Labs Data & AI Innovation Day

Labs Overview

The purpose of these labs is to build a Retrieval Augmented Generation (RAG) system. This system can then be used to ask questions using Natural Language Queries.

A computer screen shot of a diagram

Description automatically generated

# Components

The following components are used to build the application:

|  |  |
| --- | --- |
| Azure OpenAI service | Chat and embedding models inferencing |
| Azure Cosmos DB for NoSQL | Store products data and vector representations |
| Streamlit + Python | Chatbot front-end |

# Lab 1 – Create environment

Purpose of the lab:

* Prepare a Python virtual environment for all labs
* Create a database (one per team)
* Create a products container (one per team)

# Lab 2 – Ingest data from json files

Purpose of the lab:

* Deploy Azure AI models for chat and embeddings
* Ingest json files (product.json + customer.json) into Azure Cosmos DB for NoSQL, using a Python notebook in Visual Studio Code
* Verify that data was properly ingested

# Lab 3 – Create a web app using Azure AI Studio

Purpose of the lab:

* Use Azure AI Studio to deploy a web app that allows a user to query its own data
* Configure of Azure AI Studio to add your own data source
* (Optional) add voice recognition to the Azure AI Studio
* Deploy a web app front-end application

# Lab 4 – Create a Python chatbot

Purpose of the lab:

* Create a Python chatbot (Streamlit application with Langchain) to query the Azure Cosmos DB for NoSQL products container, using natural language

# Lab 5 – Get statistics on the Open AI Usage

Purpose of the lab:

* Create a report with powerBI desktop who will analyze the usage of your application develop in previous LAB