

# Generative AI Development Training

3 days (21 hours)

## Presentation

Our Developing Generative AI training course will enable you to discover and master how to create and develop AI. You'll learn how to design, train and operate models capable of generating text, images, sound and code.

You'll start by laying the foundations: how neural networks work, the architecture of Transformers and the discovery of emblematic models such as GPT, BERT or Stable Diffusion. You'll understand how these AIs "learn" and produce content from simple data.

You will then be guided through the collection, cleaning and preparation of your datasets, before moving on to training your own models. Fine-tuning, controlled generation, evaluation and alignment will hold no secrets for you.

You'll also explore the power of multimodal AI: generating images from descriptions, transforming text into music, or writing code in natural language. A real playground for creative and technical minds.

As with all our training courses, this one will feature the latest advances in Generative AI.

## Objectives

- Understand the theoretical and technical foundations of generative AI
- Identify relevant use cases for generative AI development
- Mastering essential tools and frameworks
- Design, train and refine your own generative models
- Building complete projects using generative AI
- Evaluate the quality and relevance of the content generated

- Adopt a posture of reasoned innovation by integrating generative AI into a technical or creative environment

## Target audience

- Developer
- AI Developer
- project manager
- IA Engineer

## Prerequisites

- Basics of machine learning or deep learning
- Python skills

## Program of our Generative ia development training course

### Introduction to Generative AI

- Definition and distinction from discriminative AI
- Classical neural networks to transformers
- The emergence of Deep Learning and Self-Attention
- Successive revolutions: GANs, VAEs, Transformers
- Conversational assistants
- Creative content generation
- Voice / music / visual synthesis
- Professional applications: marketing, prototyping, video games, dev

### Deep Learning Fundamentals

- Artificial neuron, activation, propagation
- Loss function and gradient descent
- CNN, RNN, LSTM
- Encoder-Decoder
- Batching, epochs, learning rate
- Regularization and overfitting

### Transformers and Modern NLP

- Architecture: Attention, Multi-Head Attention, Position Encoding
- Difference with RNN and LSTM
- Tokenization (WordPiece, BPE, SentencePiece)
- Embeddings (word2vec, GloVe, positional)
- GPT, BERT, T5, LLaMA, Mistral

- Comparison of causal vs masked language modeling approaches

## Preparing data for training

- Textual, visual and audio data
- Quality vs. quantity: trade-offs
- Tokenization and standardization
- Filtering, balancing and bias
- TFRecords, Parquet, Datasets HuggingFace
- DataLoader and ingestion pipelines

## Training a generative model

- Training on GPU / TPU
- Choice of framework: PyTorch, TensorFlow, JAX
- Fine-tuning vs entraînement from scratch
- Transfer learning
- LoRA, QLoRA, PEFT
- Perplexity, BLEU, FID, etc.
- Logging with Weights & Biases, TensorBoard
- Checkpoints and recovery

## Text generation with LLM

- Sampling, top-k, top-p (nucleus), temperature
- Efficient and prompt engineering
- Dialogue history, short vs. long memory
- Context management and conversation turns
- Content generation, summaries, rewriting
- Autonomous agents (AutoGPT, CrewAI)

## Multimodal generation

- GANs: Generator vs Discriminator
- Broadcast models
- Text-to-speech (TTS), Voice Cloning
- Music generation (Jukebox, Broadcast)
- Codex, LLaMA Code, StarCoder
- Code recommendation and completion

## Evaluating and improving generative AI

- Originality, coherence, diversity
- Quantitative and qualitative measurements
- Red Teaming & adversarial prompts
- RLHF: Reinforcement Learning from Human Feedback

- Filtering and censorship
- Detecting algorithmic bias
- Mitigation strategies

## Industrializing and deploying AI

- Convert to REST API / gRPC
- Use with FastAPI / Flask
- Quantization, pruning, distillation
- ONNX, TensorRT, CPU inference
- Inference server (vLLM, TGI, Triton)
- Deployment on HuggingFace Hub, AWS, GCP, Azure

## Companies concerned

This course is aimed at both individuals and companies, large or small, wishing to train their teams in a new advanced computer technology, or to acquire specific business knowledge or modern methods.

## Positioning on entry to training

Positioning at the start of training complies with Qualiopi quality criteria. As soon as registration is finalized, the learner receives a self-assessment questionnaire enabling us to assess his or her estimated level of proficiency in different types of technology, as well as his or her expectations and personal objectives with regard to the training to come, within the limits imposed by the selected format. This questionnaire also enables us to anticipate any connection or security difficulties within the company (intra-company or virtual classroom) which could be problematic for the follow-up and smooth running of the training session.

## Teaching methods

Practical course: 60% Practical, 40% Theory. Training material distributed in digital format to all participants.

## Organization

The course alternates theoretical input from the trainer, supported by examples, with brainstorming sessions and group work.

## Validation

At the end of the session, a multiple-choice questionnaire verifies the correct acquisition of skills.

## Sanction

A certificate will be issued to each trainee who completes the course.

---

[Training Program Web page](#) - Appendix 1 - Training sheet

Training organization registered under number 11 75 54743 75. This registration does not imply government approval.  
Ambient IT 2015-2025. All rights reserved. Paris, France - Switzerland - Belgium - Luxembourg