FactCheckr.AI

Guarding your financial decisions with verified facts.

Submitted for CWB Hackathon 2025

Agenda Overview

01

Background

02

Problem Statement

03

Objectives

04

Technical Architecture

05

Proposed solution

06

Links & Appendices

01. Background

Misinformation and disinformation spread rapidly across digital platforms, influencing public opinion, policy decisions, and trust in institutions. AI-powered solutions can help combat false narratives by verifying content, detecting manipulation, and promoting reliable sources.

Link to details on CWB Hackathon website

02. Problem Statement

01

1. Is the Information Trustworthy?

How can users quickly and confidently determine the **credibility** and potential bias of news sources in an overwhelming digital environment?

02

2. Battling Online Misinformation

In a world saturated with online content, how can individuals **easily identify** and **flag** specific factual claims that contradict known facts or exhibit characteristics of **disinformation**?

03

3. Decoding Sentiment

How can readers efficiently ascertain the overall **sentiment** or underlying tone of a news article, image, or social media post without a deep, time-consuming analysis?

03. Objectives

To offer businesses or individuals a reliable to assess the news source and the factual claims within text and images

01

1. Is the Information Trustworthy?

To provide users with **rapid**, data-driven insights into the credibility of news sources and the objectivity of their content, by generating a justified **credibility score (1-5)** for any given **text** or **image** input

02

2. Battling Online Misinformation

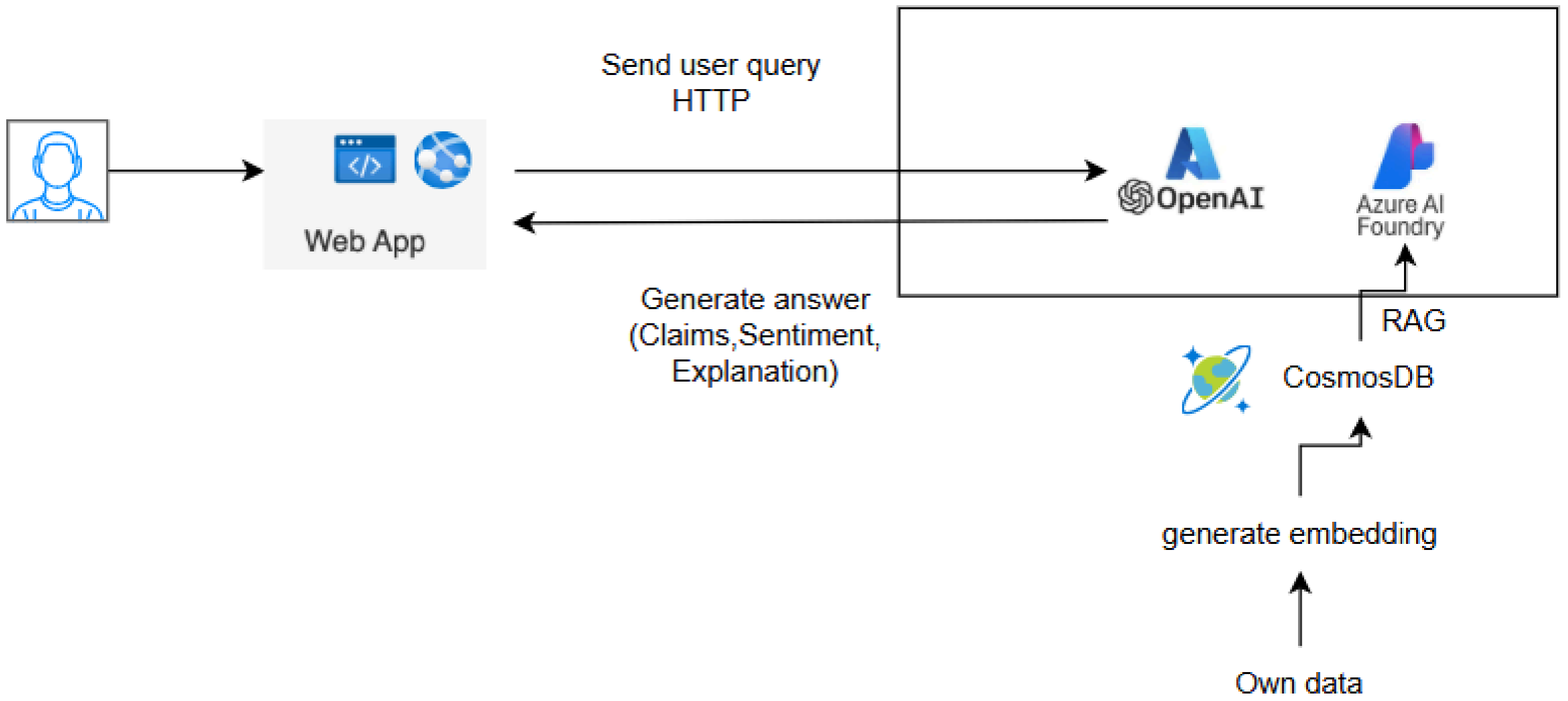
To empower users to proactively identify and understand potential misinformation, by **effectively highlighting factual claims** that **contradict** known **facts** and **flagging** characteristics indicative of **disinformation**.

03

3. Decoding Sentiment

To offer users a **quick and accurate** understanding of the emotional tone and potential biases embedded within news content, by **providing an overall sentiment assessment** (e.g., very positive, neutral, likely negative) for any given input.

04. Technical Architecture



MODEL SELECTION

gpt-4.1

INTRODUCTION

- gpt-4.1 series is the latest iteration of the gpt-4o family.
- Has the context token limit up to 1M input tokens
- Likely train on Public + licensed internet data (books, websites, code, etc.) until 2024
- Comes with Azure Content Safety API

WHY THIS MODEL?

- Latest GPT-4 model - better performance in language understanding, summarization, reasoning and conversation
- Enterprise-Grade Infrastructure - hosted on Azure
- Fine-tuning & embedding support

THINGS TO TAKE NOTE

- Cost for API usage
- Limited customization compared to other open-source models

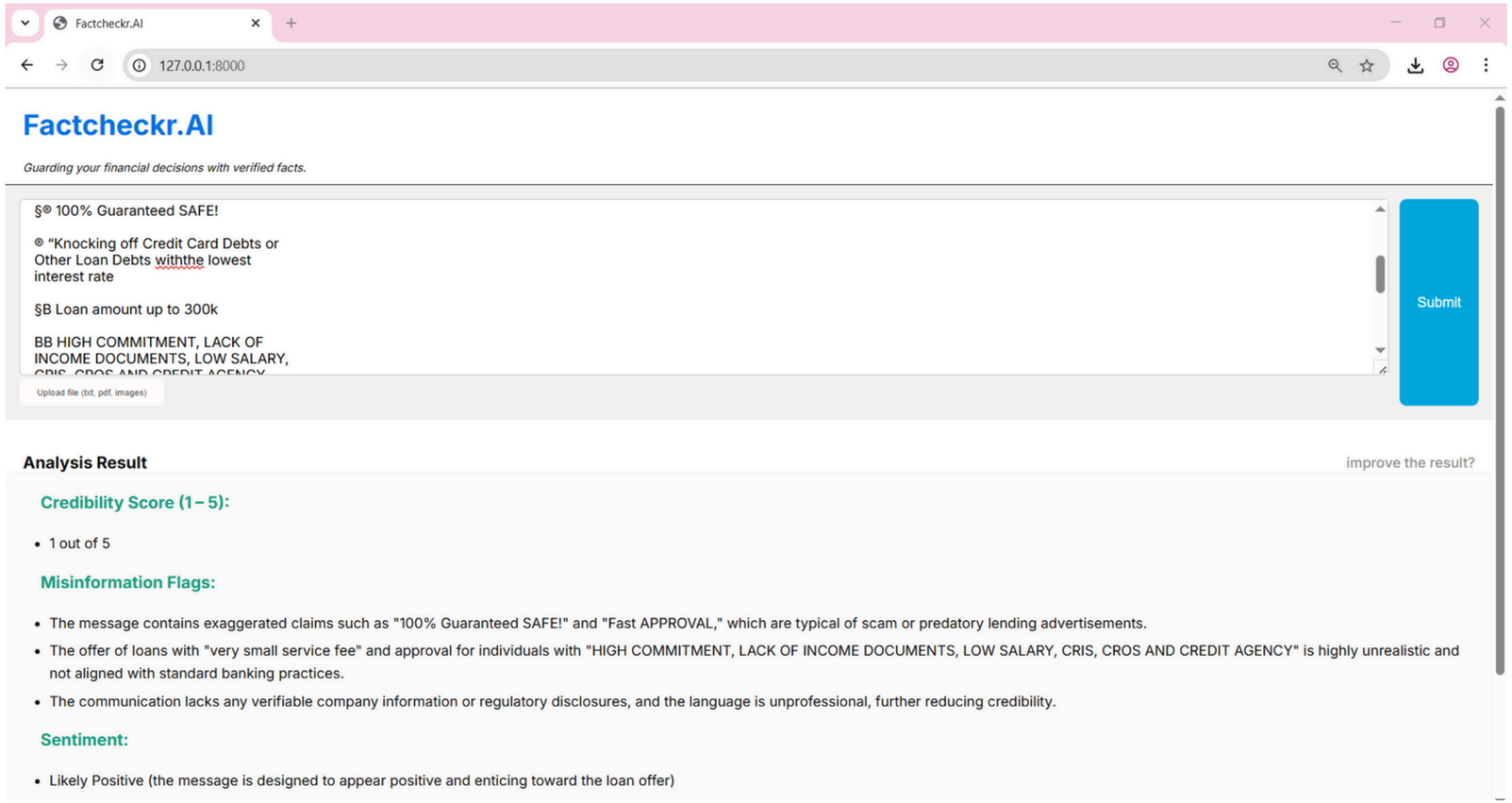
INCLUDE MORE DATA

In addition to general knowledge of GPT-4.1

- To overcome the limitations of large language models (like GPT-4.1) regarding real-time or custom data, we can extend or simulate training using methods like RAG, embeddings, or APIs.
- The embedding & docs can be stored in CosmosDB
- Sample data used: Fin-Fact
- Retrieval Augmented Generation (RAG) can be used to leverage CosmosDB embedding with GPT-4.1. RAG woks by
  - Retrieving relevant info
  - Vector search in Cosmos DB
  - Contextual prompting
  - Gen AI responses: GPT4.1 uses its general knowledge and the provided context to create an answer.
-

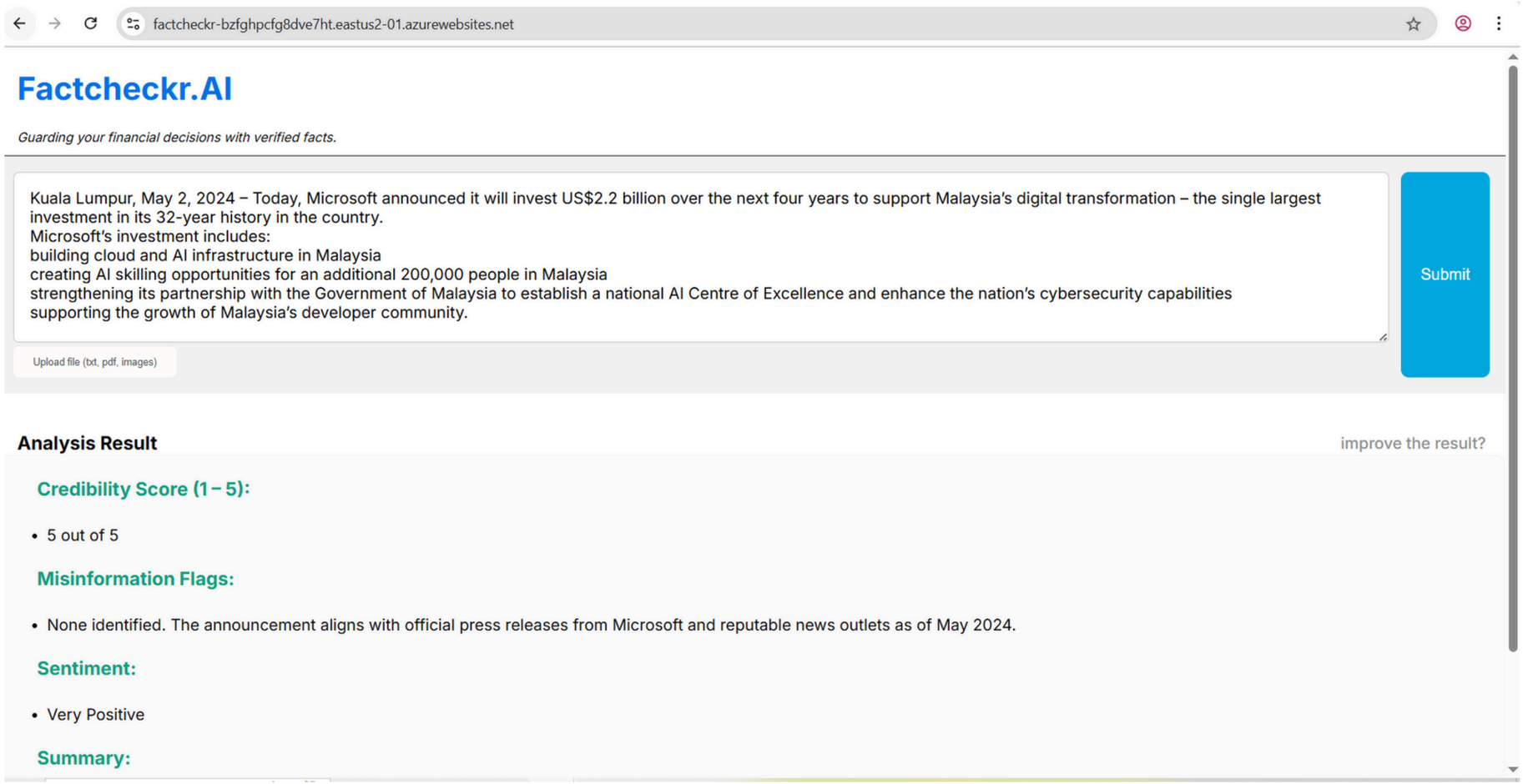
05. Proposed Solutions

On local machine



Proposed Solutions

Deployed on Azure Web App



Future Improvements

1. MORE UP-TO-DATE & ACCURATE DATA

- Improve the results with more up-to-date & accurate labeled data

2. FEATURES TO ALLOW MORE USER INTERACTIONS

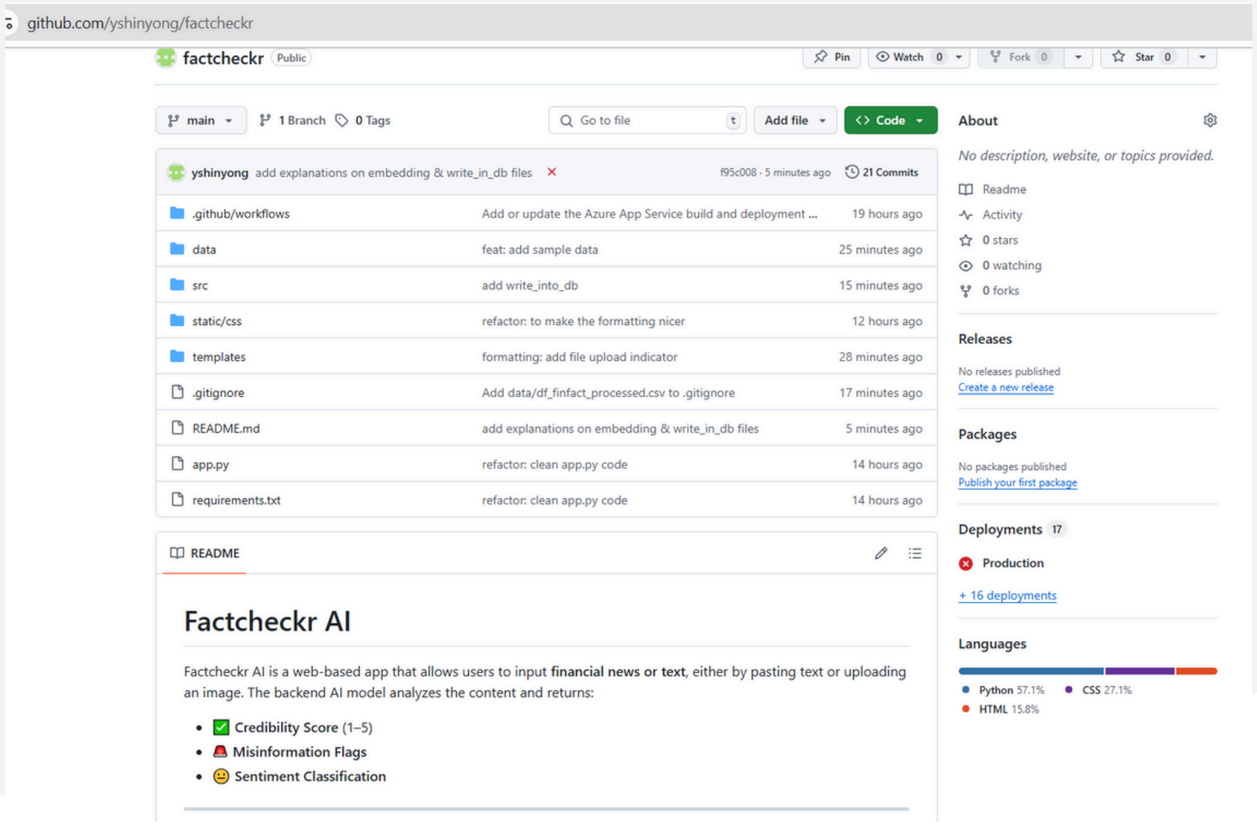
- Allow users to label and tag data for model improvement
- Support ongoing conversational interaction with the user

3. PRIORITIZE HIGH-RISK FINANCIAL MISINFORMATION

- Given the overwhelming volume of financial information published daily, I recommend prioritizing improvements in categories most prone to fraud — such as investment opportunities, market manipulation, and loan-related misinformation

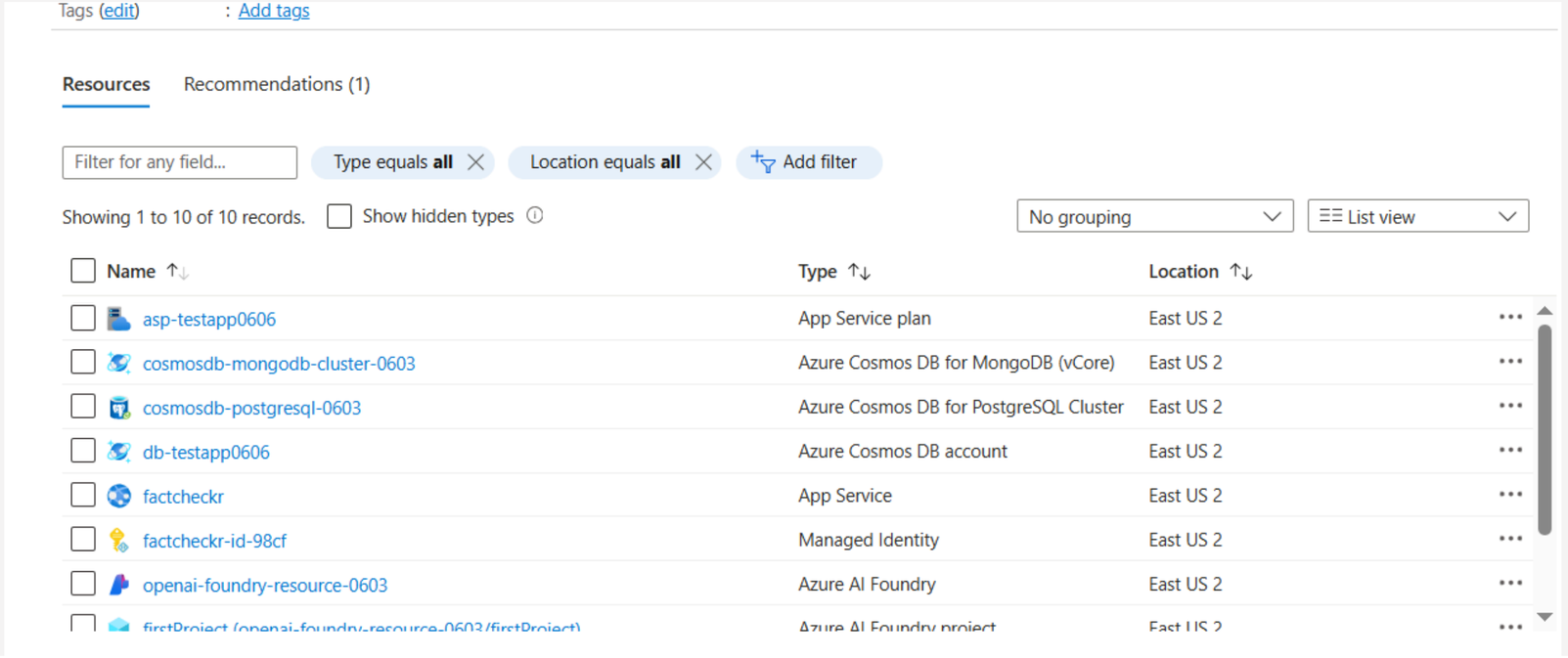
06. LINKS & APPENDICES

1. GitHub repository



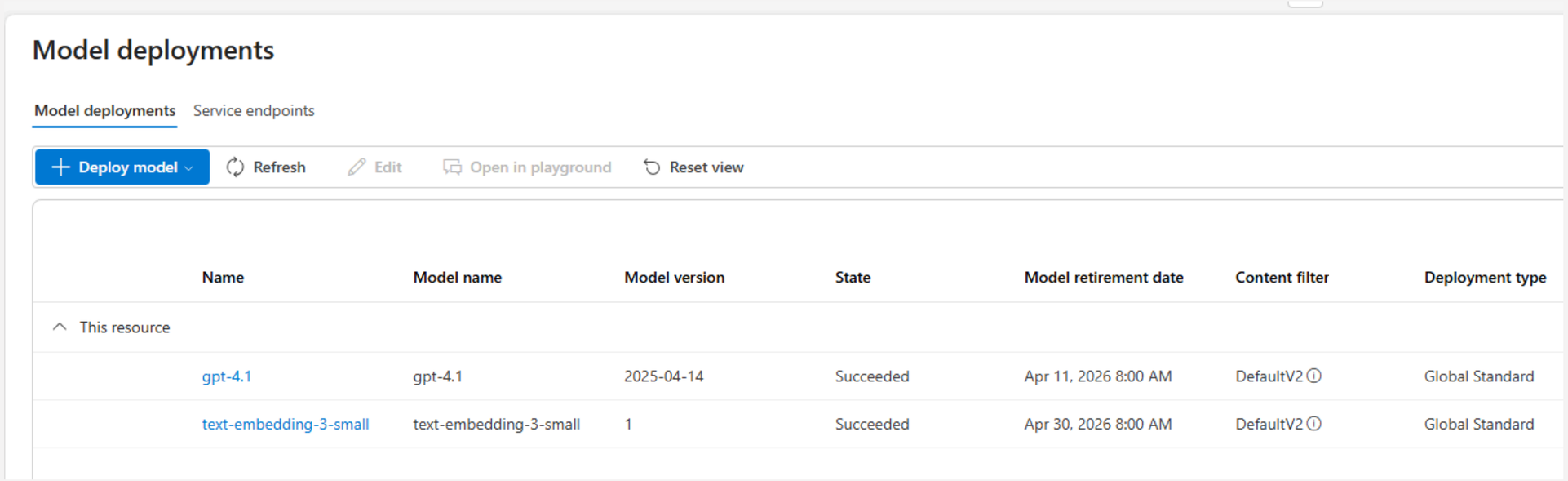
06. LINKS & APPENDICES

Services used on Azure Portal



06. LINKS & APPENDICES

Models Deployed



THANK YOU