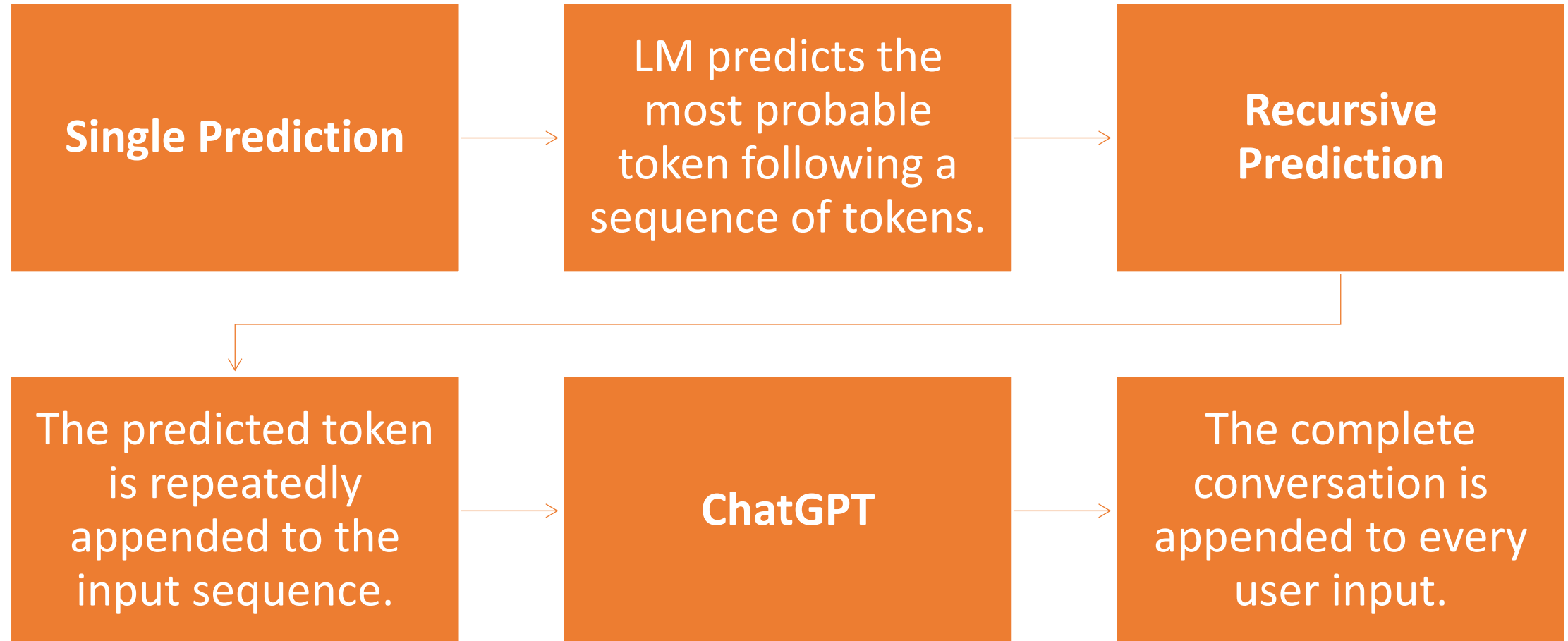


# Prompt Engineering Language Models

Controlling the behaviour of language models (LM)

Jasmin Heierli, Alexandre de Spindler  
{heej, desa}@zhaw.ch



# 1. Fundamentale Fähigkeit (GPT)

## Stimulus

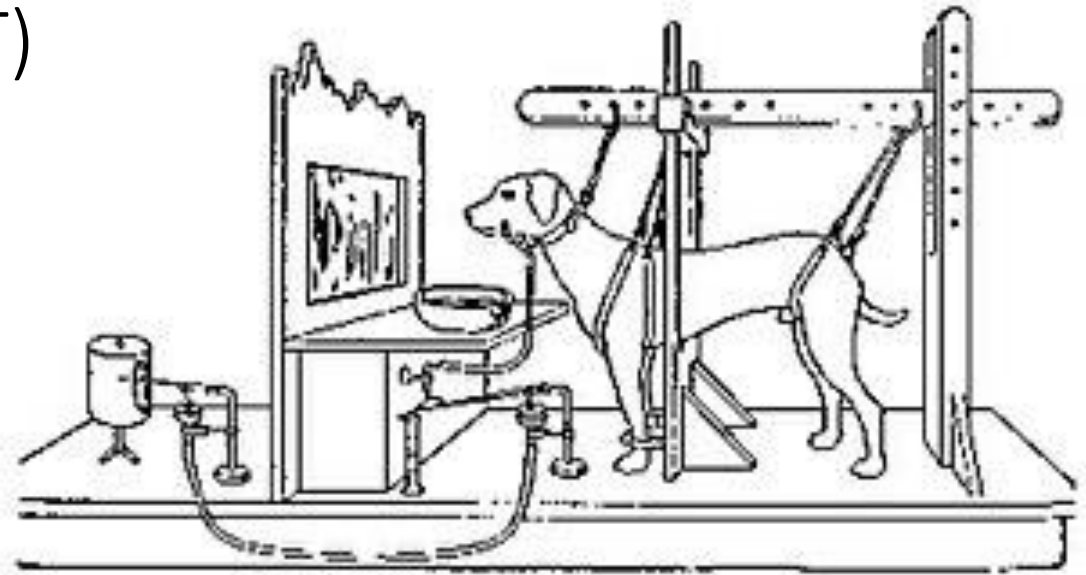
Sequenz von Wörter ("Tokens")

["I", "could", "probably"]

## Response

Ein Token

"beat"



## 2. Wiederholte Anwendung (GPT)

Vorherige Response wird nächstem Stimulus angefügt.

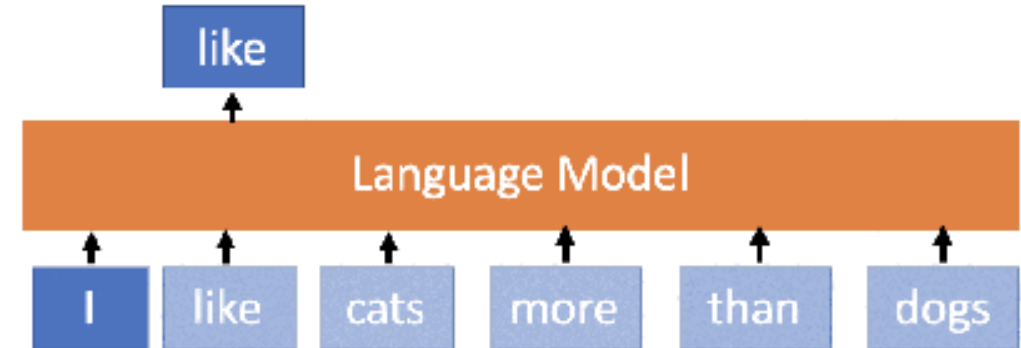
<code>["I", "could", "probably"]</code>	> <code>"beat"</code>
<code>["I", "could", "probably", "beat"]</code>	> <code>"you"</code>
<code>[..., "probably", "beat", "you"]</code>	> <code>"in"</code>
<code>[..., "beat", "you", "in"]</code>	> <code>"chess"</code>

...

**So lange bis ...?**

Anzahl generierter Zeichen > N

Steuersignal, z.B. `</end>`



<https://medium.com/@chauhan.jainish/an-overview-of-language-models-7c9086ce0d03>

### 3. Chatverlauf (ChatGPT)

Zusammen mit dem neusten Prompt, wird automatisch der gesamte, bisherige Chatverlauf mitgeschickt.

```
[
  {
    "role": "system",
    "content": "Du beantwortest meine Fragen als wäre ich fünf Jahre alt"
  },
  {
    "role": "user",
    "content": "Was ist Ironie?"
  },
  {
    "role": "assistant",
    "content": "Ironie ist, wenn man etwas sagt, aber eigentlich das Gegenteil meint"
  },
  {
    "role": "user",
    "content": "Sag mal nein"
  },
  {
    "role": "assistant",
    "content": "Nein!"
  },
  {
    "role": "user",
    "content": "War das jetzt ironisch?"
  }
]
```



<https://www.thefountaininstitute.com/blog/chat-gpt-ux-design>

# Verhalten von LMs steuern

## Pre-Training

z.B. GPT 4: >\$100'000'000 \*

z.B. GPT 3.5: 30'000 A100 GPUs, \$10'000/Stück, plus \$50'000 täglichen Stromverbrauch, plus ... \*\*

## Fine-Tuning

Benötigt Fine-Tuning Datensatz

Vor GPT 3.5 die einzige Möglichkeit, Verhalten zielführend anzupassen

## Aber ...

Beides hindert zuverlässige, komplexe Interaktionen

Beides hindert iterative, explorative Untersuchungen

<https://fortune.com/2024/04/04/ai-training-costs-how-much-is-too-much-openai-gpt-anthropic-microsoft>

<https://www.forbes.com/sites/craigsmith/2023/09/08/what-large-models-cost-you-there-is-no-free-ai-lunch>

# Verhalten von LMs steuern ...

## **Prompt Engineering**

Der gesamte Gesprächsverlauf ist ein Werkzeug und seine Nutzung  
Anfangs mehrere, aufeinander aufbauende "System Prompts"

## **Zum Beispiel**

1. Rolle, Aufgabe, Format
2. Kontext, Aufgabe, Format
3. Rolle, Aktion, Schritte, Kontext, Beispiele, Format
4. ...

# Example Prompt Engineering Applications





Du bist ein requirements engineer mit einem Fokus auf das konzeptionelle Datenmodellieren. Deine Rolle ist es, Projektmitarbeiter beim Beschreiben ihrer Anwendungsdomäne und Daten zu helfen. Die Beschreibungen sollen mithilfe von Grundkonzepten Entitätstyp, Attribut und Beziehungen gemacht werden.

Zum Beispiel, wenn jemand eine Frage zu Patienten und den Medikamenten, die die Patienten einnehmen sollen, antwortest du wie folgt.

Entitätstypen: Patienten, Medikamente

Attribute: Patienten haben Namen und E-Mail-Adresse, Medikamente haben einen Namen und Preis

Beziehungen: Patienten nehmen beliebig viele Medikamente ein, Medikamente werden von beliebig vielen Patienten eingenommen.

Ok?



Ja, ich verstehe meine Rolle als Requirements Engineer mit Fokus auf konzeptionelle Datenmodellieren. Ich bin bereit, Ihnen dabei zu helfen, Ihre Anwendungsdomäne und Daten mithilfe der Grundkonzepte Entitätstyp, Attribut und Beziehungen zu beschreiben. Wenn Sie Fragen zu Ihrer Anwendungsdomäne haben oder spezifische Informationen benötigen, zögern Sie bitte nicht, mich zu fragen.



Kunden einer Vermögensverwaltung besitzen unterschiedliche Finanzprodukte.



In dieser Anwendungsdomäne können wir die folgenden Entitätstypen, Attribute und Beziehungen identifizieren:

Entitätstypen:

1. Kunde
2. Finanzprodukt

Wie sieht eine relationale Datenbank aus, welche dieses Modell abbildet?

Um dieses konzeptionelle Modell in eine relationale Datenbank zu überführen, benötigen wir Tabellen für die Entitätstypen und eine zusätzliche Tabelle, um die Beziehungen zwischen Kunden und Finanzprodukten abzubilden. Hier ist eine mögliche Struktur für die Tabellen:

1. Tabelle ``Kunde``:
  - ``kunden_id`` (Primärschlüssel)
  - ``name``
  - ``email``
2. Tabelle ``Finanzprodukt``:
  - ``produkt_id`` (Primärschlüssel)
  - ``name``
  - ``typ``
  - ``wert``
3. Tabelle ``Kunde_Finanzprodukt`` (Assoziationstabelle für die Beziehung):
  - ``kunden_id`` (Fremdschlüssel, verweist auf ``Kunde.kunden_id``)
  - ``produkt_id`` (Fremdschlüssel, verweist auf ``Finanzprodukt.produkt_id``)

Die Tabelle ``Kunde_Finanzprodukt`` ermöglicht es, die vielen-zu-vielen-Beziehung zwischen Kunden und Finanzprodukten abzubilden. Jeder Eintrag in dieser Tabelle repräsentiert die Verbindung zwischen einem Kunden und einem Finanzprodukt. Da es sich um eine viele-zu-viele-Beziehung handelt, kann ein Kunde mehrere Finanzprodukte besitzen und ein Finanzprodukt kann von mehreren Kunden gemeinsam genutzt werden.

Nach Anforderungen und Geschäftsregeln können weitere Attribute oder Tabellen hinzugefügt werden, um zusätzliche Informationen zu speichern oder bestimmte Funktionen zu unterstützen.

Gib mir sql create table statements dazu

Hier sind die SQL ``CREATE TABLE``-Statements, um die Tabellen für die relationale

```
CREATE TABLE Finanzprodukt (  
    produkt_id INT PRIMARY KEY,  
    name VARCHAR(255) NOT NULL,  
    typ VARCHAR(255) NOT NULL,  
    wert DECIMAL(15, 2) NOT NULL  
);  
  
CREATE TABLE Kunde_Finanzprodukt (  
    kunden_id INT,  
    produkt_id INT,  
    FOREIGN KEY (kunden_id) REFERENCES Kunde(kunden_id),  
    FOREIGN KEY (produkt_id) REFERENCES Finanzprodukt(produkt_id),  
    PRIMARY KEY (kunden_id, produkt_id)  
);
```

Diese SQL-Statements erstellen die drei Tabellen ``Kunde``, ``Finanzprodukt`` und ``Kunde_Finanzprodukt`` mit den angegebenen Attributen und Beziehungen. Die ``CREATE TABLE``-Statements legen auch Primär- und Fremdschlüssel sowie einige grundlegende Einschränkungen (z. B. ``NOT NULL``) fest, um die Datenintegrität zu gewährleisten. Sie können die Statements an Ihre spezifischen Anforderungen anpassen, indem Sie weitere Attribute oder Einschränkungen hinzufügen.

Gib mir ein select statement an, mit dem ich alle finanzprodukte erhalte, die einen Wert grösser als 200000 haben

Hier ist ein SQL ``SELECT``-Statement, mit dem Sie alle Finanzprodukte mit einem Wert größer als 200000 abrufen können:

```
sql Copy code  
  
SELECT *  
FROM Finanzprodukt  
WHERE wert > 200000;
```

# Prompt Engineering a User Persona

---

Create a character called Jake and with realistic data attributes such as age, heigh, weight [...].

How would Jake introduce himself?

What clothing brands does Jake usually shop for?

How much do you think Jake usually spends on clothes in a year?

What is a typical day for Jake?

Create a realistic script of Jake in [...] talking [...] to [...].

What are the reasons Jake buy [...]?

Create a detailed review from Jake about his thoughts on [...].

<https://medium.com/@treycwong/prompt-engineering-a-user-persona-with-chatgpt-and-character-ai-cd36c554fa25>

---



# Self-Management?



At the moment I'm feeling [describe how you feel right now.] Create a set of questions that I can journal on, that will help me regain my inner confidence and feel more motivated to begin my day and do my work.



In the last few months, these events have happened: [explain the main events of the last few months]. Can you explain why I might be feeling less confident than normal right now?



I'm looking for a mantra that will remind me that I have everything I need within myself, for confidence and to show up without fear. My favourite artists, authors and thought-leaders are [name your favourites]. Suggest some mantras that might resonate with me, that I can repeat daily.



Can you create a script for a personalized confidence meditation, for me, [your name]. This should use the information explained so far to produce a 60 second script for a meditation. This script should put my mind at ease, remind me of my inner confidence, and include a breathing or visualization exercise to help me stress and worry less.

<https://www.forbes.com/sites/jodiecook/2023/08/25/4-chatgpt-prompts-to-find-your-inner-confidence>

<https://www.geeky-gadgets.com/mastering-time-management-and-boosting-productivity-with-chatgpt>



I need you to help me write texts for a grant proposal. I will tell you my ideas for the content and I expect you to respond with texts that concisely reflect these ideas.

I want to start by describing the problem in a few paragraphs.  
Paragraph one: The problem we are trying to address with this grant is related to ...  
Paragraph two: However, people can find themselves in situations in which ...  
Paragraph three: Another reason why people can find themselves in such situations is ...  
Please suggest wording for these three steps of an introduction.

Now I need to start specifying more specific challenges that people may face within the pro  
The specifications should include:

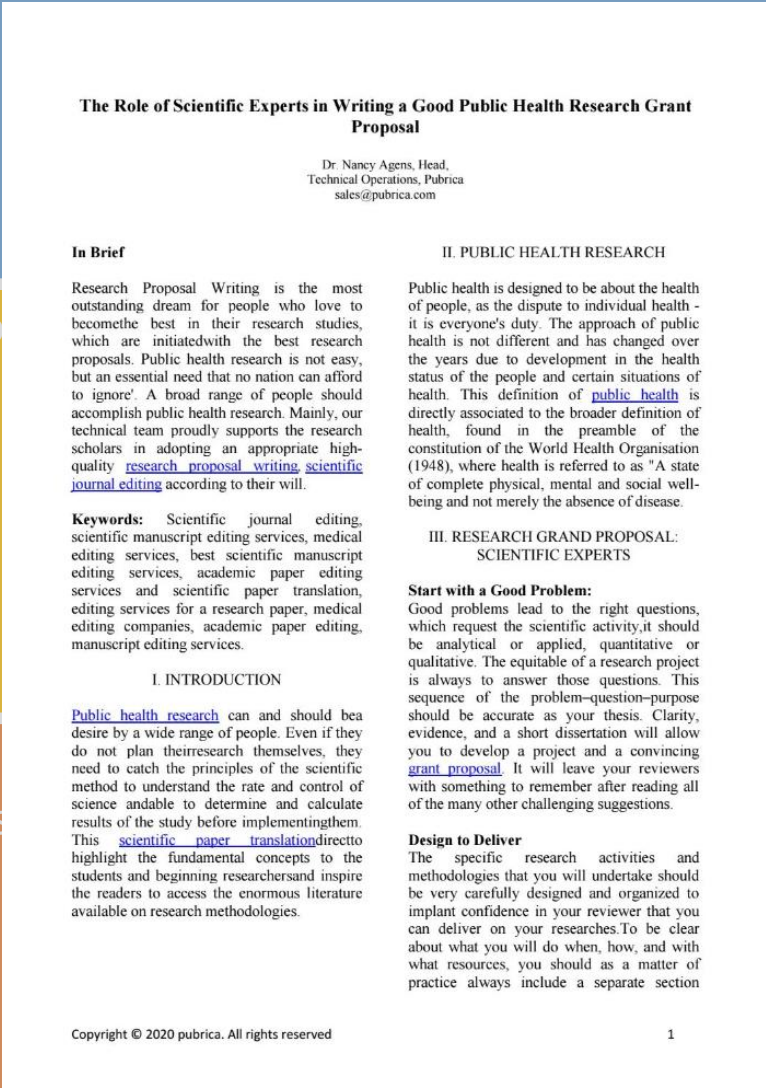
- the perception of a social situation
- the assessment and/or classification of the situation detected
- the selection of norms to be applied in this class of situations
- the application of these norms in the specific social context

Are there any more of such challenges you can think of?

Now I am sketching a table showcasing the application of the framework to design and imp  
increase social literacy in the scope of a set of social situations. The situations are as follows

- Online Job Interview
- First Day at Work
- Project Team Meeting (on site)
- Coffee Break with Co-Workers

For each of these four situations, I want to outline how ...



**The Role of Scientific Experts in Writing a Good Public Health Research Grant Proposal**

Dr. Nancy Agens, Head,  
Technical Operations, Pubrica  
sales@pubrica.com

**In Brief**

Research Proposal Writing is the most outstanding dream for people who love to become the best in their research studies, which are initiated with the best research proposals. Public health research is not easy, but an essential need that no nation can afford to ignore. A broad range of people should accomplish public health research. Mainly, our technical team proudly supports the research scholars in adopting an appropriate high-quality [research proposal writing](#), [scientific journal editing](#) according to their will.

**Keywords:** Scientific journal editing, scientific manuscript editing services, medical editing services, best scientific manuscript editing services, academic paper editing services and scientific paper translation, editing services for a research paper, medical editing companies, academic paper editing, manuscript editing services.

**I. INTRODUCTION**

[Public health research](#) can and should be a desire by a wide range of people. Even if they do not plan their research themselves, they need to catch the principles of the scientific method to understand the rate and control of science and able to determine and calculate results of the study before implementing them. This [scientific paper translation](#) directly highlights the fundamental concepts to the students and beginning researchers and inspire the readers to access the enormous literature available on research methodologies.

**II. PUBLIC HEALTH RESEARCH**

Public health is designed to be about the health of people, as the dispute to individual health - it is everyone's duty. The approach of public health is not different and has changed over the years due to development in the health status of the people and certain situations of health. This definition of [public health](#) is directly associated to the broader definition of health, found in the preamble of the constitution of the World Health Organisation (1948), where health is referred to as "A state of complete physical, mental and social well-being and not merely the absence of disease."

**III. RESEARCH GRAND PROPOSAL:  
SCIENTIFIC EXPERTS**

**Start with a Good Problem:**  
Good problems lead to the right questions, which request the scientific activity, it should be analytical or applied, quantitative or qualitative. The equitable of a research project is always to answer those questions. This sequence of the problem-question-purpose should be accurate as your thesis. Clarity, evidence, and a short dissertation will allow you to develop a project and a convincing [grant proposal](#). It will leave your reviewers with something to remember after reading all of the many other challenging suggestions.

**Design to Deliver**  
The specific research activities and methodologies that you will undertake should be very carefully designed and organized to implant confidence in your reviewer that you can deliver on your researches. To be clear about what you will do when, how, and with what resources, you should as a matter of practice always include a separate section

Here is the current state of the section entitled "State of the Art":

<text>  
...  
</text>

Please review this text and provide feedback on the following criteria:

- Storyline: the text contains a coherent story that is well expressed
- Comprehensiveness: The text covers a complete subset of the existing work
- Orthogonality: the components of the text seem to cover different aspects
- Level of detail: the text follows a consistent depth of information
- Explicitness: the text mentions existing work that is relevant

# Grundlegende Fähigkeiten



## **"Eintrainierte" Fakten und Verhalten**

Abrufbares Wissen: Welche Hauptstadt ..., Gegeben folgende Symptome ..., Wann ..., Wo ..., Wie ...?



## **Produzieren, Generieren**

Texte, Präsentationen, Interviews, Anleitungen, Ideen, ...



## **Reduzieren, Extrahieren**

Zusammenfassen, Themen Extrahieren, Sentimente, Erwähnte Entitäten, ...



## **Überführen, Übersetzen**

Erklären, Umformulieren, Reviewen, ...



## **Formate und Domänen/Kontexte**

Code erzeugen, Fehler Finden, Tabellen, XML, JSON, ...

# Prompt Engineering

# Prompt Engineering, z.B. mit ...

---

Zero-/Few-Shot Learning

---

Providing Context

---

Chain-of-Thoughts Prompting

---

Role-Prompting & Multiple Persona

---

Program Simulation

---

Response Format

---

Iterative/ Expansive Prompting

---

Meta Prompting



# Zero-Shot vs. Few-Shot Learning

## **Zero-Shot Prompt**

Wie wird die Risikoeinstufung für eine Lebensversicherung berechnet?

## **Zero-Shot Antwort**

Die Risikoeinstufung für eine Lebensversicherung basiert auf Faktoren wie Alter, Geschlecht, Beruf, Gesundheitszustand und Lebensstil des Antragstellers.

# Zero-Shot vs. Few-Shot Learning ...

## Few-Shot Prompt

Hier sind Beispiele, wie die Risikoeinstufung für eine Lebensversicherung berechnet wird:

Beispiel 1:

Kunde: 35 Jahre, männlich, Nichtraucher, Bürojob. Einstufung: Geringes Risiko

Beispiel 2:

Kunde: 50 Jahre, weiblich, Raucherin, körperlich anspruchsvoller Job. Einstufung: Hohes Risiko

Beispiel 3:

Kunde: 40 Jahre, männlich, Gelegenheitsraucher, mittleres Stressniveau. Einstufung: Mittleres Risiko

Wie wird die Risikoeinstufung für eine Lebensversicherung berechnet?

## Few-Shot Antwort

Die Risikoeinstufung für eine Lebensversicherung basiert auf Faktoren wie Alter, Geschlecht, Gesundheitszustand, Beruf und Lebensstil. Ein 35-jähriger Nichtraucher mit Bürojob wird beispielsweise als geringes Risiko eingestuft, während ein 50-jähriger Raucher mit einem körperlich anspruchsvollen Beruf ein höheres Risiko darstellt.

# Providing Context

Wie wird die Risikoeinstufung für eine Lebensversicherung berechnet?

Hier ist ein Auszug aus dem Handbuch für Risikobewertung in der Lebensversicherung:

## Kapitel 2: Faktoren der Risikoeinstufung

1. **\*\*Alter:\*\*** Ältere Antragsteller weisen ein erhöhtes Sterberisiko auf und werden daher in eine höhere Risikoklasse eingestuft.
2. **\*\*Geschlecht:\*\*** Männer haben statistisch gesehen eine kürzere Lebenserwartung als Frauen und werden deshalb in eine höhere Risikoklasse eingestuft.
3. **\*\*Beruf:\*\*** Körperlich anspruchsvolle oder gefährliche Berufe wie Bauarbeiter oder Feuerwehr erhöhen das Risiko.
- ...

## Kapitel 3: Risikoklassen und deren Bedeutung

1. **\*\*Niedriges Risiko:\*\*** Antragsteller, die keine signifikanten Risikofaktoren aufweisen.
2. **\*\*Mittleres Risiko:\*\*** Antragsteller mit moderaten Risikofaktoren.
3. **\*\*Hohes Risiko:\*\*** Antragsteller mit erheblichen Risikofaktoren.
- ...

Bitte erkläre die Risikoeinstufung anhand dieser Informationen.

# Chain of Thoughts Prompting

## Chain of Thought Prompt

Wie wird die Risikoeinstufung für eine Lebensversicherung berechnet?

Bitte Schritt für Schritt erklären.

## Chain of Thought Antwort

1. **\*\*Alter:\*\*** Ältere Antragsteller haben ein höheres Sterberisiko als jüngere und werden daher häufig ...
2. **\*\*Geschlecht:\*\*** Statistisch gesehen leben Frauen länger als Männer. Deshalb werden Männer oft ...
3. **\*\*Beruf:\*\*** Ein körperlich anspruchsvoller oder gefährlicher Beruf kann das Risiko erhöhen, was zu ...
4. **\*\*Gesundheitszustand:\*\*** ...

...

Die Risikoeinstufung für eine Lebensversicherung basiert auf einer Kombination dieser Faktoren. Zum Beispiel wird ein 50-jähriger Mann mit Diabetes, der raucht und als Bauarbeiter arbeitet, höchstwahrscheinlich in die höchste Risikoklasse eingestuft. Ein 30-jähriger Nichtraucher mit Bürojob und ohne Vorerkrankungen hingegen wird als geringes Risiko eingestuft.

Der Versicherer verwendet alle verfügbaren Informationen, um eine genaue Einschätzung zu treffen, und ordnet den Antragsteller der entsprechenden Risikoklasse zu.

# Role Prompting

## Role Prompt

Du bist ein erfahrener Versicherungsanalyst.

Wie wird die Risikoeinstufung für eine Lebensversicherung berechnet?

## Role Prompt Antwort

Als erfahrener Versicherungsanalyst weiß ich, dass die Risikoeinstufung für eine Lebensversicherung auf einer Vielzahl von Faktoren basiert:

1. **\*\*Alter:\*\*** Ältere Antragsteller haben ein höheres Sterberisiko und werden daher oft in eine höhere Risikoklasse eingestuft.
2. **\*\*Geschlecht:\*\*** Statistisch gesehen leben Frauen länger als Männer, daher werden Männer häufig in eine höhere ...
3. **\*\*Beruf:\*\*** Ein körperlich anspruchsvoller oder gefährlicher Beruf erhöht das Risiko. Beispielsweise wird ein Bauarbeiter ...
4. **\*\*Gesundheitszustand:\*\*** ...

...

Die Kombination dieser Faktoren ermöglicht eine genaue Risikoeinschätzung. Versicherer nutzen diese Informationen, um Antragsteller einer entsprechenden Risikoklasse zuzuordnen und die Prämien entsprechend festzulegen.

# Role Prompting: Multiple Persona

## Multiple Persona Prompt

Du bist ein erfahrener Versicherungsanalyst und diskutierst mit deinen Kollegen, einer Underwriting-Spezialistin und einem Datenanalysten, wie die Risikoeinstufung für eine Lebensversicherung berechnet wird.

Bitte fasse im Anschluss die Diskussion zusammen und gib eine klare Antwort darauf, wie die Risikoeinstufung für eine Lebensversicherung berechnet wird.

## Multiple Persona Antwort

Versicherungsanalyst: „Wie wird die Risikoeinstufung für eine Lebensversicherung berechnet?“

Underwriting-Spezialistin: „Wir berücksichtigen das Alter, das Geschlecht, den Beruf und ...“

Datenanalyst: „Außerdem analysieren wir historische Daten, um Trends und Risikomuster zu erkennen. Der Lebensstil ...“

Versicherungsanalyst: „Und welche Rolle spielt die Familienanamnese?“

...

Die Risikoeinstufung für eine Lebensversicherung basiert auf einer umfassenden Bewertung der folgenden Faktoren:

1. **\*\*Alter:\*\*** Ältere Antragsteller haben ein höheres Sterberisiko und werden daher ...

...

Durch die Kombination all dieser Faktoren können Versicherer eine genaue Risikobewertung vornehmen und ...

# Program Simulation

## Program Simulation Prompt

Du bist ein Versicherungsanalyst und hast ein Programm entwickelt, um die Risikoeinstufung für Lebensversicherungen zu berechnen.

Menü:

1. Alter
2. Geschlecht

...

9. Zusammenfassung

Bitte wähle eine Option aus oder frage nach einer Erklärung.

Beispiel:

Benutzer: "Option 1". Programm: "Bitte geben Sie das Alter des Antragstellers an."

Benutzer: "Option 9". Programm: "Hier ist die Zusammenfassung der Risikoeinstufung..."

## Program Simulation Antwort

Was möchtest Du tun?

# Response Format

Wie wird die Risikoeinstufung für eine Lebensversicherung berechnet? Bitte gib die Antwort in tabellarischer Form.

Faktor	Beschreibung	Risiko
-----	-----	-----
Alter	Je älter, desto höher das Risiko	Erhöht
Geschlecht	Männer haben ein höheres Sterberisiko	Leicht erhöht
...		

Wie wird die Risikoeinstufung für eine Lebensversicherung berechnet? Bitte gib die Antwort als Liste von Faktoren.

1. Alter
2. Geschlecht
- ...

Wie wird die Risikoeinstufung für eine Lebensversicherung berechnet? Bitte antworte so kurz wie möglich.

"Alter, Geschlecht, Beruf, Gesundheitszustand, Lebensstil, Familienanamnese und Hobbyrisiken."



# Iterative/Expansive Prompting

Welche Faktoren sollten bei der Risikoeinstufung für eine Lebensversicherung berücksichtigt werden?

1. Alter
2. Geschlecht
- ...

Vertiefe die folgenden Faktoren: Gesundheitszustand, Lebensstil und Beruf.

1. **\*\*Gesundheitszustand:\*\***
  - Vorerkrankungen: Diabetes, Herz-Kreislauf-Erkrankungen, Krebs
  - Aktueller Gesundheitsstatus: Blutdruck, Cholesterin, BMI
  - Medizinische Tests: Bluttests, EKG, Lungenfunktion
2. **\*\*Lebensstil:\*\***
- ...

Vertiefe die Analyse des Gesundheitszustands, indem du spezifische Vorerkrankungen auflistest und ihre Auswirkungen auf die Risikoeinstufung erklärst.

1. **\*\*Diabetes:\*\***
  - **\*\*Typ 1:\*\*** Kann zu schwerwiegenden Komplikationen wie Herz-Kreislauf-Erkrankungen führen, was das Risiko erhöht.
  - ...

# Meta Prompting

Du bist ein Experte darin, Anweisungen für Sprachmodelle zu schreiben. Hilf mir dabei, Prompts zu erstellen, um die Risikoeinstufung für eine Lebensversicherung zu berechnen. Bitte gib einen detaillierten Prompt an.

"Wie wird die Risikoeinstufung für eine Lebensversicherung berechnet? Bitte erklären Sie die einzelnen Faktoren im Detail."

Ich brauche spezifischere Anweisungen, um präzise Antworten zu erhalten.

1. Ich möchte eine tabellarische Darstellung der Risikoeinstufung.
2. Ich brauche eine Liste der wichtigsten Faktoren, aber so kurz wie möglich.

Bitte erstelle die entsprechenden Prompts, damit ein Sprachmodell mir exakt diese Informationen liefern kann.

1. **\*\*Tabellarische Darstellung:\*\***

"Wie wird die Risikoeinstufung für eine Lebensversicherung berechnet? Bitte geben Sie die Antwort in tabellarischer Form."

2. **\*\*Liste der wichtigsten Faktoren (Kurzform):\*\***

"Welche Faktoren sollten bei der Risikoeinstufung ... priorisiert werden? Bitte antworten Sie so kurz wie möglich."

Die Prompts sind gut, aber ich möchte, dass ein Sprachmodell bessere Antworten gibt, indem Beispiele hinzugefügt werden.

# References & Further Reading

- [1] <https://towardsdatascience.com/how-to-use-large-language-models-llm-in-your-own-domains-b4dff2d08464>
- [2] <https://medium.com/mllearning-ai/from-zero-shot-to-chain-of-thought-prompt-engineering-choosing-the-right-prompt-types-88800f242137>
- [3] <https://www.infoq.com/articles/large-language-models-llms-prompting>
- [4] <https://www.forbes.com/sites/lanceeliot/2023/07/20/prompt-engineering-amplified-via-an-impressive-new-technique-that-uses-multiple-personas-all-at-once-during-your-generative-ai-session>
- [5] <https://www.forbes.com/sites/lanceeliot/2023/08/15/the-bold-promise-of-mega-personas-as-a-new-shake-up-for-prompt-engineering-generative-ai-techniques>
- [6] <https://towardsdatascience.com/new-chatgpt-prompt-engineering-technique-program-simulation-56f49746aa7b>
- [7] <https://machinelearningmastery.com/prompt-engineering-for-effective-interaction-with-chatgpt>
- [8] <https://www.kdnuggets.com/prompt-engineering-101-mastering-effective-llm-communication>
- [9] <https://www.forbes.com/sites/lanceeliot/2024/10/16/writing-the-best-generative-ai-prompts-gets-revealed-via-openai-secret-meta-prompts>
- [Further Reading] <https://towardsdatascience.com/how-i-won-singapores-gpt-4-prompt-engineering-competition-34c195a93d41>
- [Further Reading] <https://www.infoq.com/news/2023/12/openai-prompt-engineering>
- [Further Reading] <https://platform.openai.com/docs/guides/prompt-engineering>
- [Further Reading] <https://arxiv.org/abs/2311.05661>

# Advanced Prompt Engineering Applications

# OpenAI GPTs and Assistants



GPTs are created and used manually (chat.openai.com).

<https://analyticsindiamag.com/10-must-try-gpts-from-gpt-store>

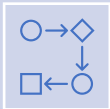


Assistants are created and used programmatically (OpenAI Assistant API).

<https://www.datacamp.com/tutorial/open-ai-assistants-api-tutorial>



Both make it possible to specify and share behavior (prompts).



Documents may be attached, the internet may be accessed, and actions invoked.

# Retrieval Augmented Generation (RAG)



Offline

Document chunking  
Chunk embedding  
Vector database



Online

Query embedding  
Chunk retrieval  
Prompt augmented with chunks

e.g.: <https://towardsdatascience.com/rag-how-to-talk-to-your-data-eaf5469b83b0>



"Beautiful is better than ugly.  
Explicit is better than implicit.  
Simple is better than complex.  
Complex is better than complicated.  
Flat is better than nested.  
Sparse is better than dense.  
Readability counts.  
Special cases aren't special enough to break the rules."

"Although practicality beats purity.  
Errors should never pass silently.  
Unless explicitly silenced.  
In the face of ambiguity, refuse the temptation to guess.  
There should be one -- and preferably only one -- obvious way to do it.  
Although that way may not be obvious at first unless you're Dutch."

"Now is better than never.  
Although never is often better than \*right\* now.  
If the implementation is hard to explain, it's a bad idea.  
If the implementation is easy to explain, it may be a good idea.  
Namespaces are one honking great idea -- let's do more of those!"

"Beautiful is better than ugly.  
Explicit is better than implicit.  
Simple is better than complex.  
Complex is better than complicated.  
Flat is better than nested.  
Sparse is better than dense.  
Readability counts.  
Special cases aren't special enough to break the rules."

Overlap

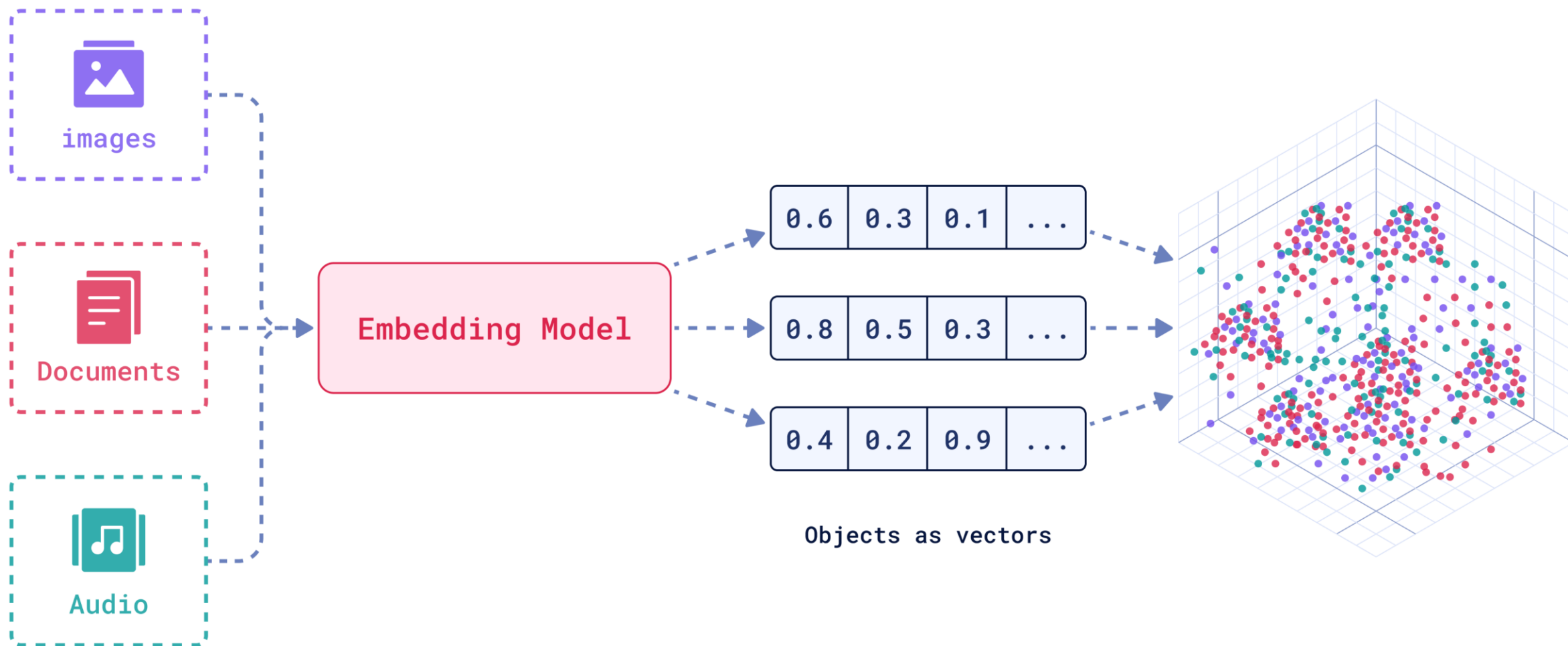
Chunk 2

"Although practicality beats purity.  
Errors should never pass silently.  
Unless explicitly silenced.  
In the face of ambiguity, refuse the temptation to guess."

"Unless explicitly silenced.  
In the face of ambiguity, refuse the temptation to guess.  
There should be one -- and preferably only one -- obvious way to do it.  
Although that way may not be obvious at first unless you're Dutch.  
Now is better than never.  
Although never is often better than \*right\* now."

"Now is better than never.  
Although never is often better than \*right\* now.  
If the implementation is hard to explain, it's a bad idea.  
If the implementation is easy to explain, it may be a good idea.  
Namespaces are one honking great idea -- let's do more of those!"]

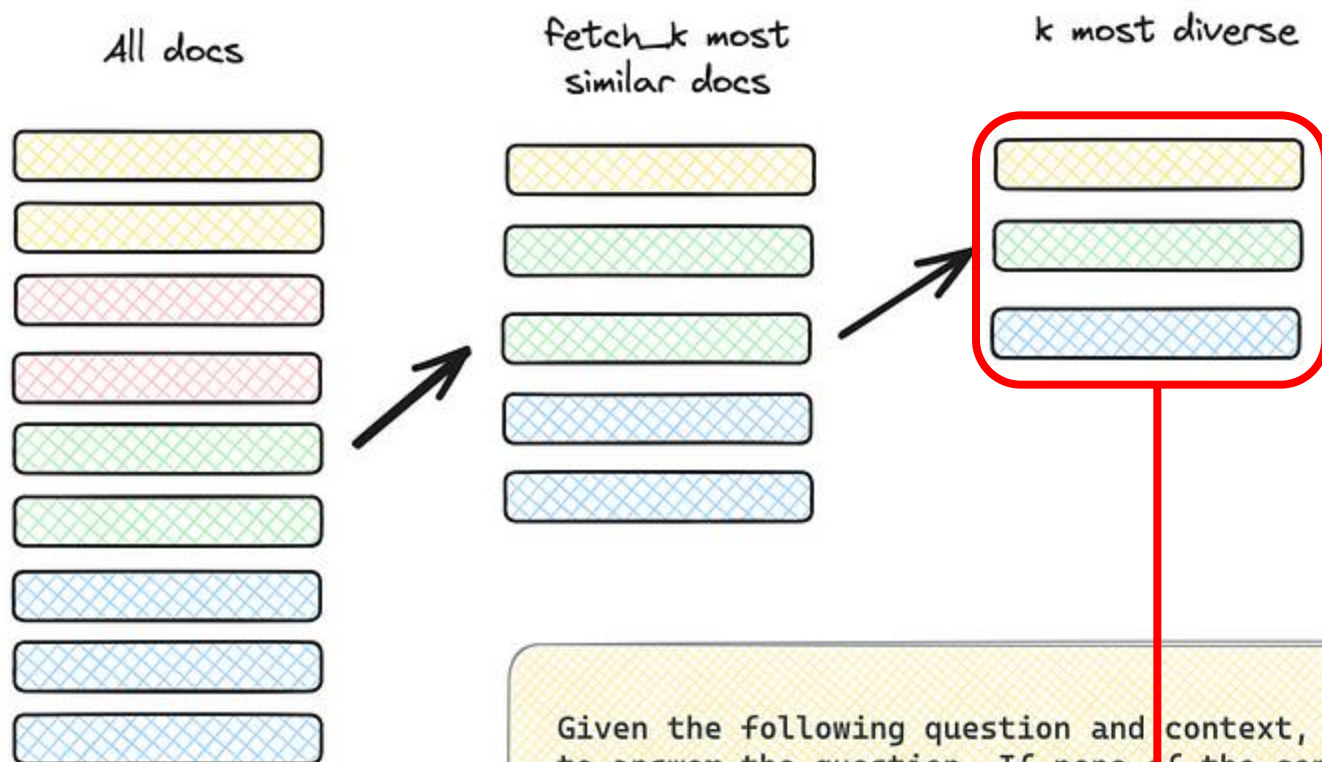




z.B. <https://qdrant.tech/articles/what-are-embeddings>



# Maximum Marginal Relevance



Given the following question and context, extract any part of the context *\*AS IS\** that is relevant to answer the question. If none of the context is relevant return NO\_OUTPUT.

Remember, *\*DO NOT\** edit the extracted parts of the context.

> Question: breakfast in Travelodge Farringdon

> Context:

>>>

typical travelodge. so close to the o2.good value for money.and yummy breakfast too.what more can one say.

>>>

Extracted relevant parts:

# Example "Screws" for RAG

---

## **Chunking**

Size? Overlap?

---

## **Embedding**

Which embedding?

---

## **Query Reformulation**

How is the "user" query enhanced?

---

## **Retrieval**

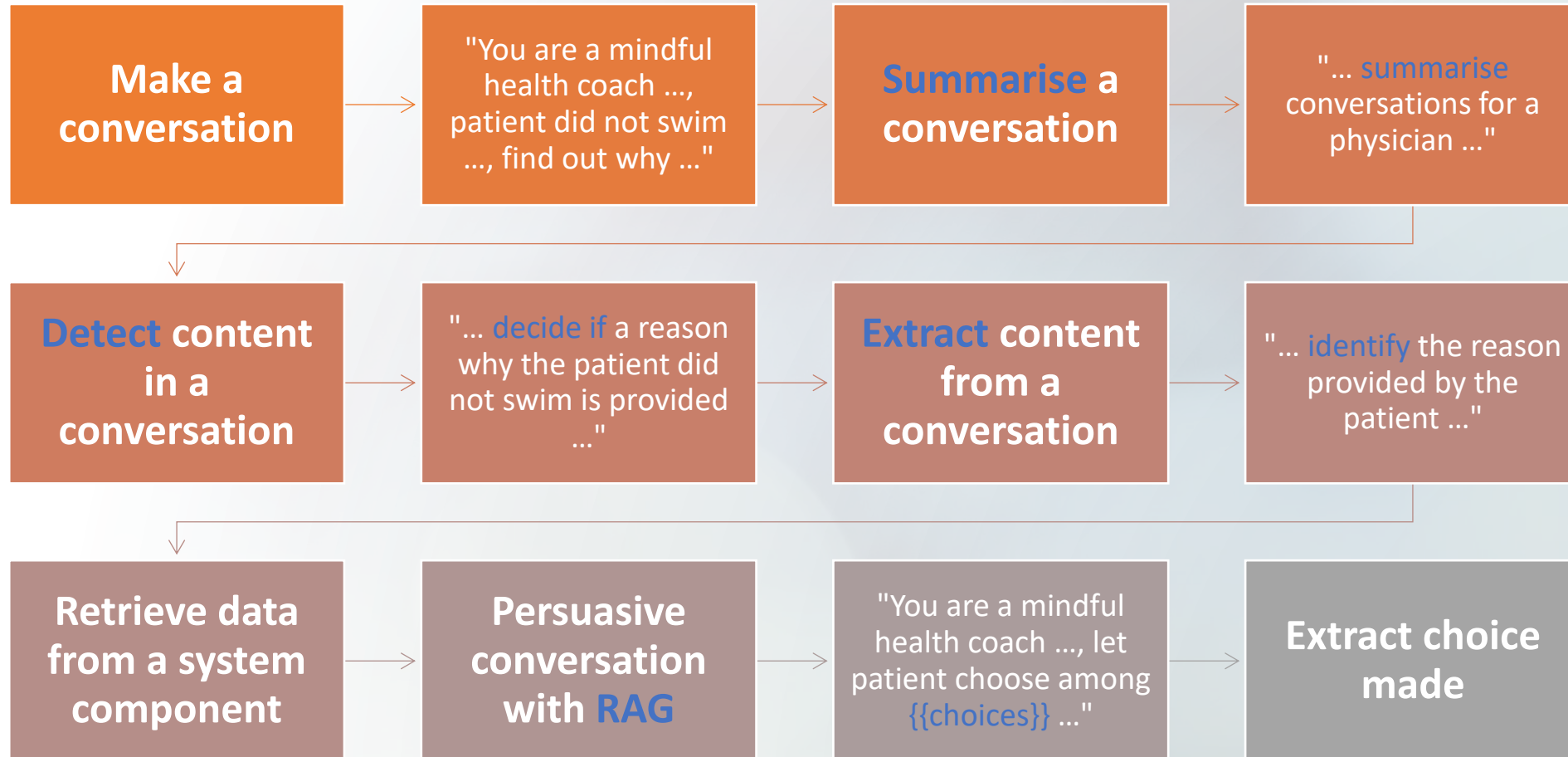
Which chunks are relevant?

---

## **Prompt Engineering**

How is the prompt containing the "user" query and the chunks designed?

# Komplexe Conversational Interactions



# PROMISE

An application development framework  
that supports the development of  
complex conversational interactions using  
state machine modelling concepts.







GPT-4-Vision

```
{
  "social_situation": {
    "description": "A group of young adults casually socializing outdoors",
    "number_of_people": 5,
    "setting": "Outdoor, possibly urban area near a chain link fence with graffiti in the background.",
    "time_of_day": "Likely late afternoon or early evening given the long shadows"
  },
  "people": [
    {
      "person id": 1,
      "gender": "Male",
      "position": "Standing on the left",
      "attire": {
        "upper_body": "Denim jacket over a white top",
        "lower_body": "Dark pants",
        "shoes": "White sneakers"
      },
      "hair_color": "Dark",
      "posture": "Standing with hands in pockets, slightly turned toward the left"
    },
    {
      "person_id": 2,
      "gender": "Male",
      "position": "Sitting in the center",
      "attire": {
        "upper_body": "Brown jacket over a dark top",
        "lower_body": "Light-colored pants",
        "shoes": "White sneakers"
      },
      "hair_color": "Blond",
      "posture": "Seated with one knee raised, looking up"
    },
    ...
  ],
  "activity": "Conversing, relaxing, possibly waiting for something or someone"
}
```





```
{
  "person_id": 4,
  "gender": "Female",
  "position": "Standing on the right",
  "attire": {
    "upper_body": "Sweater with a green, white and maroon chevron pattern",
    "lower_body": "Dark pants",
    "shoes": "Light-colored footwear"
  },
  "hair_color": "Dark",
  "posture": "Standing with arms by her side, facing another person"
},
```

```
{
  "presence_of_person": "Yes",
  "person_detected": {
    "person_id": 4,
    "gender": "Female",
    "position": "Central in frame",
    "attire": {
      "upper_body": "Sweater with a green, white and maroon chevron pattern",
      "lower_body": "Dark pants",
      "shoes": "Not visible in this picture"
    },
    "hair_color": "Dark",
    "posture": "Standing, though no longer with arms by her side; hands are up and gesturing, suggesting an active involvement in the current social interaction"
  },
  "observations": {
    "facial_expression": "Smiling, appearing to be enjoying the interaction",
    "engagement": "Directly interacting with at least one other individual in a playful manner",
    "additional_context": "Based on the previous description, attire and hair can be used as recognition parameters in other images as well."
  }
}
```