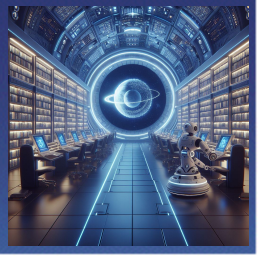
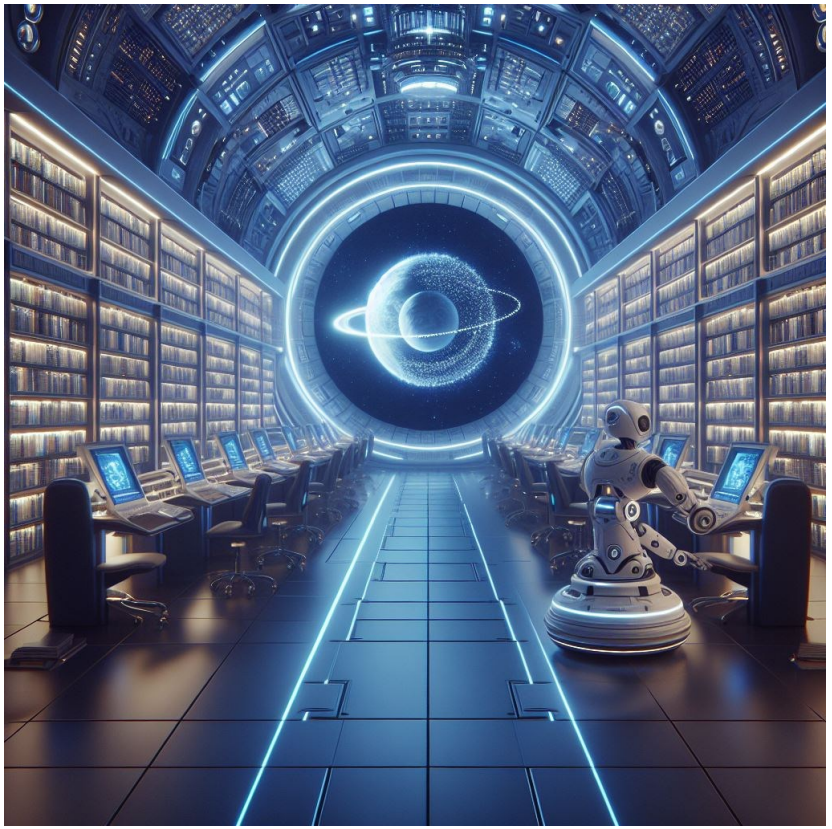


Foundations of Generative AI: Practical Prompting & Advanced Applications

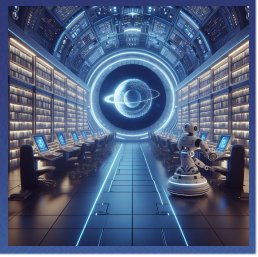


NOTE: most of the images in this presentation have been generated with Microsoft's Image Creator

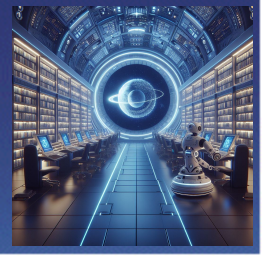
Maurizio Martignano
Spazio IT – Soluzioni Informatiche s.a.s
Via Manzoni 40
46051 San Giorgio Bigarello, Mantova
https://spazioit.com

January 2025

Generative AI



Generative AI

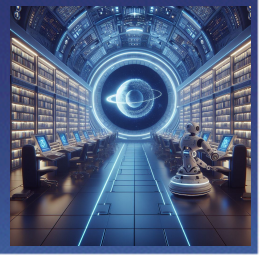


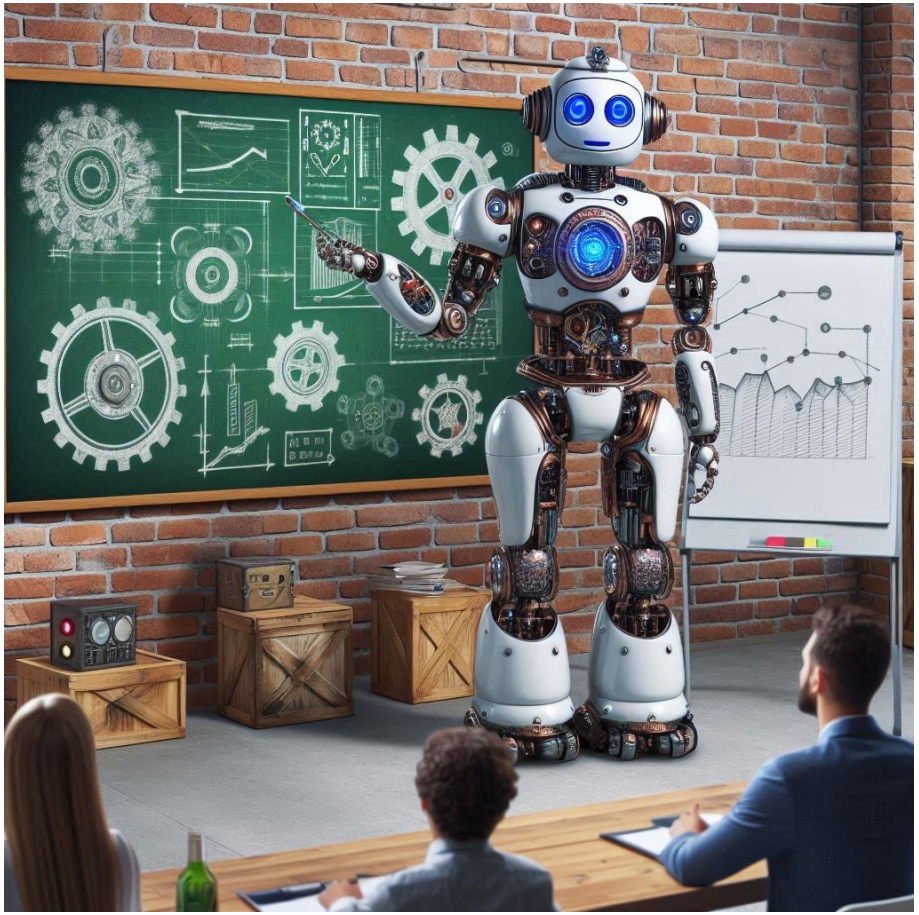
- Generative AI and Large Language Models (LLMs) are related but distinct concepts in the field of artificial intelligence.

- The term **Generative AI** refers to any AI system whose primary function is to **generate content**. This could include a variety of AI models that generate different types of content, such as images, text, code, audio and video. Generative AI emphasizes the content-creating function of these systems.

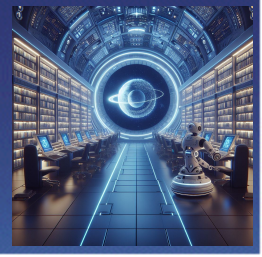
- On the other hand, **Large Language Models** (LLMs) are **a specific type of AI system that works with language**. They are designed to analyze and produce text. LLMs are a form of generative AI, but they specifically deal with text-based content.

Course Introduction



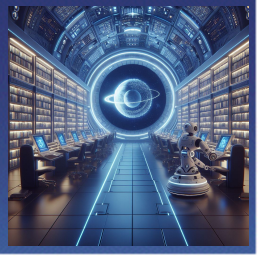


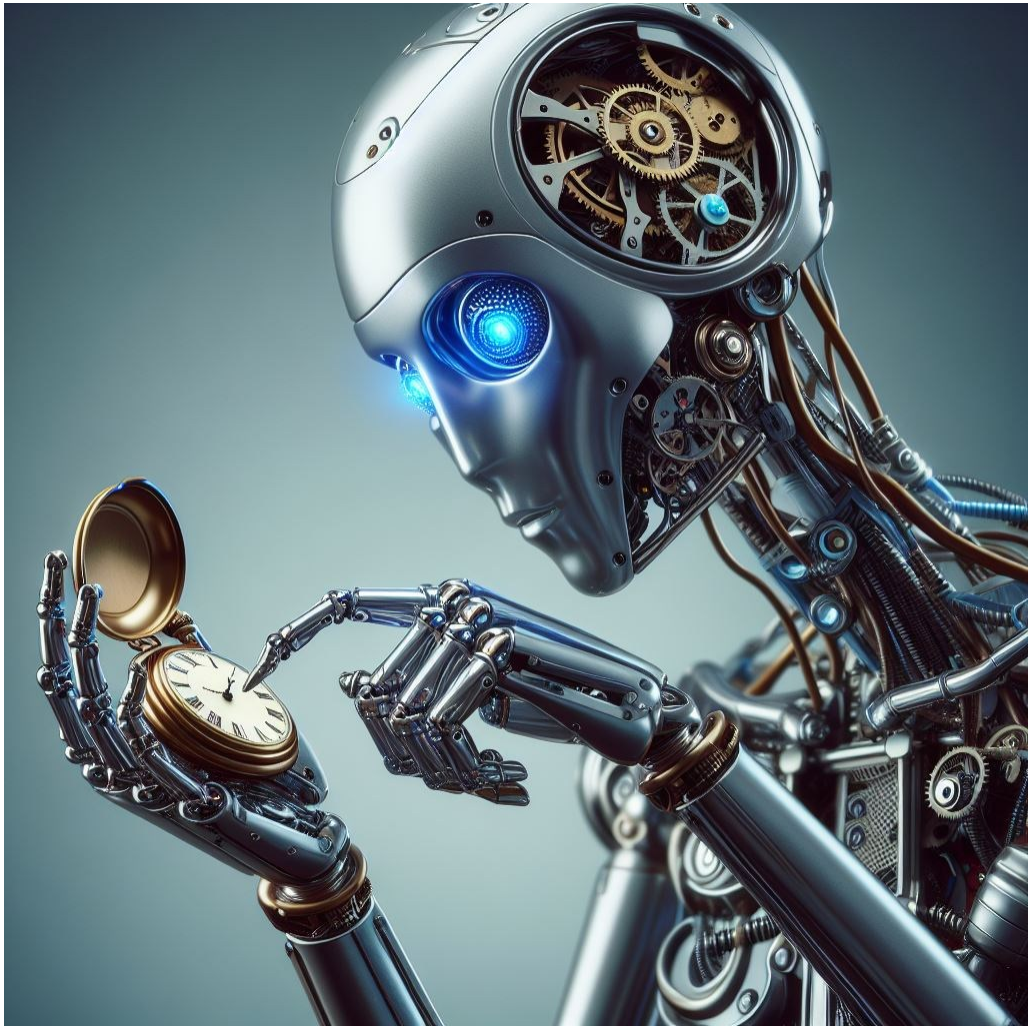
Course Introduction



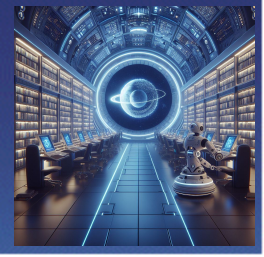
- **Title:** Foundations of Generative AI: Practical Prompting & Advanced Applications
- **Duration:** 2 days (9:00 AM - 5:30 PM)
- **Audience**:
  - The primary audience includes data scientists, analysts, software engineers, and developers. However, business analysts, consultants, product managers, marketing professionals, researchers, entrepreneurs, and startups may also find the course valuable.
  - No prior knowledge of Generative AI is required, and the course is suitable for both beginners and experienced professionals looking to expand their understanding of the field.

Agenda



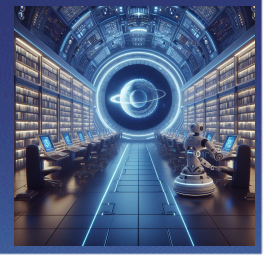


Agenda – Day 1: Foundations and Practical Prompting



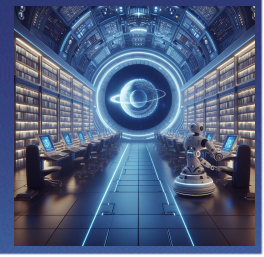
- **9:00 AM - 10:30 AM**: **Introduction and History of Transformers and Large Language Models (LLMs)**
  - **Course Introduction**
  - **Overview of AI Evolution**: Brief history of AI development, leading up to the introduction of neural networks and deep learning.
  - **Transformers**: Explanation of transformer architecture, self-attention mechanisms, and why transformers revolutionized NLP.
  - **Development of LLMs**: Evolution from early NLP models to large-scale models (GPT, BERT, etc.).
  - **Key Milestones**: Major breakthroughs in LLM capabilities, such as GPT-3, GPT-4, and other industry advancements.
- **10:30 AM - 10:45 AM**: **Break**
- **10:45 AM - 12:15 PM**: **Most Popular and Widely Used Generative Chatbots**
  - **Overview of Key Chatbots**: Introduction to popular models like ChatGPT, Claude, Bard, etc.
  - **Feature Comparison**: Capabilities and limitations of each model, including fine-tuning, response generation quality, and API availability.
    Use Cases and Applications: Examples of real-world applications (customer service, content creation, research).
  - **Ethics and Challenges**: Brief overview of ethical concerns (e.g., bias, misuse) and technical challenges in maintaining AI reliability.
- **12:15 PM - 1:15 PM**: **Lunch Break**
  -

Agenda – Day 1: Foundations and Practical Prompting



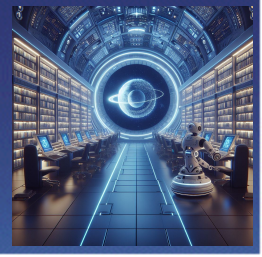
- **1:15 PM - 3:00 PM**: **Hands-On Exercise: Generating Proper Prompts for Text**
  - **Summarizing, Analyzing, and Improving Text**: Introduction to prompt engineering techniques for text-based tasks. **Text Summarization**: Crafting prompts to condense information while retaining meaning.
  - **Text Analysis**: Developing prompts to extract insights or key points from unstructured text.
  - **Improving Text Quality**: Using AI to enhance writing quality, clarity, and flow.
  - **Practice Session**: Guided exercises to build confidence in generating prompts for various text tasks.
- **3:00 PM - 3:15 PM**: **Break**
- **3:15 PM - 5:00 PM**: **Hands-On Exercise: Generating Proper Prompts for Code**
  - **Code Generation**: Writing prompts that guide models to generate code snippets, solve specific coding problems, or write functions.
  - **Code Analysis and Improvement**: Techniques for prompting code reviews, debugging, or enhancing code efficiency.
  - **Practice Session**: Participants engage in exercises, getting feedback on prompt adjustments to achieve the desired output.

Agenda – Day 2: Advanced Applications, Privacy, and Deployment



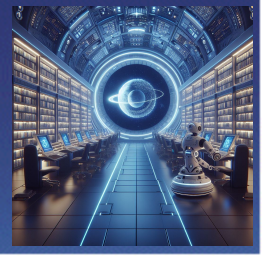
- **9:00 AM - 10:30 AM**: **Hands-On Exercise: Generating Proper Prompts for Data Extraction**
  - **Extracting Data from Unstructured Sources**: Methods for generating prompts to extract structured data (e.g., names, dates, specific details) from unstructured text.
  - **Techniques for Accurate Data Extraction**: Crafting prompts for precision, managing ambiguity in language, and maintaining context.
  - **Practice Session**: Real-world examples where participants practice data extraction from various types of unstructured text.
- **10:30 AM - 10:45 AM**: **Break**
- **10:45 AM - 12:15 PM**: **Hands-On Exercise: Generating Prompts for Image Generation**
  - **Overview of Image Generation**: Introduction to popular image-generative models (e.g., DALL-E) and prompt structures for effective image output.
  - **Prompting Techniques**: Crafting prompts to generate images with specific styles, compositions, or attributes.
  - **Practice Session**: Guided practice on writing prompts and refining results based on initial outputs.
- **12:15 PM - 1:15 PM**: **Lunch Break**

Agenda – Day 2: Advanced Applications, Privacy, and Deployment



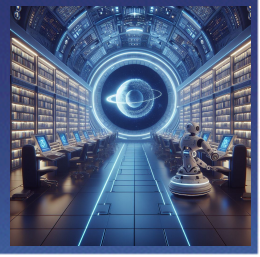
- **1:15 PM - 2:45 PM**: **Privacy Protection and Hallucinations**
  - **Understanding Privacy Concerns**: Overview of data privacy risks in using generative AI models, including inadvertent data exposure.
  - **Understanding Hallucinations**: Overview of hallucinations, i.e. instances when generative models produce nonsensical outputs that appear credible.
  - **Balancing Innovation with Responsibility:** Properly using Generative AI.
- **2:45 PM - 3:00 PM**: **Break**
- **3:00 PM - 4:00 PM**: **Local Execution of AI Models on Disconnected Machines**
  - **Setting Up Local Environments**: Step-by-step guide on configuring local environments for AI model execution (using tools like Docker, Anaconda).
  - **Running AI Models Offline**: Demonstration of running generative models in a local environment for secure, offline use.
  - **Practice Session**: Participants follow along with setup steps or observe demonstrations
- **4:00 PM - 4:15 PM**: **Break**
- **4:15 PM - 5:00 PM**: **Introduction to Retrieval-Augmented Generation (RAG)**
  - **Overview of RAG**: Explanation of combining retrieval and generation techniques for more accurate, context-aware answers.
  - **RAG Use Cases**: Applications in customer support, knowledge bases, and complex question answering.
  - **Practice Session**: Participants try simple RAG tasks, retrieving information to augment generated content.

Agenda – Day 2: Advanced Applications, Privacy, and Deployment



- **5:00 PM - 5:30 PM**: **Improving RAG by Reducing Non-Determinism**
  - **Enhancing RAG Accuracy**: Techniques for tuning RAG systems for higher quality retrieval and response generation.
  - **Strategies to Reduce Non-Determinism**: Methods to manage the variability of responses, including prompt engineering and model adjustments.
  - **Practice Session**: Exercises for refining prompts and model settings to produce consistent outputs.
- **5:30 PM - 6:00 PM**: **Q&A and Wrap-Up**
  - **Summary of Key Takeaways**: Review of the major concepts and skills covered.
  - **Open Floor for Questions**: Opportunity for participants to clarify concepts or revisit any topics.
  - **Feedback and Next Steps**: Discussion on further learning resources and practical steps post-course.

Who am I?



Maurizio Martignano – Spazio IT
https://spazioit.com
https://www.linkedin.com/in/mauriziomartignano

Active since the 1990 (sigh) in Avionics Software Development and Verification
https://spazioit.com/pages_en/sol_inf_en/code_quality_en/

Providing software consultancies in various application domains, e.g. Healthcare, Cybersecurity, Data Protection,…

Not a developer but a **user** of AI Technologies (especially in Healthcare and Software Verification Applications)

Who are you?