

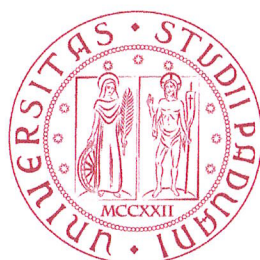


Verbale Riunione 2024-05-22

Gruppo Argo — Progetto ChatSQL

Informazioni sul documento

Versione	1.0.0
Approvazione	Marco Cristo
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	2024-06-01	Marco Cristo	Marco Cristo	Approvazione del documento
0.0.1	2024-05-24	Raul Pianon	Marco Cristo, 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	Presentazione <i>Streamlit</i> e confronto con altre tecnologie	4
2.2.2	Discussione delle metriche di qualità	5
2.2.3	Funzione di traduzione di txtai	6
3	Todo / In Progress	7

1 Informazioni

- **Inizio incontro:** 12:00
- **Fine incontro:** 13:00
- **Pianificazione incontro:** Mail
- **Tipo incontro:** remoto (Zoom)

1.1 Descrizione

Durante il meeting, il gruppo ha presentato alla *Proponente*_e le caratteristiche delle tecnologie individuate per sviluppare l'applicazione. Il discorso si è poi concentrato su una discussione più approfondita di casi d'uso e metriche di qualità e sulla risoluzione dei dubbi del team.

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 _e 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 *framework*_e *Streamlit*_e;
- Discussione delle metriche di qualità;
- Discussione sulla feature di traduzione di *txtai*_e.

2.2 Argomenti e temi dell'incontro

2.2.1 Presentazione *Streamlit*_e e confronto con altre tecnologie

Descrizione: I programmatori elencano le caratteristiche di *Streamlit*_e, il *framework*_e individuato dal team per realizzare il prototipo dell'applicazione. Vengono indicati i vantaggi e gli svantaggi della suddetta tecnologia, quali:

- Il back-end di *Streamlit* è sviluppato in *Python*_e;
- Il front-end, invece, sfrutta le funzionalità della libreria *React*_e e dell'estensione *JSX*_e;
- Tecnologia open source;
- Utilizzato principalmente per applicazioni di data science e machine learning, che operano analisi e visualizzazione di dati, interazione con *LLM*_e ed elaborazione del linguaggio naturale;
- Disponibilità di componenti interattivi facilmente integrabili nell'applicazione;
- Comunità ampia e molto attiva, specialmente nella risoluzione dei problemi di personalizzazione.

Risposta: La Proponente non impone vincoli sulla scelta delle tecnologie, ma invita il gruppo, per la prossima riunione, a soffermarsi su come il framework individuato possa risolvere i problemi posti dal capitolato, piuttosto che sulle sue specifiche tecniche. Viene inoltre spiegato che adottando una tecnologia si accettano i rischi che ne conseguono, e tale decisione spetta al team di fornitori.

Descrizione: Il gruppo propone quindi di mostrare i risultati ottenuti tramite l'interazione tra *txtai*_e e *Streamlit* durante il prossimo incontro. Dichiaro inoltre di aver già iniziato la fase di integrazione e implementato i primi *casi d'uso*_e, ma che questi non coprono ancora i requisiti minimi. Nel frattempo, il team ha anche analizzato i pro e contro di framework più consolidati, sia back-end (*Flask*_e, *Django*_e) che front-end (*Vue.js*_e, *Next.js*_e). La coppia *Flask-Vue.js* è stata considerata dal gruppo come un'ottima alternativa, e pertanto verrà approfondita nel prossimo sprint.

Risposta: La Proponente conferma la proposta del team, chiedendo di essere contattata qualora ci siano degli avanzamenti significativi nello sviluppo dei casi d'uso principali.

2.2.2 Discussione delle metriche di qualità

Domanda: La Proponente ha individuato delle metriche di qualità significative che il prodotto dovrà rispettare?

Risposta: Relativamente all'interfaccia grafica, la Proponente non avanza richieste specifiche, lasciando che sia il gruppo a individuare metriche di qualità significative. Si esprime invece sulle metriche che riguardano la parte funzionale del prodotto, ossia l'interazione tra l'applicazione e gli LLM_o. In particolare, la Proponente invita il team a definire una batteria di test per verificare quante volte il risultato ottenuto (inteso come lista di tabelle pertinenti alla richieste dell'utente) è vicino a quello desiderato. Di seguito sono riportate alcune casistiche che la Proponente ritiene debbano essere verificate:

- Frasi in cui è coinvolta una sola tabella;
- Query con la clausola JOIN per combinare due o più tabelle;
- Frasi che richiedono un ordinamento;
- Query in cui vengono applicati dei filtri;
- Query che richiedono funzioni di aggregazione.

Per ciascuno dei casi sopracitati, l'ideale sarebbe disporre di un numero sufficiente di test per capire quanti danno una risposta soddisfacente e quanti no.

Domanda: Il team chiede alla Proponente se abbia già definito un range di tollerabilità per tale metrica.

Risposta: La Proponente non esprime una misura specifica, in quanto l'individuazione di un valore desiderabile è parte integrante del processo di sviluppo e testing.

Domanda: Lo scopo della metrica, quindi, è capire quanto i modelli siano in grado di fornire dei risultati soddisfacenti?

Risposta: La Proponente sottolinea come gli LLM vengano ormai venduti come prodotti sicuri e affidabili; lo scopo del progetto è capire se e quanto questo corrisponda alla realtà.

Domanda: C'è un valore desiderabile riguardo la velocità e le prestazioni della traduzione?

Risposta: La Proponente spiega che i gruppi del primo lotto non hanno riscontrato particolari lacune o difficoltà in merito alla velocità di esecuzione. Un aspetto da tenere in considerazione, invece, è la dimensione del *prompt*_o generato, poiché i ChatBOT impongono dei limiti, seppur grandi, in termini di *token*_o. Tuttavia, la riduzione della dimensione del prompt deve essere effettuata con accortezza, al fine di evitare l'esclusione di tabelle rilevanti. Si deve quindi trovare il giusto equilibrio tra dimensione e accuratezza del prompt, per garantire una corretta costruzione della *query*_o SQL.

Domanda: Il team avanza un ulteriore dubbio riguardo la quantità delle metriche da applicare.

Risposta: La Proponente ritiene che non sia facile effettuare una stima di questo tipo. A tal proposito spiega che un'eccessiva quantità di metriche non è sempre

vantaggiosa, e anzi rischia di complicare la gestione del progetto. Indica però la presenza di 5 metriche che possono essere considerate fondamentali:

- Tempo, ossia il tempo che il team di sviluppo impiega per dare una risposta a una certa domanda o per completare una determinata attività;
- Sforzo;
- Costo;
- Qualità;
- Dimensione;

La Proponente si sofferma sulla metrica "Dimensione", intesa come "grandezza di un programma". Se si misurano solo le linee di codice e si indica come "codice di qualità" un codice lungo, potrebbe risultare invece che quest'ultimo non sia sufficientemente ottimizzato. Un esempio riguarda gli algoritmi Merge sort e Quicksort; il primo, seppur più lungo in termini di linee di codice, è meno efficiente del secondo. Infine, la Proponente suggerisce di discutere delle metriche anche con il Committente.

2.2.3 Funzione di traduzione di txtai

Domanda: Per una richiesta in italiano, è più opportuno generare il prompt nella lingua del dizionario dati o predisporre una funzione di traduzione che restituisca il prompt in italiano?

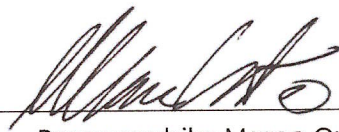
Risposta: La Proponente spiega che l'obiettivo del progetto è quello di fornire un prompt da copiare e incollare su ChatGPT (o altri ChatBOT). La query finale non cambia in base alla lingua della richiesta, poiché il linguaggio rimane l'SQL. Un utente non deve necessariamente leggere il contenuto del prompt, che può quindi mantenere la stessa lingua del dizionario dati.

3 Todo / In Progress

Ticket	Incarico	Incaricato/a	Scadenza
ARGO-68	Scelta definitiva del modello	Riccardo Cavalli	2024-05-28
ARGO-69	Definizione test di correttezza del prompt generato	Tommaso Stocco, Martina Dall'Amico, Riccardo Cavalli	2024-06-03
ARGO-71	Stesura verbale esterno del 2024-05-22	Marco Cristo	2024-05-28
ARGO-65	Modalità di visualizzazione dei risultati del debug	Martina Dall'Amico	2024-05-29
ARGO-66	Funzionalità di debug lato back-end	Tommaso Stocco, Riccardo Cavalli	2024-06-01
ARGO-77	Test di correttezza della struttura del <i>dizionario dati</i> .	Mattia Zecchinato	2024-06-03

Luogo e Data:
Padova (PD) 2024-05-22

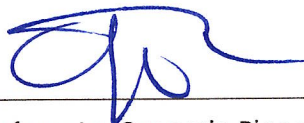
Firma: _____



Responsabile: Marco Cristo

Per approvazione:

Firma: _____



Referente: Gregorio Piccoli (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 2024-05-22

◆ v 1.0.0

7 di 7