Best Practice Guide

Multi-GPU AI Train the Trainer Workshop

Training, Twinning, and Mentoring

Organisers: CASTIEL2, EuroCC2

Contents

Executive Summary 3
Objectives 3
Background and Context 3
Planning 4
Preparation 5
Delivery 5
Support 5
Improvement 6
Impact 7
Summary and Conclusions 7
Appendix: Captures of Training Delivery 9

Executive Summary

This document outlines the best practices established and lessons learned during the *Multi-GPU AI Train the Trainer Workshop*, held virtually from *January 30th* to *February 5th, 2026.* The workshop provided participant-trainers with remote access to the EuroHPC Leonardo supercomputer, hosted by CINECA in Bologna, Italy. The programme was directed toward National Competence Centre (NCC) and Centre of Excellence (CoE) trainers, HPC centre instructors, university educators, researchers, and advanced users with significant prior experience in Artificial Intelligent (AI) and High-Performance Computing (HPC) in Europe. Key contributions of this guide include a revised planning timeline for a seamless registration process, robust frameworks for ensuring GDPR compliance, and the implementation of real-time support mechanisms, such as Discord, to enhance trainer engagement. Following an overview of the guide's objectives and background, the report details the critical organisational considerations made during each phase of the training lifecycle; Planning, Preparation, Delivery, and Improvement; as illustrated in Figure 1. The document concludes by highlighting the impact and results of the event, followed by a summary and conclusion.

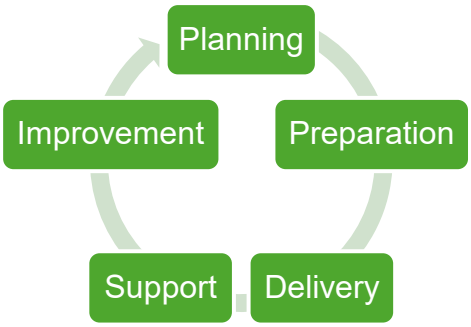

Figure 1: Phases of training lifecycle

Objectives

The primary objective of this best practice guide is to highlight the strategies frameworks implemented over all the training lifecycle phases to deliver a successful train the trainer HPC focused event within the CASTIEL 2 and EuroCC 2 partnership. Furthermore, this document investigates potential improvements and establishes actionable guidelines for future Train the Trainer initiatives. These recommendations are based on the analysis of participant feedback and lessons-learned, ensuring that subsequent HPC educational events maintain a standard of excellence and technical rigor.

Background and Context

The main goal of the Multi-GPU AI Train the Trainer Workshop was to equip future trainers with the pedagogical and technical skills required to teach advanced multi-GPU AI topics across European HPC centres. The programme provided specific focus on distributed AI, Large Language Models (LLMs), MLOps, and HPC workflows. Beyond technical proficiency, the workshop was designed to facilitate knowledge transfer, arming trainers with necessary resources to replicate this high-level course at their respective national and institutional levels.

To this end, the workshop was structured as a full time, five-day intensive programme featuring a progressive curriculum delivered by partner NCCs. Although the event was hosted exclusively online, participant-trainers were granted remote access to the EuroHPC Leonardo supercomputer at the CINECA[1] premises. This hands-on environment was supplemented by real-time support and discussions. Following the workshop, access to a comprehensive suite of training materials including recorded sessions[2], and GitHub repositories was provided to ensure the long-term utility of the workshop for the training community.

Planning

Figure 2 presents the Multi-GPU AI Train the Trainer Workshop timeline across the planning, preparation, delivery, and support stages. The planning phase represents the most extensive and resource-intensive stage, encompassing activities 1-5. During this period, the curriculum was designed with a broad focus on diverse topics and skill level to ensure a comprehensive learning experience.

By adapting this approach, the planning of the curriculum targeted specific courses that cover this entire range, effectively bridging introductory topics with in-depth intermediate and advanced modules. From this foundation, prerequisites were determined and published well in advance. Providing this information to future participants prior to registration could significantly increase participant satisfaction during the delivery of the workshop, as attendees would be better aligned with the technical demands of the course.

Next, the preparation phase (activities 6-8 Figure 2) established the first contact with participants through the registration process. It is noted that the actual registration period occurred during a window where both participants and organisers are often unavailable due regularly scheduled holidays. However, this timing was necessary due to the urgency of finalising the course development before the conclusion of the Castiel2 project, future iteration could benefit from a flexible schedule. Finally, following the successful workshop delivery (activity 9 Figure 2), a post workshop survey was conducted, and feedback was discussed among organisers (activity 10 Figure 2) to inform future best practices.

	Activity	Start	End	Days
		2025 Q4 September	2025 Q4 October	2025 Q4 November
1	Define course scope & requirements	2-sep-25	23-oct-25	51
2	Decide on cluster system (JUWELS/LUMI/Leonardo)	7-oct-25	13-ene-26	98
3	Repository hosting & access setup	23-oct-25	2-dic-25	40
4	Licensing & legal compliance	4-nov-25	2-dic-25	28
5	Curriculum development & content integration	7-oct-25	26-ene-26	111
6	Registration & participant selection	2-dic-25	26-ene-26	55
7	HPC resource reservation & account creation	13-ene-26	30-ene-26	17
8	Communication setup (Discord, GitHub, emails)	13-ene-26	30-ene-26	17
9	Course delivery (Workshop Week)	30-ene-26	5-feb-26	5
10	Post-workshop review & future planning	17-feb-26	28-feb-26	11

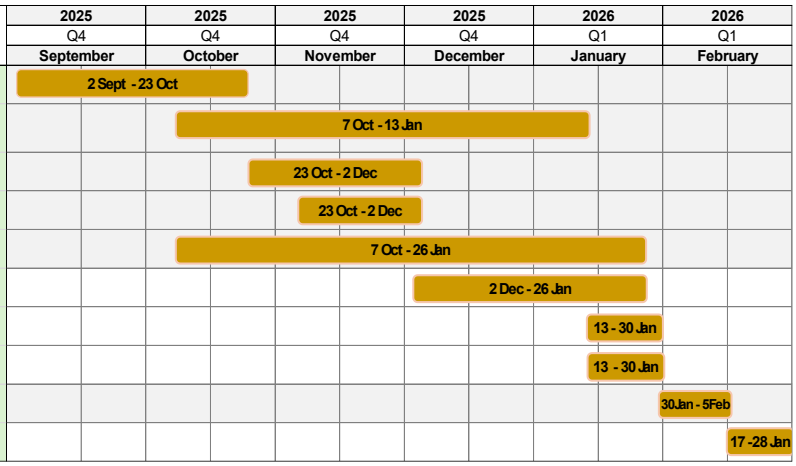

Figure 2: Actual training the trainer workshop planning timeline.

[1] https://www.hpc.cineca.it/systems/hardware/leonardo/

[2] https://hpc-portal.eu/training/short-courses/castiel2-multi-gpu-ai-train-the-trainer-workshop

Preparation

From first contact with participant-trainers in the preparation phase, GDRP regulations are ensured by obtaining consent of personal information at registration. To ensure GDPR compliance of the training for all participant-trainers, sessions are recorded such that no personal information of attendees is captured, allowing for easy compliant dissemination afterwards.
Similarly, the workshop requires access to the HPC resources, preparation of credentials and account setup was conducted in a dedicated onboarding session ahead of the main training delivery. This ensured a smoother delivery by resolving technical access issues before the curriculum begins.

Delivery

The delivery of the training was conducted online over a 5-day period. The full schedule is shown in Table 1. To accompany the live instruction, Discord[3] was used as a tool to provide real-time support and communication. In addition, a separate Discord channel for non-relevant chatting was established to avoid crosstalk and foster engagement between participants. The *Appendix: Captures of Training Delivery* provides image captures of some training sessions throughout the workshop.

Table 1: Multi-GPU AI Train the Trainer Workshop schedule

	Day 1	**Day 2**	**Day 3**	**Day 4**	**Day 5**
Morning 9:30 – 13:00 CET *NCC*	Intro to GPU architecture and access to HPC infrastructure *NCC Italy*	PyTorch Distributed Data Parallel *NCC Netherlands*	Model parallelism with PyTorch *NCC Hungary*	Computer Vision *NCC Romania*	Ray + Retrieval-Augmented Generation (RAG). *NCC Italy*
Lunch Break: 13:00 – 14:00 CET					
Afternoon 14:00 – 17:30 CET *NCC*	Intro to deep learning *NCC Sweden*	Model parallelism with PyTorch *NCC Hungary*	LLMs, finetuning HuggingFace Accelerate, DeepSpeed *NCC Belgium, NCC Italy*	MLOps. *NCC Belgium*	Hyperparameter tuning, Q&A *NCC Finland*

Support

In addition to providing real-time support during the event via Discord, assistance was extended to participant-trainers post-workshop through the systematic sharing of session recordings on YouTube[4] and the continued availability of all training materials. Maintaining these resources in a centralised, open-access format ensures that the knowledge remains a permanent asset for the community.
For better understanding the participants experience, a comprehensive survey was conducted to obtain quantitative comments and suggestions. Such a survey serves as

[3] https://discord.com/
[4] https://youtube.com/playlist?list=PL6l5OWh2pdwTfkyY3xPWpJoUNPFKDdJxZ&si=ZRCqyWu1MRevTtLJ

a vital opportunity for participants to voice technical doubts and for organisers to identify which aspects of the workshop were most successful and which might require fine-tuning. For this specific event, feedback was highly positive, and the constructive suggestions provided by participants could be utilised organiser feedback meeting to refine the training lifecycle for future iterations.

Improvement

Through participant-trainers feedback and internal discussions, a revised timeline is proposed in Figure 3. While the planning phase (activities 1-5 Figure 3) remains mostly unchanged, significant improvement could be made starting with activity 6. The revised timeline proposes a shortened registration period to minimise conflicts with holiday schedules, ensuring a smoother process for both participant-trainers and organisers. Furthermore, activities 7 and 8 in Figure 3 could be decoupled so that communication setup occurs later in the preparation stage, allowing it to align more naturally with training delivery.
Additionally, course delivery (activity 10 Figure 3) could be revised to separate HPC infrastructure access (activity 9 Figure 3) from the instructional sessions. This division could significantly reduce technical frictions and accommodate common troubleshooting before the curriculum begins.

	Activity	Start	End	Days
1	Define course scope & requirements	2-sep-25	23-oct-25	51
2	Decide on cluster system (JUWELS/LUMI/Leonardo)	7-oct-25	13-ene-26	98
3	Repository hosting & access setup	23-oct-25	2-dic-25	40
4	Licensing & legal compliance	4-nov-25	2-dic-25	28
5	Curriculum development & content integration	7-oct-25	26-ene-26	111
6	Registration & participant selection	2-dic-25	20-dic-25	18
7	HPC resource reservation & account creation	13-ene-26	30-ene-26	17
8	Communication setup (Discord, GitHub, emails)	1-feb-26	15-feb-26	15
9	HPC Infrastructure Access	25-feb-26	25-feb-26	1
10	Course delivery (Workshop Week)	25-feb-26	3-mar-26	5
11	Post-workshop review & future planning	5-mar-26	28-feb-26	13

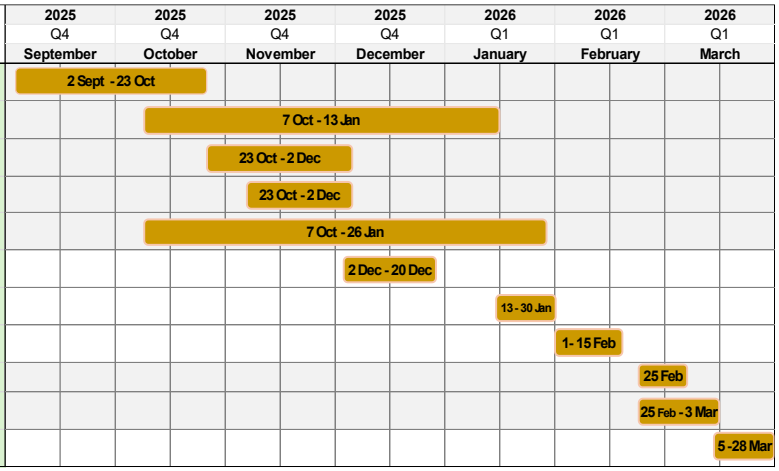

Figure 3: Revised train the trainer workshop planning timeline based on lessons-learned through completion of the workshop.

In terms of the workshop delivery, feedback and discussion generated a suggestion to migrate to a modular framework. This approach could allow participants to tailor their workshop experience to their specific skill levels in each knowledge area, while still allowing the event to cover diverse topics. While modularity might increase the planning load during the registration, as participants would need to declare their intended modules in advance, it could greatly increase the personalisation of the workshop and encourage deeper engagement.
Another approach to achieving higher depth could be to narrow the focus exclusively to advanced topics, shortening the workshop duration from five days to three. The morning session would be dedicated to theoretical frameworks and the afternoon sessions to practical, hands-on applications. This specific split could allow for a more in-depth study of advanced concepts while ensuring participants have sufficient time to apply what they have learned in the HPC environment.

Impact

The global impact of the Multi-GPU AI Train the Trainer Workshop can be summarised by 77 participant-trainers from 50 different European institutions, supported by 14 partner NCC instructor-trainers representing 10 countries. On a broader scale, all course information and pedagogical materials have been made open and accessible online. The provided code is maintained in a public GitHub[5] repository, while high-definition recordings of each session are available on YouTube and the HPC in Europe Portal[6] for repeated reference. By adopting such an open-access model, future initiatives could ensure a legacy that extends far beyond the live event.
Further, this workshop was designed to align with specific professional competencies laid out in the EVITA skill tree[7]. The skills obtained by the participating trainers by the end of the workshop are outlined in Figure 4. As a measure of quality assurance, certificates were granted to participant-trainers who had a measured participation of over 80%, using the certificate template proposed under the Castiel2 project.
The skills obtained in this workshop contribute directly to the HPC application developer career path as defined in the Training Baseline 2[8]. By mapping course content to these established frameworks, future organisers could provide participants with clear certification pathways and recognised professional development milestones.

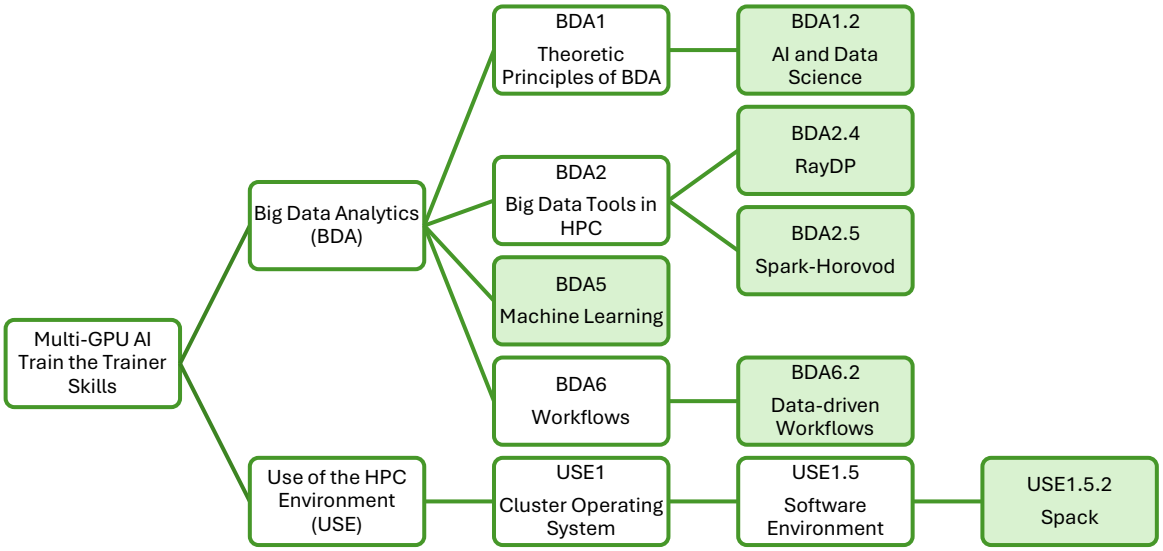

Figure 4: Skills obtained by participating trainers from EVITA skill tree.

Summary and Conclusions

The Multi-GPU AI Train the Trainer Workshop achieved high levels of reported satisfaction among its international cohort. The event was delivered through an online curriculum, with participants granted remote access to the Leonardo Supercomputer. This structure enabled direct progression from theoretical AI frameworks to high-performance computational execution. Key highlights of the workshop included the

[5] https://enccs.github.io/castiel-multi-gpu-ai/
[6] https://hpc-portal.eu/training/material/workshop-multi-gpu-ai-train-the-trainer-day-1-introduction-gpu-architectures-deep
[7] https://www.evitahpc.eu/skill-tree/
[8] https://hpc-portal.eu/training_baseline/training-baseline-2

seamless integration of HPC resources, the efficacy of real-time support via Discord, and long-term utility provided by open access training materials.
Based on the workshop outcomes, future iterations will occur on an annual basis in the *train the trainer* format with updated curriculum to be relevant and forward thinking. Future workshops can be further optimised by shifting the planning cycle to avoid fixed scheduling hurdles and holiday conflicts. Additionally, the curriculum can evolve toward a modular framework or a more specialised focus on advanced topics to better serve the needs of expert trainers. Ultimately, this event underscores how thorough planning, rigorous GDPR compliance, technical preparation and a continuous feedback loop serves as the foundation for scalable, high-quality AI training initiatives within the European HPC community.

Appendix: Captures of Training Delivery

Figure 5: Intro slide for the Multi GPU AI Train the Trainer Workshop with dates, schedule, and organiser logos.

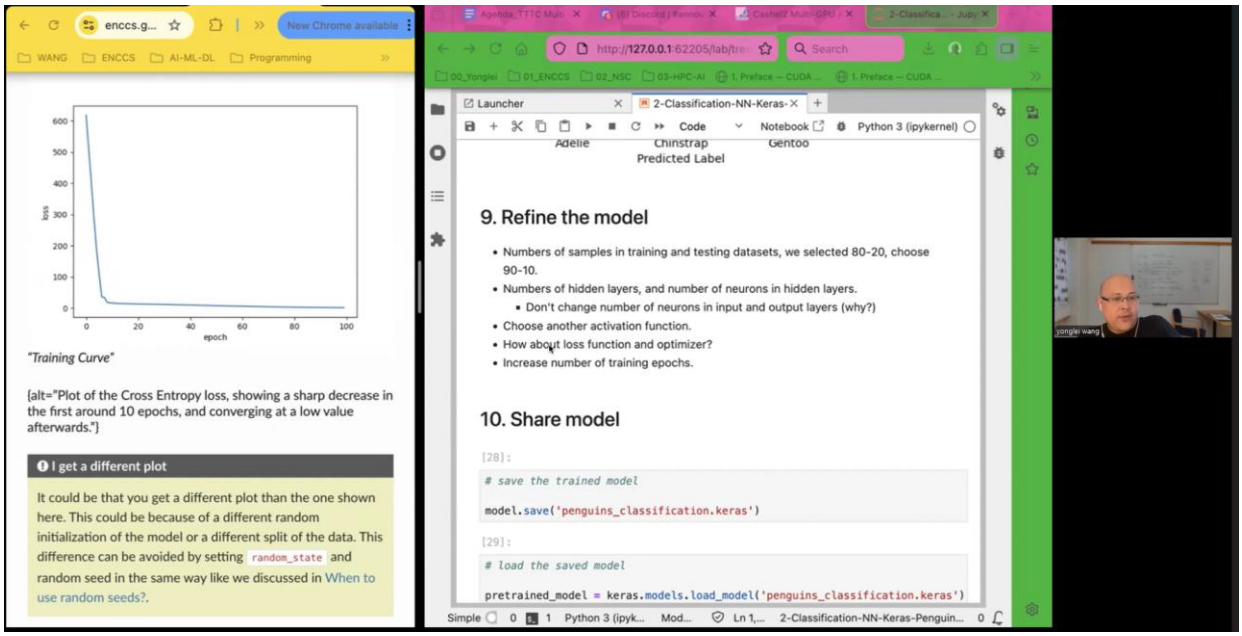

Figure 6: Training session on Introduction to Deep Learning delivered by ENCCS on Day 1.

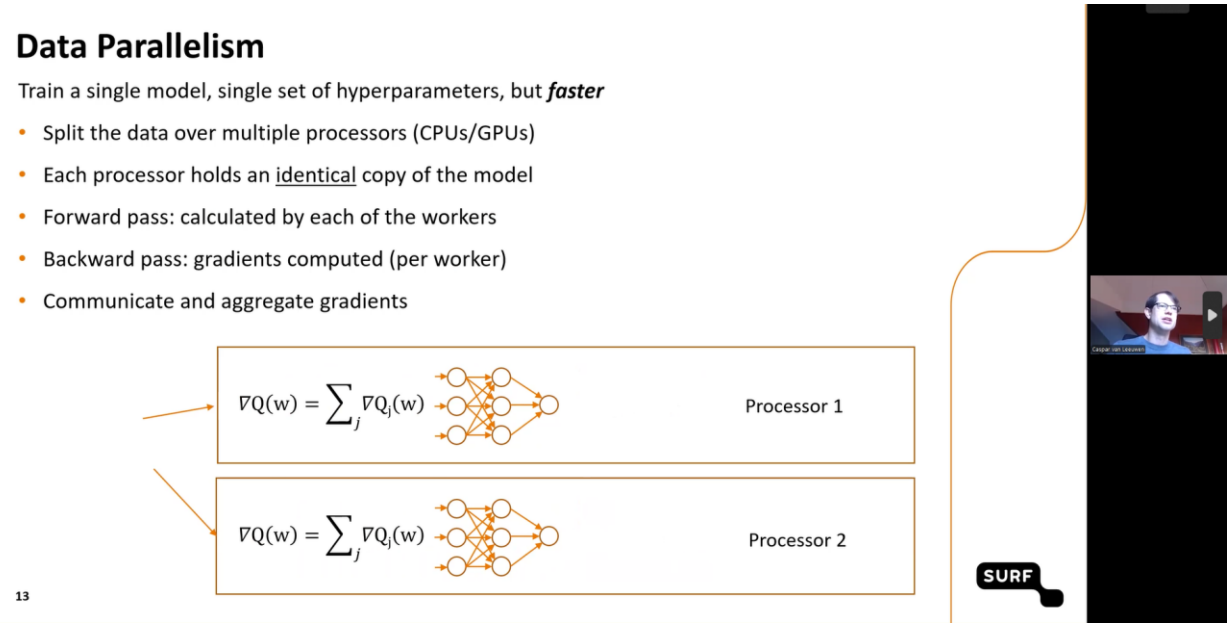

Figure 7: Training session on Pytorch Distributed Data Parallelism delivered by NCC Netherlands on Day 2.

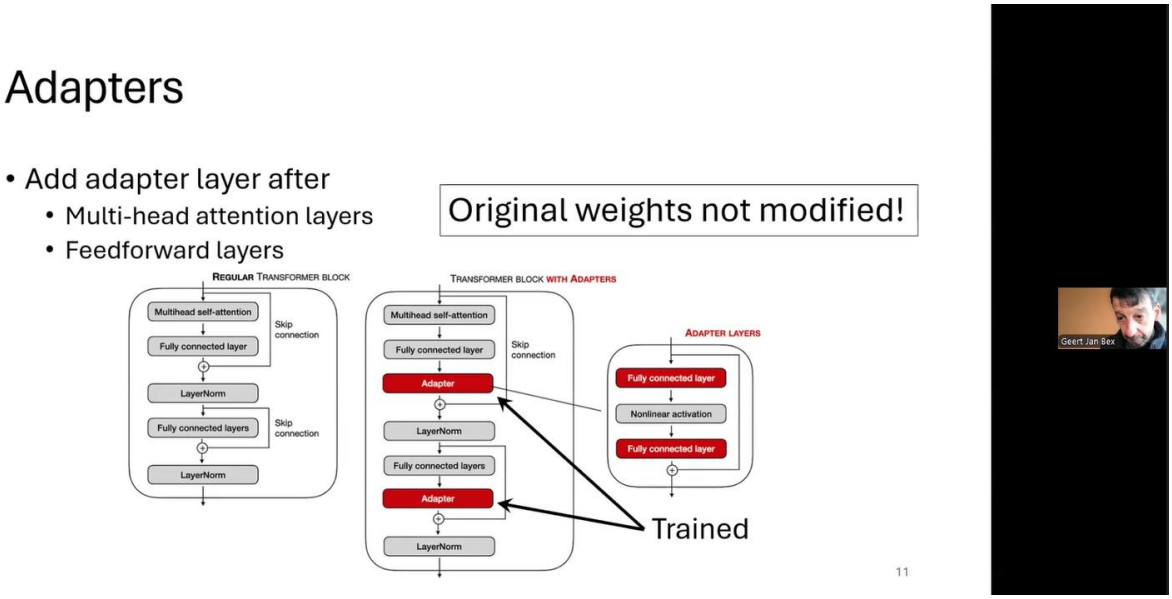

Figure 8: Training session on LLM & Finetuning delivered by NCC Belgium on Day 3.