

# Reguläres Meeting

---

Adrien Klose

`adrien.klose@student.uni-halle.de`

Martin-Luther-Universität Halle-Wittenberg

Institut für Informatik

31.05.2024

# Tätigkeiten

- ▶ Triple für BioASQ auflösen
  - ▶ Unbrauchbar (Code zu Name) oder nicht auflösbar (Endpoints existieren nichtmehr)
- ▶ 10 jeweils mit 3 Triple "willkürlich" ausgewählt
- ▶ Baseline, AIONR, AIONLR getestet

# Experiment

- ▶ GPT3.5 über ChatGPT Interface
- ▶ CoT nich sinnvoll machbar

Baseline: You are a helpful medical expert who answers with either yes or no to questions.

Question: question

AIONR: You are a helpful medical expert who answers with either yes or no. Consider the information in triple quotes as factual knowledge.

""""s has relationshsip p to o""""

Question: question

AIONLR: You are a helpful medical expert who answers with either yes or no. Consider the information in triple quotes as factual knowledge.

""""s, p, o as a natural sentence""""

Question: question

# Experiment Ergebnisse

Ansatz	GT is No	GT is Yes
Baseline	8/10	7/10
AIONR	7/10	6/10
AIONLR	9/10	7/10

- ▶ AIONR und AIONLR haben unterschiedliche Ergebnisse
- ▶ AIONR manchmal besser, manchmal schlechter
- ▶ Unterschied AIONR und AIONLR ggf. noch größer wenn weniger eindeutige Relationen

# Experimentaufbau

- ▶ nutze API, system "You are only allowed to give yes or no as an answer"
- a) Baseline: nur die Frage geben
- b) All in One None Resolve: questions + triples in raw
- c) All in One NL Resolve: questions + triples each as NL sentence
- d) CoT: give question, new prompt "given past + new triple reevaluate your answer"
- ▶ Möglichkeit genauere Informationen/Beschreibung zu Entitäten und Relationen in den System Prompt mit zu packen

# Problem

- ▶ GPT hat keinen Free Trial
  - ▶ Research Access
  - ▶ LLama 2 7B lokal (ggf. auch groq)
  - ▶ Mixtral über groq API (Adrian)
  - ▶ ChatGPT Interface
- ▶ Triple Auswahl
  - ▶ Random
  - ▶ Semantic Similarity zu Relationen
- ▶ Entity Extraction

# Next Steps

- ▶ Automatisierung des Experimentes für größere Datenmengen mit groq
- ▶ Meeting 10.6?
- ▶ Erreichbarkeit während Dienstreise?