



Fantasy or Philosophy?

The Philosopher's Spellbook



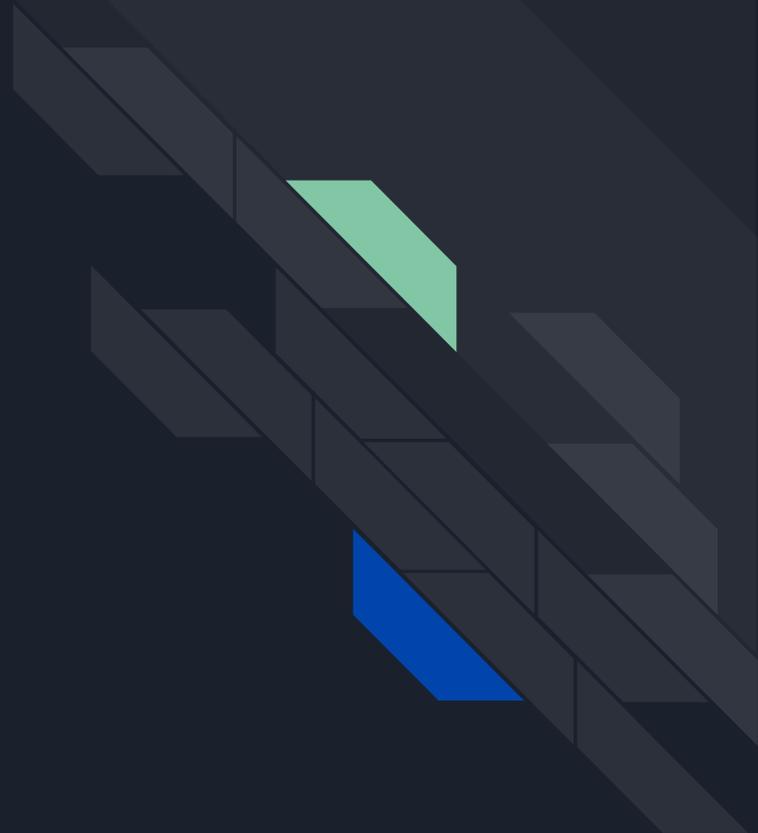
Hermine Kleiner, Sarah Anna Uffelmann, Sebastian Loftus
03.07.2024



Overview

- 1) Data
- 2) RAG-Pipeline and Architecture
 - *Demonstration* –
- 3) Evaluation and Results
- 4) Challenges and Future Work

Data





Data

Harry Potter

- Seven books by J. K. Rowling, 1997-2007
- **PDF**-Downloads from the Internet (English Corpus)
- **Number of Pages: 3.407**
- + Harry Potter and the Cursed Child

Ludwig Wittgenstein

- *Nachlass*
- Philosophical Works and Diaries, 1914-1951
- XML Downloads from the *Wittgenstein Archives*, transformed into **PDF**
- mostly German, some English
- **Number of Pages: >20.000**



Pre-Processing

Harry Potter

- Remove line breaks
- Remove Page Numbers
- Separate Chapter Number + Headlines
- Separate Chapter Headings from the first line of the text

Ludwig Wittgenstein

- Remove links
- Remove editorial notes, footnotes and annotations
- Remove superfluous whitespaces

RAG-Pipeline and Architecture

The background features a series of dark grey, 3D-style rectangular blocks arranged in a diagonal pattern from the top right towards the bottom left. Two blocks are highlighted: one in light green and one in blue, positioned in the middle-right section of the arrangement.



Models

LLM

- Mistral-7B (Ollama)
- 7.25 Billion Parameters
- Token Limit: 8192

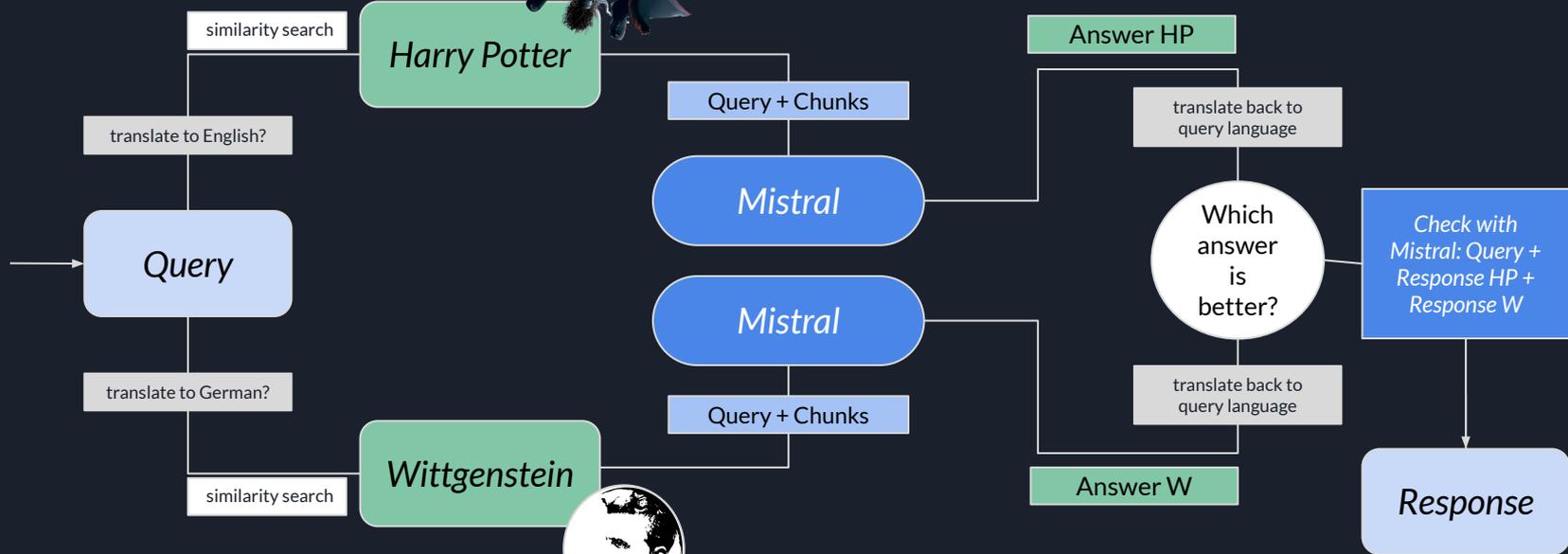
Embeddings

- Multilingual Embeddings
- Hugging Face:
sentence-transformers/stsb-xlm-r-multilingual



Libraries

- **RAG:** LangChain (Vector Store: DB Chroma)
- **Translation:** translate (Google)
- **Language Detection:** papluca/xlm-roberta-base-language-detection
- **PDF-Reader:** PyPDFD
- **Evaluation:** Evaluate
- **GUI:** Gradio



RAG-Pipeline



Prompts for Question-Answering

System Prompt

You are an AI language model assistant for {topic}. You must answer in {lang}. Your task is to answer a question ONLY based only on the following context: {context}

You will be punished if you use your own knowledge.

Topic: Harry Potter OR Wittgenstein

Language:

if Harry Potter: English

if Wittgenstein: German

Prompt

Here is the question to answer: {question}

Keep in mind, you are only allowed to use the provided context!

Question: Query entered by the User



Prompts for Deciding Which Answer Is Better

System Prompt

Your task is to decide which of the two generated answer options is more suitable to answer the given question: {options}

You are not allowed to use any information besides these two options. You will be punished if you use your own knowledge.

Context: Answer for Harry Potter +
Answer for Wittgenstein

Prompt

Here is the question to answer with the help of the provided options: {question}

If you think the Harry Potter answer is better, return: "9 3/4". If you prefer the Wittgenstein answer, you return "227". If you like both, return "42". Don't give any reasoning. You will be punished if you answer more than stated.

Question: Query entered by the User

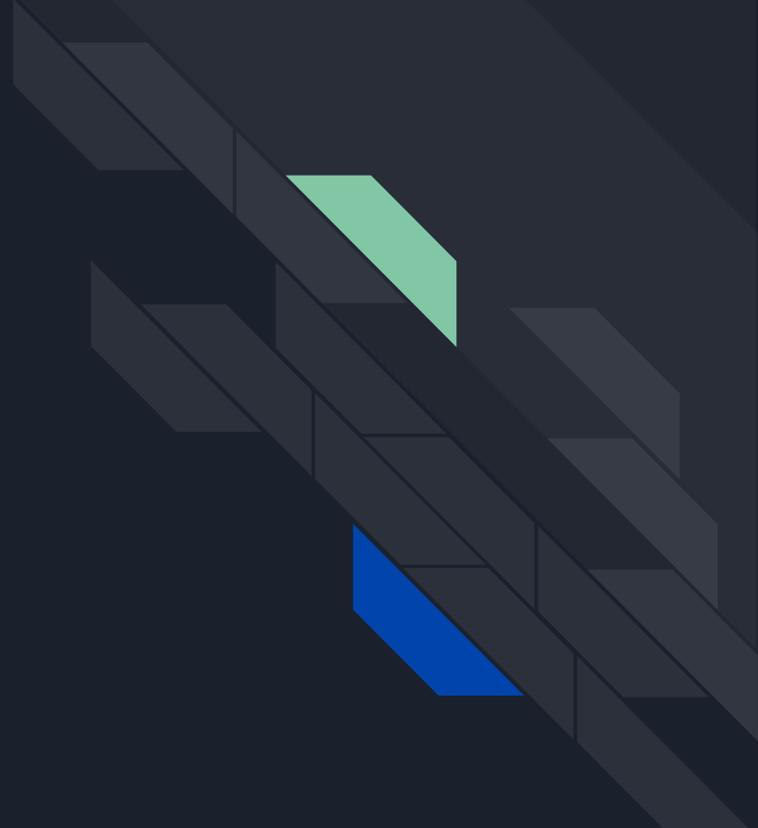
Try it yourself:

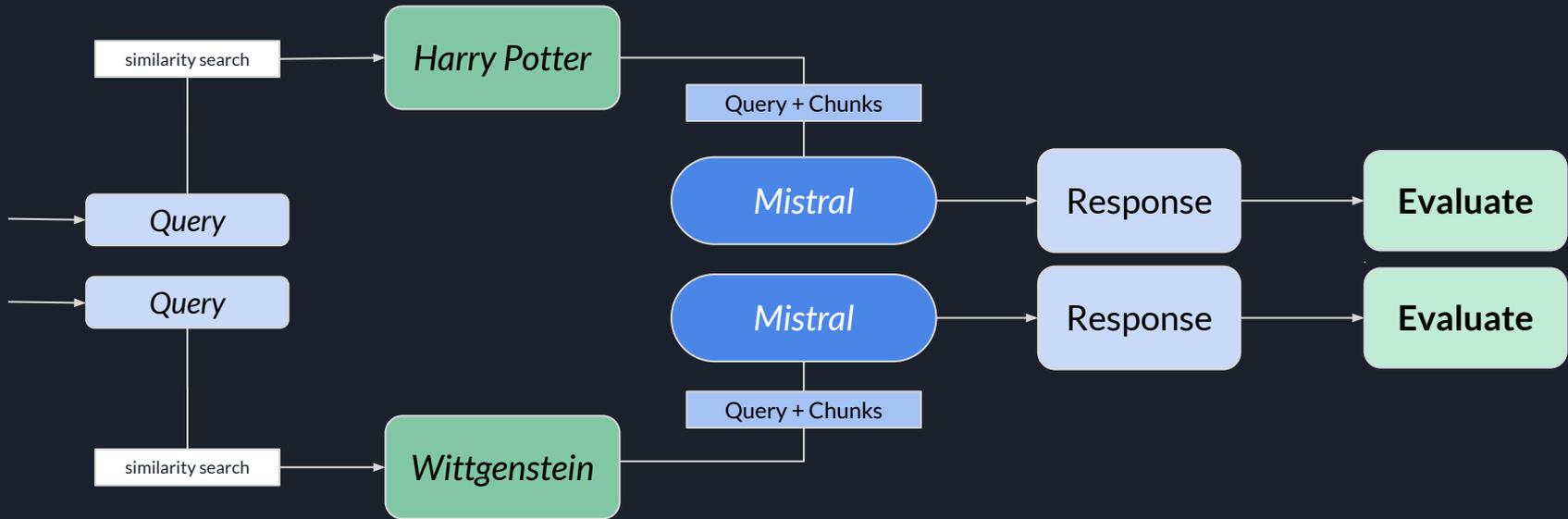
<https://645d7b37bb491fe84c.gradio.live>

(not active anymore)



Evaluation and Results





Pipelines for Evaluation



Metrics

Metrics

- **Bleu** - *Weighted geometric mean of modified n-gram precision with n=4 (default)*
- **RougeN** - *Based on recall of ngrams (with n=1 and n=2)*
- **RougeL** - *Based on recall of Longest Common Subsequence*
- **Meteor** - *Based on combination of unigram-precision, unigram-recall, and a measure of alignment*



Parameters

Chunk Size (tokens)

- 100
- 200
- 300

Chunks included in the Prompt (K)

- 20
- 15
- 10



Gold Standard

- 40 question-answer pairs for each corpus
- Answers:
 - List of 4 correct answer variations
- Context:
 - List of as many text passages as contain the answer
(in most cases just 1 text passage)



Gold Standard Example

```
{  
  "question": "Who killed Bellatrix Lestrange?",  
  "answers": ["Bellatrix Lestrange was killed by Molly Weasley.",  
              "Molly Weasley killed Bellatrix Lestrange in the Battle of Hogwarts.",  
              "Bellatrix Lestrange was a death eater who was killed by Molly Weasley.",  
              "The evil witch Bellatrix Lestrange was killed by Molly Weasley."],  
  "context": ["Bellatrix was still fighting ..."]  
}
```

Results

| | chunk size | k | Bleu | Rouge1 | Rouge2 | RougeL | Meteor |
|--------------|------------|----|------|--------|--------|--------|--------|
| HP gold | - | - | 0.23 | 0.58 | 0.43 | 0.52 | 0.67 |
| Wittg gold | - | - | 0.23 | 0.54 | 0.38 | 0.50 | 0.61 |
| Harry Potter | 200 | 15 | 0.09 | 0.33 | 0.19 | 0.27 | 0.43 |
| Wittgenstein | 200 | 15 | 0.07 | 0.25 | 0.13 | 0.20 | 0.35 |
| Harry Potter | 100 | 20 | 0.09 | 0.35 | 0.22 | 0.31 | 0.48 |
| Wittgenstein | 100 | 20 | 0.04 | 0.22 | 0.10 | 0.18 | 0.32 |
| Harry Potter | 300 | 10 | 0.10 | 0.36 | 0.22 | 0.31 | 0.45 |
| Wittgenstein | 300 | 10 | 0.06 | 0.25 | 0.13 | 0.21 | 0.35 |

Challenges and Future Work

An abstract graphic on the right side of the slide. It features several overlapping, dark grey, 3D-style rectangular blocks arranged in a diagonal pattern from the top right towards the bottom left. Two of these blocks are highlighted with different colors: one is light green and another is bright blue. The background is a dark, solid blue.



Technical Challenges

- Hardware Restrictions
- Finding an open-source LLM with good performance
- Retrieval didn't work well at first - better with multilingual embeddings



Challenges Related to the Wittgenstein Corpus

For the Wittgenstein Corpus it was challenging to find suitable question-answer pairs for the evaluation. In most cases the context is rather short.



Challenges Related to the *Harry Potter* Corpus

Implicit meaning is key in literature, yet RAG models struggle with retrieving answers to questions when the information is not explicitly stated in the text.



Example

Question: Who left the sword of Gryffindor in the forest for Harry and Hermione to find?

Context: Now, Severus, the sword! Do not forget that it must be taken under conditions of need and valor — and he must not know that you give it! If Voldemort should read Harry’s mind and see you acting for him —”
“I know,” said Snape curtly. He approached the portrait of Dumbledore and pulled at its side. It swung forward, revealing a hidden cavity behind it from which he took the sword of Gryffindor.

Generated Answer: Based on the provided context, it is not possible to definitively determine who left the sword of Gryffindor in the forest for Harry and Hermione to find. The sword was taken by Snape from behind a hidden cavity within the portrait of Dumbledore, but no information is given about its origin or the person who placed it there.



Future Work

- Compare performance for different chunk sizes
- Increase number of chunks for generation
- Compare performance on a wider range of domains
- Include re-ranking in the retrieval pipeline



Let's conclude with GPT

*Thank you for exploring the magic and
meaning with Harry Potter and Wittgenstein!*