



Gruppo "torchlight"

[torchlight.swe2324@outlook.com](mailto:torchlight.swe2324@outlook.com)

# Verbale esterno del 13 Novembre 2023

## Ordine del giorno

Incontro con Zucchetti S.p.A. incentrato sul sistema di filtraggio del dizionario dati

<b>Redattori</b>	Ye Tao Ren Federico Agafitei Ciprian Filippini Giovanni
<b>Verifica</b>	Meneghini Fabio Cappellari Marco
<b>Approvazione</b>	Pluzhnikov Dmitry De Laurentis Arianna Pia
<b>Uso</b>	Esterno



**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

# Indice

<b>1</b>	<b>Informazioni generali</b>	<b>3</b>
1.1	Svolgimento dell'incontro . . . . .	3
1.2	Registro delle presenze . . . . .	3
1.3	Presenze esterne . . . . .	3
<b>2</b>	<b>Temi affrontati</b>	<b>4</b>
2.1	Word Embedding . . . . .	4
2.2	Modelli di sentence similarity . . . . .	4
2.3	Pianificazione attivita' della settimana . . . . .	4

Scienze Interdisciplinari  
Via Colombo, 1 - 20133 Lodi  
Tel. 0371 3316100 - Fax 0371 3316101  
28100 Via G. Cesare, 1 - 03100 Viterbo  
P. 139 e 140 - Fax 0371 3316101

## 1 Informazioni generali

### 1.1 Svolgimento dell'incontro

- **Luogo:** Videochiamata Zoom
- **Data:** 13 Novembre 2023
- **Orario d'inizio dell'incontro:** 16:00
- **Orario di fine dell'incontro:** 19:00

### 1.2 Registro delle presenze

- **Membri presenti:**
  - Agafitei Ciprian
  - Cappellari Marco
  - De Laurentis Arianna Pia
  - Filippini Giovanni
  - Meneghini Fabio
  - Pluzhnikov Dmitry
  - Ye Tao Ren Federico
- **Presenze totali:** 7

Nessun membro del gruppo e' assente dalla riunione

### 1.3 Presenze esterne

- **Nome Cognome:** Piccoli Gregorio, Zucchetti S.p.A.

## 2 Temi affrontati

L'incontro si è suddiviso in due parti: nella prima il Dottor Piccoli ci ha dato degli spunti su tecnologie e metodologie da utilizzare per l'implementazione del sistema di filtraggio del dizionario dati a partire da un input in linguaggio naturale. Successivamente il gruppo ha svolto una riunione interna di discussione sul feedback ricevuto dall'azienda e a lezione per poi definire le attività da svolgere nella seconda settimana.

### 2.1 Word Embedding

Si è iniziato ponendo i dubbi sorti riguardo l'implementazione della procedura di filtraggio del dizionario dati, in particolare come gestire l'uso di termini sinonimi, errori ortografici e/o di battitura.

Il Dottor Piccoli ci ha introdotto la tecnica del "word embedding", una rappresentazione che permette di memorizzare le informazioni sia semantiche che sintattiche delle parole costruendo uno spazio vettoriale in cui i vettori delle parole sono più vicini se le parole sono riconosciute come semanticamente più simili.

### 2.2 Modelli di sentence similarity

In seguito abbiamo visto in azione la libreria "textai" e modelli di sentence similarity che applicano la tecnica del "word embedding". Si è potuto constatare come questa particolare classe di modelli siano non solo in grado di gestire i casi particolari individuati, ma anche di svolgere la ricerca semantica indipendentemente dalla lingua usata per porre l'interrogazione, ponendoci il quesito se implementare un sistema multilingue.

### 2.3 Pianificazione attività' della settimana

L'incontro con l'azienda è terminato alle 17:30. Il gruppo è rimasto in chiamata per discutere sulle attività da svolgere nei prossimi giorni. L'obiettivo principale sarà iniziare l'individuazione degli Use case e l'analisi dei requisiti.