

## AI-110: Intro to Large Language Models (LLMs) and LLM-based apps.

**Planned Course Length:** 8 training hours

### Course Description:

Artificial intelligence (AI) has become an extremely important area for IT professionals and engineers in the past 10-20 years with the scientific breakthroughs and practical applications of deep learning and more recently of generative AI systems, especially with Large Language Models (LLM) such as ChatGPT. That's why understanding the concepts, and practical usage of AI systems generally and LLMs specifically is becoming essential for all IT and other technical professionals as well as for managers with technical background.

This training focuses on Large Language Models (LLMs) and applications made possible by LLMs and gives an insight into their theory and practice, namely:

- Introduction to LLM based applications
- The Transformer Architecture, Base for all LLMs
- The 3-phase training process of LLMs (Pre-training, Fine-tuning, RLHF)
- Prompt engineering
- Using our own data sources by Retriever Augmented Generation (RAG)
- Creating LLM chains with LangChain
- Web interfaces for LLMs (Gradio or Streamlit)
- Debugging and Evaluating LLM-based apps (Langsmith, Langfuse and TrueLens)

Besides gaining a basic understanding of the theory of Large Language Models (LLMs) as well as other technologies used in LLM-based applications, students will be able to examine their features and play with them during instructor's demonstration and lab exercises.

This training is part of the AI portfolio of Component Soft which explores essential AI topics, such as:

- AI-110 Intro to Large Language Model (LLMs) and LLM-based apps.
- AI-212 Deep Learning Foundations with PyTorch
- AI-434: GPT and Open-source LLM Application Developer

**Structure:** 50% lecture, 25% demonstration by the instructor, 25% hands on lab exercises

**Target audience:** Technical managers as well as IT and telco professionals who want to familiarize themselves with Large Language Models (LLMs) and LLM based applications.

**Prerequisites:** General understanding of and experience in IT systems and/or IT development

## Detailed Course Outline

### **Module 1. Introduction to LLM based applications**

- Main usage areas of LLM-based applications
- Main types of LLM-based applications
- Building blocks of LLM-based applications
- Lab: Testing a simple LLM-based application

### **Module 2. The Transformer Architecture**

- Intuition of the transformer model
- Main elements of transformers: tokenizer, embeddings, encoder, decoder
- Variations on the transformer architecture
- Popular transformer models
- Lab: Testing popular LLM foundational models

### **Module 3. The 3-phase training process of LLMs (pre-training, fine-tuning, RLHF)**

- **Pre-training of LLMs**
- How does pre-training basically work?
- Training data set, computational and financial challenges
- **LLM Fine-tuning techniques.**
- How does fine-tuning basically work?
- Parameter efficient fine-tuning (PEFT) with LoRA and quantized parameters
- **Reinforcement Learning with Human feedback (RLHF)**
- Why do we need RLHF in the first place?
- Methods and main steps of RLHF
- Lab: Examining an LLM family before and after fine-tuning and RLHF

### **Module 4. Prompt engineering**

- What is prompt engineering?
- Prompt engineering terminology and concepts
- The “Just Ask” Principle, Zero-shot prompts
- Prompts with Few-shot learning
- Prompt Chaining
- Chain of Thought Prompting
- Prompts with Personas
- Lab: Practicing basic prompt techniques

### **Module 5. Retriever Augmented Generation (RAG)**

- What is Retriever Augmented Generation (RAG)?
- How does RAG work?
- Syntactic vs. Semantic Similarity
- Text embedding
- Vector Databases
- Lab: Demonstration of the usage of Retriever Augmented Generation (RAG) in an LLM app

**Module 6. Creating LLM chains with LangChain**

- What are LLM chains?
- LangChain architecture
- Main Building Blocks: Models, Prompts and Output Parsers
- Building LLM chains from building blocks
- LangChain Memory
- LangChain Agents
- Lab: Demonstration of the usage of LangChain in an LLM app.

**Module 7. Web interfaces for LLMs (Gradio or Streamlit)**

- Main features of LLM Web interfaces
- Example web app for simple LLM tasks
- Creating our own chatbot with a web interface
- Lab: Demonstration of a simple LLM web interface with Gradio or Streamlit