



torchlight.swe2324@outlook.com

Manuale Utente

Versione 1.0.0

| | |
|---------------------|--|
| Redattori | Meneghini Fabio Pluzhnikov Dmitry |
| Verifica | Agafitei Ciprian Cappellari Marco Filippini Giovanni De Laurentis Arianna Pia |
| Approvazione | De Laurentis Arianna Pia |
| Uso | Esterno |
| Destinatari | Prof. Tullio Vardanega Prof. Riccardo Cardin Zucchetti S.p.A. |

Registro delle Modifiche

| Ver. | Data | Descrizione | Autore | Verifica |
|-------|------------|--|-----------------------------|-----------------------------|
| 1.0.0 | 2024/04/14 | Verifica finale ed approvazione del documento | De Laurentis Arianna Pia | Cappellari Marco |
| 0.2.4 | 2024/04/13 | Aggiunte le immagini al capitolo §4, iniziato e completato il capitolo §5 | Pluzhnikov Dmitry | De Laurentis Arianna Pia |
| 0.2.3 | 2024/04/13 | Ultimata la stesura delle sezioni §4.2.2 e §4.2.2.1, §4.2.3 e §4.2.4 | Pluzhnikov Dmitry | De Laurentis Arianna Pia |
| 0.2.2 | 2024/04/12 | Iniziato lo sviluppo della sezione §4.2 e delle sottosezioni §4.2.1 e §4.2.2 | Pluzhnikov Dmitry | Agafitei Ciprian |
| 0.2.1 | 2024/04/11 | Iniziato lo sviluppo del capitolo §4, sviluppo della sezione §4.1 | Pluzhnikov Dmitry | Agafitei Ciprian |
| 0.2.0 | 2024/04/11 | Verifica dei contenuti fino ad ora prodotti | Pluzhnikov Dmitry | Cappellari Marco |
| 0.1.5 | 2024/04/10 | Sviluppato il capitolo §3 con le istruzioni di installazione | Pluzhnikov Dmitry | Cappellari Marco |
| 0.1.4 | 2024/04/10 | Compilata la sezione §2.3, terminata la stesura del capitolo §2 | Pluzhnikov Dmitry | Cappellari Marco |

| Ver. | Data | Descrizione | Autore | Verifica |
|-------|------------|--|-------------------|-----------------------------|
| 0.1.3 | 2024/04/09 | Sviluppo delle sezioni §2.1 e §2.2 | Pluzhnikov Dmitry | De Laurentis Arianna Pia |
| 0.1.2 | 2024/04/06 | Inizio compilazione del capitolo §2 | Meneghini Fabio | Filippini Giovanni |
| 0.1.1 | 2024/04/05 | Stesura del capitolo §1 | Meneghini Fabio | Filippini Giovanni |
| 0.1 | 2024/04/03 | Creazione del documento e stesura della struttura dello stesso | Meneghini Fabio | Filippini Giovanni |

Tabella 1: Registro delle modifiche

Indice

| | | |
|----------|---|-----------|
| 1 | Introduzione | 1 |
| 1.1 | Scopo del documento | 1 |
| 1.2 | Scopo del prodotto | 1 |
| 1.3 | Miglioramenti al documento | 1 |
| 2 | Requisiti | 2 |
| 2.1 | Requisiti software | 2 |
| 2.2 | Requisiti hardware | 2 |
| 2.3 | Browser supportati | 3 |
| 3 | Installazione del programma | 4 |
| 4 | Guida all'utilizzo | 5 |
| 4.1 | Interrogazione del modello | 6 |
| 4.2 | Sezione tecnico | 7 |
| 4.2.1 | Autenticazione per accedere alla sezione dedicata | 7 |
| 4.2.2 | Gestione dei file | 8 |
| 4.2.2.1 | Eliminazione | 9 |
| 4.2.2.2 | Caricamento | 9 |
| 4.2.3 | Debug dei dizionari | 11 |
| 4.2.4 | Logout | 12 |
| 5 | Supporto tecnico | 13 |

Elenco delle figure

| | | |
|----|---|----|
| 1 | Come si presenta l'applicativo appena viene aperto. | 5 |
| 2 | Selezione di un file. | 5 |
| 3 | Invio di una richiesta da parte dell'utente. | 6 |
| 4 | Accesso alla sezione tecnico: campo mancante. | 7 |
| 5 | Accesso alla sezione tecnico: credenziali errate. | 8 |
| 6 | Accesso alla sezione tecnico: login avvenuto con successo. | 8 |
| 7 | Eliminazione di un file: conferma dell'eliminazione. | 9 |
| 8 | Caricamento di un file: non conformità con lo schema. | 10 |
| 9 | Caricamento di un file: formato errato. | 10 |
| 10 | Caricamento di un file: caricamento corretto. | 11 |
| 11 | Come output verranno mostrati i punteggi assegnati dal modello alle tabelle presenti nel file. | 12 |
| 12 | Logout dalla sezione tecnico: conferma dell'avvenuta uscita dalla sezione. . . | 12 |

Elenco delle tabelle

| | | |
|---|---|-----|
| 1 | Registro delle modifiche | iii |
| 2 | Moduli richiesti per il corretto funzionamento del software | 2 |
| 3 | Requisiti hardware | 2 |
| 4 | Browser supportati per far funzionare l'applicazione | 3 |

1 Introduzione

1.1 Scopo del documento

Questo documento ha lo scopo di illustrare le istruzioni per l'utilizzo e le funzionalità fornite dall'applicativo. L'utente sarà quindi a conoscenza dei requisiti minimi necessari per il corretto funzionamento dello stesso, di come installarlo in locale e di come farne un utilizzo consapevole.

1.2 Scopo del prodotto

Nell'ultimo anno vi è stato un cambiamento repentino nello sviluppo e nell'applicazione dell'Intelligenza Artificiale all'elaborazione e raccomandazione dei contenuti alla generazione di essi, come immagini, testi e tracce audio.

Il capitolato C9, "ChatSQL", pone come obiettivo la realizzazione di un'applicazione che permetta, partendo dalla descrizione di un database, la generazione di un prompt in risposta ad un'interrogazione in linguaggio naturale. Questo processo dovrà essere gestito da un LLM (Large Language Model) specificatamente addestrato in questo scopo.

Tale applicazione sarà fruibile attraverso un'applicazione web, dove l'utente potrà caricare un file in formato JSON e richiedere, con linguaggio naturale, la produzione di prompt riguardanti il documento caricato.

1.3 Miglioramenti al documento

La maturità e i miglioramenti sono aspetti fondamentali nella stesura di un documento. Questo permette di apportare agevolmente modifiche in base alle esigenze concordate tra i membri del gruppo e il proponente, nel corso del tempo. Di conseguenza, questa versione del documento non può essere considerata definitiva o completa, poiché è soggetta a evoluzioni future.

2 Requisiti

Per poter utilizzare l'applicazione è necessario soddisfare i seguenti requisiti minimi.

2.1 Requisiti software

Per far sì che le operazioni di installazione e avvio del software avvengano correttamente e che si possa aver accesso a tutte le funzionalità, è necessario avere nella propria macchina i seguenti software.

| Software | Versione | Download |
|------------|----------|---|
| Python | 3.11.4 | python.org |
| JSONSchema | 4.20.0 | Python Package Index |
| Streamlit | 1.29.0 | Python Package Index |
| txtai | 6.2.0 | Python Package Index |
| pytest | 7.4.2 | Python Package Index |

Tabella 2: Moduli richiesti per il corretto funzionamento del software

Installando l'immagine docker i moduli sono già presenti e non necessitano di ulteriori installazioni.

2.2 Requisiti hardware

Per avere delle prestazioni accettabili dell'applicazione è preferibile avere almeno i seguenti componenti hardware:

| Componente | Requisito |
|------------|-------------------|
| CPU | Quad-Core 2,8 GHz |
| RAM | 8GB DDR4 |

Tabella 3: Requisiti hardware

2.3 Browser supportati

L'applicazione è stata testata e quindi resa compatibile con le ultime versioni dei browser riportati di seguito:

| Browser | Versione |
|-----------------|----------|
| Google Chrome | 120 |
| Microsoft Edge | 120 |
| Apple Safari | 16 |
| Mozilla Firefox | 120 |

Tabella 4: Browser supportati per far funzionare l'applicazione

3 Installazione del programma

Come prima cosa bisogna verificare che docker sia installato correttamente sulla macchina, scrivendo il comando `docker --version` all'interno del terminale.

Nel caso vi siano problemi, controllare di avere l'applicazione Docker Desktop aperta e in esecuzione (fosse necessario scaricarla, è disponibile a [questa pagina](#)¹).

Una verificato il corretto funzionamento del software, sempre nel terminale, scrivete il comando `docker pull ghcr.io/torchlight-swe2324/chatsql-docker:latest`, e inizierà il download dell'immagine in locale. La dimensione totale, che include tutte le dipendenze, è di circa **3.10GB**.

Una volta completato il download, sempre rimanendo all'interno del terminale, va scritto il comando `docker run -p 8501:8501 NOMEIMMAGINE`, dove `NOMEIMMAGINE` va sostituito con il nome dell'immagine `ghcr.io/torchlight-swe2324/chatsql-docker:latest` o con l'ID generato da docker che trovate nell'applicativo.

Una volta inviato il comando, comparirà un messaggio con il link per vedere l'applicativo. Per terminare esecuzione utilizzate la combinazione di tasti `ctrl/cmd+c` all'interno del terminale, o premendo il tasto `stop` del container dall'applicativo docker desktop.

IMPORTANTE: nel terminale vedrete l'indirizzo `0.0.0.0:8501`, ma non è quello corretto, infatti dovrete usare `localhost:8501`. Questo è un problema dovuto alla gestione delle porte di docker.

¹Visitato il 2024/04/10

4 Guida all'utilizzo

Appena avviata l'applicazione, si aprirà una pagina web con un'interfaccia simile a quella mostrata di seguito.

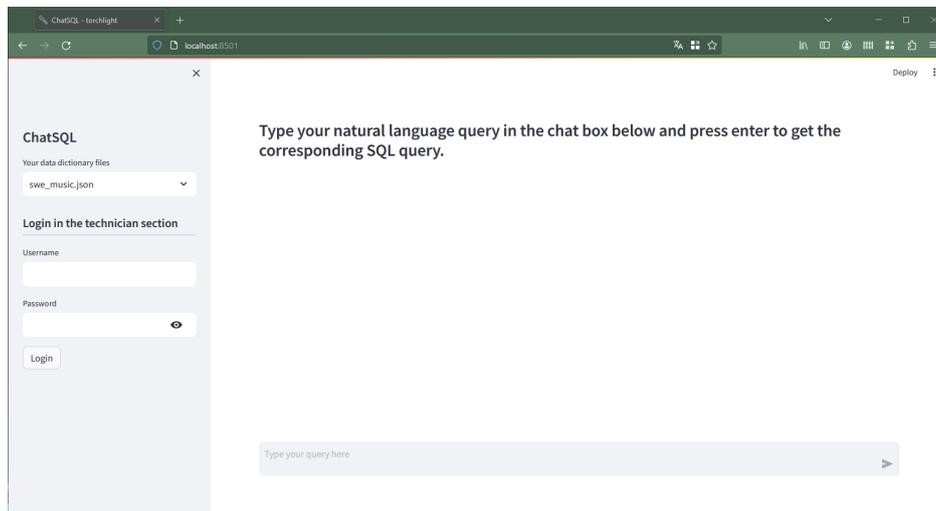


Figura 1: Come si presenta l'applicativo appena viene aperto.

Da qui è possibile, se sono presenti file all'interno dell'applicativo, selezionare quello da cui si desidera generare il prompt. I dizionari sono ordinati dall'ultimo caricato.

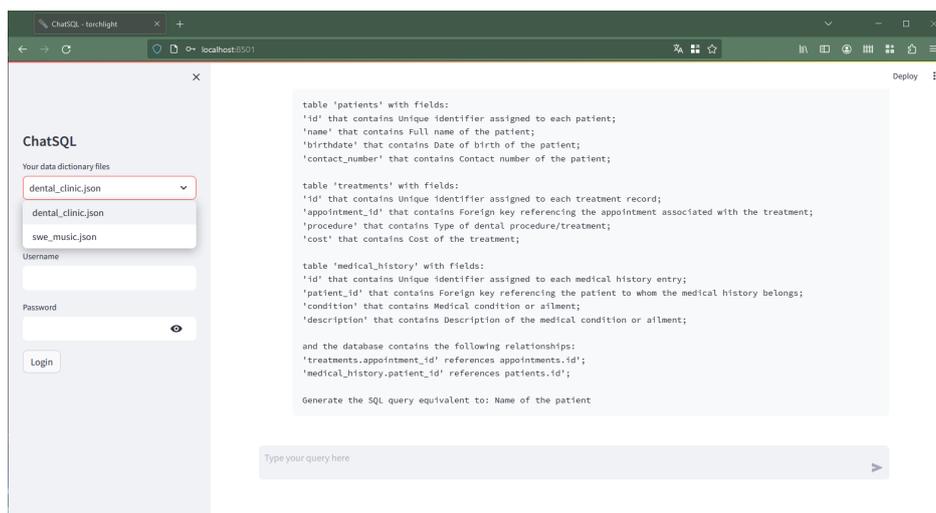


Figura 2: Selezione di un file.

4.1 Interrogazione del modello

Nel campo di inserimento dell'interrogazione in basso alla pagina, si può scrivere ciò che la query dovrebbe ottenere, e una volta inviata la richiesta il modello risponderà con il prompt che avrà al suo interno le tabelle con i campi maggiormente simili a ciò che è stato richiesto, una lista delle relazioni dei campi recuperati, l'interrogazione in linguaggio naturale inserita e il comando di generare la query SQL equivalente.

Durante la generazione del messaggio da parte del modello, comparirà una barra di caricamento, in modo da indicare lo stato di avanzamento dell'elaborazione.

Sarà possibile copiare il messaggio dalla box premendo il tasto copia in alto a destra della box di risposta per inserirlo dentro ad un GPT (ad esempio ChatGPT, Gemini, Claude, etc.) per ottenere l'interrogazione in SQL.

Da notare che la prima volta in assoluto che si avvia l'applicativo dopo aver scaricato e avviato l'immagine docker, alcune funzionalità potrebbero essere lente in quanto l'applicazione in background sta scaricando il LLM da Hugging Face.

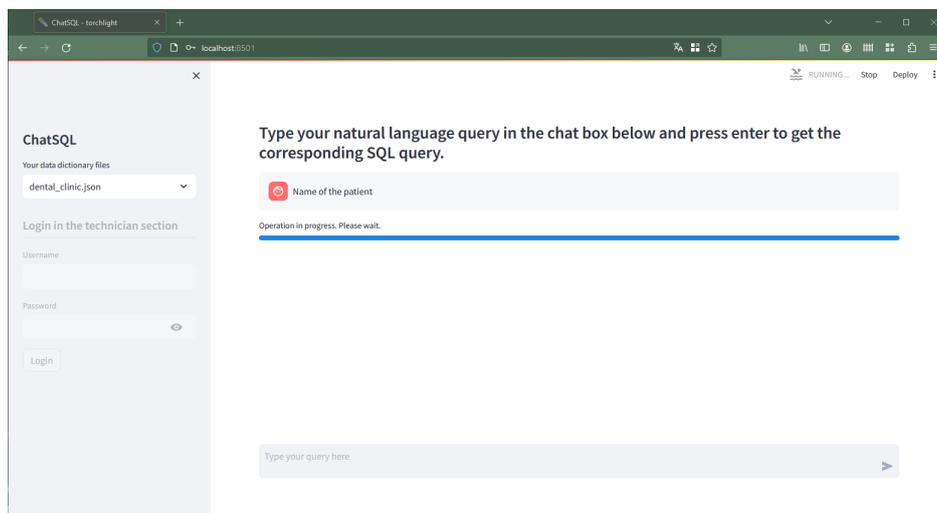


Figura 3: Invio di una richiesta da parte dell'utente.

4.2 Sezione tecnico

4.2.1 Autenticazione per accedere alla sezione dedicata

Nella colonna di sinistra (di seguito "sidebar") è presente un modulo di login per permettere al tecnico di gestire i file caricati ed eseguire il debug di essi.

Per accedere alla sezione riservata al tecnico, va scritto "admin" in entrambi i campi di nome utente e password.

Nel caso non sia stato inserito alcun valore all'interno dei campi, o le credenziali non siano corrette, verrà mostrato un messaggio in alto a destra in base alla tipologia di errore commesso. Nel caso di mancato inserimento di una o entrambe le credenziali, verrà visualizzato un messaggio con caratteri arancioni di avviso.

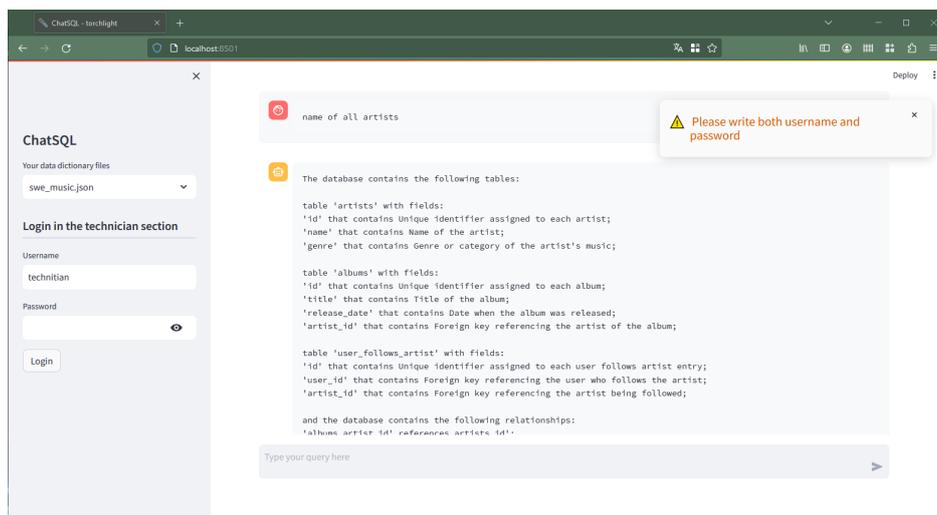


Figura 4: Accesso alla sezione tecnico: campo mancante.

Invece, in caso di credenziali errate, sarà mostrato un errore dal colore rosso.

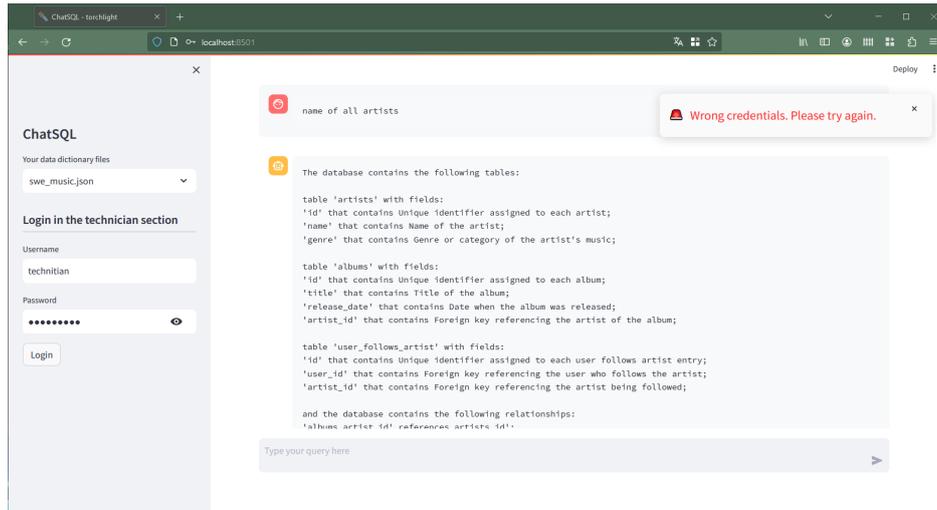


Figura 5: Accesso alla sezione tecnico: credenziali errate.

Una volta inserite le credenziali corrette, verrà mostrato il relativo messaggio di conferma e verranno visualizzate le funzionalità disponibili solo al Tecnico. Da segnalare che si perderanno tutti i messaggi presenti nella chat.

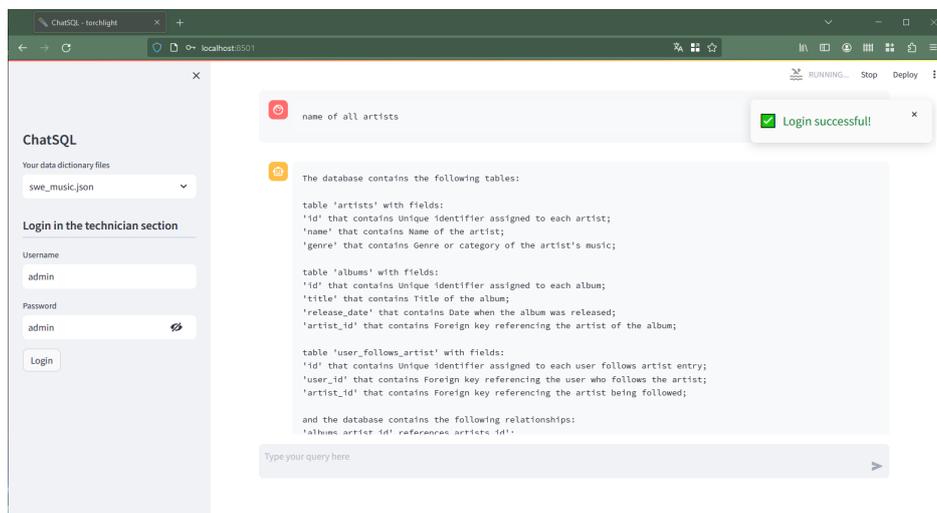


Figura 6: Accesso alla sezione tecnico: login avvenuto con successo.

4.2.2 Gestione dei file

All'interno della sidebar è possibile, oltre a poter selezionare il file su cui lavorare come avviene nella parte dell'utente non autenticato, eliminare un dizionario o caricarne uno nuovo e naturalmente uscire dalla sezione tecnico.

4.2.2.1 Eliminazione

Per eliminare un file dall'applicativo è sufficiente, una volta selezionato il file dallo stesso menu a tendina usato per la selezione del dizionario corrente, premere il pulsante "Delete selected file". Verrà mostrato un messaggio in verde per confermare l'avvenuta cancellazione dall'applicativo.

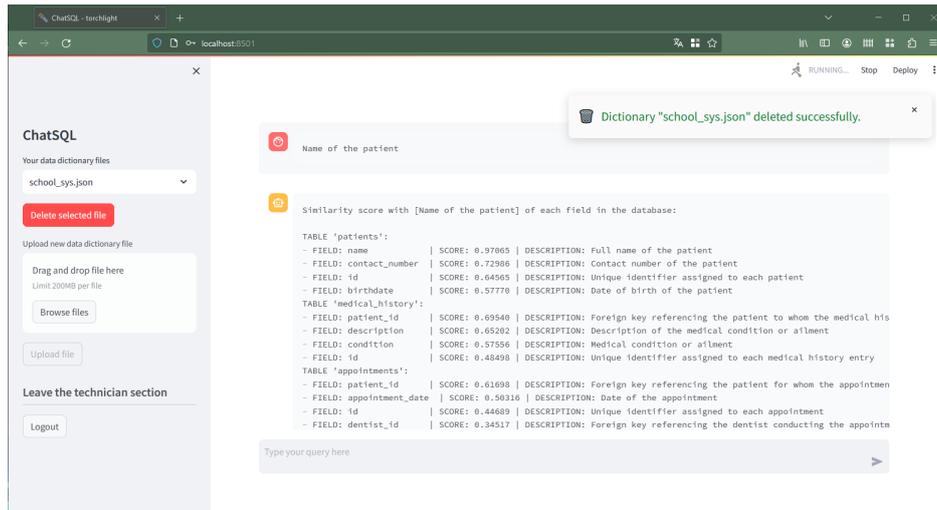


Figura 7: Eliminazione di un file: conferma dell'eliminazione.

4.2.2.2 Caricamento

È possibile caricare un nuovo file all'interno dell'applicativo. Basta trascinarlo sul box di caricamento o premendo il pulsante "browse file". Una volta inserito, si attiverà il pulsante "Upload file" per caricarlo all'interno dell'applicativo.

Per poterli caricare, i file devono essere in formato JSON, essere conformi ad uno schema per permettere il corretto funzionamento dell'applicativo, non deve già esserci a sistema un dizionario con lo stesso nome e l'applicativo può salvare al massimo 5 file in contemporanea. Nel caso in cui almeno una delle condizioni sopra citate non fosse soddisfatta, verrà mostrato un messaggio d'errore esplicativo di caricamento fallito.

Qui sotto alcuni esempi dei messaggi che potranno apparire.

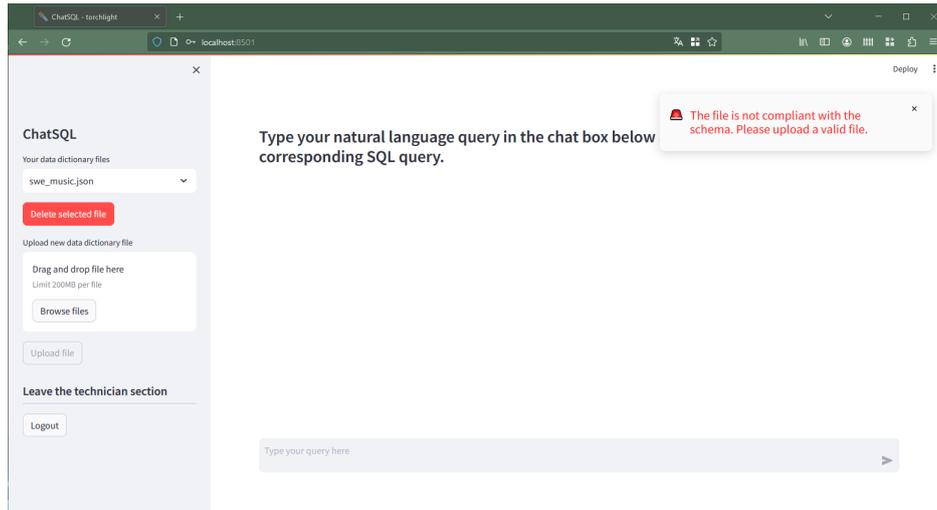


Figura 8: Caricamento di un file: non conformità con lo schema.

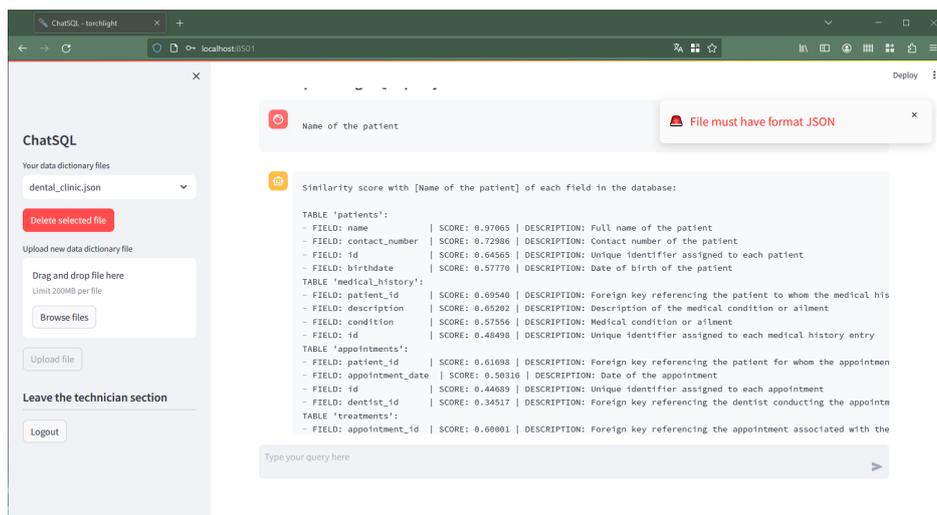


Figura 9: Caricamento di un file: formato errato.

Una volta caricato correttamente il file, verrà mostrato un messaggio di conferma, e in automatico sarà inserito come dizionario attualmente selezionato e potrà essere subito usato per le interrogazioni.

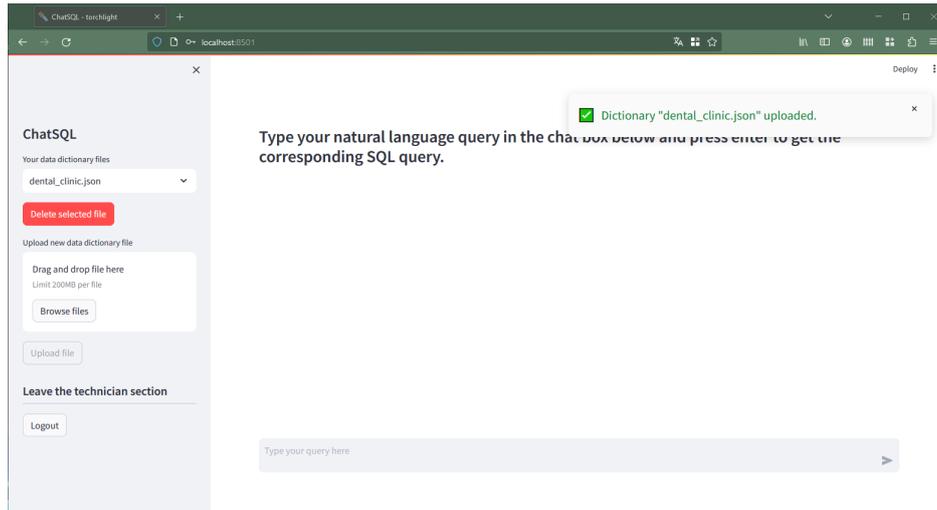


Figura 10: Caricamento di un file: caricamento corretto.

4.2.3 Debug dei dizionari

Per eseguire il debug dei dizionari è sufficiente trovarsi nella sezione tecnico, selezionare un file dalla lista ed iniziare a scrivere le domande in linguaggio naturale. Si noterà che, a differenza della sezione utente, nel messaggio generato sono mostrate tutte le tabelle con i loro campi, ognuno con il suo punteggio di similarità tra la descrizione del campo, stampata anch'essa nel messaggio, e l'interrogazione fatta.

Questo permette di capire se il file è stato strutturato in maniera corretta, suggerendo eventuali modifiche e miglioramenti nelle descrizioni dei campi nel file stesso.

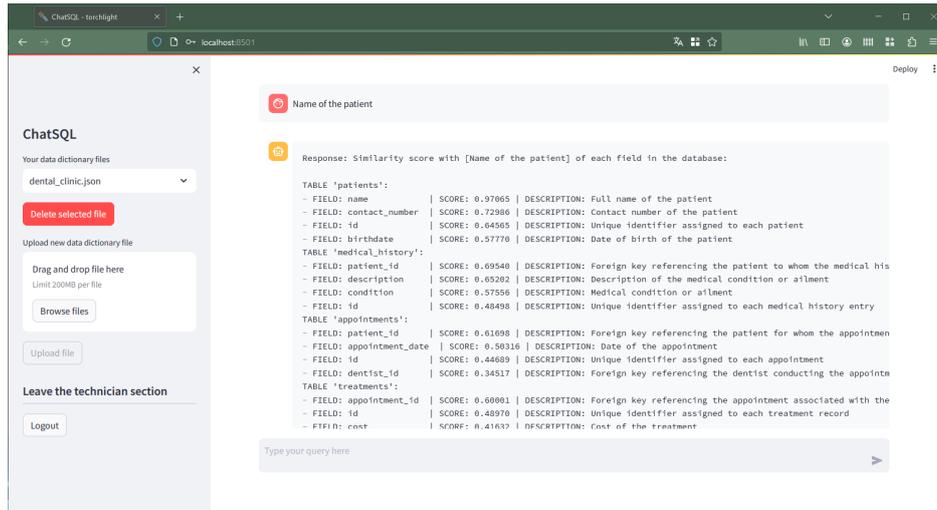


Figura 11: Come output verranno mostrati i punteggi assegnati dal modello alle tabelle presenti nel file.

4.2.4 Logout

Per uscire dalla sezione tecnico è sufficiente premere il pulsante "Logout" posto in fondo alla sidebar. Questo comporterà il ritorno alla schermata principale dell'utente, dove è possibile interrogare il modello e selezionare il file su cui eseguire l'interrogazione. Da notare che come per l'autenticazione, avverrà una eliminazione dei messaggi presenti nella chat.

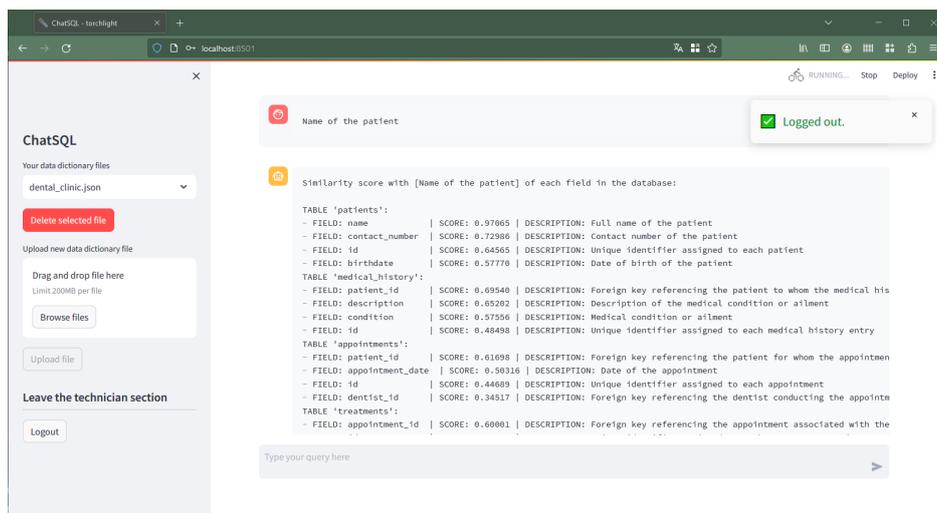


Figura 12: Logout dalla sezione tecnico: conferma dell'avvenuta uscita dalla sezione.

5 Supporto tecnico

Il gruppo è disponibile per offrire assistenza in caso di eventuali inconvenienti o domande relative al nostro progetto. Se si dovessero riscontrare difficoltà tecniche, dubbi sulle funzionalità o sull'installazione, siamo raggiungibili all'indirizzo: torchlight.swe2324@outlook.com

Ricordiamo, inoltre, che per offrire un supporto mirato ed efficiente è preferibile descrivere in maniera dettagliata il problema, includendo eventuali messaggi di errore ricevuti.

I membri del gruppo si impegneranno al massimo per risolvere prontamente qualsiasi problematica e per fornire tutte le informazioni necessarie per utilizzare al meglio il nostro prodotto.