

Guida alla Creazione di un RAG Chatbot Avanzato

Questa guida ti accompagnerà passo dopo passo nella creazione di un chatbot RAG avanzato usando Azure AI Foundry e Azure AI Search.

Prerequisiti

- Una sottoscrizione Azure attiva
 - Python installato
 - Conda installato
 - [Azure CLI](#) installata
 - Codice del progetto (con `requirements.txt`, `config.py`, `create_indexer.py`, `chat_with_docs.py`, etc.)
-

1. Creazione delle risorse su Azure

1. Accedi al [portale di Azure](#).
 2. Crea un nuovo **Resource Group**.
 3. All'interno del Resource Group, crea una risorsa **AI Foundry**.
-

2. Deploy dei modelli su AI Foundry

1. Vai nel portale di **AI Foundry**.
 2. Crea un nuovo **progetto AI Foundry** dentro l'hub appena creato.
 3. Esegui il **deploy del modello gpt-4o-mini**, assegnando al deployment lo **stesso nome del modello** (`gpt-4o-mini`).
-

3. Crea una risorsa AI Search

Nel Resource Group, crea una risorsa **Azure AI Search**.

4. Collega Azure AI Search ad AI Foundry

1. Vai al tuo progetto AI Foundry.
 2. Dal menu a sinistra, seleziona **Management Center**.
 3. In **Connected resources**, verifica se c'è già un collegamento ad Azure AI Search.
 4. Se non c'è, clicca su **New connection > Azure AI Search**.
 5. Seleziona il tuo servizio di AI Search e aggiungi la connessione.
 6. Usa **API key** per l'autenticazione.
-

5. Crea una risorsa Azure OpenAI

1. Crea una risorsa **Azure OpenAI** nello stesso Resource Group.
 2. Dal portale di AI Foundry, entra nella risorsa OpenAI.
 3. Esegui il **deploy del modello text-embedding-ada-002**, usando il nome del modello come nome del deployment.
-

6. Configura l'accesso ai modelli da Azure AI Search

1. Vai al **servizio di Azure AI Search**.
2. Seleziona **Identity** dal menu a sinistra.
3. Attiva **System assigned identity** e salva.
4. Vai alla risorsa **Azure OpenAI**.
5. Apri il pannello **Access control (IAM)**.
6. Aggiungi il ruolo:
 - **Ruolo:** Cognitive Services OpenAI User

- **Assegnato a:**

- La **managed identity** del tuo servizio Azure AI Search
- Il tuo **utente personale**

7. Conferma con **Review + Assign**.

7. Carica i documenti nel container

1. All'interno dello **storage account**, crea un **container** chiamato:
documents-pdf-files
 2. Carica i tuoi documenti PDF all'interno.
-

8. Configura le variabili d'ambiente

Nel file `.env`, inserisci i valori corretti relativi alle risorse del tuo progetto.

9. Installa le dipendenze e configura l'ambiente

</> Shell

```
1 pip install -r requirements.txt
2 az login
3 python config.py
```

10. Crea e avvia l'indexer

Esegui lo script per creare l'indice:

</> Python

```
1 python create_indexer.py
```

Se non ci sono errori, l'indice è stato **creato e popolato**. Attendi qualche minuto per avere la certezza che il processo sia terminato.

Puoi verificare lo stato dell'operazione di indicizzazione tramite accedendo alla risorsa **Azure AI Search** nel portale.

11. Avvia il chatbot

Esegui lo script principale:

```
1 python chat_with_docs.py
```

</> Python

Ora puoi interagire col tuo chatbot eseguendo query sui tuoi documenti.

Debug: Attiva il logging dettagliato

Per abilitare i log di debug nel file `config.py`, cambia questa riga:

```
1 logger.setLevel(logging.INFO)
```

</> Python

in:

```
1 logger.setLevel(logging.DEBUG)
```

</> Python

Vedrai a terminale:

- L'intento riconosciuto
- La query ottimizzata
- I chunk più rilevanti ritornati da Azure AI Search

