

Università di Padova

Corso di Ingegneria
del Software



Verbale esterno

29 novembre 2023

Versione 1.0



1 Informazioni della riunione

La riunione è stata svolta da remoto, tramite piattaforma Google Meet, dalle ore 16:00 alle ore 17:10, in presenza di Carlo Davanzo ed Emilio Bendotti, rappresentanti dell'azienda *AzzurroDigitale_G*.

1.1 Lista partecipanti

Partecipante	Presenza
Andrea Cecchin	✓
Marco Dolzan	✓
Francesco Ferraioli	✓
Francesco Giacomuzzo	✓
Leonardo Lago	✓
Giovanni Menon	✓
Anna Nordio	✓

In rappresentanza di *AzzurroDigitale* :

Partecipante
Carlo Davanzo
Emilio Bendotti

2 Ordine del giorno

2.1 Discussioni

1. Presentazione di una prima versione del *PoC_G*;
2. Discussione sulle tecnologie da utilizzare;
3. Discussione sulla presentazione del *PoC*;



3 Resoconto

3.1 Discussioni

1. L'incontro si è aperto con la presentazione di una prima versione del Proof of Concept, privo di front-end ed eseguito interamente da riga di comando, ai rappresentanti dell'azienda proponente. La demo esposta, mirata a dimostrare che le tecnologie individuate dal gruppo sono coerenti agli obiettivi e requisiti del progetto, ha raggiunto il suo scopo svolgendo le funzionalità richieste dai proponenti.

È stato interrogato un *chatbot_G* sul contenuto di un documento, fornito nei giorni precedenti da AzzurroDigitale, e quest'ultimo, almeno in parte, ha risposto alle domande in modo corretto. Quando gli veniva fatta una domanda la cui risposta era contenuta nel documento, il chatbot forniva la risposta corretta; quando la domanda non era pertinente con le informazioni fornite, il chatbot lo faceva presente non rispondendo. È stato possibile osservare che il grado di precisione delle risposte è fortemente legato a quale *LLM_G* viene utilizzato, sia nella fase di *embedding_G* che di retrieve.

Quando il PoC veniva eseguito utilizzando il modello di *OpenAI_G*, la precisione della risposta era innegabilmente maggiore rispetto a quelle ottenute utilizzando un modello *open-source_G* di *Ollama_G*, con quest'ultimo spesso incapace di fornire le risposte corrette a differenza della controparte.

2. A seguito della presentazione del preview del PoC, sono state discusse alcune tecnologie utili al progetto. Interpellati su quale fosse l'utilizzo migliore delle collezioni di vettori in *ChromaDB_G*, i proponenti hanno consigliato al gruppo di approfondire l'utilizzo dei metadata dei documenti embeddizzati, così da poter giungere ad una maggiore efficienza andando ad effettuare un retrieve mirato dei soli documenti utili ad ogni domanda basandosi sulle informazioni contenute nei metadata, utilizzando quindi un'unica collezione di vettori per tutti gli embeddings dei documenti. Viene inoltre consigliato al gruppo di testare le performance del PoC sviluppato integrando una funzione di text splitter, per osservare se il tempo necessario ad embeddizzare i documenti ed interrogare il chatbot diminuisce.

La discussione si è conclusa con il forte consiglio, da parte dei rappresentanti della proponente, di optare per l'utilizzo di un framework per il back-end del prodotto finale del progetto.

3. L'incontro si conclude con la conferma, da parte di AzzurroDigitale, di arrivare alla presentazione di un Proof of Concept per la settimana seguente. Il gruppo ha l'obiettivo di preparare la parte front-end della prima versione demo mostrata, andando anche a testare altri modelli open-source. La ricerca di un



modello da utilizzare in locale, senza ricorrere a chiamate ai servizi di OpenAI, è rafforzato dalla volontà dei proponenti di giungere ad un prodotto sicuro in termine di protezione dei dati sensibili, dunque attento alle esigenze dei clienti target del prodotto del progetto.

3.2 Prossima riunione

È stata fissata una riunione per mercoledì 6 dicembre.



4 Storia del documento

Versione	Data	Autori	Verificatori	Descrizione
1.0	2023/12/05	Francesco Giacomuzzo		Approvazione documento
0.1	2023/11/29	Andrea Cecchin	Leonardo La- go	Stesura sezio- ni 1, 2, 3

Presa visione del proponente

Luca Devec