. The computational power needed by them is directly correlated with the size and complexity of the problem under hand. As such, HPC infrastructure is essential for their work and subsequently they are mostly informed and intermediate familiar with HPC. More specifically, almost all the E&T and NS answers (87% and 81%) have at least a basic understanding of HPC. Interestingly, NS presents a narrow distribution of answers around the basic information, whereas E&T results consist of answers from HPC developers, experts, and trainers too.

On the other hand, other fields consist of single answers and no trend or conclusion may be derived

[7] Groups compared to original categories
Natural Sciences: [NATURAL SCIENCES, Physical Sciences, Chemical Sciences, Earth and Related Environmental Sciences, Biological Sciences]
Engineer & Tech: [ENGINEERING & TECHNOLOGY, Chemical Engineering, Mechanical Engineering, Materials Engineering, Electrical Engineering, Electronic Engineering, Information Engineering]

from the survey.

3.2.3 HPC Familiarity by Skill Type

The current section presents the HPC familiarity distribution across the different domains and associated areas of expertise. As seen in Figure 8, where the percentages of HPC familiarity in categories are presented as a heatmap, the participants that are "Not familiar" or "Not familiar but would like to know more" come in their vast majority from the **Computer and Information Science sector** (80%, 62%).

On the other hand, **Engineer & Tech** has the highest distribution for "Basic information" (43%), "Intermediate" (47%), "Advanced" (86%) and "Highly Advanced" familiarity (33%). These areas of expertise comprise the largest percentage of participants (40.6%, see Section 3.1.1), indicating that the most experienced and familiar users are among them. **Natural Sciences** present a similar distribution with participants belonging to this expertise being the second most familiar in the majority of the familiarity answers.

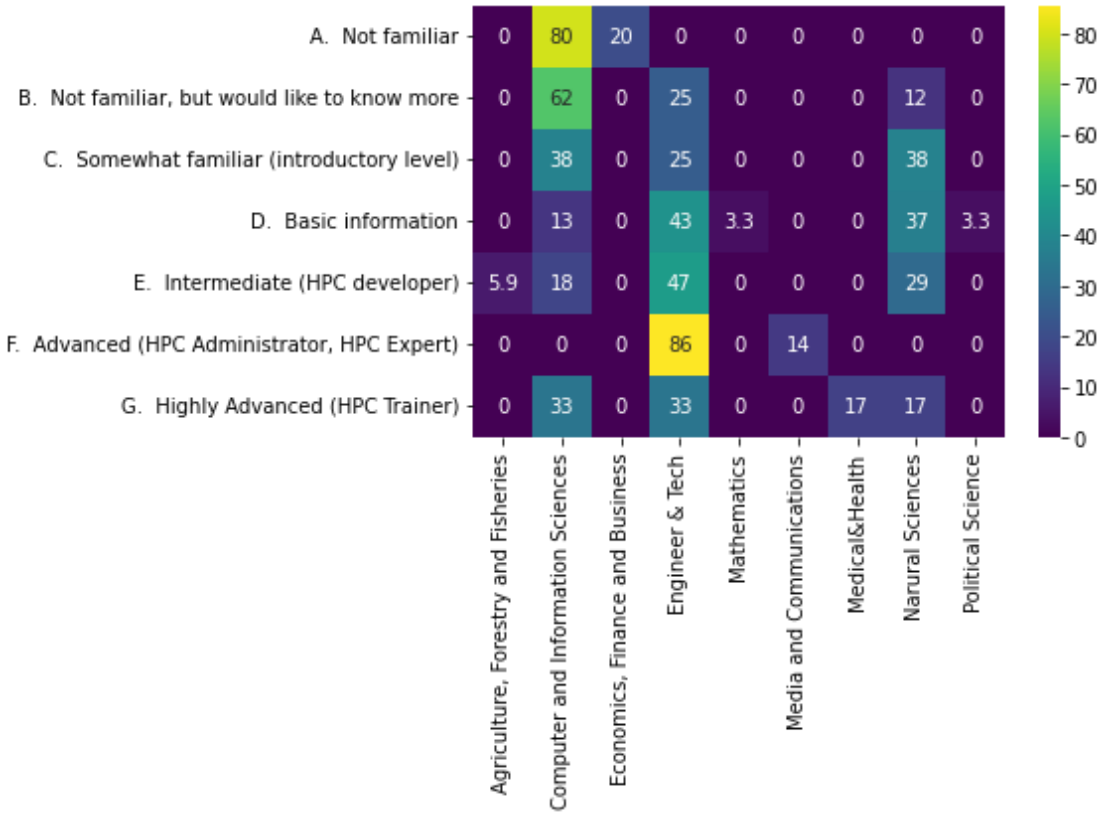

Figure 8: Map of level of skill level per field of grouped expertise[8] [N=81]. The x-axis is the field and the y-axis the skill level. The coloring goes from dark purple (0%) to yellow (100%) depending on the counts in

[8] *Groups compared to original categories*
Natural Sciences: [NATURAL SCIENCES, Physical Sciences, Chemical Sciences, Earth and Related Environmental Sciences, Biological Sciences]
Engineer & Tech: [ENGINEERING & TECHNOLOGY, Chemical Engineering, Mechanical Engineering, Materials Engineering, Electrical Engineering, Electronic Engineering, Information Engineering]

the sample.

3.3 HPC Training Needs

Aligned with the dual scope of this survey study, from the one side (1) to assess the profile, field and familiarity of the users in Greece and from the other side (2) to estimate the training needs of the academia, industry and miscellaneous ecosystem, we now focus on the latter aspect. Hence, a series of training categories were offered to the participants of the questionnaire to choose. The results are presented in Figure 9 as a heatmap plot.

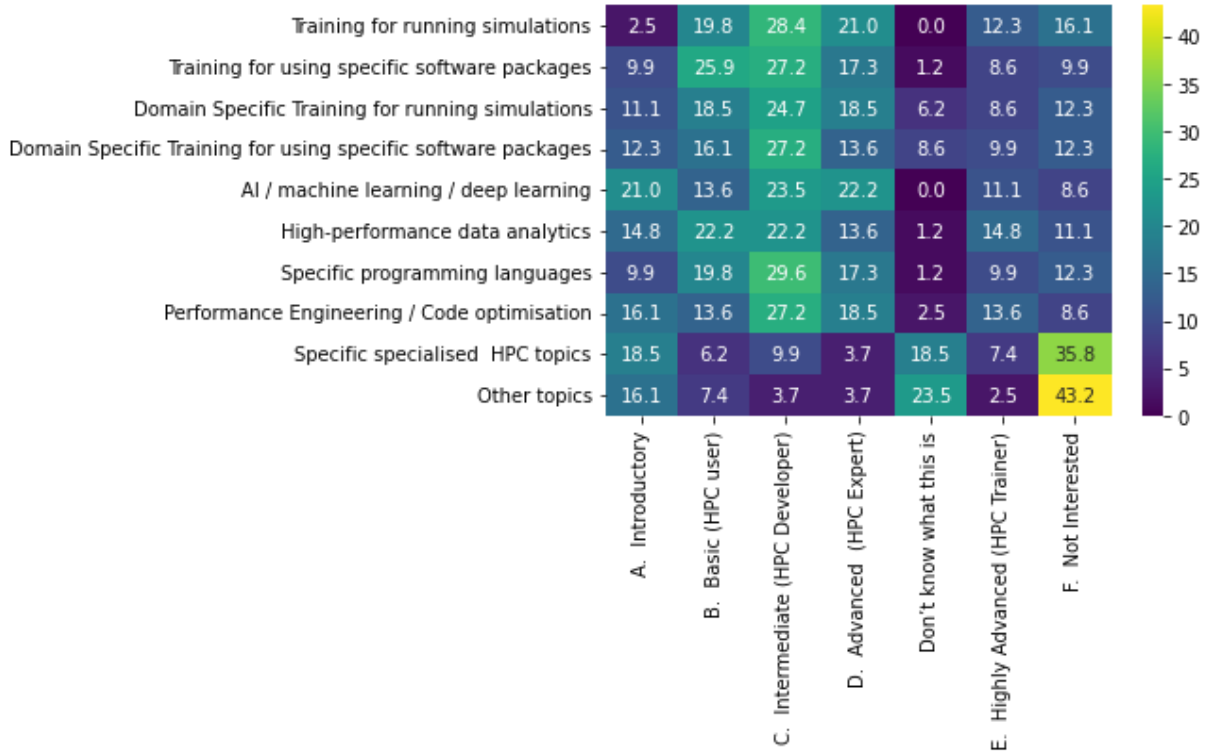

Figure 9: Map of training needs and level of experience [N=81]. The x-axis is the level of training difficulty and y-axis the training category. The coloring goes from dark purple (0%) to yellow (100%) depending on the counts in the sample.

At first glance, the **"intermediate" level training** (Level 3: HPC developer, Application Developer) accounted for a quarter of the answers (~25%) for all the categories. The only exception is "Specific, specialized HPC topics", and "Other topics" categories, that present high percentages of "not interested" answers. In the same categories, a small percentage of participants appear to be interested in introductory training. The profile of participants in the survey (presented in Section 2.1) is mostly HPC users coming from Engineering & Technology, Natural Sciences and Computer and Information Sciences. As such, they mostly utilize HPC infrastructure as a research tool or to solve industrial problems. The latter doesn't require a deepening of HPC knowledge or other related topics. For this reason, we can subsequently explain our findings for no interest in other HPC areas.

As data becomes a high value commodity, data analytics and AI are topics that are becoming more well-known, and more people have started exploring the possibility to integrate them into their production (academia or industry). The latter is captured in the current survey. There is high demand for introductory and basic AI/ machine learning / deep learning and High-Performance Data Analytics (HPDA) training.

Overall, it is found that HPC users, or potential users, demand low to intermediate level of training, with advanced and highly advanced not fitting their needs.

For assessing the training needs for the major stakeholder segment categories under investigation; industry, academia and public sector, a dedicated analysis was performed, that is presented in the sections that follow (see section 3.3.1, 3.3.2, 3.3.3).

3.3.1 HPC Training Needs: Industry

Our study findings indicate that the diffusion of HPC services in industry is relatively low (40% "Not familiar" to "Somewhat familiar" and 50% "Basic information", see Figure 6). As such, the heatmap illustration of the results (see Figure 10, below) reveals that the training needs are high for basic level (as HPC users - Skill Level 2: Beginner) for all cases and especially for high performance data analytics (HPDA) and specific programming languages. However, the sample is low (N=10) and the outcomes should be considered as an accurate indication that can lead to valid conclusions.

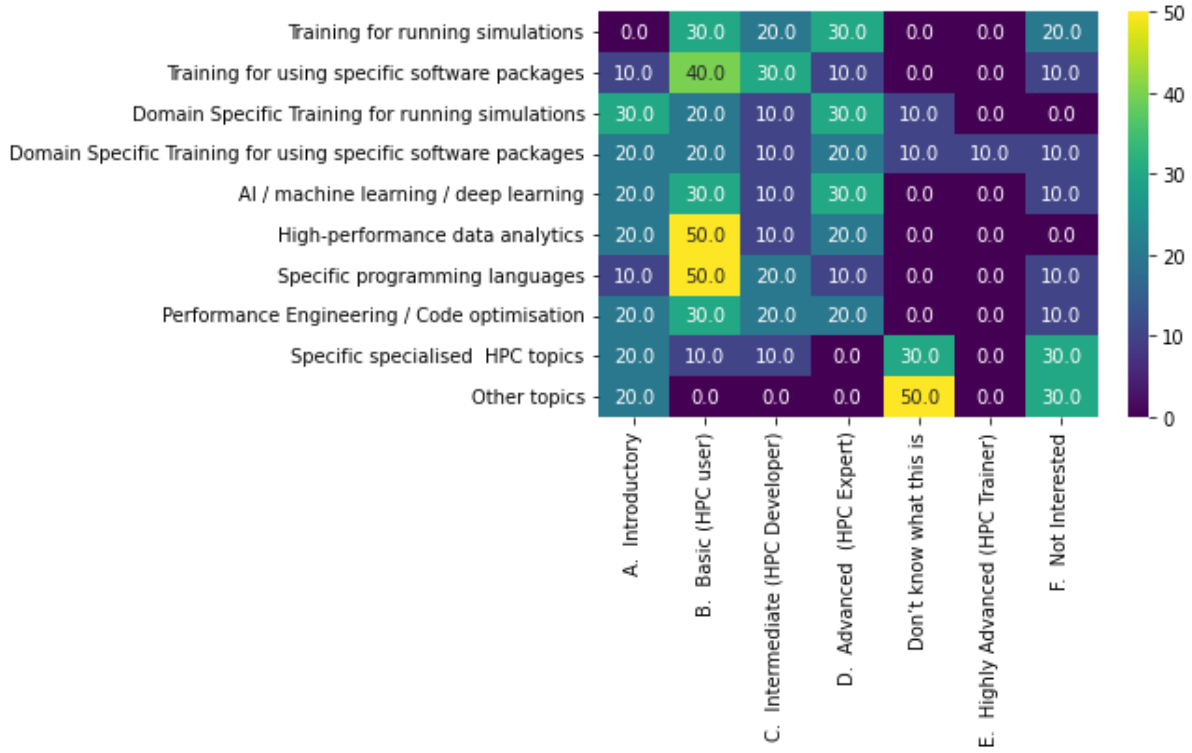

Figure 10: Map of training needs and level of experience for industry sector [N=10]. The x-axis is the level of training difficulty and the y-axis the training category. The coloring goes from dark purple (0%) to yellow (100%) depending on the counts in the sample.

3.3.2 HPC Training Needs: Academia

For the academia sector the results are qualitatively identical with the overall analysis presented in the previous section (Section 2.3). The latter is expected as the sample consists of 75% of academia participants. As such, in most cases close to a quarter of participants prefers basic or intermediate level of training. The categories of "Specific specialised HPC topics" and "Other topics" (~38%) are, by far, of no interest to academia, whereas "AI/machine learning/deep learning" and "High-performance data analytics" present almost equally the same demand in all levels of training (~20%)

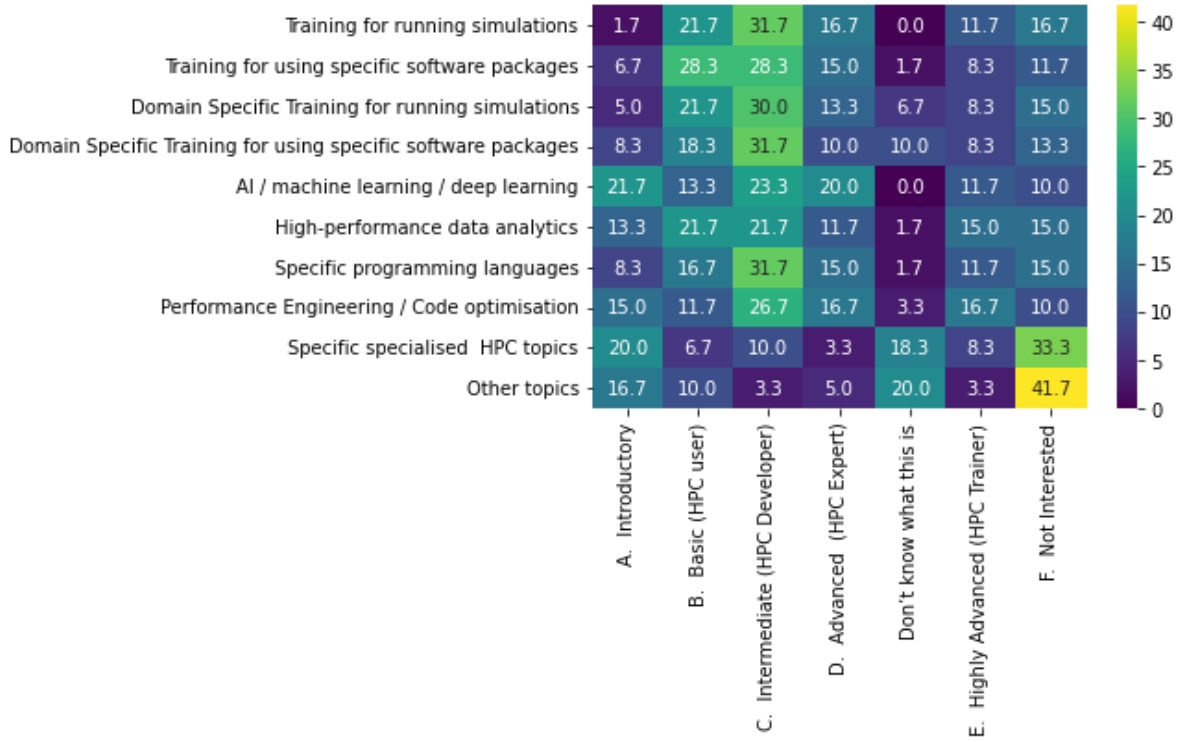

Figure 11: Map of training needs and level of experience for academia [N=60]. The x-axis is the level of training difficulty and the y-axis the training category. The coloring goes from dark purple (0%) to yellow (100%) depending on the counts in the sample.

3.3.3 HPC Training Needs: Public Sector

The answers gathered by participants representing the public sector were very few (N= 3). Hence due to the small sample, we were unable to perform any further analysis.

3.4 User-suggestions

The last part of our online survey included a user suggestion area (see section 4, in the Annex 1). As such, in the current section the user suggestion for course types and duration are illustrated and analysed.

In general, the score of all types of courses is high (above 25%). **Certification** is highly demanded with more than 50% to agree and strongly agree to the provision of certified programs in general (31% agree, strongly agree 26%) (see Figure 12). More specifically, **certified courses on specific themes** depict a very high demand, with more than 80% agreeing and strongly agreeing to such courses (44% agree, strongly agree 37%). Finally, the demand for certified seminars in the areas of HPC/HPDA/AI seem to be equally high, with around 79% of respondents agreeing and strongly agreeing to the provision of such seminars (51% agree, 28% strongly agree).

Interestingly, **"Live demonstration / code sessions" and "Hands-on sessions and guided exercises"** are highly preferred forms of training, denoting a horizontal **need to provide applied HPC training.** In particular, almost 90% prefer hands-on sessions and guided exercises (33% agree and 56% strongly agree) and 80% prefer live demonstration- code sessions (40% agree and 40% strongly agree). Surprising is the fact that hackathons, workshops and discussion appear to be far less preferred among Greek participants with around 52% to prefer such training formats (25% agree and 27% strongly agree).

Finally, a quite high percentage (74%) of the survey participants seems to prefer the **use of self-learning material** including HPC/HPDA/AI resources, notes and other relevant material (41% agree and 33% strongly agree).

In addition, the **"self-paced" choice** (60%) scores very high but lower than live demonstration, which is indicative of the complexity and specificity that HPC field presents. A generalized self-paced course may not be suited to include the specific needs of (potential) users of HPC services, compared to a live session where queries and questions will be resolved via the interaction and live conversation.

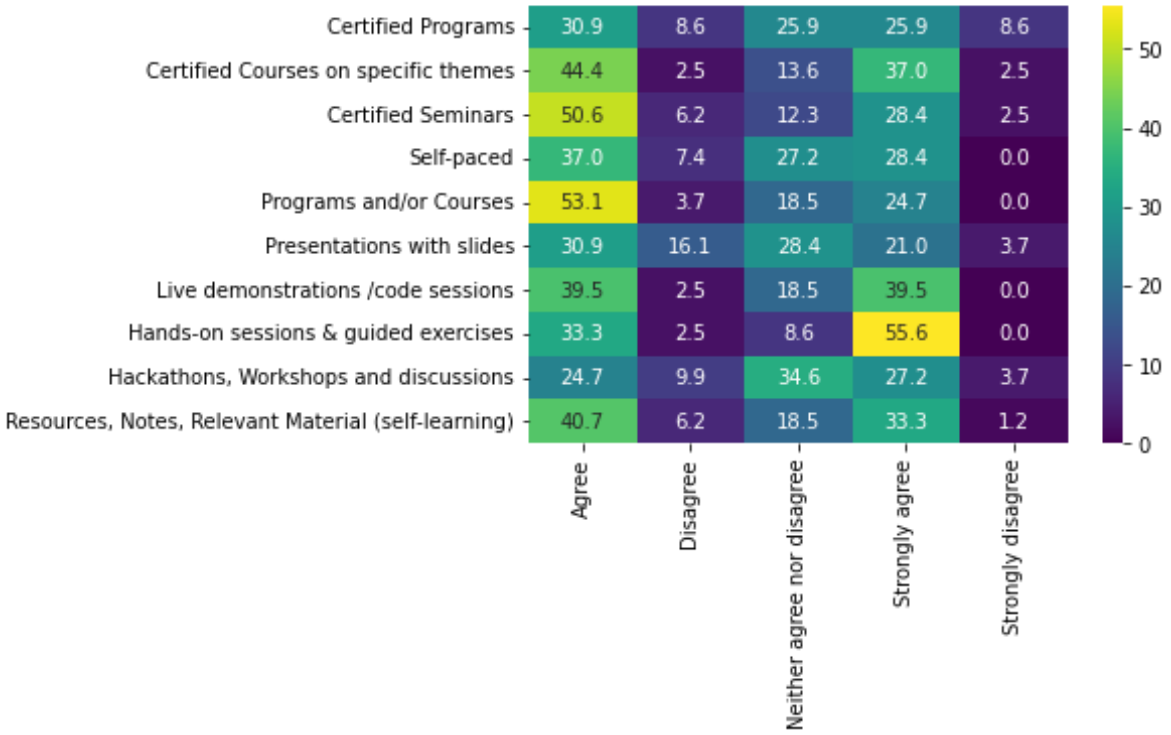

Figure 12: Preferences on type of courses and training [N=81]. The x-axis is the level of agreement and the y-axis the type of training. The coloring goes from dark purple (0%) to yellow (100%) depending on the counts in the sample.

Training Duration

Independent of the type of training, more than half of the participants of the survey (59.6 %) prefer **short sessions 2-3 hours.** Almost half agree on 1 to some days (23.5% & 24.7%), whereas longer durations present a relatively low demand among the survey respondents. Figure 13 illustrates the duration preferences of survey participants.

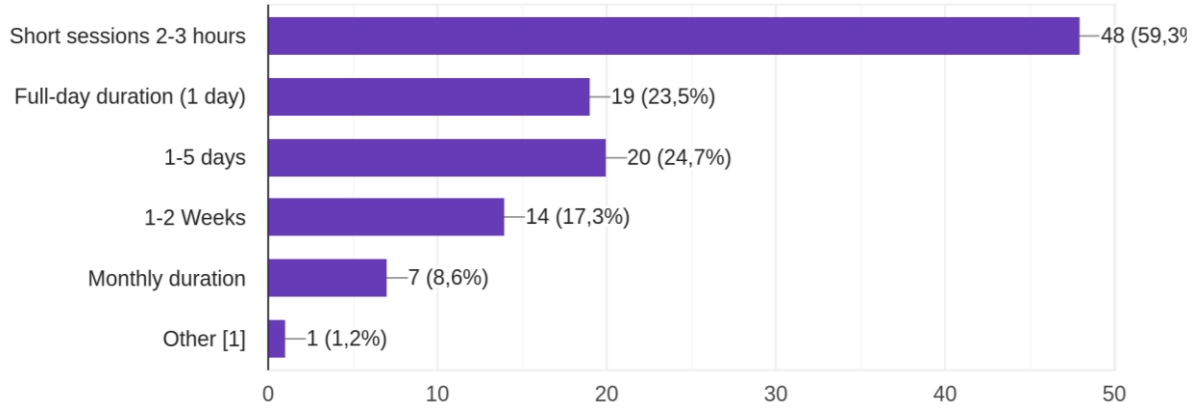

Figure 13: Preferences on duration of courses and training

4 Conclusive Remarks

In the current report, the results of the first HPC Training Needs Analysis (TNA) in Greece are presented and analyzed. The scope of this survey study has been to provide information about the HPC landscape in Greek academia, industry, and public sector, and at the same time identify the Greek HPC training needs as those were depicted by the survey respondents during 2021.

This online survey revealed that **HPC services in Greece are used mostly by academia** especially by Postgraduate students and University professors of all levels. On the other hand, **the diffusion of HPC in industry in Greece, appears to be quite low.**

The most common fields that utilize HPC infrastructures in Greece, based on our findings, are **engineers, natural and computer scientists** compared to other fields. The latter may be a result of the sample of participants and/or the kind of research applied to other fields (as seen in Figure 8 which illustrates, in a heatmap plot, the state of familiarity of HPC users).

The current state of **training "demand" for HPC courses in Greece**, is presented in Figure 14 and it is in accordance with the needs depicted in the survey. Most of the courses are relevant to **academia** and for users that are familiar with HPC. The survey confirms the gap in provided HPC training on familiarity level and type, but it is not conclusive for the interest of the industry in HPC services and training as the sample consists mainly by participants working in academia.

Type of Learner		Not Interested	Type of HPC User (including HPDA/AI)				
			Informational User (Pre-HPC User)	HPC User (Application HPC User)	HPC Developer (Application HPC Developer)	HPC Adminstrator (System HPC Administrator)	HPC Expert
			Skill Level 1: Introductory	Skill Level 2: Beginner	Skill Level 3: Intermediate	Skill Level 4: Advanced	Skill Level 5: Highly Advanced
Student, Academia, Research (BSc, MSc, PhD, post-doc, Professort)		**19%**	**12%**	**18%**	**25%**	**14%**	**11%**
Industry	**Industrial Researcher** (Data Scientists)	**14%**	**31%**	**0%**	**7%**	**17%**	**31%**
	Other	**15%**	**19%**	**32%**	**16%**	**19%**	**0%**
Public Sector		**0%**	**0%**	**0%**	**0%**	**0%**	**0%**

Figure 14: Mapping demand of HPC training in Greece for 2021 [N=81][9]

[9] **Groups**
A. Academia: [A. Undergraduate student, B. Postgraduate student, C. PhD student, D. Postdoc Researcher, E. Assistant, professor, F. Associate professor G. Full professor]
B. Industry: [I. Employee in a SME company, J. Researcher/developer in small/medium company, K. Researcher/developer in large company, L. Start-up owner/partner, M. SME owner/partner]
C. S/W Engineer Data Science: [N. Research software engineer, O. Data scientist/machine learning engineer in small/medium company]

Nonetheless, the answers of participants derived from the **industrial sector** indicate that basic trainings are in demand (as HPC users) and mostly for data analytics and specific programming languages. Overall, it is found that a: (1) fast-pace, (2) entry level, (3) applied HPC training but (4) not focused only on HPC, that will (5) provide some kind of certification, is preferred, at least from academia.

As study limitations, we consider the relatively low response rate (N=81), which does not allow us to generalise our findings (Sivo et al., 2006). However, we expect that as this is the first study of its kind in Greece future surveys are expected to have higher response rate associated with higher awareness level.

References

Regmi, P. R., Waithaka, E., Paudyal, A., Simkhada, P., & van Teijlingen, E. (2016). Guide to the design and application of online questionnaire surveys. *Nepal journal of epidemiology*, *6*(4), 640.

Sivo, S. A., Saunders, C., Chang, Q., & Jiang, J. J. (2006). How low should you go? Low response rates and the validity of inference in IS questionnaire research. *Journal of the association for information systems*, 7(1), 17, link.

Statista Research Department (2022), "Amount of data created, consumed, and stored 2010-2025", Published Mar 18, 2022, https://www.statista.com/statistics/871513/worldwide-data-created/ last accessed March 2022.

Annexes

Annex 1: HPC Training Needs in Greece Questionnaire

HPC Training Needs in Greece by EuroCC@Greece

Dear participant,

The EuroCC@Greece is one of 33 national nodes of the EuroCC project, which started during September 2020, built in the framework of the European High Performance Computing Joint Undertaking (EuroHPC JU). The aim of EuroCC is to advance competitiveness in research, improve effectiveness of government services and promote innovation in industry.

In order to implement this objective, we will design and deliver training activities in order to address the different user needs, enabling them to scale up their HPC workloads and enhance their HPC competencies.
The results from this survey will help us identify and understand the distinct needs of the Greek users as well as those that want to become HPC users and create training portfolios for the coming years, enhancing this way High Performance Computing in Greece.

Your participation is important. Let's co-create together the future HPC training events in Greece!

If you want to keep up to date with the EuroCC@Greece activities, upcoming training events, workshops, you can:

- visit our web page: https://eurocc-greece.gr/
- Sign up for our EuroCC newsletter https://eurocc-greece.gr/newsletter/
- follow us on Twitter @EuroCC_Greece and LinkedIn https://www.linkedin.com/company/eurocc-greece/
- YouTube channel https://www.youtube.com/channel/UCuPZ7fUxtrsjQbZbo7pkByw

Contact us at: contact@eurocc-greece.gr

Notice
Participation in this research study conducted by EuroCC@Greece is entirely voluntary and anonymous.
You are not required to divulge your personal identity, nor will you or your organisation be identified in any published documents. For more detailed information, see our privacy policy.

A. Generic Information

1. Please specify your current position *

Επιλογή

2. Please specify your main scientific discipline or field of work *

Επιλογή

B. HPC Familiarity & Usage

3. Please specify your familiarity with HPC *

Επιλογή

Please specify your current HPC usage *

- [ ] I do not use HPC
- [ ] I use existing Software Package(s)
- [ ] I develop my own HPC software
- [ ] I use both existing software packages and my own HPC software
- [ ] Other

C. HPC Training Needs

5. What type of training are you interested in attending and at what level? (multiple answers) *

	Introductory	Basic (HPC user)	Intermediate (HPC Developer)	Advanced (HPC Expert)	Highly Advanced (HPC Trainer)	Not Interested	Don't know what this is
Training for running simulations, etc	◯	◯	◯	◯	◯	◯	◯
Training for using specific software packages	◯	◯	◯	◯	◯	◯	◯
Domain Specific Training for running simulations, etc	◯	◯	◯	◯	◯	◯	◯
Domain Specific Training for using specific software packages	◯	◯	◯	◯	◯	◯	◯
AI / machine learning / deep learning	◯	◯	◯	◯	◯	◯	◯
High-performance data analytics	◯	◯	◯	◯	◯	◯	◯
Specific programming languages	◯	◯	◯	◯	◯	◯	◯
Performance Engineering / Code optimisation	◯	◯	◯	◯	◯	◯	◯
Specific specialised HPC topics [1]	◯	◯	◯	◯	◯	◯	◯
Other topics [2]	◯	◯	◯	◯	◯	◯	◯

6. Which is your preferred type of training activity? Please specify your preference (multiple options) *

	Strongly agree	Agree	Neither agree nor disagree	Disagree	Strongly disagree
Certified Programs (Master level, Bachelor, etc)	○	○	○	○	○
Certified Courses on specific themes	○	○	○	○	○
Certified Seminars	○	○	○	○	○
Self-paced	○	○	○	○	○
Programs and/or Courses	○	○	○	○	○
Presentations with slides	○	○	○	○	○
Live demonstrations /code sessions	○	○	○	○	○
Hands-on sessions & guided exercises	○	○	○	○	○
Hackathons, Workshops and discussions	○	○	○	○	○
Resources, Notes, Relevant Material (self-learning)	○	○	○	○	○
Other [1]	○	○	○	○	○

8. The optimal duration for training events for you is: *

- [ ] Short sessions 2-3 hours
- [ ] Full-day duration (1 day)
- [ ] 1-5 days
- [ ] 1-2 Weeks
- [ ] Monthly duration
- [ ] Other [1]

[1] Please specify your preferred duration:

Short answer text

Thank you for your participation ! Please share with us any suggestions that you may have, as to how we could help with your training needs?

Long answer text

Privacy Policy

Your privacy is very important to us. The information below will provide you all the necessary information in relation to this research study that takes place in the course of the EuroCC project.

Full title of Project: EuroCC
The project has received funding from the European High-Performance computing Joint Undertaking (JU) under grant agreement No 951732 and the Greek Secretariat for Research and Technology.
[Research Activity] Title: HPC Training Needs in Greece
[Research Activity] Organiser: EuroCC@Greece and contact@eurocc-greece.gr.
[Type of Survey] Anonymous

For any query regarding your personal data, your rights (access, information, edit, deletion, rectification, portability and further privacy issues you may contact with EuroCC@Greece via contact@eurocc-greece.gr, at all times

Things you need to know
1. Purpose and objectives of research:Identify the HPC Training Needs in Greece
2. Voluntary Participation: I understand that my participation is voluntary. I do not have to take part if I do not want to. I may leave the group at any time for any reason. I may also and/or revoke my consent at any time for any reason.
3. Risks: By participating in this research there are no direct risk.
4. Benefits: There are no direct benefits for participating in this research activity.
5. Third parties: Your data won't be shared with organisations or persons outside of the study, nor will it be published anywhere or used for non- study purposes, unless we get a separate permission from you, or unless we are legally required to do so. Your data will not be transferred to any country outside the European economic area.
6. Storage: EuroCC@Greece shall store my anonymous data as long as required to achieve the EuroCC study objectives. Thereafter, we may still keep your data for five years after the study has ended, with the understanding that during this period we may only use the data to demonstrate the scientific validity of the study outcomes and to demonstrate that the study was executed in accordance with our legal obligations.

Your participation is purely voluntary, and you may at all times withdraw your consent by contacting us, by sending an e-mail to contact@eurocc-greece.gr

By participating in this research study conducted by EuroCC@Greece you agree, and provide specific and informed consent for our research purposes.

Training Needs for HPC, HPDA, AI, HPAI in Greece

2022 Report