Training services need mapping in the Czech Republic

Best practice Guide

Karina Pešatová

December 2022

VSB TECHNICAL UNIVERSITY OF OSTRAVA | IT4INNOVATIONS NATIONAL SUPERCOMPUTING CENTER

1. Introduction

The National Competence Center in HPC of the Czech Republic (NCC) is represented by IT4Innovations National Supercomputing Center at VSB – Technical University of Ostrava (IT4I) a leading research, development, and innovation center active in the field of high-performance computing (HPC), high-performance data analysis (HPDA) and artificial intelligence (AI). IT4Innovations operates the most powerful supercomputing systems in the Czech Republic.

In addition to its publicly available computational services to academia and industry, NCC extends its training services and ensures they meet the needs of its community. Therefore, would like to understand, how the NCC can help members of the community with their projects by offering education, knowledge transfer, consultation, and other training-related support.

A survey was developed and launched in mid-2021 to gather relevant feedback. The following is the analysis of the results obtained in the period of July 2021 up to December 2022. In the period 95 participants filled in the survey which was available online in both Czech and English language. There were 3 sections of the survey each containing several questions – about you, training needs, and training format.

2. Analysis of the results

2.1 Section 1 – About you

The first section aimed at gathering information about the respondent such as contact details, scientific field, institution, position within the institution, and data relevant to industrial and public sector respondents like the size of the company they work for.

The community was divided into three groups academia, industry, and the public sector. Figure 1 is the graph of the division of respondents into the three different groups. There were 81 responses from academia, 12 from industry, and 6 from the public sector.

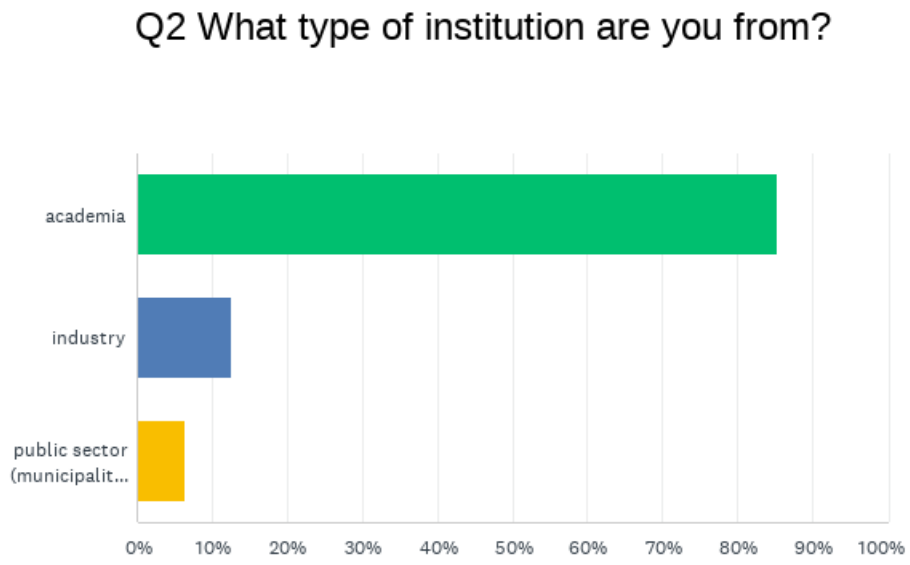

Figure 1 – Respondents by the group.

The respondents are holding different positions. 41% of them have indicated they are a researcher, 27% a student, and 6,5 % a developer. The rest of the positions mentioned included manager, team leader, professor, and some others.

The academia-specific question about the scientific domain in which the respondents work showed the following results: 38% computational sciences, 21% physics, 14% mathematics, 13% chemistry, 10% engineering, and 4% earth sciences.

The 6 respondents from the industry have indicated the following business sectors: 3 times IT, heavy industry, aerospace, and energy. 2 were from a large company (>2000 employees), 2 from a micro company (<10 employees), and one from each medium company (51-250 employees) and medium-large company (250-2000 employees).

2.2 Section 2 – Training needs

In this survey, we were mapping the interest and level of expertise of the participants in the 4 following domains: High Performance Computing (HPC), Artificial Intelligence/Machine Learning or Deep Learning (AI), High Performance Data Analytics (HPDA), and Visualisation and Virtual reality. The domains were in the survey described as follows.

HPC comprises large scale highly optimized computation and data processing that would not be possible with workstations or even public cloud services. One of the main advantages in comparison to the public cloud is high-speed interconnects between computation nodes (within computation clusters), an optimized storage hierarchy, and a physical locality of resources for the highest efficiency and scalability.

HPDA is supported by HPC to process large amounts of data (so-called Big Data) for analytical purposes to (pre) process, curate, optimize or visualize. Its advantage is the handling of data which would not be possible on ordinary computation systems due to memory and storage

limitations, and a short latency to obtain the final data products (time-to-solution).

As **AI** we hereafter consider **Machine and Deep Learning** with the use of HPC. These fields of AI are currently sought after due to their high analytical power in combination with the latest hardware architectures. IT4Innovations offers dedicated systems optimized for this domain, like the latest NVIDIA GPUs, DGX-2 systems, and software stacks, making use of the full advantages of HPC, like large scalability and efficient storage technologies.

IT4Innovations also offers solutions for the **visualization** of scientific results. This ranges from photorealistic rendering, and real-time interactive visualizations up to virtual reality simulations.

The interest in the domains is depicted in Figure 2. HPC is the domain that interests the community the most, followed by AI.

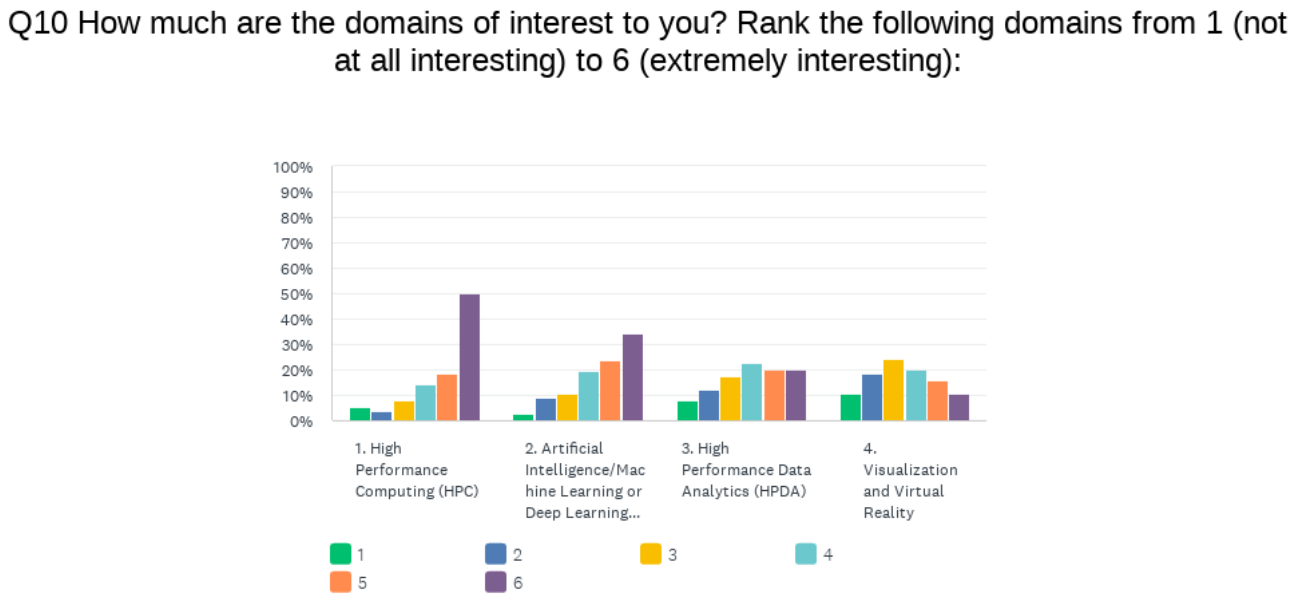

Figure 2 – Indication of interest in the four domains.

When it comes to expertise in these domains. Respondents claim to have mostly none or beginner expertise in HPDA and Visualisation and Virtual Reality. Most of the respondents are considering themselves beginners in AI. In HPC around 30% are beginners and intermediate-level practitioners and nearly 20% consider themselves to be advanced users. This relates to 14 participants. Whereas only 1 respondent considers himself advanced in HPDA, 4 in AI, and 3 in Visualisation and Virtual reality. These results are shown in Figure 3.

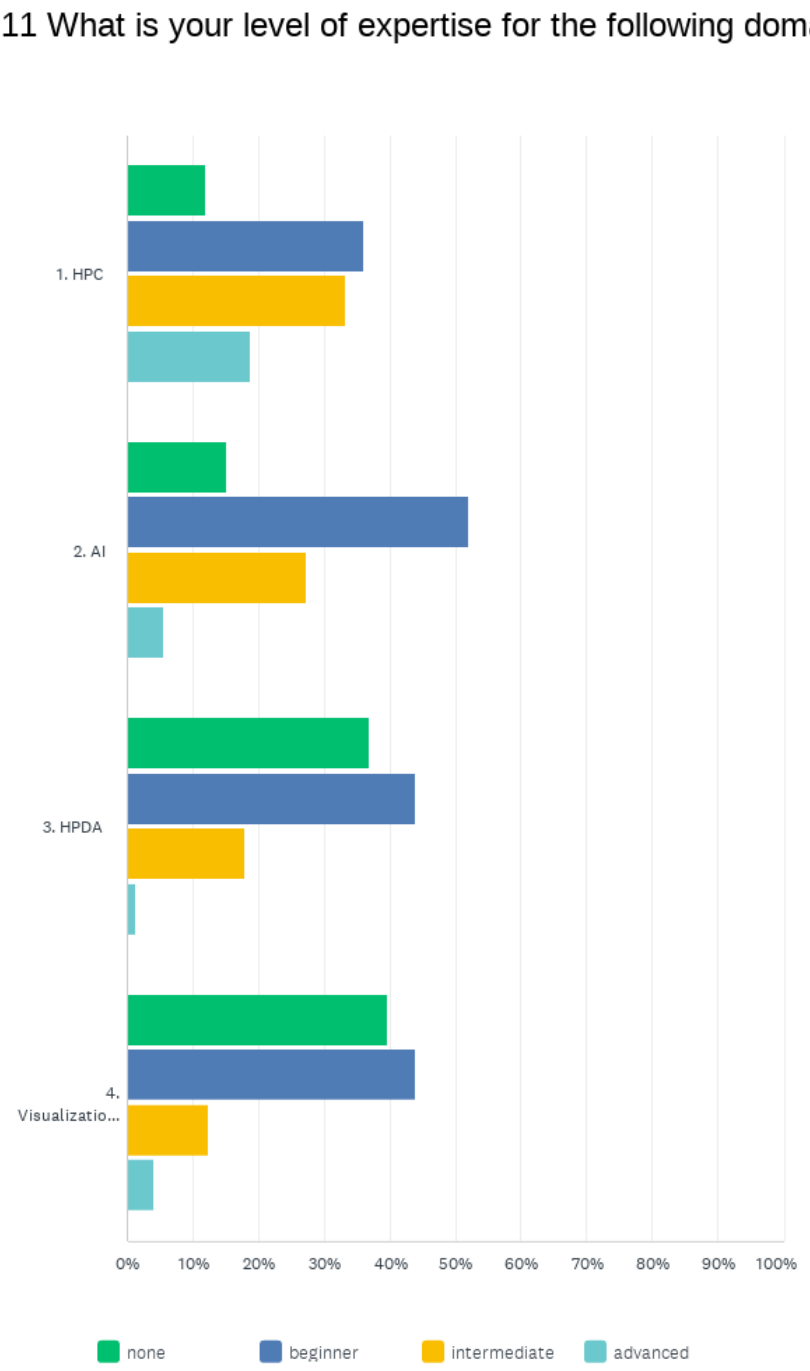

Figure 3 – level of expertise of the respondents in each domain.

Within the domains, we have identified several sub-domains and have asked about the interest in those to get more specific topics for possible future training events. In the HPC domain, the following subtopics were marked as extremely interesting: Access to HPC resources, Parallel Programming (MPI, OpenMP, vectorization), GPU Accelerated Computing, Optimization of computation, and using existing HPC applications. In the AI subdomain, the most interesting topics were: (Deep) Neural Network Architectures and the Selection of Machine Learning algorithms. The responses for HPDA did not show any significant interest in any of the sub-domains offered. In Visualisation and Virtual Reality, the topic of Visualisation Tools was the

most popular.

In the next question, we asked the participants to select three top priorities for the subdomains. The results are shown in Figure 4.

	Topic	1.	2.	3.	Total score
1.	Optimization of computations, performance analysis, profiling	5	11	1	38
2.	Parallel programming (MPI, OpenMP)	9	6	2	41
3.	GPU accelerated computing (NVIDIA, AMD)	2	5	5	21

Figure 4 – Top three training topic priorities identified by the respondents.

The next open question was which applications and tools from the area of HPC, AI, HPDA, and visualizations are they using or intend to use in the future. The following were the most mentioned: TensorFlow, Blender, ParaView, PyTorch, and Ansys.

2.3 Section 3 – Training Format

To understand how to offer the training services best to the community several questions were asked aiming at the format, ways of training, and length. The most popular format is a lecture covering a broader range, with hands-on. See the full results in Figure 5.

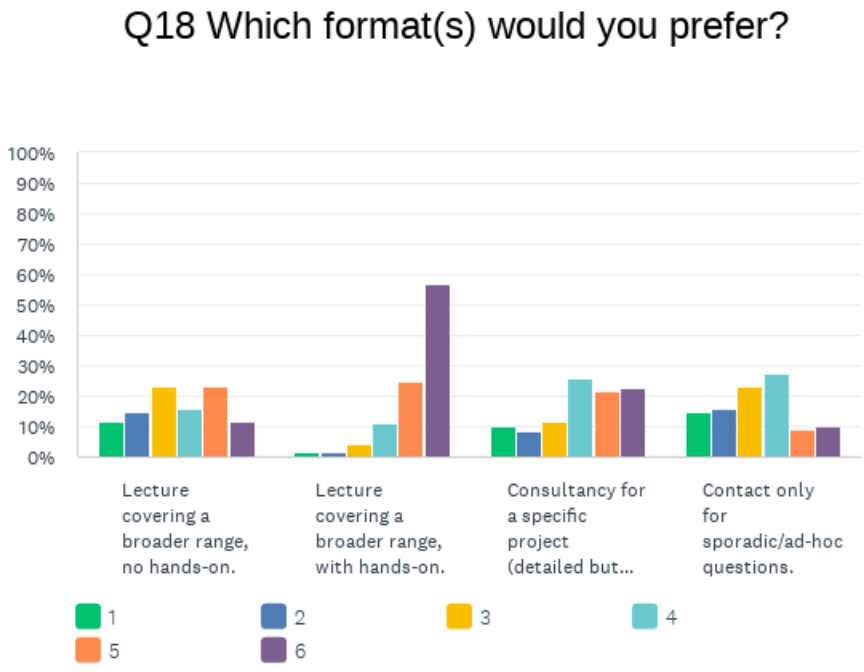

Figure 5 – The most popular format of training.

Until the year 2019, all training events provided by IT4I were face-to-face. With the pandemic the center has started to offer online training and since 2022 also hybrid form. The preference of

the respondents is to continue offering training online, but the difference between online and onsite is minimal. The results are shown in Figure 6.

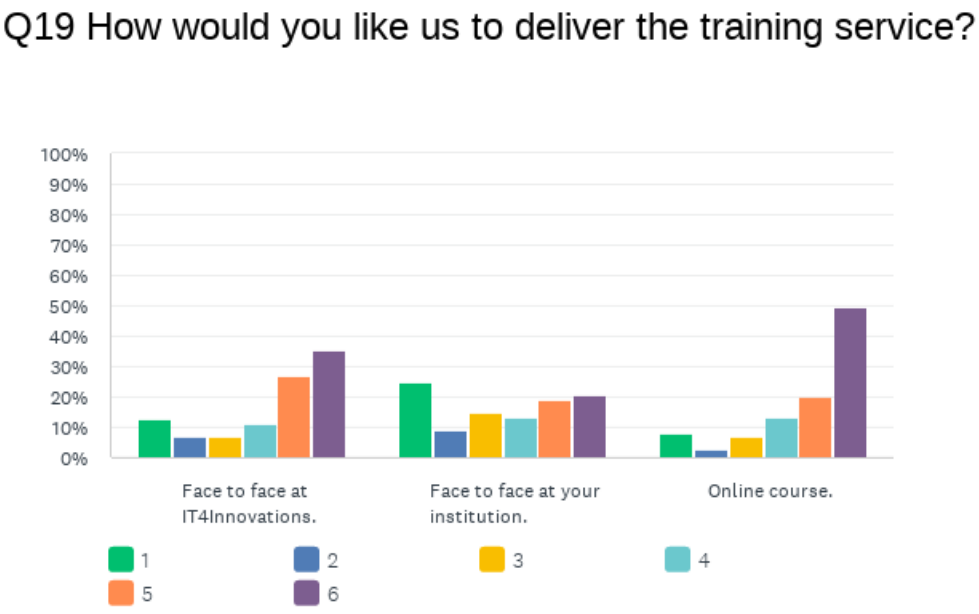

Figure 6 – Preferred ways of delivering training.

The preferred length of the training is one day but with only a small difference to two days. The results for longer training events such as schools are not clearly defined nearly an equal number of respondents find it convenient or not. The results are shown in Figure 7.

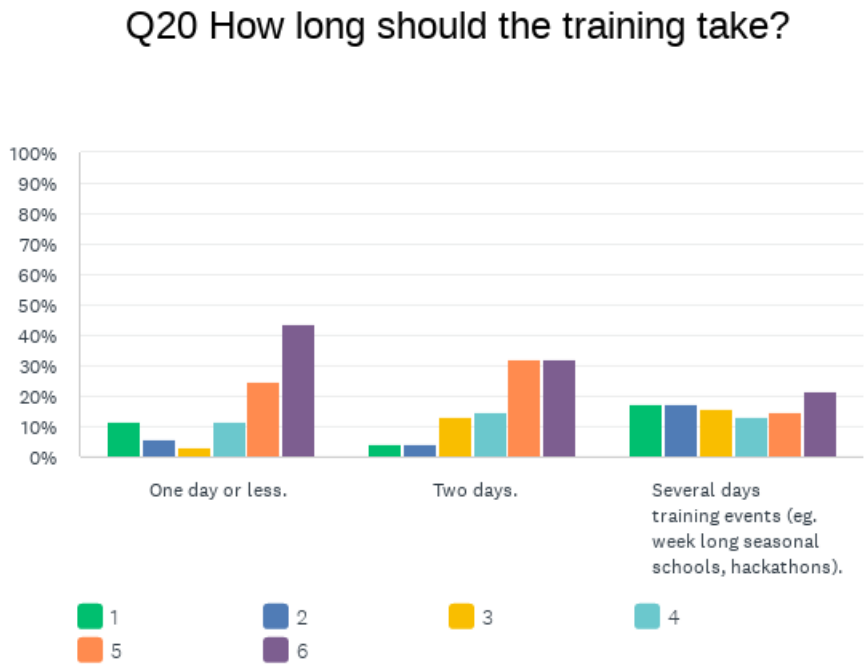

Figure 7 – Preferred length of training events.

3. Conclusions and outlook

The result of the Training needs survey is very valuable and is used as input to plan the courses and workshops which the NCC is providing. The plan reflecting the responses obtained within the survey is delivered every half a year.

The plan is to re-run the survey at the beginning of 2024. The domain of Quantum Computing in which IT4I is newly involved in is to be added.

This 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 program 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, the Republic of North Macedonia, Iceland, and Montenegro. This project has received funding from the Ministry of Education, Youth and Sports of the Czech Republic (ID:MC2101).