

# Szenario Ideen

## Inhalt

Aufgabe 1 – Erstellung eines Harry-Potter-Experten durch Prompt Engineering .....	2
Aufgabe 2 - Dokument laden .....	3
Aufgabe 3 – Embeddings .....	4
Aufgabe 4 – Text-Splitter.....	5
Aufgabe 5 -Retriever.....	6

## Harry Potter Experte

Harry Potter hat einen Hype erfahren, du kannst aber nicht mitreden, weil du die Bücher nie gelesen hast. Zeit hast du dafür auch nicht.

Damit du mitreden kannst, willst du dir in diesem Workshop einen Experten für das Harry Potter Universum zu schaffen mithilfe von KI. Als Wissensgrundlage werden dir von uns vorbereitete Dateien dienen.

Wenn das nächste Mal Harry Potter thematisiert wird, bist du der Profi. Die KI soll dir ermöglichen, dass dir das Wichtigste zusammengefasst wird.

## Aufgabe 1 – Erstellung eines Harry-Potter-Experten durch Prompt Engineering

Dein Ziel ist es, einen Chatbot zu entwickeln, der sich ausschließlich durch die Gestaltung des System Prompts als umfassender Experte für das Harry-Potter Universum auskennt. Es dürfen keine zusätzlichen Dokumente oder externen Informationen verwendet werden.

Schaue dir den initialen System Prompt an und baue ihn so um, dass der Chatbot in der Lage ist, detaillierte, präzise und informative Antworten auf alle Fragen zu den sieben Büchern der Hauptreihe von J.K. Rowling zu geben.

Der Chatbot soll sich wie ein enthusiastischer und allwissender Harry-Potter-Fan verhalten und in der Lage sein, auf spezifische Ereignisse, Charaktere, Orte, Zaubersprüche, magische Gegenstände und historische Hintergründe einzugehen. Achte darauf, dass der Prompt den Chatbot anweist, Spekulationen zu vermeiden und sich strikt an den Kanon der Bücher zu halten.

Nutze dazu die Techniken Role Prompting und Shot Prompting. Nachdem du den System Prompt angepasst hast, stelle dem Chatbot Fragen und prüfe das Ergebnis.

Anleitung:

1. Optimierte den System Prompt gem. Szenario
2. Erstelle eine Anweisung im System Prompt
3. Ergänze den System Prompt um eine Rolle (Role Prompting)
4. Ergänze Informationen, wie der Chatbot antworten soll (z.B. enthusiastisch)
5. Ergänze den System Prompt um Beispiele (Few Shot Prompting)
6. Stelle dem Chatbot Fragen

Ziel: Code und den Chatbot Client kennenlernen

## Aufgabe 2 - Dokument laden

Der Chatbot kennt zwar den Inhalt der Hauptreihe, aber nun möchtest du dem Chatbot neues Wissen hinzugeben, welches er noch nicht kennt. Dazu sind in der Datei `HP_erfundene_aussagen.md` zwei frei erfundene Aussagen hinterlegt, die etwas über Harry Potter erzählen. Nachdem du den Code so angepasst hast, dass das Dokument geladen und auch übergeben wird, schaue dir den Inhalt der Datei an und frage den Chatbot etwas ganz Spezielles, was er nur wissen kann, wenn er den Inhalt von `HP_Kindheit.md` kennt. So kannst du überprüfen, ob der Chatbot das Dokument auch wirklich kennt.

P.S.: Denk daran, damit der Chatbot die neuen Informationen auch hat, musst du die Anwendung einmal neustarten

Anleitung:

1. Lade das Dokument „`HP_erfundene_aussagen.md`“ mit dem Document Loader von Langchain
2. Erweitere den System Prompt um Context und passe es so an, dass der Context aus dem bereitgestellten Dokument priorisiert wird.
3. Stelle sicher durch Anpassung des System Prompt: Wenn eine Frage nicht beantwortet werden kann, soll der Chatbot das offen zu geben und nichts erfinden.
4. Füge in der "predict function" das Dokument als Input hinzu
5. Frage etwas ganz Spezielles, was er nur wissen kann

(Optional) Challenge: Weitere Document Loader inkl. zusätzlichem Dokument ausprobieren, z.B. PyPDF

## Aufgabe 4 – Embeddings

Nun bekommt dein Chatbot nicht nur ein Dokument, sondern gleich mehrere. Und manche davon enthalten nicht nur ein paar Sätze, sondern durchaus etwas mehr. Zudem bekommt der Chatbot nun nicht mehr den kompletten Inhalt der Dateien, sondern ein Embedding jeweils. Schaue dir die Dokumente an und stelle dem Chatbot gezielt Fragen zu den Dokumenten. Überprüfe, ob der Chatbot deine Fragen gut beantwortet anhand der Informationen aus den Dokumenten.

Achte zudem darauf, dass du deine query für die Similarity Search so anpasst, dass nur Dokumente genutzt werden, die etwas mit Harry Potter zu tun. Das Dokument „HdR\_erfundene\_Aussagen“ sollte dein Chatbot nicht kennen, hier geht es nämlich um Herr der Ringe. Frage den Chatbot etwas über dieses Dokument, um zu prüfen, ob er die Informationen daraus hat. Sofern er die Informationen kennt, schaue deine query für die Similarity Search erneut an und passe sie an, sodass dieses Dokument nicht miteinbezogen wird.

Schaue dir das Dokument „HP\_der\_verschwundene\_kuchen.md“ an. Hier stehen einige Sätze, die nichts mit Harry Potter zu tun haben. Teste, ob der Chatbot dir Fragen hierzu beantworten kann.

Beispiele:

- Kennst du Simon?
- Weißt du, was Glück ist?

P.S.: Denk daran, damit der Chatbot die neuen Informationen auch hat, musst du die Anwendung einmal neustarten

Anleitung:

1. Erstelle aus den Dokumenten einen Vector Store
2. Erstelle einen Prompt für die Similarity Search
3. Ergänze die relevanten Dokumente als Context
4. Rufe in der PDF Aufgabe 3 auf und bearbeite diese.

## Aufgabe 5 – Text-Splitter

Für diese Aufgabe kannst du die gleichen Dokumente verwenden. Schaue, ob durch den Text-Splitter sich die Qualität der Antworten verändert. Lasse dir zudem die Anzahl der Chunks ausgeben, um zu sehen, in wie viele Stücke die Dokumente zerlegt wurden.

Schaue dir das Dokument „HP\_der\_verschwundene\_kuchen.md“ erneut an. Teste, ob der Chatbot dir Fragen zu den Sätzen, die nichts mit Harry Potter zu tun haben, weiterhin beantworten kann.

Beispiele:

- Kennst du Simon?
- Weißt du, was Glück ist?
- Was ist mein Lieblingsgericht?

P.S.: Denk daran, damit der Chatbot die neuen Informationen auch hat, musst du die Anwendung einmal neustarten

## Aufgabe 5 -Retriever

In dieser Aufgabe kannst du ebenfalls die gleichen Dokumente nutzen, wie zuvor.

Wenn du bisher alles richtig gemacht hast, sollte dir aufgefallen sein, dass nach der Implementierung des Text-Splitters der Chatbot die Fragen zu den zufälligen Sätzen aus dem Dokument „HP\_der\_verschundene\_kuchen.md“ nicht mehr beantworten konnte. Das liegt daran, dass diese mit der von dir formulierten Query und der Similarity Search aussortiert wurden.

Nachdem du nun einen Langchain Retriever eingebaut hast, achte darauf die zuvor genutzte Query und die Similarity Search auszubauen, diese brauchen wir nun nicht mehr.

Stelle nun dem Chatbot erneut Fragen zu den zufälligen Sätzen aus „HP\_der\_verschundene\_kuchen.md“ und schaue, ob diese beantwortet werden können.

Beispiele:

- Kennst du Simon?
- Weißt du, was Glück ist?
- Was ist mein Lieblingsgericht?

P.S.: Denk daran, damit der Chatbot die neuen Informationen auch hat, musst du die Anwendung einmal neustarten