



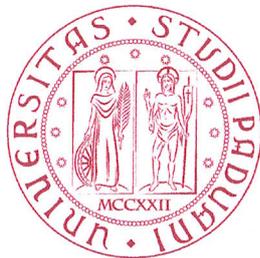
ARGO

Verbale Riunione 2024-06-07

Gruppo Argo – Progetto ChatSQL

Informazioni sul documento

Versione	1.0.0
Approvazione	Martina Dall'Amico
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	2024-06-24	Martina Dall'Amico	Martina Dall'Amico	Approvazione del documento
0.0.1	2024-06-09	Martina Dall'Amico	Riccardo Cavalli	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	Dimostrazione di una demo del software	4
2.2.2	Cambio di tecnologie	6
2.2.3	Range valori delle metriche di prodotto	6
3	Todo / In Progress	7

1 Informazioni

- **Inizio incontro:** 14:30
- **Fine incontro:** 15:30
- **Pianificazione incontro:** Mail
- **Tipo incontro:** remoto (Zoom)

1.1 Descrizione

Durante la riunione, il gruppo ha presentato alla *Proponente*_g una demo del software, valutando al contempo l'aggiunta di nuove funzionalità e *casi d'uso*_g. Inoltre, il team ha comunicato il cambio di tecnologie e discusso i valori desiderati delle *metriche di qualità*_g.

1.2 Partecipanti

- **Argo:**
 - Tommaso Stocco → 1 ora
 - Marco Cristo → 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 di una demo del software;
- Confronto sul cambio di tecnologie;
- Range valori delle *metriche di prodotto*.

2.2 Argomenti e temi dell'incontro

2.2.1 Dimostrazione di una demo del software

Descrizione: I programmatori hanno mostrato a schermo una demo al fine di capire se le funzionalità sviluppate fino a quel punto fossero state comprese nella loro interezza e, di conseguenza, fossero state implementate correttamente. Le funzionalità che sono state discusse inizialmente sono le seguenti:

- La visualizzazione del contenuto del *dizionario dati*, in linguaggio naturale, cosìché l'utente possa formulare una richiesta in modo più intuitivo;
- La gestione *CRUD* del dizionario dati;
- La generazione di un *prompt* a seguito dell'inserimento di una richiesta in linguaggio naturale.

Risposta: La Proponente ha approvato la struttura generale del prototipo.

Descrizione: Il gruppo ha mostrato la funzionalità di generazione di un *prompt*, che avviene dopo la scelta di un *dizionario dati* e l'inserimento di una richiesta. I *modelli* impiegati dal team sono stati "all-MiniLM-L6-v2", "all-MiniLM-L12-v2" e "paraphrase-multilingual-MiniLM-L12-v2".

Risposta: La Proponente ha consigliato di provare "sentence-BERTino" come *modello*, in lingua italiana.

Descrizione: Il gruppo ha già utilizzato "sentence-BERTino" in precedenza, ma i punteggi restituiti dalla funzione *search* di *txtai*, per quanto superiori rispetto ad altri modelli, sembravano essere troppo uniformati. Tuttavia, il team ha impiegato questo modello solo nelle fasi iniziali del progetto, quando ancora il dizionario dati era incompleto.

Risposta: La Proponente ha invitato il team a provare anche altri modelli, sempre in lingua italiana, come ad esempio "mmarco-sentence-BERTino". Inoltre, ha consigliato al gruppo di inserire descrizioni in italiano all'interno del dizionario dati.

Descrizione: Il team ha testato la funzionalità di generazione del *prompt*.

Risposta: La Proponente ha approvato la funzionalità e ha proposto dei test ulteriori che sono stati eseguiti durante la riunione. Questi test prevedono di copiare il *prompt* generato su chat.lmsys.org per verificare la costruzione delle *query* SQL con soluzioni alternative a ChatGPT.



Descrizione: Il prompt è stato testato con 7 LLM₆ diversi (gpt, yi-1.5-34b, qwen2-72b, claude-3-haiku, command-r, llama-3-8b, openhermes-2.5-mistral-7b). I risultati emersi hanno confermato la correttezza del prompt, in quanto i modelli hanno restituito la medesima query SQL. L'unica eccezione è stata l'uso della lettera iniziale maiuscola per i nomi propri da parte di alcuni modelli, anche se non era specificato nella frase in linguaggio naturale.

Risposta: La Proponente ha suggerito di cambiare il DBMS₆ e di non utilizzare solo MariaDB, ma anche PostgreSQL o affini.

Descrizione: Dopo aver testato con successo il cambio del DBMS, il team ha proposto l'implementazione di tale funzionalità e l'aggiunta di un caso d'uso simile alla selezione della lingua.

Risposta: La Proponente ha manifestato interesse per questa funzionalità e ne ha appoggiato l'implementazione.

Descrizione: Il team ha testato altre richieste, analizzando la struttura del prompt.

Risposta: La Proponente ha evidenziato una lacuna nel prompt restituito, poiché quest'ultimo riportava i nomi delle tabelle e dei campi, ma non le descrizioni in linguaggio naturale. Nei database aziendali, i nomi delle tabelle e dei campi sono spesso poco significativi e necessitano di una descrizione esplicativa.

Domanda: Alla luce della nuova struttura del prompt, è possibile ridurre la dimensione del prompt restituendo solo le descrizioni delle colonne rilevanti?

Risposta: La Proponente ha sottolineato come ChatGPT gestisca 128.000 token, mentre Gemini 1 milione. Inoltre, risulterebbe difficile stabilire un criterio con cui identificare le colonne rilevanti, specialmente nel caso di chiavi esterne. Pertanto, la Proponente ha consigliato al team di restituire le descrizioni di tutti i campi delle tabelle.

Descrizione: Il gruppo ha discusso la possibilità che vengano inserite nel prompt delle tabelle non pertinenti alla richiesta dell'utente.

Risposta: La Proponente ha evidenziato la scelta del team di privilegiare la recall del sistema rispetto alla precisione. Si tratta di una decisione che la Proponente si aspettava e che rispetta i requisiti del *capitolato*₆.

Descrizione: Per la funzionalità di *debug*₆, il team ha creato un file di *log*₆ che illustra il processo di generazione del prompt. Il file è diviso in due sezioni; la prima elenca le tabelle restituite dalla funzione *search* di *txtai* con il relativo punteggio. Per ogni tabella, è riportata una classifica di importanza dei termini della descrizione. La sezione successiva, invece, spiega il motivo per cui una tabella è stata inserita nel prompt oppure no.

Risposta: La Proponente ha espresso un giudizio positivo sull'implementazione della funzionalità di *debug*.



2.2.2 Cambio di tecnologie

Descrizione: Una volta terminata la dimostrazione della demo, il team ha comunicato alla Proponente la decisione di adottare nuove tecnologie per lo sviluppo dell'applicazione.

Risposta: La Proponente ha ribadito la sua neutralità riguardo alle scelte tecnologiche, sottolineando però che la decisione del team sembra essere stata maturata dopo una valutazione approfondita delle problematiche che si sarebbero potute presentare con il *framework* precedente.

2.2.3 Range valori delle metriche di prodotto

Domanda: Dopo aver seguito i gruppi del primo lotto coinvolti nello sviluppo di ChatSQL, le conclusioni raggiunte sono state sufficientemente precise per definire un metodo di valutazione delle metriche di qualità del prodotto?

Risposta: La Proponente ha evidenziato come gli LLM migliorino rapidamente e ciò che mesi fa sembrava impossibile, ora probabilmente è già stato risolto. Definire dei range per la qualità del prodotto può essere più agevole rispetto a prima, in quanto è stato dimostrato come sia possibile generare codice SQL corretto a partire da un prompt. La Proponente ha invitato quindi il gruppo a testare interrogazioni più complesse, magari con query annidate.

Domanda: La Proponente suggerisce dei test ulteriori da eseguire?

Risposta: La Proponente ha suggerito di effettuare dei test con un modello in locale tramite *LM Studio*. Lo scopo dei test è verificare che mantenendo lo stesso prompt e chiedendo a un LLM di riscrivere la frase dell'utente per un certo numero di volte, la query SQL resti semanticamente immutata.

3 Todo / In Progress

Ticket	Incarico	Incaricato/a	Scadenza
ARGO-100	Configurazione nuovi framework back-end e front-end	Mattia Zecchinato, Sebastiano Lewental	2024-06-10
ARGO-130	Modifica dizionario dati (descrizioni tabelle e chiavi esterne)	Riccardo Cavalli	2024-06-09
ARGO-131	Aggiornare il prompt (aggiunta descrizioni in linguaggio naturale, chiavi primarie, chiave esterne composte)	Riccardo Cavalli	2024-06-09
ARGO-132	Riorganizzazione del contenuto del file di log	Riccardo Cavalli	2024-06-09
ARGO-134	Test con modelli di efederici	Riccardo Cavalli	2024-06-10
ARGO-136	Test con modelli locali tramite LM Studio	Riccardo Cavalli	2024-06-13
ARGO-142	Stesura verbale riunione	Martina Dall'Amico	2024-06-10

Luogo e Data:
Padova (PD) 2024-06-07

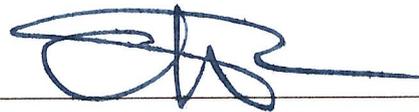
Firma: _____

Dall'Amico Martina

Responsabile: Martina Dall'Amico

Per approvazione:

Firma: _____



Referente: **Zucchetti S.p.A.** (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

Verbale Riunione 2024-06-07

◆ v 1.0.0

7 di 7