

## Verbale riunione interna

Versione	Data	Sez.modificate	Autore	Verificatore	Validatore
1.0.0	2024-12-15				Annalisa Egidi
0.1.0	2024-12-12			Michele Nesler	
0.0.1	2024-12-10	Prima stesura	Filippo Bellon		

### Abstract

Il giorno 09/12/2024 dalle ore 16:00 alle 18:21 il gruppo Rod2Cod si riunisce.

## Contents

<b>1 Presenze</b> .....	<b>2</b>
<b>2 Ordine del Giorno</b> .....	<b>2</b>
<b>3 Discussione</b> .....	<b>2</b>
3.1 Stesura verbale .....	2
3.2 Stesura diario di bordo .....	2
3.3 Progressi individuali .....	2
3.4 Confronto Casi D'uso .....	3
3.5 Approvazione WOW e aggiunta su gh-pages .....	4
3.6 Lavoro individuale da svolgere per la settimana .....	4
3.7 Incontro con l'azienda .....	4
<b>4 Argomenti da discutere in futuro</b> .....	<b>4</b>

## 1 Presenze

Nominativo	Presenze	Ruolo
Alessandro Rossi	P	Analista
Alberto Maggion	P	Analista
Luca Calzetta	P	Amministratore
Filippo Bellon	P	Verificatore
Annalisa Egidi	P	Responsabile
Michele Nesler	P	Analista

Presenze	
P	Presente
A	Assente

## 2 Ordine del Giorno

1. Chi fa il verbale
2. Chi fa il diario di bordo settimanale
3. Condivisione con il team sui progressi settimanali individuali
4. Confronto Casi d'Uso
5. Approvazione WoW e aggiunta su gh-pages
6. Confronto situazione per stesura Milestone e Baseline
7. Lavoro individuale da svolgere per la settimana

## 3 Discussione

### 3.1 Stesura verbale

Per la stesura di questo verbale è stato designato **Filippo Bellon**.

### 3.2 Stesura diario di bordo

Per la stesura del diario di bordo da presentare la prossima settimana è stato nominato **Alberto Maggion**.

### 3.3 Progressi individuali

Durante la riunione sono poi stati discussi e condivisi i **progressi** svolti individualmente da ciascun membro del gruppo nel corso della scorsa settimana.

- **Luca Calzetta:**

Ha impostato e testato l'utilizzo di un **database** di prova in linguaggio **PostgreSQL**, come consigliato dal proponente, il tutto tramite la piattaforma **Docker**. Il sistema si compone di due parti distinte: la prima riguardante il database vero e proprio; e l'altra, un'interfaccia per la gestione grafica del database. Ha poi redatto un README per l'utilizzo di questi strumenti e questa settimana si occuperà di aggiungere il suo contenuto all'interno del **Manuale Sviluppatore**.

- **Filippo Bellon:**

Ha completato una stesura iniziale dell'**analisi dei rischi**, adattata ai casi descritti dal capitolato e discussi con il proponente, il tutto prendendo spunto dai documenti appartenenti a progetti di studenti presentati lo scorso anno.

- **Annalisa Egidi:**

Ha studiato la tecnologia **RAG** come metodologia per dare **risposte** più accurate possibili e con meno **allucinazioni** possibili rispetto alla domanda presentata. Essa viene principalmente

utilizzata all'interno di **LLM**. Si basa principalmente sull'**input** e sulle **indicazioni dell'utente**, a cui poi si aggiunge altro materiale **esterno** utilizzato atto a migliorare la risposta finale. Ha poi fatto notare come, passando prima da una risposta generica basata solo sull'input, ad una risposta basata solo su dati inseriti dall'utente si riesca chiaramente ad ottenere una risposta corretta dal punto di vista appunto dell'**utente**. Tuttavia questo non implica automaticamente che la risposta sia **universalmente** corretta. Tramite l'applicazione della metodologia RAG è quindi possibile ovviare a questo problema, e **combinare** i dati inseriti dall'utente e le fonti esterne per produrre una **risposta** il più possibile **completa e corretta**. Di conseguenza ne è stato proposto l'uso all'interno nel nostro **algoritmo** per il **confronto** delle risposte ma solo appena si riuscirà ad ottimizzarla al meglio per mitigare allucinazioni e migliorare i dati su cui si basa.

- **Alberto Maggion:**

Ha approfondito **Bleurt**, un modello utilizzato per la valutazione di testo generato in linguaggio naturale. Esso, nella fase di allenamento, va a perturbare frasi, rimuovervi parole ed eventualmente aiutarsi tramite la tecnologia **Bert** per andare a generare testo finale di confronto. Successivamente, in fase operativa, riceve in **input** coppie di domanda e risposta e dopo il **confronto** fra la risposta generata e quella avuta in input, vi assegna un valore indicante la **somiglianza**, di solito compreso fra 0 e 1. Esso esegue inoltre un confronto di accordanza tra il modello e le valutazioni fatte da persone. Un **problema** presente, si presenta quando a confronto troviamo frasi di struttura simile: ad esse viene assegnato un indice di somiglianza buono anche se le frasi hanno senso opposto.

- **Alessandro Rossi:**

Sempre nell'ambito di ciò che è stato presentato da Alberto Maggion, ha descritto **Bleurt** come la **tecnica migliore** da poter utilizzare, in quanto il suo modello assomiglia molto al giudizio umano; tuttavia viene fatto notare che su domini molto specifici potrebbe sbagliare. Ha poi consigliato di informarsi e studiare più **metriche** per andare successivamente a **scegliere** quella che più si adatta al nostro caso. Ha poi esposto un Power Point per una spiegazione più mirata ed un **confronto** fra le varie tecniche e metriche di valutazione di frasi generate da algoritmi di Machine Learning. Tra i principali: **BLEU**, basato su similarità lessicali; **NIST**; **TER**; **GTM**; **METEOR**, basato sul confronto e ricerca di sinonimi; **ROUGE**, che opera tramite la semplice sovrapposizione di parole/frasi tra il testo corretto e quello generato; **BERTscore**, basato su **BERT**, il cui algoritmo cerca, per ogni parola, di associarvi il contesto corretto leggendo il resto della frase; **BARTscore**; **BEER**, il quale però si presenta con poca documentazione; ed infine **BLEURT**, che utilizza tecniche di deep learning e si basa su BERT. Infine, è stato individuato come caso specifico del capitolato e quindi dell'algoritmo, l'**Identificazione di parafrasi**.

- **Michele Nesler:**

Ha **risolto** i problemi presenti all'interno della repo di Github riguardanti la creazione e l'aggiornamento della **webpage** e dei suoi contenuti.

### 3.4 Confronto Casi D'uso

Era stato previsto un confronto sui casi d'uso preparati singolarmente fino ad ora. Tuttavia, vista la mancanza di tempo, e la decisione di analizzare i requisiti senza fretta, si è deciso di rimandare il confronto.

#### Decisioni:

- Organizzazione di una riunione aggiuntiva in settimana.

#### Motivazioni:

- Necessità di discutere i requisiti senza fretta e terminarne l'analisi.

### 3.5 Approvazione WOW e aggiunta su gh-pages

Si è discusso poi della possibilità di **approvare** quanto redatto fino ad ora all'interno del **Way of Working** e del suo **caricamento** all'interno della repo di Github, con conseguente aggiunta all'interno della webpage del progetto.

**Decisioni:**

- Approvazione e caricamento all'interno della repo, del Way of Working allo stato attuale.

**Motivazioni:**

- Necessità di esporre quanto redatto fino ad ora all'interno del Way of Working.

### 3.6 Lavoro individuale da svolgere per la settimana

È stato svolto un controllo sulle issue attualmente **assegnate** ed è stata attuata la **creazione** e successiva assegnazione di nuove issue in relazione ai compiti da svolgere.

### 3.7 Incontro con l'azienda

È stato deciso di chiedere un incontro al proponente per parlare dello Stato di Avanzamento Lavori e per confrontarsi sui Casi D'uso scelti e sulle metriche studiate fino ad ora per l'algoritmo di confronto.

## 4 Argomenti da discutere in futuro

- Confronto situazione per stesura Milestone e Baseline