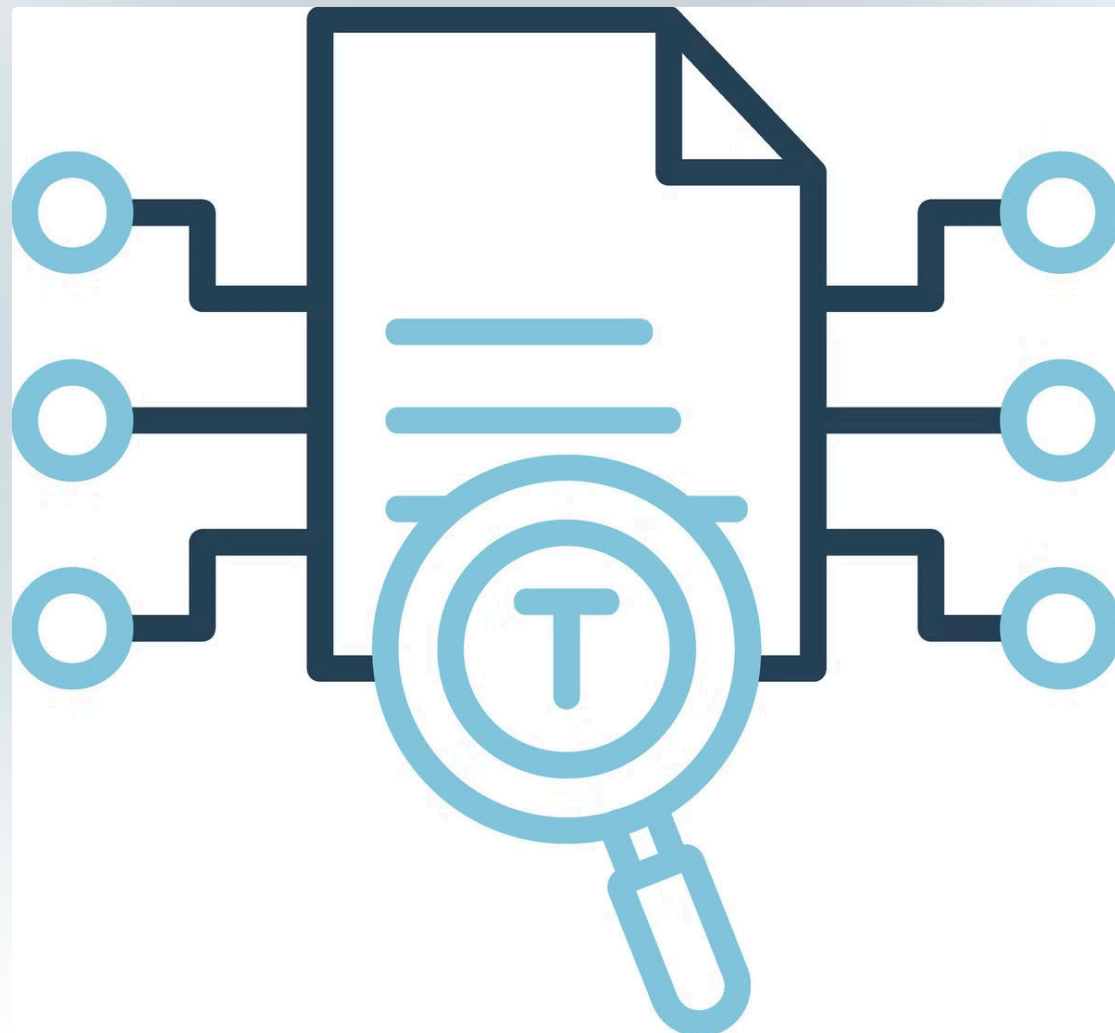


# PROGETTO TEXT MINING

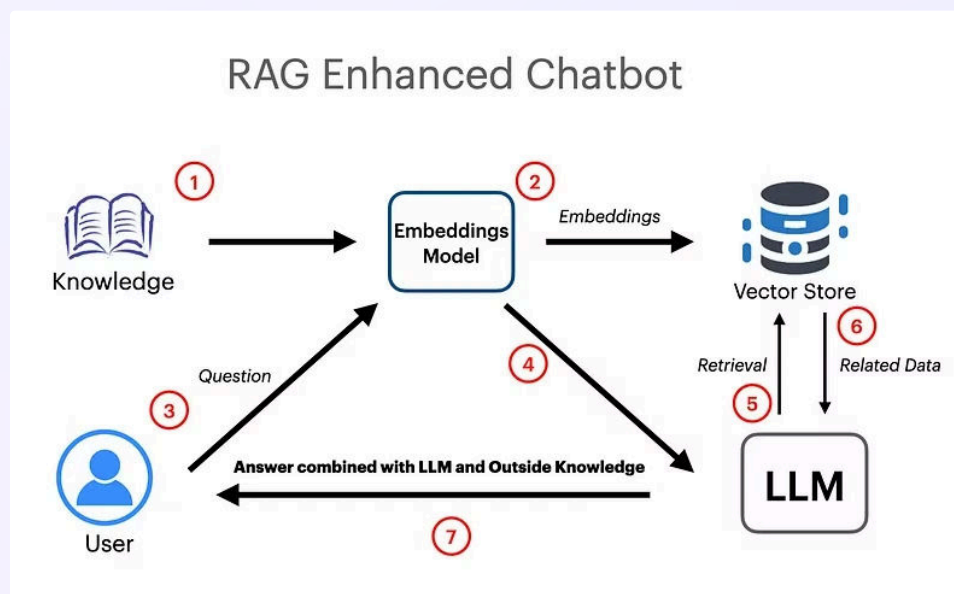
## Gruppo SGPT

- 1 Romano Vito  
M63001504
- 2 Cimmino Marco  
M63001528
- 3 Marano Stefano  
M63001428
- 4 Russo Paolo  
M63001426



# Introduzione

Un chatbot è un software progettato per simulare conversazioni con utenti umani, cercando di utilizzare un linguaggio naturale. Ormai fanno parte della vita quotidiana, basti pensare ad un qualsiasi Servizi Clienti, il quale è capace di rispondere alle domande più frequenti. Il funzionamento del chatbot è paragonabile ad un ciclo if-else, in quanto sono presenti alcune regole predefinite e una corrispondenza della parole chiavi. I primi chatbot erano abbastanza statici, in quanto si basavano su un database di risposte predefinite. Sostanzialmente un'utente inseriva una query, il chatbot cercava le regole associate alla domanda e forniva la risposta associata. Oggi esistono chatbot basati su modelli di linguaggio di grandi dimensioni **LLM** (Large Language Model), fornendo iterazioni più avanzate e dinamicità. Esistono di due tipi:



- **Chatbot basati su RAG**

Un chatbot **RAG** (Retrieval-Augmented Generation) combina tecniche di recupero delle informazioni con modelli generativi per fornire risposte precise e pertinenti.

Estrae dati da un database e utilizza un modello linguistico avanzato per generare risposte contestuali e coerenti, risultando particolarmente efficace per compiti che richiedono dati aggiornati o specifici.

- **Chatbot non basati su RAG**

Utilizza un modello di linguaggio pre-addestrato, come GPT-4, per generare risposte basate sulla comprensione del contesto e delle informazioni apprese durante l'addestramento.

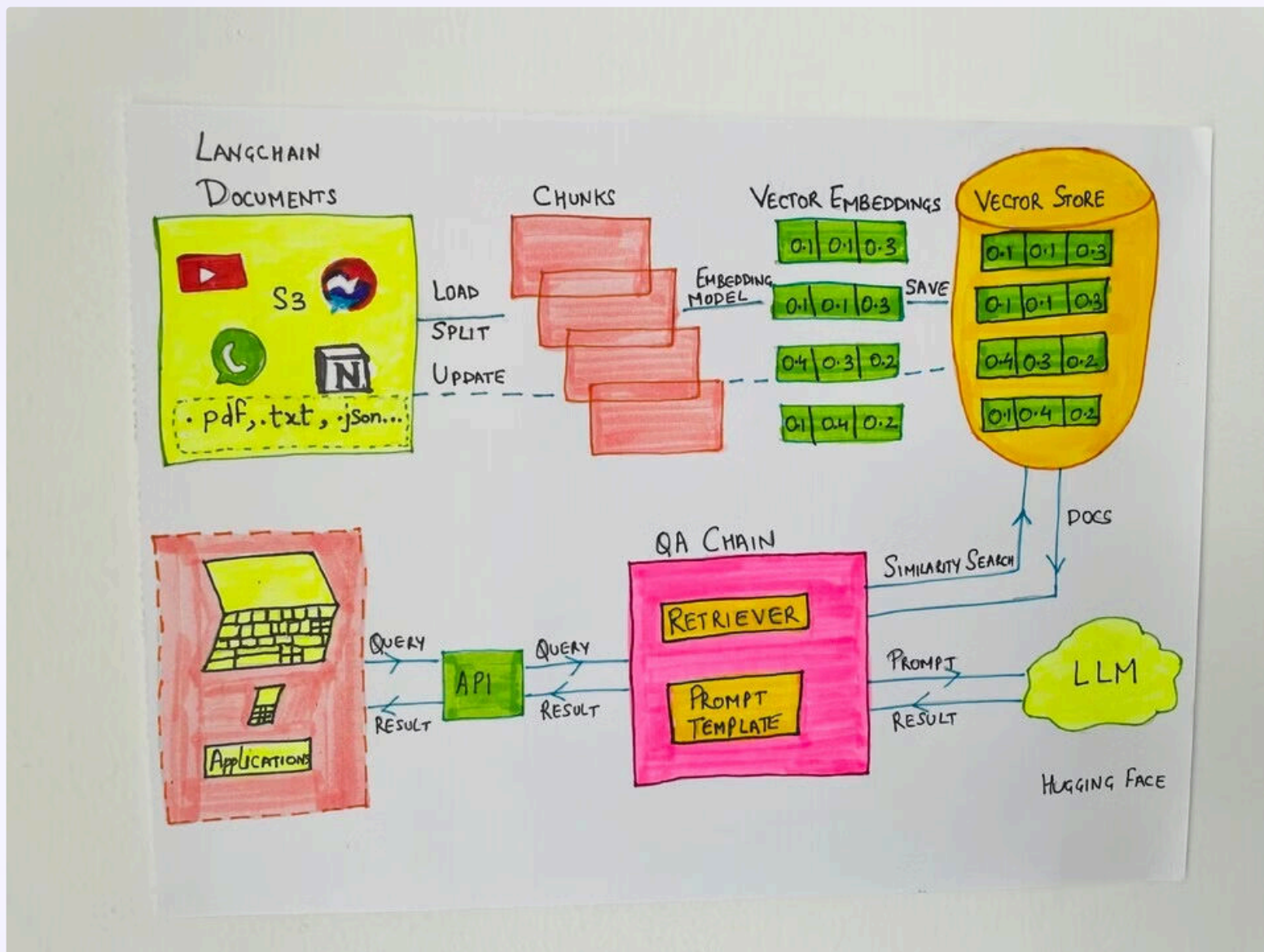
Le risposte sono create esclusivamente sulla base del testo d'ingresso e delle conoscenze interne del modello.

Il chatbot presentato in questo elaborato utilizza RAG e sarà descritto nel punto successivo.

# FUNZIONAMENTO

Il **Chatbot RAG** (Retrieval-Augmented Generation) è un tipo avanzato di chatbot che utilizza una combinazione di tecniche di recupero delle informazioni e modelli di generazione del linguaggio per fornire risposte precise e pertinenti agli utenti. Questo approccio permette al chatbot di sfruttare una vasta gamma di fonti di dati per generare risposte contestuali e coerenti, andando oltre le semplici risposte predefinite.

Un esempio di architettura del Chatbot RAG è il seguente:



## Nella prima fase

I documenti di input, che possono includere PDF, pagine web e altre fonti di informazione, vengono suddivisi in piccoli frammenti chiamati **chunk**. Questi chunk vengono quindi convertiti in **vettori di embedding** utilizzando un modello di deep learning pre-addestrato. Questo processo di codifica trasforma il testo in rappresentazioni numeriche, che possono essere facilmente memorizzate e consultate in un database apposito.

Grazie a questa indicizzazione dei contenuti, il Chatbot RAG può rapidamente recuperare le informazioni più rilevanti per rispondere alle domande degli utenti.

Grazie all'integrazione di tecniche di recupero delle informazioni e modelli di generazione del linguaggio, il Chatbot RAG è in grado di fornire risposte accurate e articolate, superando i limiti dei chatbot tradizionali basati su risposte predefinite. Questa architettura avanzata lo rende particolarmente adatto per compiti che richiedono l'accesso a dati aggiornati e specifici.

## Nella seconda fase

Quando un utente pone una domanda, questa viene anch'essa convertita in un **vettore di embedding**. Questo vettore viene utilizzato per cercare i chunk più simili all'interno del database, in modo da individuare le informazioni più pertinenti per generare una risposta.

La risposta finale viene quindi inviata all'utente e memorizzata nello storico delle conversazioni. Questa storia può essere utilizzata per affinare le domande future e fornire risposte sempre più personalizzate e contestuali.

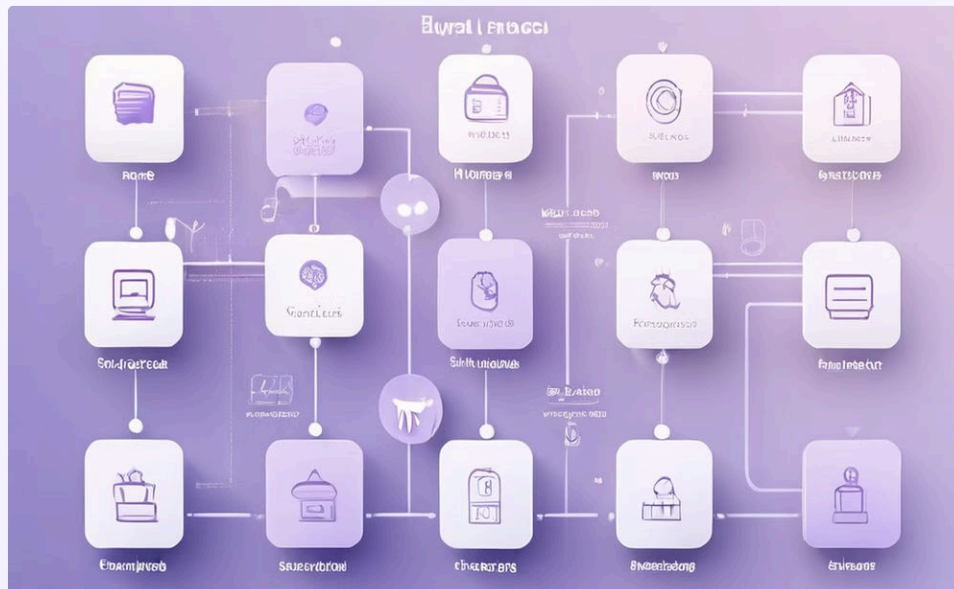


# Implementazione

Il progetto è stato sviluppato in Python e il codice, strutturato in 10 punti salienti, è stato implementato utilizzando Google Colab, pertanto:



# Punti Salienti



## Elaborazione dell'input

Il sistema analizza e tokenizza il testo inserito dall'utente, identificando le **parole chiave** e il contesto della richiesta.



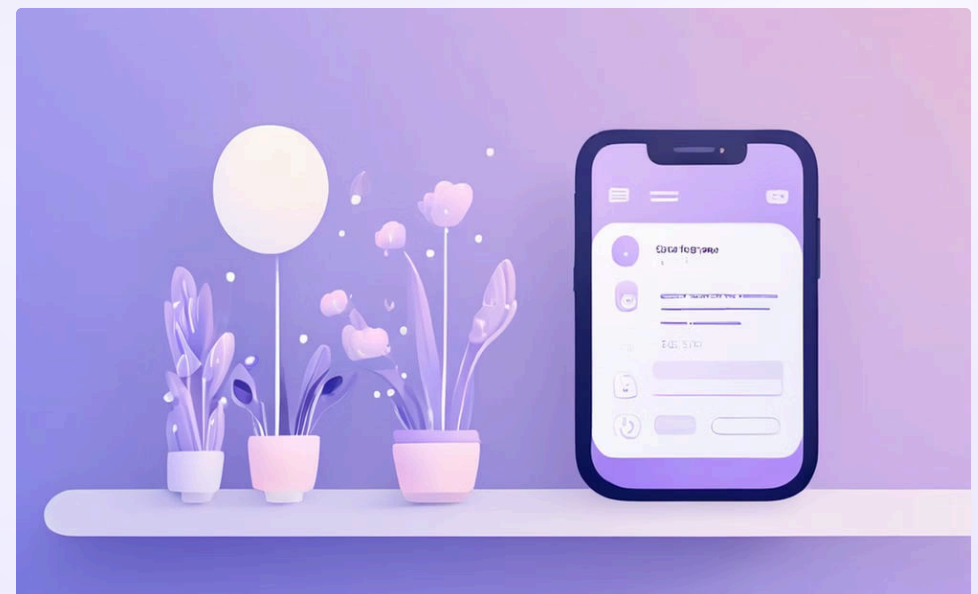
## Ricerca nel database

Utilizzando le informazioni elaborate, il sistema cerca nel database vettoriale (**Chroma**) i chunk di testo più rilevanti.



## Generazione della risposta

Il modello **LLM** (Large Language Model) genera una risposta coerente e contestualizzata.



## Output all'utente

Infine, la risposta elaborata viene presentata all'utente attraverso l'interfaccia **Gradio**.



# Test e risultati

## 7.1 Test

Per completezza sono stati riportate alcuni test eseguiti per capire l'efficienza del chatbot. I test sono stati eseguiti su un dataset con tema principale la geografia. Presentiamo l'interfaccia del chatbot:



sarà possibile fare una domanda cliccando 'Ask'; cancellare tutta la domanda cliccando 'Clear' e andare alla pagina GitHub del progetto.

Di seguito sono riportati tre casi d'uso:

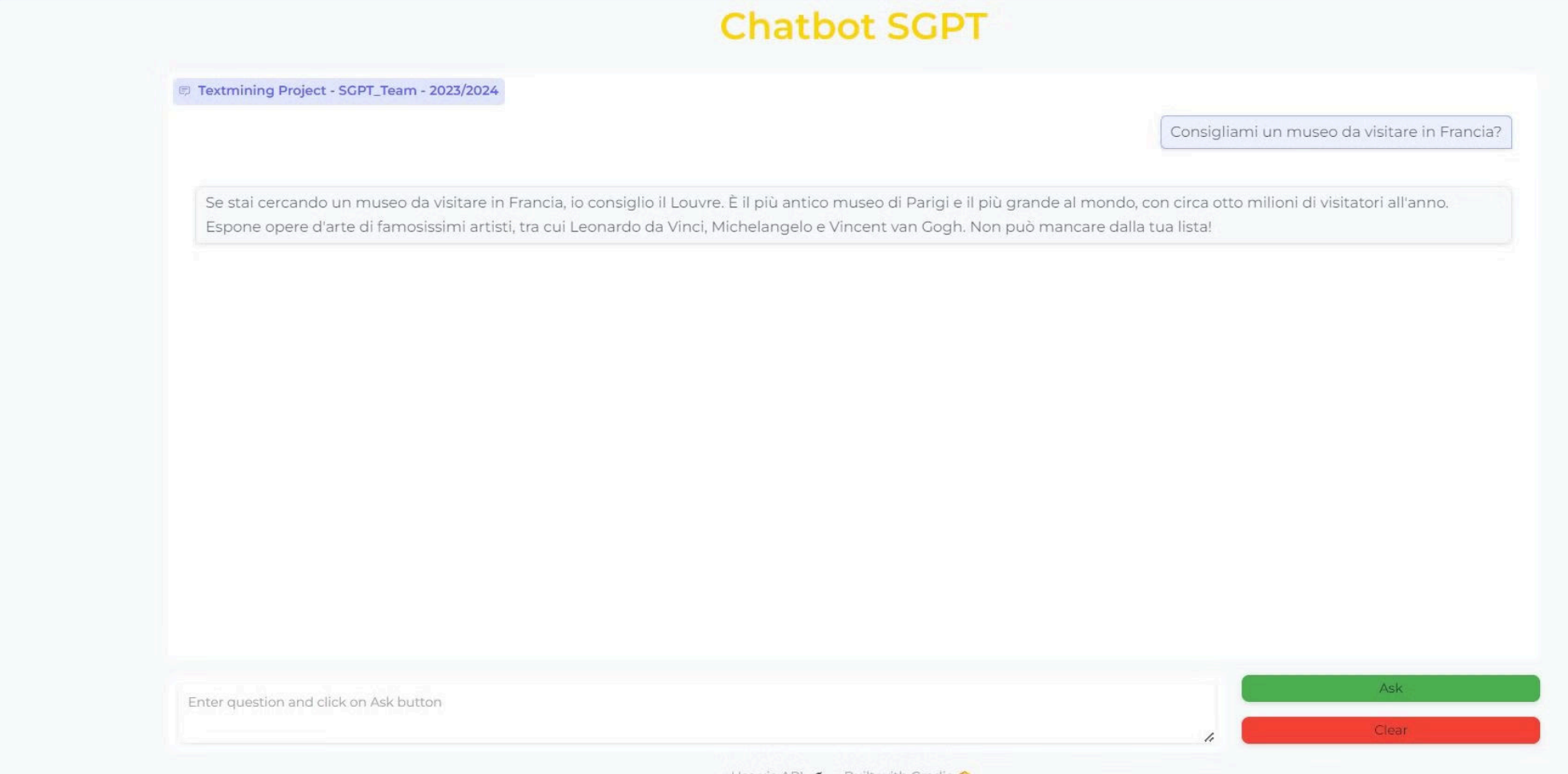
### 1. Caso d'uso



### 2. Caso d'uso



### 3. Caso d'uso



### 4. Caso d'uso



## 5. Caso d'uso

### Chatbot SGPT

Textmining Project - SGPT\_Team - 2023/2024

Differenza tra ONU e UE

La differenza principale tra l'Organizzazione delle Nazioni Unite (ONU) e l'Unione Europea (UE) è che l'ONU è un'organizzazione intergovernativa composta da stati sovrani indipendenti, mentre l'UE è un'organizzazione transnazionale formata da stati membri con poteri conferiti dall'accordo di Schengen, della zona euro e della Carta dei Diritti Fondamentali. L'ONU è principalmente dedicata alla promozione della pace e della cooperazione internazionale, mentre l'UE si concentra sulla cooperazione economica, politica e sociale tra i suoi membri. Entrambe le organizzazioni operano in diversi settori come la sicurezza, la giustizia, la cooperazione scientifica e la cooperazione sociale, ma la UE ha un mandato più specificamente economico e politico.

Enter question and click on Ask button

Ask

Clear

# Considerazioni finali

Il chatbot sviluppato si è rivelato uno strumento efficace per fornire risposte veloci e accurate alle domande degli utenti. Tuttavia, presenta ancora alcune imprecisioni su determinate questioni, dovute principalmente alla necessità di ampliare ulteriormente la conoscenza di base. Ciò può essere ottenuto attraverso l'aggiunta di nuovo materiale di apprendimento e un'intensificazione della fase di tuning.

Nel complesso, il sistema ha dimostrato il suo potenziale e con ulteriori miglioramenti potrà diventare uno strumento ancora più prezioso per gli utenti. La continua evoluzione e il perfezionamento del chatbot saranno essenziali per garantire un'esperienza ottimale e soddisfare appieno le esigenze degli utenti.