

Glossary: Building GenAI-Powered Apps with Python



Skills
Network

Welcome! This alphabetized glossary contains many of the terms you'll find within this course. This comprehensive glossary also includes additional industry-recognized terms not used in course videos. These terms are important for you to recognize when working in the industry, participating in user groups, and participating in other certificate programs.

Term	Definition
Application Programming Interface (API)	An interface that facilitates communication across applications. APIs help to extract and share data using a set of definitions and protocols.
Bootstrapping Language-Image Pre-training (BLIP)	An AI-based model, used to perform multi-modal tasks like visual question answering, image-text retrieval, and image captioning. It is a pre-training framework for unified vision-language understanding and generation.
BlenderBot	An AI-based chatbot that can converse naturally with people and takes direct feedback to improve its responses.
Chatbot	A computer program that simulates a human conversation with an end user. Though not all chatbots are equipped with artificial intelligence (AI), modern chatbots increasingly use conversational AI techniques like natural language processing (NLP) to make sense of the user's questions and automate their responses.
ChatGPT	A chatbot build developed by OpenAI that uses large language models (LLMs) to enable users to interact and get desired responses.
CodeT5	A text-to-code seq2seq model developed by Google AI trained on a large data set of text and code. CodeT5 is the first pre-trained programming language model that is code-aware and encoder-decoder based.
Cascading Style Sheets (CSS)	A computer language to lay out and structure web pages using codes.
Deep learning	A type of machine learning focused on training computers to perform tasks through learning from data. It uses artificial neural networks.
Falcon	A large language model developed by the Technology Institute of Innovation (TII). Its variant, falcon-7b-instruct, is a 7-billion-parameter model based on the decoder-only model.
Foundation models	AI models with broad capabilities that can be adapted to create more specialized models or tools for specific use cases.
Generative Adversarial Network (GAN)	A type of generative model that includes two neural networks: generator and discriminator. The generator is trained on vast data sets to create samples like text and images. The discriminator tries to distinguish whether the sample is real or fake.
Generative AI models	Models that can understand the context of input content to generate new content. In general, they are used for automated content creation and interactive communication.
Generative Pre-trained Transformer (GPT)	A series of large language models developed by OpenAI designed to understand language by leveraging a combination of two concepts: Training and transformers.
Google flan	An encoder-decoder foundation model based on the T5 architecture.
Gradio	An open-source Python package that allows the building of a demo or web application for machine learning models. It helps to create UIs to demo and deploy models and share them easily.
Hugging Face	An AI platform that allows open-source scientists, entrepreneurs, developers, and individuals to collaborate and build personalized machine learning tools and models.
HyperText Markup Language (HTML)	A standard markup language consists of elements that create web pages, structure them, and help display them. A start tag, some content, and an end tag define HTML elements.
IBM Watson	An integrated AI and data platform with a set of AI assistants designed to scale and accelerate the impact of AI with trusted data across businesses.
IBM Cloud Code Engine	A fully managed, serverless platform that is used to manage and secure the underlying infrastructure of codes, container images, and batch jobs.
Large language model (LLM)	A deep learning model trained on substantial text data to learn the patterns and structures of language. They can perform language-related tasks, including text generation, translation, summarization, sentiment analysis, and more.
Llama	A large language model from Meta AI.
LlamaIndex	A flexible data framework to connect custom data sources to large language models using a central interface.
LangChain	A framework designed to simplify the creation of applications using LLMs that help in document analysis, document summarization, chatbot building, and code analysis.
Natural Language Processing (NLP)	A subset of artificial intelligence that enables computers to understand, manipulate, and generate human language (natural language).
Named-Entity Recognition (NER)	A subtask of information extraction that helps to locate and classify named entities like first and last names, geographic location, age, address, and phone number in unstructured data sources.

Term	Definition
OpenAI Whisper	An automatic speech recognition system trained on 680,000 hours of supervised data that can transcribe speech in several languages.
Python	A high-level, general-purpose programming language that supports multiple programming paradigms, including structured, object-oriented, and functional programming.
Python Imaging Library (PIL)	A versatile Python library that adds image processing capabilities to Python interpreter that helps to perform tasks such as reading, rescaling, and saving images in different formats.
Retrieval-Augmented Generation (RAG)	An AI framework designed to retrieve facts from an external knowledge base to ground large language models (LLMs) that provide information on the latest research, statistics, or news to generative models.
Sentiment analysis	A process of analyzing digital text to determine the emotional tone of a message. Performed on textual data, helping businesses monitor brands through customer feedback.
Streamlit	An open-source framework to build and share machine learning and data science web apps. It turns data scripts into shareable web apps in minutes.
Tokenizer	A tokenizer is a tool in natural language processing that breaks down text into smaller, manageable units (tokens), such as words or phrases, enabling models to analyze and understand the text.
Training data	Data (generally, large data sets that also have examples) used to teach a machine learning model.
Transformers	A deep learning architecture that uses an encoder-decoder mechanism. Transformers can generate coherent and contextually relevant text.
Text generation	A model that is trained on code from scratch that helps to automate repetitive coding tasks.
Unsupervised learning	A subset of machine learning and artificial intelligence that uses algorithms based on machine learning to analyze and cluster unlabeled data sets. These algorithms can discover hidden patterns or data groupings without human intervention.
Variational Autoencoder (VAE)	A generative model that is a neural network model designed to learn the efficient representation of input data by encoding it into a smaller space and decoding it back to the original space.
watsonx.ai	A studio of integrated tools for working with generative AI capabilities powered by foundational models and building machine learning models.
watsonx.data	A massive, curated data repository that can be used to train and fine-tune models with a state-of-the-art data management system.
watsonx.governance	A powerful toolkit to direct, manage, and monitor your organization's AI activities.

Author(s)

Sina Nazeri

Talha Siddiqui

Richa Arora

Vicky Kuo

