



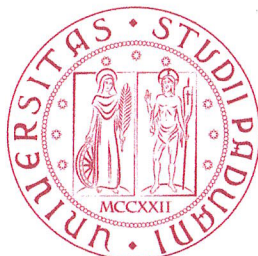
ARGO

Verbale Riunione 2024-07-09

Gruppo Argo – Progetto ChatSQL

Informazioni sul documento

Versione	1.0.0
Approvazione	Mattia Zecchinato
Uso	Esterno
Distribuzione	Zucchetti S.p.A. Prof. Tullio Vardanega Prof. Riccardo Cardin Gruppo Argo



UNIVERSITÀ
DEGLI STUDI
DI PADOVA

Registro delle modifiche

Ver.	Data	Redazione	Verifica	Descrizione
1.0.0	2024-07-10	Mattia Zecchinato	Mattia Zecchinato	Approvazione del documento
0.0.1	2024-07-09	Raul Pianon	Tommaso Stocco, Martina Dall'Amico	Stesura del documento

Indice

1	Informazioni	3
1.1	Descrizione	3
1.2	Partecipanti	3
1.3	Glossario	3
2	Riunione	4
2.1	Ordine del giorno	4
2.2	Argomenti e temi dell'incontro	4
2.2.1	Presentazione del Proof of Concept	4
2.2.2	Conclusioni	5
3	Todo / In Progress	6

1 Informazioni

- **Inizio incontro:** 17:00
- **Fine incontro:** 18:00
- **Pianificazione incontro:** Mail
- **Tipo incontro:** remoto (Zoom)

1.1 Descrizione

Il gruppo ha organizzato un incontro con la *Proponente*_g per esporre il lavoro svolto nelle ultime settimane e presentare la versione del *Proof of Concept*_g che verrà portata alla revisione *RTB*_g.

1.2 Partecipanti

- **Argo:**
 - Tommaso Stocco → 1 ora
 - Riccardo Cavalli → 1 ora
 - Raul Pianon → 1 ora
 - Martina Dall'Amico → 1 ora
 - Sebastiano Lewental → 1 ora
- **Zucchetti S.p.A.:**
 - Gregorio Piccoli

1.3 Glossario

Allo scopo di evitare incomprensioni relative al linguaggio utilizzato nella documentazione di progetto, viene fornito un *Glossario*, nel quale ciascun termine è corredato da una spiegazione che mira a disambiguare il suo significato. I termini tecnici, gli acronimi e i vocaboli ritenuti ambigui vengono formattati in corsivo all'interno dei rispettivi documenti e marcati con una lettera _g in pedice. In questo documento viene formattata solamente la prima ricorrenza di un termine definito nel *Glossario*.

2 Riunione

2.1 Ordine del giorno

- Presentazione del PoC_o realizzato con le nuove tecnologie.

2.2 Argomenti e temi dell'incontro

2.2.1 Presentazione del Proof of Concept

Il gruppo espone il motivo dell'incontro, ovvero presentare la versione finale del Proof of Concept, realizzata con *Vue.js_o* come strumento per il *front-end_o* e *FastAPI_o* per il *back-end_o*. In primo luogo, il team illustra i cambiamenti effettuati dall'ultima riunione, specialmente quelli riguardanti la gestione delle chiavi esterne composte. Vengono poi menzionati i modelli in lingua italiana utilizzati per eseguire i test di generazione del *prompt_o*. In particolare, sono stati testati i modelli "sentence-BERTino" e "mmarco-sentence-BERTino" di efederici. I test in locale sono stati effettuati con *LMstudio_o* utilizzando "Llama 3". Questi test sono stati eseguiti mantenendo lo stesso *prompt*, e chiedendo a un *LLM_o* di riscrivere la frase dell'utente per un certo numero di volte. Le *query_o* risultanti sono state salvate in un documento di testo al fine di realizzare un confronto.

In seguito, il gruppo presenta il Proof of Concept tramite il quale la Proponente può osservare le funzionalità implementate, tra cui:

- Processo di autenticazione per accedere alla sezione riservata all'admin;
- Generazione di un *prompt*;
- Copia del *prompt*.

Successivamente, il team mostra il *prompt* generato a partire dalla seguente richiesta: "the surname of users who paid for all their orders with PayPal". Il risultato viene copiato e incollato su ChatGPT e chat.lmsys.org per confrontare le *query SQL_o* generate da modelli differenti.

La Proponente si sofferma sulla *query* restituita da ChatGPT, che utilizza soltanto un operatore "NOT EXISTS". Il gruppo evidenzia che, sulla base dei test svolti, ChatGPT ritorna spesso *query* più contenute rispetto ad altri modelli. La Proponente chiede di testare il *prompt* con altri modelli, utilizzando chat.lmsys.org. Ad ogni tentativo, la Proponente esamina con il gruppo il significato e la correttezza delle *query*, mostrando come alcuni modelli sembrino omettere dei controlli o costruire *query* eccessivamente complesse. La Proponente domanda al gruppo se sono stati svolti dei test anche su un *database_o* locale; il gruppo afferma di aver svolto dei test con la stessa richiesta, senza verificare, però, il caso in cui un utente non abbia effettuato alcun ordine. La Proponente sottolinea come non vi siano errori strutturali evidenti nella *query*. Tuttavia, da questo esempio si evince che potrebbero verificarsi condizioni di errore; pertanto, viene chiesto al gruppo di svolgere ulteriori test con database reali.



Al termine dei test, il gruppo riporta l'attenzione al Proof of Concept per illustrare le funzionalità rimanenti:

- Debug della generazione del prompt, con la possibilità di scaricare un file in formato testuale;
- Gestione *CRUD*_e dei dizionari dati;
- Persistenza degli *indici*_e;
- Processo di logout;
- Pagina della chat, con una sezione per le impostazioni (selezione del dizionario dati, del *DBMS*_e e della lingua);
- Visualizzazione di un'anteprima del dizionario dati;
- Eliminazione della cronologia della chat.

Infine, il team illustra le impostazioni di accessibilità con le quali è possibile cambiare la lingua, il tema e le dimensioni dell'interfaccia.

La Proponente rinnova l'invito a effettuare test su un database locale, al fine di verificare la correttezza sintattica e semantica delle query generate. Gli aspetti da valutare includono:

- Correttezza delle query;
- Posizionamento delle parentesi;
- Complessità delle query;
- Operatori utilizzati (LIKE oppure =, != oppure <>).

2.2.2 Conclusioni

Il gruppo conclude con una domanda in vista della revisione RTB.

Domanda: Ritiene che sia necessario aggiungere un flag nel dizionario dati per segnalare i campi indicizzati del database?

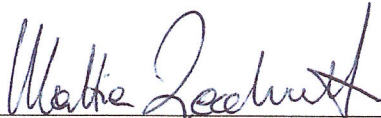
Risposta: No, non è necessario ai fini della correttezza della query SQL. Bisogna considerare inoltre il rischio derivante dall'aggiunta di questa informazione nel prompt, che potrebbe indurre un *Large Language Model*_e a generare una query non appropriata.

3 Todo / In Progress

Ticket	Incarico	Incaricato/a	Scadenza
ARGO-210	Stesura verbale riunione	Raul Pianon, Martina Dall'Amico, Tommaso Stocco	2024-07-11
ARGO-213	Test con un database locale	Marco Cristo, Riccardo Cavalli	2024-07-20

Luogo e Data:
Padova (PD) 2024-07-09

Firma: _____



Responsabile: Mattia Zecchinato

Per approvazione:

Firma: _____



Referente: Gregorio Piccoli (Zucchetti S.p.A.)

Verbale Riunione 2024-07-09
v 1.0.0

Zucchetti S.p.A.
Via Solferino, 1 - 26900 LODI
Tel. 0371.5945700 - Fax 0371.5945753
Sede Op.: Via G. Cittadella, 7 - 35137 PADOVA
P. IVA e Cod. Fisc. 05006900962

6 di 6

Enochell & Co.
754 Colton St. - 2500 LOD
Tel. 531-242700 - Fax 531-242701
Sales Office: 754 Colton St. - 2500 LOD
P.O. Box 1000 - 2500 LOD