

Workshop AI (day 4)

Microsoft-AGIC-UniPadova

Document Pre-processing per migliorare i RAG

Il document pre-processing è un passaggio chiave per migliorare i sistemi RAG. In particolare ci concentriamo sull'uso di **Azure Document Intelligence** (ex Form Recognizer), un servizio consolidato per l'analisi automatica dei documenti tramite OCR, modelli pre-addestrati, modelli personalizzati e altri tool.

A differenza delle tecnologie di generative AI più generali, che non sono in grado di valutare autonomamente la qualità dei risultati prodotti, i modelli **domain-specific** (come Azure Document Intelligence) sono progettati per rispondere a esigenze più precise e formali. Questi modelli operano seguendo regole più rigide, producendo output altamente affidabili e accompagnati da un **punteggio di confidenza**, che permette di misurare in modo oggettivo la qualità del risultato. Proprio grazie a questa capacità di "auto-valutazione", risultano particolarmente adatti nei contesti in cui è necessario estrarre informazioni in maniera strutturata e verificabile.

Azure Document Intelligence

Il primo passo per utilizzare questa tecnologia è la creazione di una **risorsa Azure Document Intelligence**. Non ci sono vincoli legati alla regione, trattandosi di un servizio consolidato e **disponibile a livello globale**. È possibile iniziare con un piano gratuito, che consente di elaborare solo le prime due pagine di ciascun documento. Per scenari più estesi, è disponibile un piano a pagamento basato sul consumo effettivo:

- **Read APIs:** ~<0,20€ ogni 100 pagine
- **Layout APIs:** ~<1€ ogni 100 pagine
- **Add-on capabilities:** ~<0,50€ ogni 100 pagine

Il servizio è **serverless**, quindi si paga soltanto in base all'uso reale, senza costi fissi. È inoltre possibile modificare il piano selezionato in qualsiasi momento, accedendo alla sezione "*Piano tariffario*" della risorsa. Maggiori dettagli sono disponibili nella [pagina ufficiale dei prezzi](#).

Document Intelligence Studio

Una volta creata la risorsa su Azure, si può accedere a **Document Intelligence Studio**, l'interfaccia grafica che consente di testare e configurare i modelli disponibili (evitare il modello "General Document" che è deprecato).

Prebuilt Models

Questi modelli sono pronti all'uso e ottimizzati per analizzare **formati di documenti standard**, come scontrini, biglietti da visita, fatture, documenti d'identità, contratti, ecc. Sono ideali per applicazioni rapide in cui non è necessario un addestramento personalizzato.

Custom Models

Questi modelli possono essere **addestrati su documenti specifici**, adattandosi a esigenze aziendali particolari. Hanno il vantaggio di richiedere un numero molto ridotto di documenti per l'addestramento iniziale. Le funzionalità principali includono:

- **Extraction:** estrazione di informazioni.
- **Classification:** classificazione dei documenti in categorie. È buona prassi prevedere anche una categoria "**other**", che il modello possa utilizzare per gestire documenti non riconducibili alle classi principali.

Per utilizzare i modelli personalizzati, è necessario rispettare alcuni [requisiti](#).

Durante il workshop abbiamo provato alcuni degli strumenti messi a disposizione da Document Intelligence:

Test: OCR / Read

Eseguendo un test della funzionalità **OCR/Read** si nota che il sistema restituisce un **confidence score** ([Confidence score: come interpretarlo](#)), espresso con un valore compreso tra 0 e 1, associato a ciascuna parola riconosciuta in un dato documento. Questo punteggio rappresenta il livello di certezza del modello rispetto alla correttezza del riconoscimento, ed è molto utile per valutare in modo formale la qualità del risultato.

Test: Layout

Quando si esegue un test utilizzando il modello **Layout**, oltre all'estrazione del testo, il sistema è in grado di riconoscere anche la **struttura visiva del documento**. Questo include elementi come **tabelle, grafici, immagini** e **selection marks**, offrendo quindi un'analisi più completa rispetto alla semplice lettura del contenuto.

Durante i test, però, si osservano ancora alcune **limitazioni nell'analisi di grafici e figure complesse**. Per questi casi più avanzati, sarà necessario affidarsi al servizio **Content Understanding**, pensato per offrire una comprensione più profonda e contestuale dei contenuti visivi all'interno dei documenti.

Chiavi ed Endpoint per l'utilizzo in modalità code-first

Su Azure, all'interno della sezione "**Keys and Endpoint**" della risorsa Document Intelligence, è possibile recuperare le informazioni necessarie per l'integrazione del servizio. In particolare, è sufficiente disporre di una **chiave di accesso** (key) e dell'**endpoint**. Questi due elementi sono tutto ciò che serve per utilizzare Document Intelligence anche in modalità **code-first**.

Altri link utili

Repository con codice di test: [ElisaPiccin/doc-intelligence-upskilling](https://github.com/ElisaPiccin/doc-intelligence-upskilling)

Add on capabilities: [Document Intelligence add-on capabilities](#)