

Workshop AI (day 5)

Microsoft-AGIC-UniPadova

Azure Content Understanding

Azure Content Understanding è un servizio attualmente in **public preview**. È una risorsa **multimodale**, in grado di processare non solo testi, ma anche contenuti visivi, audio, video e altro ancora.

Questo tool è particolarmente efficace nell'**interpretazione dei grafici** e nella comprensione del loro significato, un ambito in cui gli LLM standard fanno ancora molta fatica.

Utilizzo di Content Understanding su AI Foundry

Poiché il servizio è in **public preview**, è disponibile solo se la risorsa Azure AI Service è stata creata nelle regioni supportate (vedi la [documentazione ufficiale](#) per l'elenco aggiornato).

In **Foundry**, accedere alla sezione "**Servizi di intelligenza artificiale**", quindi selezionare l'area dedicata alla "**Comprensione dei contenuti**".

Da qui è possibile avviare la creazione di un nuovo **progetto di Content Understanding**. Il sistema provvede automaticamente alla generazione delle risorse Azure supplementari necessarie al funzionamento del servizio.

Test: analisi di un PDF

Dopo il caricamento di un file PDF, il tool suggerisce automaticamente l'utilizzo del modulo "**Document Analysis**". In questa fase è possibile definire lo **schema del documento** utilizzando il linguaggio naturale.

Successivamente, una volta eseguito lo strumento, si visualizzano i risultati nell'interfaccia grafica.

Nel caso in cui il sistema interpreti erroneamente uno dei campi, è possibile **intervenire manualmente**. Le correzioni vengono apprese dal sistema per migliorare le analisi successive.

Una volta definito lo schema, si può procedere alla creazione di un **build analyzer**, che genera un **endpoint REST**. Questo consente di eseguire da codice la stessa analisi disponibile via interfaccia utente. L'output dell'analisi è un file **JSON** contenente varie informazioni tra cui:

- Il livello di **confidence** per ciascun campo
- I **valori estratti**
- La struttura dello schema utilizzato

Poiché il servizio è ancora in **preview**, al momento **non è disponibile un SDK dedicato**. L'interazione programmatica avviene quindi tramite chiamate alle API REST.

Test: analisi audio

Dopo il caricamento di un **file audio**, invece, Foundry propone alcuni **template predefiniti**, tra cui:

- *Speech Transcript Analysis*
- *Conversation Analysis*
- *Post Call Analytics*

Durante il workshop abbiamo provato il template **Post Call Analytics**.

Lanciando l'analisi, lo strumento riesce a **comprendere il contenuto** e ad **estrarre correttamente lo schema** (che può essere predefinito o personalizzabile), includendo elementi come il **riassunto della telefonata**,

l'**argomento della conversazione**, l'**emozione della telefonata** (classify) e altri campi.

Anche in questo caso è possibile creare un build analyzer relativo.

Confronto tra AI Agent e Content Understanding

Abbiamo condotto un esperimento in due parti:

1. **Uso di un Agent:** abbiamo fornito ad un agent un documento contenente grafici. Alla domanda su un grafico specifico, l'agent non è stato in grado di rispondere correttamente.
2. **Uso di Content Understanding:** abbiamo analizzato lo stesso documento con l'uso di content understanding, avendo definito uno schema personalizzato:
 - **Tipo di grafico (classify)**
 - **Sintesi del contenuto**
 - **Tabella in formato Markdown**

In questo caso il tool è riuscito a **comprendere il grafico**, interpretarne il significato e fornire una risposta coerente. Questo mostra chiaramente la superiorità del servizio nel trattare contenuti strutturati e visivi rispetto a un classico agent LLM.