

## COURSE GLOSSARY

# Working with the OpenAI API

- **OpenAI:** An artificial intelligence research company that develops advanced AI models, including GPT and DALL·E.
- **OpenAI API:** A programming interface that allows developers to integrate OpenAI's models into their applications.
- **API (Application Programming Interface):** A set of rules and tools that enables communication between software applications.
- **Endpoint:** A specific URL in the API where a request is sent (e.g., Chat Completions endpoint).
- **Chat Completions Endpoint:** An OpenAI API endpoint for generating conversational text responses.
- **Request:** A call made to the API, typically including input text, parameters, and authentication information.
- **Response:** The output returned by the API after processing a request, often in JSON format.
- **API Key:** A unique identifier used to authenticate requests to the OpenAI API.
- **Authentication:** The process of verifying the identity of the user making an API request, usually through API keys.
- **JSON (JavaScript Object Notation):** A lightweight data format used for sending and receiving API requests and responses.
- **Token:** A unit of text (e.g., a word or part of a word) used by language models to process input and generate output.
- **Max Completion Tokens:** A parameter controlling the maximum length of the model's output.
- **Temperature:** A parameter that controls the randomness of model outputs (higher values = more creative, lower values = more deterministic).
- **Zero-shot Prompting:** Asking the API to perform a task without examples.
- **Few-shot Prompting:** Providing examples in the request so the model generates better-targeted responses.
- **System Role:** A type of input message in the Chat API that defines the model's behavior and constraints.
- **User Role:** The input provided by the end user when interacting with the API.
- **Assistant Role:** A role in the API that provides example responses or guides the assistant's behavior.
- **Multi-turn Conversation:** A dialogue where context is preserved across multiple exchanges between user and assistant.
- **Cost of Tokens:** The pricing model for API usage, calculated based on the number of input and output tokens processed.
- **Rate Limits:** Restrictions on how many requests can be sent to the API within a given timeframe.
- **Shot Prompting:** The practice of structuring API prompts with zero-shot, one-shot, or few-shot examples to improve results.