

**Università degli Studi di Torino**

SCUOLA DI SCIENZE DELLA NATURA

Corso di Laurea Magistrale in Informatica



Tesi di Laurea Magistrale

**Design, ingegnerizzazione e  
realizzazione di un sistema di dialogo  
basato su LLM nel dominio delle  
tecnologie assistive**

RELATORE

**Prof. Alessandro Mazzei**

CANDIDATO

**Stefano Vittorio Porta**

859133

Anno Accademico 2023/2024



### **Dichiarazione di Originalità**

*Dichiaro di essere responsabile del contenuto dell'elaborato che presento al fine del conseguimento del titolo, di non avere plagiato in tutto o in parte il lavoro prodotto da altri e di aver citato le fonti originali in modo congruente alle normative vigenti in materia di plagio e di diritto d'autore. Sono inoltre consapevole che nel caso la mia dichiarazione risultasse mendace, potrei incorrere nelle sanzioni previste dalla legge e la mia ammissione alla prova finale potrebbe essere negata.*

## **Ringraziamenti**

Todo

## **Abstract**

Todo.

## **Parole chiave**

«classificazionexNLU», «data annotation/augmentation», «NLGBasataSuParafrasi»,  
«ingegnerizzazione»



# Indice

<b>1 Introduzione</b>	<b>1</b>
1.1 Contesto generale .....	1
1.2 Motivazioni e obiettivi della tesi .....	1
1.3 Struttura del documento .....	1
<b>2 Natural Language Understanding</b>	<b>2</b>
2.1 Come AIML gestisce la comprensione .....	2
2.1.1 Criticità e limiti di AIML .....	2
2.2 Dataset .....	3
2.2.1 Etichettatura automatica .....	3
2.2.2 Data augmentation .....	3
2.2.3 Etichettatura manuale .....	3
2.3 NLU: Capire cosa si sta dicendo .....	4
2.3.1 Metodi classici con Spacy .....	4
2.3.2 Classificazione con LLM .....	4
2.3.3 Valutazione e performance .....	4
2.4 NLU: Capire di cosa si sta parlando .....	5
2.4.1 NER e Slot-filling .....	5
2.4.1.1 Il ritorno di Spacy .....	5
2.4.2 Valutazione e performance .....	5
<b>3 Data Retrieval</b>	<b>6</b>
3.1 Retrieval tramite query .....	6
3.2 Retrieval basato su script .....	6
3.3 Retrieval automatico guidato dalle LLM .....	6
<b>4 Natural Language Generation</b>	<b>7</b>
4.1 Come vengono fornite le risposte in AIML .....	7
4.2 Generazione di risposte tramite LLM .....	7
4.2.1 Parafrasi .....	7
4.2.2 Prompting .....	7
4.3 Qualità delle risposte .....	7
4.3.1 Valutazione automatica .....	7
4.3.2 Valutazione umana .....	7
<b>5 Ingegnerizzazione</b>	<b>8</b>
5.1 Composizione del sistema .....	8
5.2 Compilatore .....	8
5.2.1 Pipeline .....	8
5.3 Runner .....	8
<b>Bibliografia</b>	<b>8</b>





# **1 Introduzione**

## **1.1 Contesto generale**

dare un'idea di chatbot, cos'è NovaGraphS, perché è importante, come si inserisce nel contesto delle tecnologie assistive

## **1.2 Motivazioni e obiettivi della tesi**

## **1.3 Struttura del documento**

## **2 Natural Language Understanding**

### **2.1 Come AIML gestisce la comprensione**

#### **2.1.1 Criticità e limiti di AIML**

## **2.2 Dataset**

### **2.2.1 Etichettatura automatica**

### **2.2.2 Data augmentation**

### **2.2.3 Etichettatura manuale**

## **2.3 NLU: Capire cosa si sta dicendo**

### **2.3.1 Metodi classici con Spacy**

### **2.3.2 Classificazione con LLM**

### **2.3.3 Valutazione e performance**

## **2.4 NLU: Capire di cosa si sta parlando**

### **2.4.1 NER e Slot-filling**

#### **2.4.1.1 Il ritorno di Spacy**

#### **2.4.2 Valutazione e performance**

## **3 Data Retrieval**

### **3.1 Retrieval tramite query**

### **3.2 Retrieval basato su script**

### **3.3 Retrieval automatico guidato dalle LLM**

## **4 Natural Language Generation**

### **4.1 Come vengono fornite le risposte in AIML**

### **4.2 Generazione di risposte tramite LLM**

#### **4.2.1 Parafrasi**

#### **4.2.2 Prompting**

### **4.3 Qualità delle risposte**

#### **4.3.1 Valutazione automatica**

#### **4.3.2 Valutazione umana**

## **5 Ingegnerizzazione**

### **5.1 Composizione del sistema**

### **5.2 Compilatore**

#### **5.2.1 Pipeline**

### **5.3 Runner**

## **Bibliografia**