

# **Das Erstellen von Wissensgraphen aus unstrukturierten Texten mit Hilfe von großen Sprachmodellen**

**Houda El Abbassi**

**Bachelor of Science**

**Principal Supervisor:  
Prof. Dr. Martin Simon**

**2024**

# DECLARATION

Signature: \_\_\_\_\_

2024

# Abstract

Die Verwendung von Large Language Models (LLM) wie GPT-3.5 zur Erstellung von Wissensgraphen bietet mehrere Vorteile:

**Textverständnis und Generierung:** LLMs haben die Fähigkeit, menschenähnliche Sprachmuster zu verstehen und zu generieren. Durch die Verarbeitung großer Mengen von Text können sie Beziehungen zwischen verschiedenen Konzepten erkennen und diese in einem strukturierten Format wiedergeben.

**Automatisierung:** LLMs ermöglichen eine teilweise oder vollständige Automatisierung des Prozesses zur Erstellung von Wissensgraphen. Das bedeutet, dass Sie große Mengen von unstrukturierten Daten verarbeiten und in einem organisierten Wissensgraphenformat ausgeben können, ohne dass menschliche Expertise für jeden Schritt erforderlich ist.

**Skalierbarkeit:** Durch die Nutzung von LLMs können Wissensgraphen auf große Mengen von Daten angewendet werden, was die Skalierbarkeit des Systems verbessert. Dies ist besonders nützlich, wenn Sie mit umfangreichen und sich ständig ändernden Datenmengen arbeiten.

**Echtzeitaktualisierung:** Da LLMs in der Lage sind, neue Informationen zu verarbeiten, können Wissensgraphen schnell aktualisiert werden, wenn sich der zugrunde liegende Wissensstand ändert. Dies ermöglicht eine dynamische und aktuelle Darstellung von Informationen.

**Semantische Suche und Analyse:** Ein Wissensgraph, der mithilfe von LLMs erstellt wurde, ermöglicht eine semantische Suche, bei der die Bedeutung von Begriffen und Beziehungen verstanden wird. Dies verbessert die Präzision von Suchanfragen

und ermöglicht fortgeschrittene Analysen.

Anwendungen in KI-Systemen: Wissensgraphen, die von LLMs erstellt wurden, können als Grundlage für KI-Systeme dienen, die komplexe Aufgaben wie maschinelles Lernen, Sprachverarbeitung und Entscheidungsfindung durchführen.

**Schlüsselbegriffe:** Wissensgraphen, Künstliche Intelligenz, NaLLM Framework

# Table of Contents

<b>Declaration</b>	<b>i</b>
<b>Abstract</b>	<b>ii</b>
<b>Table of Contents</b>	<b>iv</b>
<b>List of Tables</b>	<b>vi</b>
<b>List of Figures</b>	<b>vii</b>
<b>Chapter 1 Einführung</b>	<b>1</b>
1.1 Motivation . . . . .	1
1.2 Aufgabenstellung und Zielsetzung . . . . .	1
1.3 Ziel der Arbeit . . . . .	1
1.4 Aufbau der Arbeit . . . . .	1
<b>Chapter 2 Stand der Technik</b>	<b>2</b>
2.1 Was ist ein Graph? . . . . .	2
2.2 Was ist ein Wissensgraph . . . . .	2
2.3 Kennzeichnung von Wissensgraphen . . . . .	2
2.4 Konstruktion von Wissensgraphen . . . . .	2
2.5 Framework NaLLM . . . . .	2
<b>Chapter 3 Implementierung</b>	<b>3</b>
3.1 NaLLM funktionsweise . . . . .	3

3.2	Wissensgraphen aus Reports . . . . .	3
3.3	Wissensgraphen auswerten vergleichen . . . . .	3
3.4	Zusammenfassung . . . . .	3

# List of Tables

# List of Figures



# Chapter 1

## Einführung

### 1.1 Motivation

### 1.2 Aufgabenstellung und Zielsetzung

### 1.3 Ziel der Arbeit

Ist der erstellte Wissensgraph qualitativ gut im Vergleich zu einem manuell erstellten Wissensgraphen? Wie geht das Framework mit redundanten Daten um oder mit einem nicht wissensreichen Text? Wie sieht der Wissensgraph aus bei Nachtragung? Ist das Erstellen von Wissensgraphen mit Hilfe von großen Sprachmodellen zuverlässig?

### 1.4 Aufbau der Arbeit

# Chapter 2

## Stand der Technik

### 2.1 Was ist ein Graph?

### 2.2 Was ist ein Wissensgraph

### 2.3 Kennzeichnung von Wissensgraphen

### 2.4 Konstruktion von Wissensgraphen

Wissensgraphen Konstruktion erzielt das extrahieren von strukturierten Daten von unstrukturierten Texten, mit verschiedenen Methoden Named Entity Recognition (NER) (Chiu and Nichols, 2016), Relation Extraction (RE) (Zeng et al., 2015), Event Extraction (EE) (Chen et al., 2015), Entity Linking (EL) (Shen et al., 2015), and Knowledge Graph Completion (Lin et al., 2015).

NER hat das Ziele verschieden Entitäten zu Identifizieren. RE extrahiert die Beziehung zu den gegebenen Entitäten EE EL versucht zwischen Entitäten eine Beziehung zu finden

### 2.5 Framework NaLLM

# Chapter 3

## Implementierung

3.1 NaLLM funktionsweise

3.2 Wissensgraphen aus Reports

3.3 Wissensgraphen auswerten vergleichen

3.4 Zusammenfassung