

# Prompt Engineering mit ChatGPT 3.5 für Midjourney am Beispiel eines Use Cases für Wildlife Photography

Jonathan Kreiss

Hochschule Pforzheim  
Tiefenbronner Straße 65  
75175 Pforzheim

E-Mail: [kreissjo@hs-pforzheim.de](mailto:kreissjo@hs-pforzheim.de)

Frank Morelli

Hochschule Pforzheim  
Tiefenbronner Straße 65  
75175 Pforzheim

E-Mail: [frank.morelli@hs-pforzheim.de](mailto:frank.morelli@hs-pforzheim.de)

## ABSTRACT

Das vorliegende Paper beleuchtet das Phänomen des Prompt Engineerings anhand der Anwendung von ChatGPT. In praktischen Anwendungen wird analysiert, was einen effektiven Prompt ausmacht und wie ChatGPT gezielt genutzt werden kann, um präzise Bildbeschreibungen für Midjourney zu generieren. Auf Grundlage dieser Erkenntnisse werden zudem Variationen für verschiedene Use Cases abgeleitet.

## SCHLÜSSELWÖRTER

Prompt Engineering, Midjourney, ChatGPT

### 1. Theoretische Grundlagen Prompt Engineering Problemstellung und Zielsetzung

Das Konzept des Prompt Engineerings gewinnt mit der zunehmenden Verbreitung von Large Language Models (LLMs), wie beispielsweise OpenAIs ChatGPT, immer mehr an Bedeutung. Prompt Engineering zielt darauf ab, präzise Anweisungen, sogenannte Prompts, für LLMs zu entwickeln, um LLMs gezielt zu steuern und dadurch die Qualität der generierten Antworten zu verbessern. Die Forschungslage bezüglich der Merkmale eines präzise formulierten Prompts ist bis dato noch begrenzt. In diesem Zusammenhang hat der folgende Artikel das Ziel, einen Überblick über die Methoden zur Entwicklung effektiver Prompts zu liefern.

#### Die Bedeutung effektiver Prompts

Ein präziser formulierter Prompt steigert die Wahrscheinlichkeit nützlicher Antworten von Large Language Models (LLMs) zu erhalten. Um Prompt Engineering auf unterschiedliche Use Cases anzuwenden zu können, erfordert es Expertise im Prompt Engineering. Diese Fähigkeit ermöglicht es Nutzern, maßgeschneiderte Anweisungen für unterschiedliche Szenarien zu erstellen und so die Leistung von LLMs zu maximieren. Prompt Engineering ist hierbei entscheidend, um das volle Potenzial von LLMs auszuschöpfen. Nachfolgend werden die hierzu gewonnenen Erkenntnisse aus einer Bachelor-Thesis zusammengefasst.

### Open AI Best Practices

OpenAI, das für die Entwicklung von ChatGPT verantwortliche Unternehmen, hat einige Best Practices zum Thema Prompt Engineering mit ChatGPT veröffentlicht, die im nachfolgenden kurz vorgestellt werden:

#### 1. Trennung von Anweisung und Kontext

Zu empfehlen ist es, in einem Prompt zu Beginn präzise Anweisungen und anschließend den Kontext zu platzieren. Zur klaren Trennung zwischen Anweisung und Kontext können Trennzeichen wie "###" oder "====" eingesetzt werden.

#### 2. Detaillierte und spezifische Anweisungen

Klare Anweisungen formulieren, etwa in Bezug auf:

*Länge:* Begrenzen der gewünschten Länge des Texts, beispielsweise auf maximal 3 Sätze, 300 Wörter oder 1000 Zeichen.

*Format:* Spezifizieren des gewünschte Formats, sei es in Form von Stichpunkten, einer Tabelle, Absätzen oder mit Zwischenüberschriften.

*Schreibstil:* Präzisieren des gewünschten Schreibstils, ob wissenschaftlich, geschäftlich, kreativ, humorvoll, formal oder in einem anderen Stil.

#### 3. Ausgabeformat angeben

Einige Beispiele für mögliche Ausgabeformate sind:

*Tabellenform:* Die gewünschte Antwort wird in tabellarischer Form präsentiert, um Informationen übersichtlich und strukturiert darzustellen.

**Stichworte:** Die Antwort wird in Form von Stichworten oder Schlagworten verfasst, um Schlüsselinformationen hervorzuheben.

**Sätze mit Zwischenüberschriften:** Die Antwort erfolgt in gegliederten Sätzen, die durch Zwischenüberschriften strukturiert sind, um eine klare Struktur in der Antwort zu schaffen.

**Pro/Kontra Liste:** Die Antwort wird in Form einer Liste von Argumenten aufgeteilt, um Vor- und Nachteile zu verdeutlichen.

**Schritt-für-Schritt Erklärung:** Die Antwort wird schrittweise erläutert, um komplexe Konzepte oder Abläufe verständlich darzustellen.

#### 4. Show don't tell

Statt ausführliche Instruktionen zu schreiben, ist es oft zielführender, ein oder mehrere Beispiele direkt in den Prompt zu integrieren. Die verschiedenen Arten des Promptings werden im nächsten Abschnitt ausführlicher beleuchtet.

#### Promptingarten

In der Forschung und Anwendung des Promptings haben sich drei Hauptvarianten herausgebildet: Zero-Shot, One-Shot und Few-Shot Prompting. Diese Varianten unterscheiden sich in der Anzahl der Beispiele, die in den Prompt integriert werden:

- **Zero-Shot Prompting:** In dieser Variante sind keinerlei Beispiele im Prompt enthalten. Der Prompt besteht lediglich aus einer reinen Aufforderung ohne vorherige Beispiele. Eignet sich besonders für einfache Instruktionen wie zum Beispiel Fragen.
- **One-Shot Prompting:** Bei dieser Methode wird präzise ein einziges Beispiel in den Prompt integriert, welches als Grundlage für die gewünschte Antwort dient. Eignet sich für komplexere Instruktionen.
- **Few-Shot Prompting:** Beim Few-Shot Prompting hingegen werden mindestens zwei, oft aber mehrere Beispiele in den Prompt aufgenommen. Diese Beispiele dienen dazu, dem Modell zusätzlichen Kontext und Informationen zur Verfügung zu stellen, um die Qualität und Relevanz der generierten Antwort zu steigern. Eignet sich vor allem für sehr komplexe Instruktionen.

#### Chain-of-Thought Prompting

Chain-of-Thought (CoT)-Prompting ist besonders geeignet für komplexe und neuartige Aufgabenstellungen. Im Vergleich zum herkömmlichen Prompting werden

bei CoT-Prompting gezielt Zwischenschritte in die Anweisung miteingebaut. CoT-Prompting ermöglicht es, komplizierte Sachverhalte in mehrere einfachen Schritte aufzuteilen.

#### Promptingmatrix

Die folgende Matrix soll eine Hilfestellung geben, bei welcher Aufgabenstellung, welche Promptingart verwendet werden kann (siehe Abb. 1).

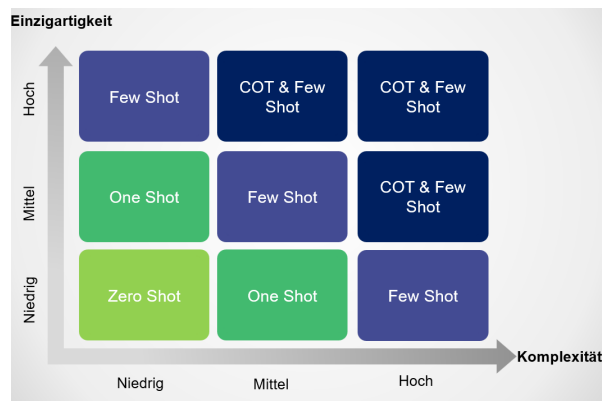


Abb. 1: Promptingmatrix

- Y-Achse: "Einzigartigkeit" – Diese Achse gibt das Ausmaß der Einzigartigkeit einer gegebenen Anweisung an. Eine höhere Position auf dieser Achse deutet darauf hin, dass die Anweisung für Aufgaben verwendet wird, die neuartig sind und wahrscheinlich nicht im Trainingsdatensatz des Large Language Models (LLM) enthalten sind.
- X-Achse: "Komplexität" – Die X-Achse stellt die Komplexität einer Anweisung dar. Anweisungen mit geringer Komplexität umfassen klar formulierte Fragen oder einfache Aufgabenstellungen, die direkt auf Informationen abzielen. Eine höhere Position auf dieser Achse weist auf komplexere Anweisungen hin die eine mehrstufige Verarbeitung erfordern können.

#### Prompt Engineering – ein Prozess

In der Praxis beginnt der Prozess des Prompt Engineering mit einer grundlegenden und simplen Anweisung. Im Verlauf wird der Prompt schrittweise verfeinert und jede erzeugte Antwort sorgfältig überprüft. Dieser iterative Prozess setzt sich fort, bis der Prompt eine derart präzise Form erreicht, dass die generierte Antwort den Erwartungen entspricht und einen konkreten praktischen Nutzen bietet.

Geringfügige Änderungen, wie beispielsweise das Austauschen von Wörtern, können erhebliche Auswirkungen auf die erzeugte Antwort haben. Die Unterscheidung zwischen effektiven und weniger effektiven Anweisungen erfordert oft experimentelles Ausprobieren und Er-

proben unterschiedlicher Formulierungen. Dieser kontinuierliche Anpassungsprozess ist entscheidend, um die Qualität und Relevanz der generierten Antworten zu optimieren und den angestrebten Mehrwert zu gewährleisten.

## 2. Prompting mit ChatGPT 3.5 für Midjourney

### Ausgangslage

Zur besseren Vergleichbarkeit zwischen den generierten Bildern und den dazugehörigen Prompts wird ein Tiger als Motiv für alle Generierungen ausgewählt. Zudem wurden bei jeder Bilderstellung dieselben Parameterwerte festgelegt, um die Vergleichbarkeit zwischen den Bildern zu gewährleisten

### Vorgehensweise

Die Generierung geeigneter Prompts für Midjourney mithilfe von ChatGPT 3.5 steht im Mittelpunkt der Untersuchung. Die Methodik dieses Vorhabens ist in mehrere aufeinanderfolgende Prozessschritte unterteilt:

1. *Prompt-Formulierung mit ChatGPT:* Entwicklung präziser Prompts, um qualitativ hochwertige textuelle Bildbeschreibungen zu generieren.
2. *Bildgenerierung in Midjourney:* Die in Schritt 1 erstellten Prompts werden in Midjourney getestet
3. *Bewertung der generierten Bilder:* Ein KI-Tool wird zur Bewertung und Analyse der erzeugten Bilder eingesetzt.
4. *Analyse und Optimierung der Prompts:* Die genutzten Prompts werden analysiert und verbessert, um die Bildqualität zu verbessern.

Der Prozess wird nach Abschluss des vierten Schrittes iterativ wiederholt, da Prompt Engineering keine lineare Methode darstellt, sondern fortlaufende Anpassungen und Verfeinerungen erfordert, um optimale Ergebnisse zu erzielen.

## 3. Praktische Umsetzung (Ergebnisse)

Die KI-basierte Bildbewertungs Website <https://aesthetics.everypixel.com/> wird genutzt, um die generierten Bilder zu bewerten und anschließend miteinander zu vergleichen. Die Einschätzung des Bildes erfolgt in Prozent, wobei das Maximum bei 100% liegt. Je näher die Bewertung an 100% heranreicht, desto ästhetischer wurde das Bild bewertet.

Es ist jedoch wichtig anzumerken, dass die ästhetische Bewertung durch die KI nicht gänzlich objektiv ist und stark von den Merkmalen des Algorithmus beeinflusst wird. Es bleibt unklar, nach welchen genauen Kriterien der Algorithmus Bilder als ästhetisch oder nicht ästhetisch einstuft. Dennoch ermöglicht der Einsatz des Algorithmus eine Bewertung und somit auch ein Vergleichbarkeit zwischen den generierten Bildern.

## 3.1 Tiger

### Struktur des finalen Prompts

Der folgende Screenshot, zeigt die Struktur des finalen Prompts. Dieser Prompt lässt sich vielfältig nutzen, um verschiedene Variationen zu generieren. Dabei können alle Parameter verändert werden, um neue Bildbeschreibungen zu generieren. Im nächsten Kapitel (3.2) werden zudem Abwandlungen für andere Tierarten vorgestellt.




Rank	Bild
1.	
99,40%	Wildlife Photography of a Tiger. Full Body Shot. Background Jungle. Captured by Sony Alpha a9 II with Sony FE 400mm f/2.8 GM OSS lens. Low-angle wide shot. Dappled sunlight. Intense.
2.	
99,20 %	Wildlife Photography of a Tiger. Full Body Shot. Background Jungle. Captured by Sony Alpha a9 II with Sony FE 400mm f/2.8 GM OSS lens. Low-angle wide shot.
3.	
99,00 %	Wildlife Photography of a Tiger. Full Body Shot. Background Jungle. Captured by Canon Sony Alpha a9 II.

Tabelle 1

Certainly, here's the modified breakdown with "Photography Type" in place of "Wildlife Photography of a Tiger":

Aspect	Description
Photography Type	Wildlife Photography of a Tiger
Subject	Tiger
Type of Shot	Full Body Shot
Background	Jungle
Camera Used	Sony Alpha a9 II
Lens Used	Sony FE 400mm f/2.8 GM OSS lens
Shot Angle	Low-angle wide shot
Lighting	Dappled sunlight
Pose	Crouching and ready to pounce

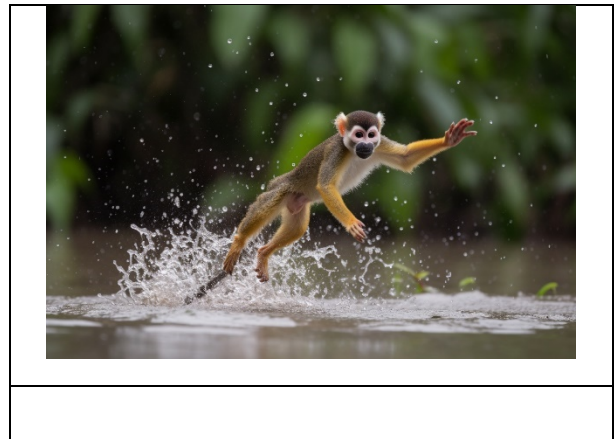
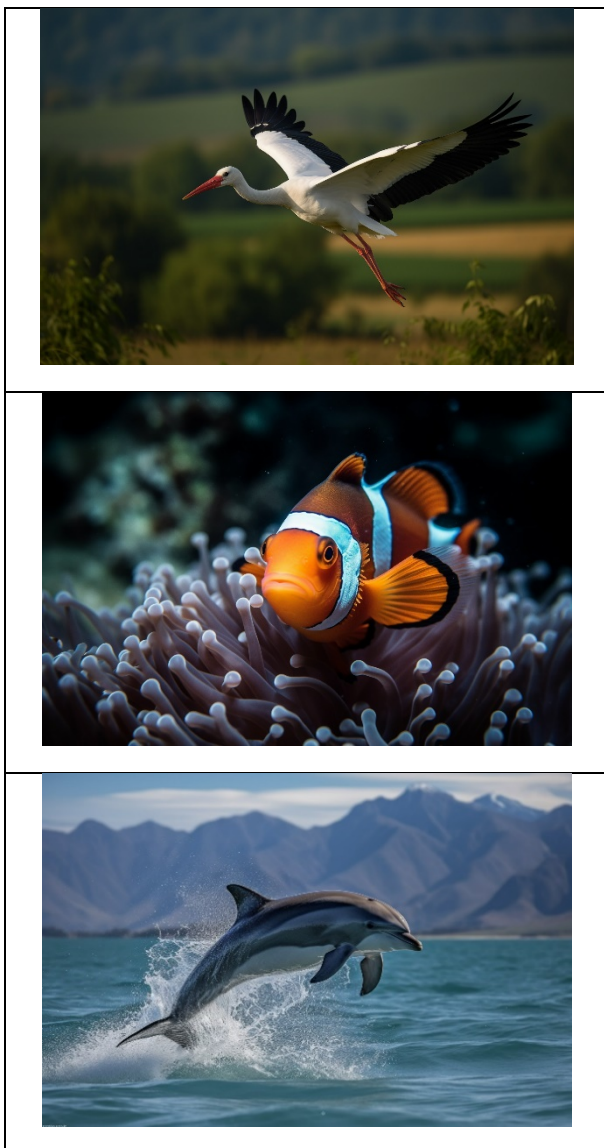


Tabelle 2

### 3.2 Abwandlungen andere Tiere



#### Fazit

Das Prompt Engineering gewinnt mit der zunehmenden Verbreitung von LLMs immer mehr an Bedeutung. Durch präzise formulierte Prompts lässt sich die Antwort von LLMs gezielt steuern, was zu verbesserten Ergebnissen führt. Ein Beispiel dafür ist das gezielte Trainieren von ChatGPT für die Generierung von Bildbeschreibungen in Midjourney. Die drei Hauptvarianten des Promptings: Zero-Shot-, One-Shot- & Few-Shot-Prompting bieten differenzierte Ansätze, die je nach Aufgabenkomplexität variieren.

Für komplexere Anweisungen wie die Generierung einer Bildbeschreibung eignen sich One-Shot bzw. Few-Shot Anweisungen mit einem oder mehreren Beispielen.

#### LITERATUR

Brown et. al (2020) Language Models are Few-Shot Learners, OpenAI

<https://help.openai.com/en/articles/6654000-best-practices-for-prompt-engineering-with-openai-api>

Brown, T, et al. (2020). Language Models are Few-Shot Learners. arXiv (Cornell University). <https://arxiv.org/pdf/2005.14165.pdf>