

Ideenfindung mit KI

Neue Wege für Kreative

» Hier geht's
direkt
zum Buch

DIE LESEPROBE

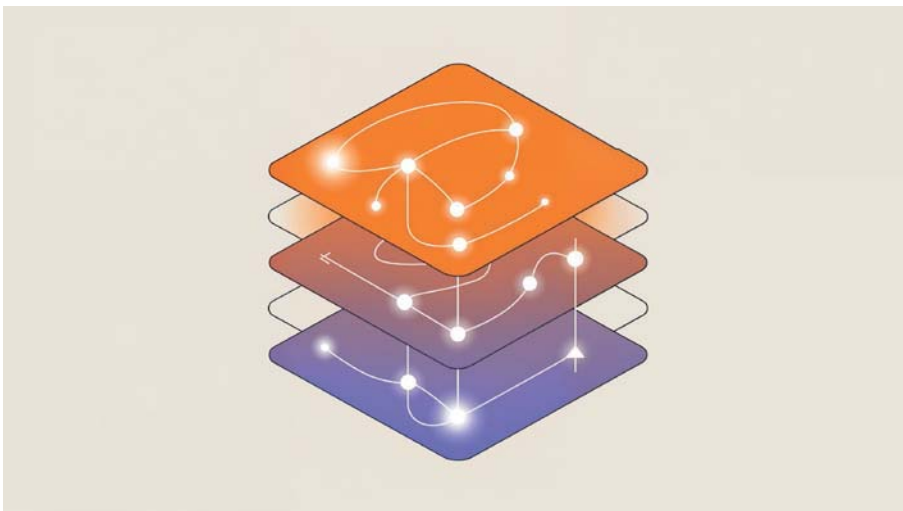
KAPITEL 4

Umsetzung in Bild und Text

*Bilder und Texte mit KI zu generieren ist einfach. Bilder und Texte, die einem Anspruch an Qualität, Zielgruppenansprache und Einzigartigkeit gerecht werden und dabei gleichzeitig den Branding-Guidelines entsprechen, sind besonders mit KI eine Herausforderung. In der Natur der KI liegt die Unvorhersehbarkeit des Ergebnisses – was bei der kreativen Ideenfindung ein Geschenk sein kann, aber wenn wir etwas ganz gezielt generieren möchten, zur Herausforderung wird. In diesem Kapitel möchte ich Ihnen einen Einblick in meine Erfahrungen mit generativer KI geben, indem ich Ihnen Techniken an die Hand gebe und Ihnen zeige, welche KI-Tools für welche Aufgaben geeignet sind. Da sich KI aber so rasant weiterentwickelt und sich die Entwickler*innen viel einfallen lassen, um auf dem Markt bestehen zu bleiben, ist dies eine Momentaufnahme der aktuellen Möglichkeiten. Umso besser, wenn Ihr Lieblingstool nun eine neue Funktion hat, die Ihnen weitere Möglichkeiten zur Nutzung eröffnet. Probieren Sie die Techniken bei verschiedenen Tools aus, und prüfen Sie, wie diese auf Ihre Eingaben reagieren. Jedes Projekt ist eine neue Möglichkeit, mehr über KI und ihre Möglichkeiten sowie Grenzen zu lernen.*

Bildgenerierung

Die Welt der KI bietet uns viele neue Möglichkeiten, unsere Ideen visuell auszudrücken – und die Zahl der Bildgeneratoren wächst stetig. Da wir in Zukunft vermehrt damit arbeiten werden, lohnt sich ein Blick hinter die Kulissen – also welche »Arten« es gibt und wie diese funktionieren. Bestimmt haben Sie schon die Erfahrung gemacht, dass sich die Ergebnisse der KI-Modelle nicht komplett bestimmen lassen und auch schwer reproduzierbar sind. Lernen wir unsere neuen Assistenzen besser kennen, um ganz gezielt damit arbeiten zu können. So lassen sich ihre Stärken bewusst nutzen und typische Schwächen besser ausgleichen. Mit dem richtigen Verständnis gewinnen Sie schlussendlich mehr Kontrolle über den Stil, die Qualität und die Aussage Ihrer Bilder.



Eine abstraktere Darstellung der Layer eines neuronalen Netzes: oben der »Input Layer«, in der Mitte die »Hidden Layer« und unten der »Output Layer«

Wie KI Bilder generiert

Im ersten Kapitel haben wir uns bereits mit der Funktionsweise von KI beschäftigt und gelernt, was ein Algorithmus ist. Auch bei Bildgeneratoren läuft dieser Algorithmus innerhalb der Schichten im neuronalen Netz der KI ab. Diese neuronalen Netze wurden entwickelt, um das menschliche Gehirn besser zu verstehen, und bilden die Grundlage für viele KI-Programme. Es gibt eine Eingabe-Schicht (Input Layer), die uns die Möglichkeit bietet, der KI eine Anfrage via Prompt zu stellen. Die Ausgabe-Schicht (Output Layer) gibt uns das Ergebnis, das der Algorithmus entwickelt hat, aus. Dazwischen sind die verborgenen Schichten (Hidden Layer), die unsere Anfrage analysieren und verarbeiten – ähnlich wie unser Gehirn. Was in diesen Neuronen der Hidden Layer genau passiert, weiß niemand – was es natürlich umso schwerer macht, sie unter unsere »Kontrolle« zu bringen. Daher wird KI auch gerne »Black Box« genannt. Gehen wir aber einmal weiter und schauen uns an, wie KI lernt und wie die Bilder entstehen, die wir jetzt in Sekundenschnelle generieren können.

KI lernt anders als wir Menschen

Um überhaupt Bilder generieren zu können, müssen diese Modelle zunächst trainiert werden. Hier unterscheiden sich Mensch und KI maßgeblich: Wir verarbeiten Eindrücke ganzheitlich, verbinden sie mit Emotionen, Erfahrungen und bringen diese in einen Kontext. Unser Gehirn vernetzt dabei Wahrnehmung (z. B. Sehen, Hören), unsere Erinnerungen und Gefühle zu einem persönlichen, lebendigen Wissen. Wenn wir ein Bild sehen, dann verarbeiten wir es auf verschiedenen Ebenen, und jeder Mensch nimmt das gleiche Bild anders wahr.

Ein gutes Beispiel ist ein Kleinkind, das zum ersten Mal eine Katze sieht. Man zeigt darauf und sagt »Katze«. Katzen gibt es aber in sehr vielen verschiedenen Rassen, Farben, Formen. Trotzdem müssen wir nicht jedes Mal von Neuem erklären, dass es sich um eine Katze handelt – das menschliche Gehirn erkennt die Ähnlichkeit, kombiniert die bisherige Erfahrung und gleicht sie mit der Realität ab. Das geht auch sehr gut mit der abstrakten Darstellung einer Katze – einem Kreis mit zwei Dreiecken darauf und vielleicht einem Herzchen in der Mitte. Egal wie wir Katzen darstellen, wenn unser Gehirn eine Beziehung zu unserer bisherigen Erfahrung herstellen kann, erkennen wir, womit wir es zu tun haben.

Bilder generiert mit Midjourney



Zum Glück gibt es so viele Katzenbilder im Netz – daher wurden auch die ersten Modelle mit Katzenbildern trainiert. Heute können wir von abstrakt bis realistisch alles an Katzenbildern generieren, was das Herz begehrt.

Bei der KI funktioniert das völlig anders – sie braucht Katzen in allen möglichen Farben, Formen, mit unterschiedlichen Hintergründen, in unterschiedlichen Posen und unterschiedlich gemalt, gezeichnet oder modelliert. Erst so kann sie lernen und das Bild einer Katze generieren.

Die KI-Modelle und das menschliche Gehirn funktionieren völlig unterschiedlich.

Nun können Sie sich vorstellen, dass alleine dafür schon eine unvorstellbare Menge an Bildern notwendig ist – jetzt versuchen Sie das auf den Rest unserer komplexen Welt zu addieren. Vielleicht können Sie nun besser verstehen, warum es deshalb noch für viele Modelle schwierig ist, Bilder mit mehreren Personen oder Hände in verschiedensten Positionen zu generieren. Denn je komplexer die Szene ist und je mehr Details zusammenkommen, desto größer wird die Herausforderung für die KI.

Wir alle sind Trainingsmaterial

Um die Qualität und Quantität an Trainingsdaten zu ermöglichen, wird sich an dem Paradies der Datensammler bedient: dem Internet. Alles, was wir veröffentlicht haben, kann im Trainingsdatenpool der KI-Entwickler*innen landen – ob wir zugestimmt haben oder nicht. Das trifft zu Recht auf Widerstand und Unmut vieler Menschen. Fairerweise muss man sagen, dass die Entwicklung wirklich rasant voranging – aber KI-Modelle werden schon seit den 80er-Jahren entwickelt und erforscht. In der Forschung gibt es oft andere Regeln als in der freien Wirtschaft. Nun verdienen aber Unternehmen ihr Geld damit, indem Sie KI-Modelle mit den Daten anderer trainieren – ohne einen Ausgleich.

Immer mehr Plattformen – unter anderem Meta mit Instagram, WhatsApp und Facebook oder LinkedIn – nutzen die Daten ihrer User*innen für das KI-Training. Hier gibt es zumindest die Möglichkeit, zu widersprechen. Doch was die Daten im Internet und die zahlreichen Bücher und Medien, die zum Training genutzt wurden, angeht, gibt es noch keine befriedigende Regelung in der Rechtsprechung.

Wie aus Daten Bilder werden

KI sieht in einem Bild zunächst einmal Daten, ähnlich einer Tabelle mit Pixelwerten. Daher ist es praktisch ein Wunder, wie gut KI mittlerweile Bilder erkennen und auswerten kann – aber es ist auch kein Wunder, wenn sie völlig falsch liegt. Um dazu in der Lage zu sein, braucht die KI nicht nur die Bilder, sondern weitere Informationen dazu. Es gibt zwar auch eine Phase des Selbstlernens, in der die KI die Bilder analysiert und kategorisiert – aber die Modelle, die wir tagtäglich nutzen, kommen nicht ohne das Training durch Menschen aus.

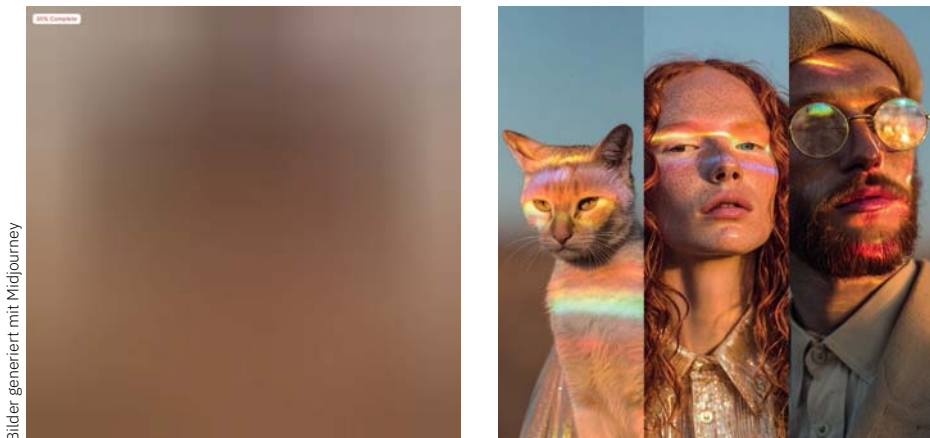
Das sind Menschen, die in Niedriglohnländern wie Madagaskar sitzen und den ganzen Tag Objekte auf Bildern markieren, Bildbeschreibungen machen oder Texte von Dokumenten abtippen. Ohne dieses Zutun würde die KI viele Bilder gar nicht auswerten können. Nehmen wir als Beispiel den Versuch, KI mit Röntgenbildern von Lungen zu trainieren. Ziel war es, dass die KI eigenständig kollabierte Lungen erkennt. Das hat auch funktioniert, aber nicht so, wie die Entwickler*innen es sich gedacht haben. Die KI hat, was sie ja sehr gut kann, Muster erkannt, die auf eine kollabierte Lunge in den Trainingsdaten hinwiesen. Aber diese hatten nichts mit der Lunge zu tun, sondern mit den Schläuchen, die sich auf den Bildern befanden. Für die KI war der Fall klar: Schläuche = kollabierte Lunge. Die Entwickler*innen mussten also noch einmal ganz neu an die Sache herangehen. Das Beispiel zeigt uns, wie komplex unser Gehirn ist. Was für uns selbstverständlich ist, kann für KI ein Hindernis sein. Dabei ist sie nicht nur auf unser Trainingsmaterial angewiesen, sondern auch auf für sie aufbereitete Daten.

Ein weiterer Teil des Trainings besteht darin, dass Menschen Anfragen an die KI stellen und die Antworten überprüfen und bewerten. Durch das Feedback werden die Modelle verbessert und justiert. Auch wir User können dabei helfen, die Modelle zu verbessern, indem wir Feedback geben. Das wird natürlich nicht sofort und ohne Überprüfung übernommen, kann aber dazu beitragen, gewisse Probleme aufzudecken oder auch zu zeigen, was gut lief.

Vom Bild zum Rauschen und zurück

Nach dem Erkennen von Mustern und dem Kategorisieren der Daten werden die modernen Bildgeneratoren nun trainiert, eigene Bilder zu generieren. Dazu wird ein Bild Schritt für Schritt verrauscht – daher auch der Name »Diffusion Model« –, bis wir ein völlig zufälliges verrauschtes Bild haben – so als würden wir den Gaußschen Weichzeichner ganz oft anwenden.

Nun soll die KI mittels Textprompt aus diesem Anfangsrauschen ein neues Bild generieren (Denoising). Jedes generierte Bild startet also mit einem Rauschen, einem einzigartigen Fingerabdruck, der sich mittels »Seed« ansteuern lässt. Der Seed ist die Kennzahl dieses Anfangsrauschens und kann in KI-Tools wie Midjourney oder Freepik verwendet werden. Nun geht der Algorithmus Schritt für Schritt vor und fügt dem Bild jedes Mal mehr Details hinzu, bis die maximale Anzahl der Schritte (ca. 25–1.000) erfolgt ist, die ein Modell während der Bildgenerierung durchläuft. Je mehr Schritte, desto mehr Details, die aber auch mehr Rechenleistung und Zeit kosten. Bei manchen Modellen kann man die Generierung auch vorzeitig stoppen, was praktisch sein kann, wenn Sie ein Bild mit weniger Details wünschen.



Bei Midjourney kann man sehr gut den Prozess beobachten: Links sehen wir das Bild bei 35 % der maximalen Schritte im Zwischenstadium. Erst bei ca. 40 % kam die Dreiteilung hinzu. Rechts das fertige Bild mit maximalem Detailgrad von Midjourney V7.

PRAXIS

Promptvariationen mit Seeds

Eine Bildgenerierung startet immer mit einem Anfangsrauschen, zu dem schrittweise Details hinzugefügt werden. Dieses erste Rauschen hat eine einzigartige Kennzahl, den Seed. Bei vielen Bildgeneratoren können Sie den Seed auslesen und für neue Generierungen nutzen. Das sehen wir uns hier am Beispiel von Midjourney an.

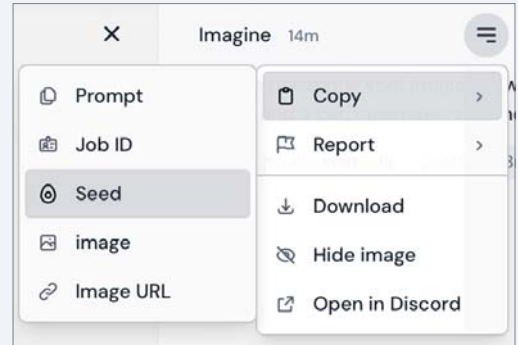
Ich generiere vier Bilder mit folgendem, eher abstraktem Prompt:

 **Prompt** photo of a person on the right, levitating in the space, around the person are machines, computers, smartphones, science fiction, realistic, high quality photoshoot, orange and purple colors



Diese Version hat mir besonders gut gefallen. Darauf würde ich gerne aufbauen, aber mit kleinen Änderungen im Prompt.

Jedes Mal, wenn ich den Prompt abschicke, erhalte ich unterschiedliche Bilder. Wenn mir eine Bilderserie gefällt und ich darauf aufbauen möchte, ist eine Möglichkeit, den Seed zu nutzen.



Um an den Seed zu kommen, können Sie das Bild in Ihrer History anklicken und oben rechts über das Menü die Option »Copy • Seed« auswählen.

Ich kopiere mir den Seed dieser Serie und passe den Prompt geringfügig an. Den Seed füge ich über einen Shortcut mit ein:

 **Prompt** photo of **a woman** on the right, levitating in the space, around the person are machines, computers, smartphones, science fiction, realistic, high quality photoshoot, orange and purple colors **--seed 885042361**

Was ich jetzt erhalte, ist keine exakte Kopie des Bildes, denn wie erwähnt ist es nur das Anfangsrauschen, das wir hier nutzen. Trotzdem ähneln sich die Bilder im Aufbau, wenn der Prompt nicht allzu weit von dem entfernt ist, was wir ursprünglich generiert haben.



Der Prompt wurde nur etwas verändert, und die Bilder ähneln sich teilweise im Aufbau.

Bemerkenswert ist das erste und zweite Bild der Serie, die sehr nah an der ersten Generierung liegen. Das liegt unter anderem daran, dass der Prompt Farben und Bildaufbau ganz gut beschreibt, wenn auch mit viel Interpretationsspielraum.

Midjourney beschreibt die Nutzung von Seeds als Experiment, um zu sehen, wie der Prompt das Ergebnis beeinflussen kann. Auch kann der Seed sich nach einer Weile ganz anders auf das Ergebnis auswirken als zu einem früheren Zeitpunkt. Warten Sie also nicht zu lange mit dem Experimentieren.

Auch wenn wir hier sehr ähnliche Ergebnisse hatten: Seeds sind nicht dafür geeignet, irgendwelche Stile zu übertragen und auch keine Garantie dafür, dass die neuen Bilder eine Variante der vorherigen werden. Ändern wir den Prompt maßgeblich, erhalten wir Bilder, die stark von den Originalen abweichen, wie im folgenden Beispiel:



Derselbe Seed, aber ein ganz anderer Prompt



Prompt photo of a cat in a space suit --seed 885042361

Seeds und Parameter

Achten Sie darauf, dass Sie keine weiteren Parameter einsetzen, die die Anwendung des Seeds »verhindern«, wie die Stylization oder andere Funktionen, die dafür sorgen, dass das Ergebnis stärker vom Prompt abweicht.

Im Laufe des Kapitels lernen Sie weitere Möglichkeiten kennen, um mehr Konsistenz in Ihrer Bildgenerierung zu erhalten. Manchmal ist es eine Kombination aus mehreren Funktionen, oder Sie müssen je nach Bild, Prompt sowie gewünschtem Ergebnis immer wieder experimentieren, welche Funktion Ihnen am meisten bringt. Ich selbst nutze Seeds im Alltag immer dann, wenn mir der Aufbau der Bilder gefällt, aber ich noch nicht mit dem Ergebnis zufrieden bin.

Einflussfaktoren auf das Ergebnis

Damit wir immer neue und spannende Bilder generieren können, sind die Modelle speziell trainiert und programmiert worden, sodass Sie auch mit nur einem Wort oder sogar einem Emoji ein (mehr oder weniger brauchbares) Bild erhalten. Das bedeutet, dass die KI einfach Details im Bild »erfindet« oder »halluziniert«. Welche Details das sind, hängt von folgenden Faktoren ab:

- **Trainingsdatenpool und Programmierung:** Die Auswahl der Trainingsdaten und die Richtung, in die das Modell trainiert wurde und unsere Eingabe interpretiert.
- **Prompt:** Je genauer die Beschreibung, desto besser können Sie das Ergebnis beeinflussen. Bei sehr einfachen Prompts nimmt sich die KI mehr Spielraum.
- **Zufälligkeit und Rauschen:** Selbst bei identischem Prompt entstehen leicht unterschiedliche Bilder. Das liegt daran, dass die Modelle mit einer gewissen Zufälligkeit (»Random Seed«) arbeiten. Dieses »Rauschen« ist gewollt – es sorgt dafür, dass jedes Ergebnis einzigartig wirkt und kreative Varianz entsteht.
- **Version und Trainingsstand des Modells:** Jedes KI-Modell hat eine eigene »Persönlichkeit«. Neue Versionen reagieren oft sensibler auf Details oder interpretieren Prompts anders, weil sie mit erweiterten Datensätzen oder veränderten Gewichtungen trainiert wurden.
- **Parameter/Einstellungen:** Manche Bildgeneratoren bieten Einstellungen an, die der KI mehr oder weniger Freiheit geben, den Prompt zu interpretieren.
- **Vorlagen/Bildprompts:** Je nach Generator können Sie Bilder als Vorlage für den Bildaufbau oder den Stil nutzen, an denen sich die KI orientieren soll. Bei sogenannten »Strukturvorlagen« wird der Bildaufbau als Vorgabe genutzt, bei Bildprompts das Bild an sich und bei Stilvorlagen der Bildstil, weniger der Inhalt.

Schauen Sie beim Bildgenerator Ihrer Wahl, welche Möglichkeiten Ihnen geboten werden. Probieren und experimentieren Sie, um ein Gefühl für das Verhalten des Modells zu bekommen. Das macht die Arbeit mit KI umso spannender, denn wir können uns immer wieder von neuen Details überraschen und inspirieren lassen.

Modelle und ihre Einsatzmöglichkeiten

Bei der Vielfalt an vorhandenen Bildgeneratoren ist es gar nicht so einfach, die passenden zu finden. Daher möchte ich Ihnen einen Überblick über die gängigsten Plattformen geben und für welche Bereiche Sie diese einsetzen können. Diese Liste hat keinen Anspruch auf Vollständigkeit, und bestimmt finden Sie noch viele weitere Möglichkeiten, KI für Ihre Projekte einzusetzen. Dabei beziehe ich mich auf meine eigenen Erfahrungen sowie die von anderen. Sehen Sie es als Starthilfe, um erste Schritte in der Bildgenerierung zu machen, oder als Inspiration, welche Modelle es auf dem Markt gibt, die Sie vielleicht einmal ausprobieren möchten.

Bildgeneratoren und Modelle

Schauen wir uns zuerst die Plattformen an, die sich auf die reine Bildgenerierung oder das Generieren von Bewegtbild spezialisiert haben.

Modell	Beschreibung	Einsatzmöglichkeiten
Midjourney <i>midjourney.com</i>	Liefert ästhetisch hochwertige, detailreiche Bilder in unterschiedlichen Stilrichtungen	Kreative Bildideen, realistische Szenen und Porträts, Moodboards, Produktvisualisierungen, Editorials
Magnific AI <i>magnific.ai</i>	Verfeinert bestehende Bilder und erhöht deren Auflösung bei gleichzeitiger Detailverstärkung, eigenes Modell zur Bildgenerierung	Upscaling, Schärfen, Nachbearbeitung, Stiltransfer, Belichtungsanpassung, realistische Bildgenerierung, Schriftzüge
Ideogram <i>ideogram.ai</i>	Kombiniert realistische Bilddarstellung mit präziser Typografieintegration, Bildgenerierung mit Charakter-Vorlage	Schrift-Bild-Kompositionen, Plakate, Social Media, Logodesigns
Stable Diffusion <i>dreamstudio.stability.ai</i>	Open-Source-Modell mit flexibler Steuerung und vielfältigen vortrainierten Modellen	Individuelle Stile, Experimente, Produkt- und Konzeptvisualisierung
Canva <i>canva.com</i>	Design-Tool mit integrierten KI-Funktionen für Bildgenerierung und Bearbeitung	Social Media, Präsentationen, Layouts, schnelle Visuals
KlingAI <i>app.klingai.com</i>	Schwerpunkt Video-Generierung, mit verschiedenen Modellen zur Bildgenerierung	KI-Videos, Motion Design, Szenen-Visualisierung, Kampagnenclips, realistische und cinematische Visuals
FLUX <i>playground.bfl.ai/image</i>	Verschiedene Modelle für realistische, aber auch visuell anspruchsvolle Motive, die Modelle »Kontext« verstehen Anweisungen wie ein Chatbot	Porträts, Produktbilder, Editorials, natürliche Lichtstimmungen, Bilderserien von Szenen, Personen und Produkten

Überblick über aktuelle Bildgeneratoren und Modelle – von klassischen Tools bis zu neuen Plattformen für Bild- und Bewegtbildgestaltung

Plattformen mit mehreren Modellen und Funktionen

In der folgenden Aufzählung finden Sie Plattformen, die Ihnen Zugang zu mehreren KI-Modellen oder sogar weiteren Funktionen bieten. Ideal, wenn Sie Flexibilität brauchen und experimentieren möchten, welche Modelle am besten zu Ihren Projekten passen.

Name	Beschreibung	Modelle
Freepik <i>freepik.com</i>	Bietet neben klassischen Grafikressourcen einen eigenen KI-Generator, Video-generierung, KI-Assistenten u.v.m.	Seedream 4K, Seedream 3, Flux 1.0, Flux 1.0 Fast, Flux 1.0 Realism, Flux 1.1, Flux Kontext [Pro], Flux Kontext [Max], Mystic 1.0, Mystic 2.5, Mystic 2.5 Flexible, Mystic 2.5 Fluid, Google Imagen 3, Google Imagen 4, Google Imagen 4 Fast, Google Imagen 4 Ultra, Google Nano Banana, Ideogram 3, GPT 1, GPT 1 HQ, Runway, Classic und Classic Fast
Adobe Firefly <i>firefly.adobe.com</i>	Bildgenerierung, Bildbearbeitung und weitere spannende KI-Funktionen	Firefly Image 5 (Vorschau), Firefly Image 4 Ultra, Firefly Image 4, Firefly Image 3, Flux Kontext Max, Flux Kontext Pro, Flux 1.1 Pro, Flux 1.1 Ultra, Flux 1.1 Ultra (Raw), Gemini 2.5 (Nano Banana), Imagen 4, Imagen 3, GPT Image, Ideogram 3.0 und Runway Gen-4 Image
Leonardo AI <i>app.leonardo.ai</i>	Besonders stark bei konsistenten Serienmotiven, realistischer Lichtführung und Stilvariationen	Lucid Origin, Nano Banana, Lucid Realism, Ideogram 3.0, GPT-Image-1, FLUX.1 Kontext Max, FLUX.1 Kontext, Phoenix 1.0, FLUX Dev, FLUX Schnell und Phoenix 0.9
Krea <i>krea.ai</i>	Bild- und Videogenerator mit verschiedenen weiteren Funktionen wie Nodes und Real Time Video	Krea 1, Flux, Nano Banana, ChatGPT Image, Seedream 4, Wan 2.2, Qwen, Flux.1 Krea, Flux Kontext, Flux Kontext Pro, Seedream 3, Imagen 4 Ultra, Imagen 4 Fast, Imagen 4, Runway Gen-4, Ideogram 3.0, Flux 1.1 Pro, Flux 1.1 Pro Ultra und Imagen 3

Diese Plattformen bieten meist umfangreiche Möglichkeiten, KI in den Projekten einzusetzen. Achten Sie auf Lizenz- und Abomodelle, um die Kosten im Blick zu haben.

PRAXIS

Die Wahl des Modells

Bei der Fülle an Möglichkeiten ist es manchmal schwer, das richtige Modell für sich zu finden. In meinem Buch »KI für Kreative« habe ich einen ähnlichen Test mit einem recht einfachen Prompt gemacht. Diesmal möchte ich Ihnen zeigen, wie kreativ die Modelle mit komplexeren Prompts umgehen. Daher nehme ich einen der Prompts aus dem vorigen Kapitel, entstanden während eines Workshops beim gemeinsamen Entwickeln eines Prompts mit Hilfe der Promptformel und ChatGPT:



Prompt

Editorial photography of an elegant older person, standing with dignity, a delicate scarf floating inside a giant soap bubble. Captured from a frog's-eye view, main subject arranged according to the golden ratio. The background shows colourful highlights of a funfair with a chain carousel and a candyfloss stand, slightly blurred to create depth. Pastel boho inspired colors in sage green, beige, cream, rose and terracotta dominate the palette, creating an earthy and artistic atmosphere. Fabric shines with a soft glow, warm candlelight adds intimacy, while the mood remains slightly surreal and mysterious.

Dabei wähle ich bei jedem Modell die Standard-einstellungen. Das bedeutet auch, dass das Ergebnis noch weiter beeinflusst werden kann, indem vortrainierte Stile angewandt oder weitere Einstellungen gemacht werden können – je nach Modell. Lassen Sie einen Prompt ruhig zwei bis dreimal durchlaufen, bevor Sie weitere Anpassungen und Einstellungen testen. Manchmal ist dann auch schon das gewünschte Ergebnis mit dabei.



Midjourney V7



Ideogram 3.0



Leonardo AI (Lucid Origin)



Stable Diffusion (Stable Image Ultra)



Magnific AI (Mystic v2.5)



OpenAI GPT 1.5 – High (via Freepik)



FLUX Kontext Max (via Adobe Firefly)



Google Nano Banana Pro (via Freepik)



FLUX Kontext Pro (via Adobe Firefly)



Google Imagen 4 (via Adobe Firefly)

Das passende Modell finden

Meine Empfehlung ist es, mehrere Modelle zur Auswahl zu haben, um zu testen, welches für die aktuelle Aufgabe am besten geeignet ist. Manchmal brauchen wir realistische Bilder, die nah am Prompt sind – bei anderen Motiven wie diesem hier braucht es manchmal eine kreati-

vere Auslegung des Prompts, um ein visuell ansprechendes Ergebnis zu erhalten.

Wie man bei diesem Beispiel und zugegeben herausfordernden Prompt sehen kann, hat jedes Modell einen eigenen Look. Obwohl hier nicht explizit von einer Frau die Rede war, wur-



Seedream 4K (via Freepik)



Firefly Image 5 (Preview)



Kling Image 2.1



Z Image (via Freepik)

den (meist hellhäutige) Frauen generiert – bis auf Adobe Firefly. Das Wort »older« hat dazu geführt, dass alle Personen graues Haar haben. Am besten sind die Ergebnisse von Midjourney, Leonardo und FLUX Kontext Max gelungen. Googles Nano Banana Pro hält sich stark an den Prompt und wirkt sehr realistisch. ChatGPT hat

leider immer noch den typischen Warmfilter, der nicht zu jedem Bild passt. Die Darstellung der Personen ist zum Großteil realistisch, trotz des außergewöhnlichen Prompts. Auch wenn Ihnen ein Ergebnis nicht gefällt, probieren Sie die Modelle mit anderen Motiven aus, und lernen Sie die einzelnen Stärken kennen.

Bildprompts mit Hilfe von KI

In den letzten Kapiteln haben Sie die Promptformel kennengelernt und auch, wie Sie kooperativ mit KI Prompts entwickeln können. Ein paar Tipps habe ich noch zu ergänzen, wie Sie KI nutzen können, um Bildprompts zu schreiben:

- ChatGPT ist sehr gut darin, Bildprompts zu schreiben. Statt mir dort ein Bild zu generieren, lasse ich mir meist einen Prompt auf Englisch, optimiert z. B. für Midjourney, schreiben. Es gibt auch GPTs, die von anderen dazu trainiert wurden, Prompts für bestimmte Modelle oder Einsatzzwecke zu generieren.
- Viele Bildgeneratoren bieten Hilfe beim Prompten an, wie die Chatfunktion in Midjourney. Andere können Ihre Prompts optimieren, wenn diese zu lang oder zu kurz sind, wie bei Ideogram, Freepik, Dreamstudio und Leonardo.
- Bei Midjourney, Freepik und Ideogram können Sie ein Bild hochladen und es beschreiben lassen. Auch alle Textmodelle wie ChatGPT können Ihnen eine Bildbeschreibung in Form eines Prompts erstellen.

Bildgenerierung via Chat

Viele klassische Textgeneratoren bieten auch die Bildgenerierung an. Das hat Vor- und Nachteile: Statt einen Prompt zu entwickeln, können Sie wie in einem Gespräch mit der KI Ihre Wünsche zum Bild äußern, besonders praktisch innerhalb eines Projekts, wenn die KI bereits einige Hintergrundinformationen aus dem Chatverlauf hat. Aber es gibt auch Bildgeneratoren, die eine Chatfunktion mit Kontextverständnis anbieten. Hier folgt eine Liste der Tools, mit denen ich bereits gute Erfahrungen gemacht habe.

Name	Beschreibung	Einsatzmöglichkeiten
ChatGPT <i>chatgpt.com</i>	Dieses LLM versteht den Kontext sehr gut und kann darauf aufbauend Bilder generieren. Leider nur eines auf einmal, ohne viele weitere Einstellmöglichkeiten, außer der Wahl von ein paar Stilen.	Visualisierung von Ideen, Prompting, komplexere Grafiken und Schaubilder, Vorlagen für Bildprompts
Gemini <i>gemini.google.com</i>	Viel Aufsehen hat Google mit dem Modell »Nano Banana« erregt. Gemini versteht den Kontext und kann über den Chat Bilder generieren oder bearbeiten.	Besonders Nano Banana Pro und Nano Banana 2 eignen sich für die Bildgenerierung und -bearbeitung.

Name	Beschreibung	Einsatzmöglichkeiten
Freepik <i>freepik.com/pikaso/assistant</i>	Neben vielen Bild- und Videofunktionen bietet Freepik den Assistant, der mittels Chat Bilder generiert und bearbeitet, aber auch vorgefertigte Workflows anbietet.	Besonders geeignet für Bilderserien von Personen und Produkten, Ausprobieren von Stilen, Bearbeiten von Bildern
Midjourney <i>midjourney.com</i>	Statt eines Prompts kann eine Chatfunktion aktiviert werden, mit der man Wünsche und Ideen zum Bild äußert und die KI den Prompt schreibt.	Sehr praktisch, wenn man sich an ein Motiv herantasten möchte oder das Prompting nicht zum gewünschten Ergebnis führt

Manchmal ist es einfacher, statt einen fertigen Prompt einzugeben, in den Dialog zu gehen – besonders bei der Ideenfindung. Die Tools können nicht nur Bilder generieren, sondern auch richtig gute Prompts schreiben und verstehen den Kontext innerhalb eines Projekts.

Bildstile entwickeln und anwenden

Für viele Projekte oder in der Unternehmenskommunikation ist ein durchgängiger Stil besonders wichtig – aber oft eine Herausforderung. Eine Mischung aus Stockbildern, Auftragsfotografien und jetzt auch noch KI macht es nicht leicht, einen durchgängigen visuellen Look beizubehalten. Doch die Anbieter von Bildgeneratoren bieten mittlerweile umfangreiche Möglichkeiten, individuelle Stile in der Generierung anzuwenden.

Textprompt als Stilvorgabe

Der Textprompt ist ein wichtiger Faktor, wenn es um Bildstile geht. Beschreiben Sie den Stil, die Farben und die Stimmung des Bildes, um den Stil zu beeinflussen. Sie können den Stil von vorhandenen Bildern mit Hilfe der Promptformel beschreiben oder ein LLM bitten, das für Sie zu tun.

Prompt Erstelle mir eine Beschreibung des Bildstils als JSON-Code. Achte dabei auf Licht, Farben, Stil und Stimmung.

ChatGPT hat mir mit Hilfe dieses Prompts eine detaillierte Analyse des Bildstils gemacht und mir einen JSON-Code ausgegeben. Im folgenden Praxisteil erfahren Sie mehr über JSON-Codes und deren Erzeugung sowie Anwendung als Bildprompt. Anschließend habe ich im selben Chat die KI gebeten, mir einen Bildprompt für einen Vogel, der in einem Strauch sitzt, zu schreiben und den Bildstil aus dem Code zu adaptieren. Sie können solche Stilbeschreibungen in Tools wie ChatGPT, Gemini und Freepik auch direkt als JSON-Code prompten. Erfahrungsgemäß funktioniert diese Art der »Stilübertragung« am besten mit dem Tool, das das ursprüngliche Bild generiert hat.