



ARGO

Verbale Riunione 2024-05-02

Gruppo Argo — Progetto ChatSQL

Informazioni sul documento

Versione	1.0.0
Approvazione	Raul Pianon
Uso	Interno
Distribuzione	Prof. Tullio Vardanega Prof. Riccardo Cardin Gruppo Argo



UNIVERSITÀ
DEGLI STUDI
DI PADOVA

Registro delle modifiche

Ver.	Data	Redazione	Verifica	Descrizione
1.0.0	2024-05-10	Raul Pianon	Raul Pianon	Approvazione documento
0.0.3	2024-05-07	Raul Pianon	Martina Dall'Amico	Aggiornamento sezione todo
0.0.2	2024-05-06	Raul Pianon	Martina Dall'Amico	Aggiornamento corpo del documento e aggiunta sezione todo
0.0.1	2024-05-03	Raul Pianon	Martina Dall'Amico, Mattia Zecchinato	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	Discussione e decisioni	4
2.2.1	Analista	4
2.2.2	Amministratore	4
2.2.3	Responsabile	4
2.2.4	Programmatore	5
2.2.5	Progettista	5
2.2.6	Ridistribuzione dei ruoli	5
2.2.7	Incontro con La Proponente	5
2.2.8	Punti da trattare nel prossimo diario di bordo	6
2.2.8.1	Done	6
2.2.8.2	Difficoltà	6
2.2.8.3	Todo	6
2.2.8.4	Dubbi	7
3	Todo / In Progress	8

1 Informazioni

- **Inizio incontro:** 10:00
- **Fine incontro:** 11:30
- **Pianificazione incontro:** Telegram
- **Tipo incontro:** remoto (Discord)

1.1 Descrizione

Durante l'incontro sono state discusse le attività svolte durante la settimana e sono stati presentati e discussi i punti per il diario di bordo del 3 maggio 2024. È stata inoltre inviata una mail alla Proponente e fissato un incontro per il 6 maggio 2024.

1.2 Partecipanti

- **Argo:**
 - Tommaso Stocco → 1 ora e 30 minuti
 - Marco Cristo → 1 ora
 - Riccardo Cavalli → 1 ora e 30 minuti
 - Mattia Zecchinato → 1 ora e 30 minuti
 - Raul Pianon → 1 ora e 30 minuti
 - Martina Dall'Amico → 1 ora e 30 minuti
 - Sebastiano Lewental → 1 ora e 30 minuti

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. Tutte le ricorrenze di un termine definito nel *Glossario* subiscono la formattazione sopracitata.

2 Riunione

2.1 Ordine del giorno

- Analisi delle attività svolte durante la settimana dai membri del gruppo;
- Discussione dei punti che verranno trattati al diario di bordo del 3 maggio 2024;
- Invio di una mail per l'organizzazione di un incontro con La Proponente per aggiornarLa sullo stato del progetto.

2.2 Discussione e decisioni

All'inizio dell'incontro, viene chiesto di esporre al gruppo le attività svolte durante la settimana in vista della fine dello *sprint*₆.

2.2.1 Analista

L'Analista ha evidenziato difficoltà nella conversione del documento di *Analisi dei Requisiti* e ha stimato a giorni la conclusione della conversione. Sono stati anche approfonditi i casi d'uso relativi al login per comprenderne la fattibilità, evidenziando la possibilità di avere un utente premium in caso di interazioni con l'*API*₆. È stata poi espressa la necessità di trovare mezzi e idee per l'approfondimento dei casi d'uso dello strumento di debug fornito al tecnico.

Sono state inoltre evidenziate delle difficoltà sull'ordinamento del documento di *Analisi dei Requisiti* date dallo strumento *Latex*₆. L'analista ha esposto la necessità di un incontro con il Professor. Riccardo Cardin per delucidazioni sulla funzione di login utente che sarebbe, con l'attuale impostazione, disponibile anche per il tecnico già loggato. Il problema rimane comunque arginabile definendo delle precondizioni per il login.

2.2.2 Amministratore

Sono avanzati i lavori sul *Piano di Progetto* e sono iniziati i lavori sul *Piano di Qualifica*. Sono inoltre stati sviluppati tabelle e grafici da inserire per preventivi e consuntivi del *Piano di Progetto* e sono iniziati i lavori sui file di configurazione per la compilazione dei file pdf in automatico.

2.2.3 Responsabile

Il Responsabile si è occupato della pianificazione del preventivo del primo *sprint*₆ e l'elaborazione di tabelle e grafici per preventivo e consuntivo degli *sprint*₆. Sono inoltre stati definiti attività e grafico *Gantt*₆ su *Jira*₆ e steso il verbale del meeting precedente.

2.2.4 Programmatore

Il Programmatore discute dello studio e dei progressi ottenuti con $txtai_e$, spiegando al gruppo la trasformazione che la tecnologia opera al fine di trasformare il linguaggio naturale in un vettore e in un indice di vettori. L'idea per il funzionamento è quella di formattare dizionario dati, $query_e$ e le informazioni delle tabelle in documenti che simulino il $database_e$. La $query_e$ verrebbe pre-processata e generare dei $token_e$ da questa, per renderla più accessibile ai meccanismi di $txtai_e$. La frase verrebbe quindi processata e passata all'indicizzatore che eseguirebbe lo stesso procedimento con il dizionario dati. Rispetto ai lavori precedenti sono stati aggiunti due nuovi moduli per la pre-processazione.

Vengono infine discussi i prossimi lavori che riguardano la creazione di un secondo indice che faccia da indice intermedio tra la $query_e$ e l'indice reale per estrarre l'informazione. Questo sarebbe utile perché arricchirebbe i risultati parziali con informazioni quali la descrizione della tabella e i vincoli comuni. Oltre a ciò c'è la volontà di testare l'inserimento del numero di colonne per confrontare e capire se il risultato finale sia ottimale.

2.2.5 Progettista

Il Progettista ha lavorato coordinandosi con il programmatore per la strutturazione di un dizionario dati in formato $JSON_e$ che potesse essere più efficiente nell'interazione con $ChatGPT_e$. Sono stati citati diversi tipi di formato in linguaggio naturale che risultano essere più adatti gli LLM_e che verranno approfonditi in seguito. Infine è stato proposto di utilizzare il linguaggio $Python_e$ per il $backend_e$ visto che il gruppo sta lavorando con questo linguaggio e si sta trovando bene. Rimangono aperte e da esplorare diverse opzioni e software in cui sviluppare il prodotto.

2.2.6 Ridistribuzione dei ruoli

Per il terzo sprint vengono distribuite nuovamente le ore ai vari membri del gruppo. Da questo sprint, le ore non si concentreranno solo su un ruolo, ma spazieranno su più ruoli per fornire sostegno a membri che necessitano di maggiori risorse e per evitare di assegnare solo un ruolo con poche ore rimanenti. In particolare:

- Sono state richieste più ore sul ruolo di progettista per dare una migliore definizione del lavoro al Programmatore. Allo stesso tempo saranno assegnate più ore al Programmatore e meno ad altri ruoli in vista di una redistribuzione più mirata;
- Vengono inoltre richieste più risorse al ruolo di Amministratore vista la mole di lavoro che altrimenti dovrebbe affrontare un solo membro. Vengono quindi assegnate altre ore a 2 membri del gruppo, in modo da offrire assistenza.

2.2.7 Incontro con La Proponente

Vengono in seguito discussi i punti da trattare con La Proponente in un meeting pianificato per il 6 maggio 2024.

- Proposta dell'utilizzo del framework *Streamlit*_e per usare *Python*_e anche per la parte relativa al *frontend*_e o in alternativa l'utilizzo di altri framework per velocizzare i lavori;
- Proposta dell'utilizzo di *Python*_e anche per il *backend*_e come da suggerimento del Progettista;
- Aggiornamento sullo stato del progetto e sulle attività future.

2.2.8 Punti da trattare nel prossimo diario di bordo

Il gruppo ha discusso i punti salienti da trattare nel diario di bordo del 3 maggio 2024.

2.2.8.1 Done

- Pianificazione e preventivo *sprint*_e 3;
- Inizio stesura del *Piano di Qualifica*;
- Conversione in *Latex*_e del documento di *Analisi dei Requisiti*;
- Inserimento di grafici nel *Piano di Progetto*;
- Progettazione della conversione da linguaggio naturale al prompt finale;
- Definizione di un dizionario dati efficiente;
- Aggiornamento *Norme di Progetto* e popolamento del *Glossario*;
- Pianificazione di un incontro con La Proponente.

2.2.8.2 Difficoltà

- Poche risorse assegnate al ruolo di amministratore
- Difficoltà date dal nuovo ruolo nella rotazione
- Caricamento automatico di documenti da un *repository*_e ad un altro.

2.2.8.3 Todo

- Creazione di un'indice intermedio per una maggior accuratezza nell'interazione tra dizionario dati e *query*_e;
- Bozza di interfaccia grafica della *web app*_e;
- Progettazione iniziale della *web app*_e;
- Avanzamento nella scrittura dei documenti (*Glossario*, *Piano di Progetto*, *Analisi dei Requisiti*, *Norme di Progetto*, *Piano di Qualifica*);
- Aggiornameto sullo stato del progetto alla proponente: qui verranno anche proposti gli strumenti di sviluppo della webapp.

2.2.8.4 Dubbi

- Dubbio sul come far valere le ore che ci si ritrova a fare in un ruolo diverso;
- A che livello di dettaglio si dovrebbe arrivare nel documento di *Norme di Progetto*.

Gli argomenti sopracitati sono stati riportati nella presentazione per il diario di bordo del 3 maggio 2024.

3 Todo / In Progress

Ticket	Incarico	Incaricato/a	Scadenza
ARGO-12	Aggiornamento <i>Norme di Progetto</i> : utilizzo di <i>Jira</i> _e	Riccardo Cavalli	2024-05-09
ARGO-25	Stesura verbale riunione	Raul Pianon	2024-05-07
ARGO-27	Caricamento file di configurazione per il deploy dei documenti	Riccardo Cavalli	2024-05-06
ARGO-30	Consuntivo <i>sprint</i> _e 2	Raul Pianon	2024-05-08
ARGO-31	Integrazione <i>Jira</i> _e - <i>GitHub</i> _e	Riccardo Cavalli	2024-05-09
ARGO-32	Bozza interfaccia grafica	Sebastiano Lewental, Riccardo Cavalli	2024-05-10
ARGO-33	Inizio sviluppo <i>web app</i> _e	Sebastiano Lewental	2024-05-10
ARGO-40	Studio del secondo indicizzatore	Marco Cristo, Raul Pianon	2024-05-10

Luogo e Data:
Padova (PD) 2024-05-02

Firma: _____



Responsabile: Raul Pianon