

Midjourney, Stable
Diffusion, Firefly,
ChatGPT, Copilot,
Dall-E, Luminar NEO

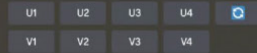


/imagine

prompt



cute robot
is painting a
landscape with
impressive
mountains,
camera behind
--ar 4:5



Midjourney & Co.

Wie du mit KI beeindruckende Bilder schaffst

- ▶ Die Kunst, KI zu nutzen – Kreativität und Effizienz mit Midjourney
- ▶ Mit guten Prompts Bilder, Grafiken und Videos generieren
- ▶ KI für alle – einfach ohne Vorwissen einsteigen

Rainer Hattenhauer



Rheinwerk
Computing

Kapitel 1

Die faszinierende bunte Welt der Bild-KIs

Du wolltest schon immer tolle Fotos und Bilder machen – aber leider fehlt dir das handwerkliche Geschick? Dann gibt es gute Nachrichten für dich: Dank leistungsfähiger KI-Werkzeuge ist es heute für alle möglich, die eigene Fantasie in spektakuläre Bilder umzusetzen.

Die Werkzeuge der künstlichen Intelligenz (kurz *KI*)¹ sind in aller Munde. Spätestens seit der Veröffentlichung der textgenerativen *KI ChatGPT* im November 2022 ging ein Ruck durch die digitale Welt. Das Erzeugen beliebiger Textwerke durch wenige, gezielte Stichworte beflügelte auch weniger Sprachbegabte.

Aber nicht nur im literarischen, sondern gerade auch im künstlerischen Bereich haben sich KI-Tools wie DALL-E und Midjourney etabliert. Hier setzt dieses Buch an. Es soll dir als Reiseführer den Weg in die wunderbare Welt der bildgenerativen KIs ebnen. Tauche ein in Landschaften, deren Horizont allein deine eigene Vorstellungskraft begrenzt.

1.1 Kunstwerke – von Profis und Laien wie von Zauberhand erschaffen

Welche Möglichkeiten bieten moderne KI-Tools im Bereich der Bilderstellung, und welche Voraussetzungen musst du selbst mitbringen, um diese produktiv zu nutzen? Der folgende Abschnitt liefert Antworten.

¹ Wichtige technische Begriffe werden im Glossar am Ende des Buches in einfacher Sprache erklärt. Dadurch ist sichergestellt, dass der Lesefluss nicht durch ein Übermaß an technischen Erklärungen gestört wird. Mehr zum didaktischen Aufbau des Buches findest du in Abschnitt 1.1.3, »Dein Fahrplan in die Kreativität«.

1.1.1 Aufbruch in eine neue Welt

Es waren zwei unterschiedliche Fotos, die die kreative Kunstszene in Aufruhr versetzten. Das erste Bild (siehe Abbildung 1.1) stammte vom Berliner Fotografen Boris Eldagsen. Er gewann damit im Jahr 2023 mit einem interessanten Foto den renommierten Fotowettbewerb *Sonys World Photo Award*. Boris Eldagsen hatte das Bild nicht mit einer Kamera, sondern mit dem KI-Bot DALL-E² generiert. Fairerweise lehnte er den Hauptpreis des Wettbewerbs – eine hochpreisige Kameraausrüstung – ab. Für seine Tätigkeit schuf er das Schlagwort *Promptografie*. An die Stelle der physikalischen Kamera tritt zur Bildzeugung ein sogenannter metaphysischer Prompt, der an eine KI übergeben wird.



Abbildung 1.1 Der Urknall bildgenerativer KI: Boris Eldagsen, »PSEUDOMNESIA | The Electrician«, promptography 2022, courtesy Photo Editon Berlin

Das zweite Werk, das uns schlagartig auch die Gefahren des unbedachten Einsatzes künstlicher Intelligenz vor Augen führte, zeigte im April 2023 den amtierenden Papst in der Daunenjacke eines bekannten Modelabels (siehe Abbildung 1.2). Was ist Realität und was ist Fiktion? Wie wird man diese in Zukunft auseinanderhalten können? Fragen, die sich diverse Forscherinnen und Forscher auch im Hinblick auf denkbare Bildmanipulationen im Kontext von Wahlen stellen.

2 DALL-E wird wie ChatGPT von der Firma OpenAI entwickelt.



Abbildung 1.2 »Der Papst im Daunenmantel«, generiert von Midjourney, war das erste bekannt gewordene Beispiel für die Manipulation eines Fotos im Bereich Prominenter.

1.1.2 Benötigte Skills

Welche Fähigkeiten sind erforderlich, um bildgenerative KI-Werkzeuge sinnvoll einsetzen zu können? In den Epochen der alten Meister waren handwerkliche Fähigkeiten unabdingbar, um klassische Gemälde wie die »Mona Lisa« oder »den Mann mit dem Goldhelm« zu realisieren. Beeindruckende Werke konnten andererseits nur dann entstehen, wenn der Künstler über eine gehörige Portion Fantasie, Zeit und oftmals auch Geld verfügte. Daran ändert sich prinzipiell auch nichts, wenn man eine KI arbeiten lässt. Handwerkliches Geschick ist dann aber nicht mehr gefragt: An die Stelle des sauber ausgeführten Pinselstrichs tritt die natürliche Sprache, in der die Prompts, die Aufträge an die KI, formuliert werden.

Das ist sehr hilfreich für Leute wie mich, die mangels Talent und/oder Fleiß im Kunstunterricht in der Schule nie über ein »befriedigend« hinausgekommen sind. Eine Bild-KI ist ein ideales Werkzeug, wenn man zwei linke Hände hat, aber gut mit dem Medium »Sprache« umgehen kann. So bringen Midjourney & Co. die Kunst unter das unbegabte Volk – wir erleben quasi die Demokratisierung der Kunst. In meiner Kindheit gab es das »Malen nach Zahlen«, bei dem eine vorgefertigte Skizze einfach mit den per Zahlencode vorgegebenen Farben ausgefüllt wurde. So könnte man heute vom »Malen nach Worten« sprechen. Das Beste daran: Man macht sich nicht einmal die Hände mit Farbe

schmutzig. Und auch im Bereich der Fotografie können beeindruckende Werke entstehen, wie etwa Abbildung 1.1 zeigt.

Betrachten wir die Evolution bei der Erschaffung von Bildern am Beispiel eines konkreten Motivs: »Pferd mit Ritter vor einem Schloss«. Die Bilderreihe in Abbildung 1.3 zeigt eine eigene einfache Skizze aus meiner Feder, gefolgt von der Umsetzung mit einer frühen Version von Midjourney. Hier kann man mit etwas Fantasie zwar Pferd und Schloss, nicht jedoch einen Ritter erkennen. Das finale Bild zeigt schließlich, was die aktuelle Version zu leisten vermag.



Abbildung 1.3 Evolution der Bildgeneration zum Thema »Pferd mit Ritter vor einem Schloss«

Zum Erzeugen von KI-Bildern ist es somit wichtig, eine grobe Vorstellung vom darzustellenden Motiv zu haben und dieses in knappen Worten beschreiben zu können.

Die Superkräfte eines KI-Künstlers

Früher benötigte man Talent und handwerkliches Geschick, heute in erster Linie Fantasie und darüber hinaus die Fähigkeit, Szenen, die sich im Kopfkino abspielen, in prägnanten Wortfolgen (= Prompts) zu beschreiben. Den Schöpfungsprozess können auch Handskizzen wie die in Abbildung 1.3 unterstützen. Diese helfen oft dabei, die eigenen Gedanken zu sortieren.

1.1.3 Dein Fahrplan in die Kreativität

Welche Zielgruppe soll das vorliegende Buch ansprechen? Du bist hier richtig, wenn ...

- ... du einen fundierten, praktischen Einstieg im Umgang mit bildgenerativen KI-Tools suchst.
- ... du anhand diverser Workshops und Praxisbeispiele lernen möchtest, was mit den aktuellen Werkzeugen machbar ist und was nicht.

- ▶ ... du lernen möchtest, wie man mit sogenannten *Superprompts* (das sind sehr mächtige Befehle für den KI-Bot) das Optimum aus einer bildgenerativen KI herausholen kann.
- ▶ ... du darüber hinaus ein wenig hinter die Kulissen von bildgenerativen *Chatbots* blicken und verstehen willst, auf welcher Basis sie ihre Ausgaben generieren.

Du bist hier falsch, wenn ...

- ▶ ... du dich für tiefgehende theoretische Grundlagen der (Sprach-)Modelle interessierst, die hinter KI-Chatbots stecken.
- ▶ ... du eine Bild-KI als Photoshop-Ersatz ansiehst, die dein geplantes Projekt zu 100 % identisch mit deiner Vorstellung im Kopf umsetzen kann und wird. Das funktioniert leider (noch) nicht.

Dein Reiseführer in die Welt der KI-Bots ist folgendermaßen aufgebaut:

- ▶ Im Rest des vorliegenden Kapitels stellst du den ersten Kontakt zu den bekanntesten bildgenerativen KI-Tools her und lernst, warum der sorgsame Aufbau von Prompts so wichtig ist, wenn du perfekte Ergebnisse erzielen möchtest.
- ▶ Kapitel 2 widmet sich rundum Midjourney. Du lernst den Goldstandard der bildgenerativen KIs kennen und beherrschen. Anhand von einfachen Beispielen wirst du in die magische Welt des Prompts eingeführt. Du lernst die Oberfläche des Programms und seine Einstellungen anhand vieler Beispiele detailliert kennen und erfährst, wie du deine Bildersammlungen optimal organisierst. Schließlich klären wir die Frage, wem das KI-generierte Bildmaterial überhaupt gehört.
- ▶ In Kapitel 3 werden wir uns auf Einsteigerniveau klassischen Themen in Form von leicht nachvollziehbaren Workshops widmen. Du wirst lernen, wie man typische Fotoshooting-Situationen einfängt oder Bilder im Stil der großen Meister mithilfe von KI generiert. Spektakuläre Landschaften werden in Szene gesetzt, und wir tasten uns Schritt für Schritt an die Parameter heran, die Midjourney zur Bildgestaltung anbietet.
- ▶ Kapitel 4 wendet sich an alle (fortgeschrittenen) Kreativitätsprofis. Klassische Aufgaben aus dem Grafikbereich, wie beispielsweise das Entwerfen von Logos oder das Erstellen von Bildmaterial für einen Flyer, werden hier umgesetzt. Auch der manchmal steinige Weg zu verlustfrei skalierbaren Vektorgrafiken wird mithilfe von KI beschritten. Du erfährst, wie man eigenes Bildmaterial, z. B. Fotos oder Skizzen, mit KI weiterverarbeiten kann. Ein Abschnitt beschäftigt sich mit dem Einfluss aktueller KI-Tools auf die Kunstszene, und du lernst, wie du witzige Ideen materialisieren kannst.

- In Kapitel 5 gehen wir etwas in die Tiefe und wenden uns fortgeschrittenen Themen zu. Du lernst, wie du Teilbereiche eines Bildes gezielt variiert und Objekte austauschst. Wir werden mit Stilen experimentieren und die Auflösung von Bildern durch Upscaling erhöhen. Anschließend werden wir uns damit beschäftigen, wie man (beispielsweise für einen Comicstrip) konsistente Charaktere erzeugen kann. Dabei lernst du den Parameter `character reference` (kurz: `--cref`) kennen. Schließlich erfährst du, welchen Weg Profis beschreiten, um optimale Prompts zu erstellen.
- In Kapitel 6 stelle ich dir weitere spezielle KI-Werkzeuge vor. Dabei begeben wir uns in den Bereich multimedialer Anwendungen und du lernst, wie du per KI Videos und dazu passende Tonspuren erstellen kannst. Anschließend erfährst du, wie du KI-Tools oder KI-erweiterte Anwendungen in Adobe Photoshop oder Skylum Luminar Neo für die Bildbearbeitung nutzen kannst. Du lernst auch, wie du eine bildgenerative KI auf dem lokalen Rechner betreibst und damit Datenschutzprobleme umgehst. Im letzten Teil des Kapitels zeige ich dir, wie du textgenerative KIs (z. B. ChatGPT) verwenden kannst, um Superprompts für Midjourney & Co. zu erstellen.
- Kapitel 7 gibt dir schließlich einen Ausblick darauf, was die bildgenerativen Tools für uns in der Zukunft bereithalten werden. Dabei werden nicht nur Chancen, sondern auch Gefahren thematisiert.

Eine Anmerkung zum pädagogischen Konzept des vorliegenden Buches: Wesentliches Ziel ist es, dir den praktischen, produktiven Umgang mit Midjourney & Co. zu vermitteln. Dabei lässt es sich nicht vermeiden, dass gelegentlich auch einmal theoretische Hintergründe zur Sprache kommen. Diese werden entweder in Form von Kästen im Fließtext behandelt oder als *Fachbegriffe* im Glossar am Ende des Buches erläutert. Wenn du also etwas tiefer in die Theorie einsteigen möchtest, dann findest du dort weiterführende Informationen. Abschließend noch einige generelle Anmerkungen zu den Inhalten:

- Bei der Erstellung der Prompts wurde auf Universalität geachtet. Die im Buch vorgestellten Prompts lassen sich in nahezu jeder bildgenerativen KI verwenden. Dennoch kann es aufgrund der rasanten Entwicklung im Bereich der KI vorkommen, dass das eine oder andere beschriebene Werkzeug eine leicht geänderte Benutzeroberfläche aufweist. In derartigen Fällen hilft dir eine Webrecherche weiter.
- Die Features von Midjourney & Co. sind einem steten Wandel unterworfen. Die Beispiele im Buch wurden in der Regel mit dem Modell v6 realisiert. Im Sommer 2024 stand aber schon das verbesserte Modell v7 in den Startlöchern. In diesem Zusammenhang möchte ich dich bitten, etwas Nachsicht zu haben, falls sich die Begriffe in den Buch-Screenshots sowie die Resultate der Prompts nicht zu 100 % mit deinen Bildern bzw. der von dir verwendeten bildgenerativen KI decken. Insbesondere das Webfrontend von Midjourney befand sich zum Zeitpunkt der Manuskripterstellung

noch in der Alpha-Phase. An der im Buch beschriebenen prinzipiellen Vorgehensweise ändert sich in der Regel nichts.

- ▶ Die Beispiele für bildgenerative KIs wie Midjourney wurden in englischer Sprache verfasst, da das Trainingsmaterial optimal auf diese Sprache zugeschnitten ist. Bei textgenerierenden Bots wie ChatGPT funktioniert das Prompten aber auch auf Deutsch.
- ▶ Viele Beispiele in diesem Buch basieren auf Prompts, die von begeisterten Midjourney-Nutzer*innen auf *youtube.com* und *medium.com* zur Verfügung gestellt wurden. Wann immer ich einen Prompt von jemand anderem verwendet habe, findest du den Link dazu in einer Fußnote. Oft sind hinter dem Link weitere interessante Prompts und Werke der kreativen Künstler*innen zu finden, sodass es sich lohnt, dem Link zu folgen.
- ▶ Einige Textteile im Buch sind unter Mitwirkung von KI-Tools entstanden. An der Erstkorrektur bezüglich der Rechtschreibung, der Grammatik und des Stils hat DeepL Write mitgewirkt. Für die Glossareinträge wurde ein Custom GPT von ChatGPT mit GPT-4 verwendet.

Nun aber genug der Vorworte. Bist du bereit für eine Reise zu neuen Horizonten der Kreativität, die das Zusammenspiel von menschlicher und künstlicher Intelligenz ermöglicht? Dann lies weiter, staune und begreife, wie sich unsere Vorstellung von Fotografie und Bildkunst dank KI grundlegend verändern wird.

1.2 Am Anfang steht der Prompt ...

Diesen Begriff hast du im letzten Abschnitt schon kennengelernt: Der Kern einer jeder Kommunikation mit einer künstlichen Intelligenz ist der sogenannte *Prompt*. Was es damit auf sich hat und wie man diesen gestaltet, um optimale Ergebnisse zu erzielen, erfährst du in diesem Abschnitt.

1.2.1 Prompts und Diffusionsmodelle

Üblicherweise kommunizieren Menschen mit Computern, indem sie kryptische Programmiersprachen wie Java, Python oder C lernen. Seitdem KI-Tools wie ChatGPT oder Midjourney auf dem Markt sind, hat sich der Umgang mit dem Computer drastisch verändert. Dank künstlicher Intelligenz können wir mit der Maschine in unserer eigenen, natürlichen Sprache kommunizieren. Für die Formulierung einer Anweisung per Prompt sind allerdings auch gewisse Regeln notwendig, damit man ein optimales Ergebnis erhält. Beginnen wir mit einem Beispiel aus dem Reich der textgenerativen

KIs. Wir geben folgenden Prompt in ChatGPT oder Microsoft Copilot (*copilot.microsoft.com*) ein und erhalten ausführlich Antwort:³

Welche Möglichkeiten bietet eine bildgenerative KI?

Eine bildgenerative KI, wie sie beispielsweise durch Modelle wie DALL-E oder Stable Diffusion realisiert wird, bietet eine Vielzahl von Möglichkeiten, die sich über diverse Anwendungsfelder erstrecken. Hier sind einige der bemerkenswertesten Anwendungen:

1. **Kunst und Design**: Künstler und Designer nutzen bildgenerative KI, um neue Kunstwerke zu schaffen, Designkonzepte zu visualisieren oder um Inspiration für ihre Arbeit zu finden. Solche Technologien ermöglichen es, einzigartige Bilder, Texturen oder sogar ganze Kunststile zu erzeugen.
 2. **Bildmanipulation und -restauration**: Diese Technologien können eingesetzt werden, um Bilder zu bearbeiten, zu verbessern oder zu restaurieren. Das umfasst beispielsweise das Auffüllen von Bildbereichen, das Entfernen von Objekten, das Verbessern der Bildqualität oder das Restaurieren alter Fotos.
 3. **Erstellung von Lehr- und Lernmaterialien**: Bildgenerative KI kann genutzt werden, um anschauliches Material für Bildungszwecke zu erstellen, wie z. B. Diagramme, Illustrationen und andere visuelle Hilfsmittel, die komplexe Konzepte verständlich machen.
- [...]
9. **Wissenschaftliche Visualisierungen**: Forscher in Feldern wie der Medizin, Astronomie oder Biologie nutzen bildgenerative KI, um komplexe Daten oder Konzepte durch Visualisierungen greifbar zu machen.
 10. **Simulation und Training**: In der Luft- und Raumfahrt, dem Militär oder der Medizin können realistische Simulationen erstellt werden, die für Trainingszwecke eingesetzt werden, ohne reale Risiken einzugehen.

Diese Technologien entwickeln sich rasant weiter, sodass ständig neue Anwendungsmöglichkeiten entstehen. Die Fähigkeit, visuelle Inhalte schnell und in hoher Qualität zu generieren, eröffnet ein weites Feld für Kreativität und Innovation in vielen Bereichen.

3 Die Antwort stammt von ChatGPT unter Verwendung des Sprachmodells GPT-4. Das Resultat des Prompts wurde in geeigneter Form gekürzt. Der Text der KI wurde 1 : 1 übernommen und nicht korrigiert.

Was genau passiert nach der Eingabe eines solchen Prompts in eine textgenerative KI? Du hast sicher schon einmal beobachtet, dass nach der Eingabe einiger Wörter auf der virtuellen Tastatur deines Smartphones neue Wortvorschläge oberhalb des Buchstabenfeldes erscheinen, die du durch Antippen für die Fortsetzung deines Textes verwenden kannst. Diese Wortvorschläge stammen aus Wörterbüchern und werden anhand deines Textes als nächstwahrscheinliche Wörter vorgeschlagen. Während das Wörterbuch für die Handytastatur jedoch nur einige hunderttausend Wörter und ihre Kontexte enthält und dafür wenige Megabyte an Speicherplatz benötigt, sind in Sprachmodellen wie *GPT* & Co. mehrere Terabyte an Wörtern gespeichert.

Ein Prompt in ChatGPT wird nach Schlüsselwörtern – sogenannten *Token* – durchsucht, und das Sprachmodell gibt nach dem Zufallsprinzip Wortfolgen oder ganze Sätze aus, die in den Kontext der durch den Prompt formulierten Frage passen. Weiterhin lernen ChatGPT & Co. im Laufe einer Kommunikation dazu, wenn du den Bot mit weiteren Informationen fütterst.

Eine bildgenerative KI wie Midjourney arbeitet in ähnlicher Weise. Zunächst wird der Prompt in eine Folge von Token zerlegt. Zu jedem Token wird dann im riesigen Trainingssatz des Modells eine Assoziation zu einem gespeicherten Bildelement hergestellt. Das Zielbild zum Prompt wird schließlich stufenweise generiert: Aus einem verschwommenen Pixelgemisch kristallisiert sich langsam ein konkretes Motiv heraus. Dieser Prozess ähnelt dem physikalischen Prinzip der Diffusion,⁴ bei dem sich unterschiedliche Gasmoleküle selbstständig vermischen, wenn diese zuvor getrennt waren. In Fall der bildgenerativen KI erfolgt die Diffusion in umgekehrter Richtung: Aus Chaos entsteht Ordnung (siehe Abbildung 1.4).



Abbildung 1.4 Bei Verwendung einer bildgenerativen KI, die auf einem Diffusionsmodell basiert, kristallisiert sich das Motiv während der Erzeugung langsam heraus.

4 Eines der bekanntesten KI-Modelle zur Bilderzeugung trägt daher den Namen »Stable Diffusion«.

1.2.2 Das erste eigene Bild

Genug der grauen Theorie! Du scharrst bestimmt schon mit den Hufen und möchtest endlich selbst KI-Bilder erstellen, und das möglichst schnell, kostenlos und ohne langwierige Anmeldeprozeduren. Dann lass uns loslegen – mit Craiyon!

Begib dich dazu auf die Seite www.craiyon.com. Dort kannst du direkt starten, indem du einen Prompt deiner Wahl in das Eingabefeld schreibst (siehe Abbildung 1.5). Ich wähle in diesem Fall als Motiv ein Einhorn vor einem Regenbogen. Die Erstellung des Bildes dauert etwa 1 Minute und beginnt, nachdem du auf die Schaltfläche DRAW geklickt hast. Lass deiner Fantasie freien Lauf und spiele ein wenig mit Craiyon! Es hat sich bewährt, die Prompts für Craiyon in Englisch zu verfassen. Dabei können dir Google Translate (<https://translate.google.com>) oder DeepL (www.deepl.com/translator) helfen.

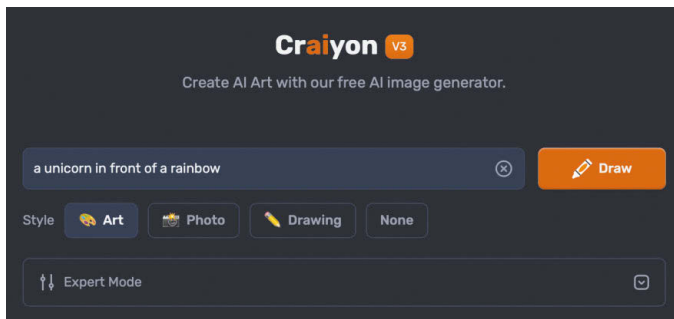


Abbildung 1.5 Die Oberfläche von Craiyon ist einfach gehalten, sodass auch Anfänger gut damit zurechtkommen.

Mit den Schaltflächen unterhalb des Eingabefeldes kannst du wählen, ob das Bild eher künstlerischen (ART), fotografischen (PHOTO) oder skizzenhaften (DRAWING) Charakter haben soll.



Abbildung 1.6 Ein ganzer Zoo voller Einhörner, einige recht exotische Exemplare sind auch dabei. Im vorliegenden Fall wurden in jedem Viererraster von links nach rechts die Stile »Art«, »Photo« und »Drawing« ausgewählt.

Wenn dir eines der Bilder im Raster besonders gut gefällt, klicke es einfach an und betätige anschließend den Button **UPSCALE**. Dadurch wird das Bild in maximaler Auflösung generiert und kann schließlich heruntergeladen werden.

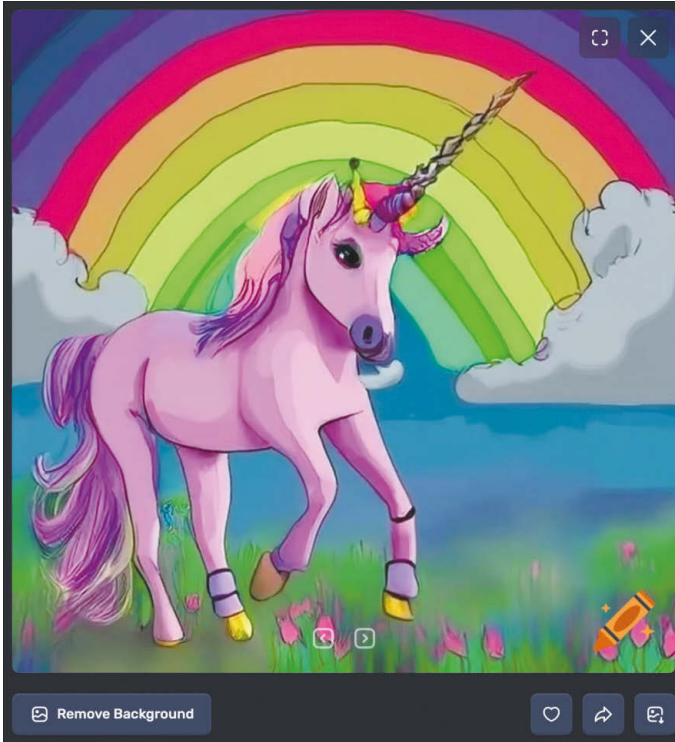


Abbildung 1.7 Das hochskalierte Bild kann über die Download-Schaltfläche unten rechts heruntergeladen werden.

1.3 KI-Werkzeuge für Profis

Die kostenlose Version von Craiyon ist für Anfänger sicher ein guter Einstieg in die Welt der bildgenerativen KIs. Wenn du aber wirklich professionelle Ergebnisse erzielen möchtest, kommst du nicht umhin, etwas Geld auszugeben. Sehen wir uns in diesem Abschnitt einmal bei kostenpflichtigen Alternativen um.

1.3.1 DALL-E

DALL-E ist der bildgenerative Sprössling aus dem Hause OpenAI. Du erinnerst dich: Das ist die Firma, die uns ChatGPT beschert hat. DALL-E ist Bestandteil der kostenpflichtigen

(Plus-)Version von ChatGPT. Über einen kleinen Umweg lässt sich DALL-E aber auch kostenlos nutzen, und zwar dann, wenn du ein Microsoft-Konto besitzt. In diesem Fall erhältst du jeden Monat einige Credits für kostenlose Bilderzeugungen via DALL-E. Möchtest du DALL-E in Verbindung mit deinem Microsoft-Konto nutzen, dann begib dich auf die Seite des Microsoft Copilots (copilot.microsoft.com) und melde dich dort mit deinen Zugangsdaten an. Danach kannst du deinen ersten Prompt absetzen.



Abbildung 1.8 Das Ergebnis, das DALL-E in Verbindung mit Microsofts Copilot erzeugt, ist sehr ansprechend. Der Prompt lautete: »Erstelle ein Bild von einem Einhorn vor einem Regenbogen.« Unten rechts wird die verbleibende Anzahl von Credits eingeblendet, die dir für die Bilderzeugung zur Verfügung stehen.

Wenn du Glückspilz bereits ein ChatGPT-Plus-Abonnement besitzt, kannst du direkt im Chatfenster Bilder erzeugen lassen. Stelle zunächst sicher, dass das Sprachmodell GPT-4 aktiviert ist.

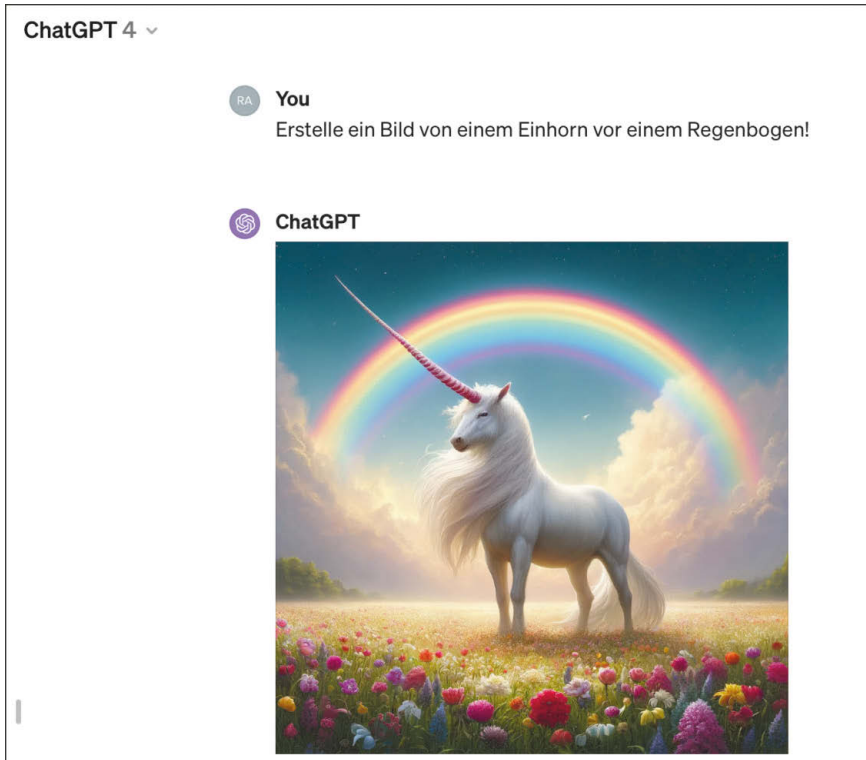


Abbildung 1.9 Interpretation des Einhornmotivs in ChatGPT mithilfe von GPT-4/DALL-E. Die Proportionierung des Horns ist verbesserungsbedürftig.

1.3.2 Midjourney

Wenden wir uns nun dem Titelgeber des vorliegenden Buches zu. Im Vergleich zu DALL-E gibt es hier eine Vielzahl von Möglichkeiten, die Bildgestaltung nicht nur über den Prompt selbst, sondern auch über Einstellungen und Parameter zu beeinflussen. Testen wir einmal die aktuelle Version von Midjourney (zum Zeitpunkt der Drucklegung war dies die Version 6). Wie du ein Konto für Midjourney anlegst, werde ich später in Kapitel 2 genau zeigen. An dieser Stelle genügt es zu wissen, dass ein Prompt in der Form `/imagine <Bildbeschreibung>` übergeben wird.



Abbildung 1.10 Midjourney, Meister der bildgenerativen KI

Beeindruckend, nicht wahr? Ich muss allerdings zugeben, dass ich ein wenig »gemoelt« habe. Der Prompt zur Erzeugung des Midjourney-Bildes wurde in englischer Sprache formuliert und lautete:

A photograph of a majestic unicorn, pure white with a shimmering, multi-colored mane, standing proudly in front of a vibrant rainbow. The unicorn is gazing towards the horizon, its horn gleaming with the same spectral colors as the rainbow behind it. The setting is a lush, green meadow dotted with wildflowers, under a clear blue sky with a few fluffy clouds. Created Using: DSLR camera, soft natural light, high dynamic range, detailed texture on the unicorn's coat, vivid saturation on the rainbow, depth of field effect, magic hour lighting --ar 16:9 --v 6.0 --style raw

Dieser Prompt ist damit deutlich komplexer als die oben verwendeten einfach strukturierten Beispiele. Andererseits ist Midjourney der einzige Kandidat unter den bislang vorgestellten KI-Tools, der derart komplexe Prompts umsetzen kann. Daher werden wir die meisten Beispiele in diesem Buch mit dieser außergewöhnlichen bildgenerativen KI realisieren. Steigen wir also im nächsten Kapitel direkt in die Grundlagen von Midjourney ein.

Kapitel 2

Einstieg in Midjourney

Ein Malkasten, der auf Abruffunktioniert und der Fantasie keine Grenzen setzt: Das ist Midjourney. Das KI-Tool hat sich mittlerweile zum Quasi-standard im Bereich der bildgenerativen KI entwickelt. Damit der Einstieg reibungslos funktioniert, sind einige Dinge zu beachten.

Eine Bemerkung vorweg: Midjourney ist kein Bildbearbeitungsprogramm im Stil von Illustrator, Photoshop & Co. Die Werkzeuge, die Midjourney zur Verfügung stellt, basieren nicht auf einem virtuellen Pinsel/Stift, Ebenen und Farbpaletten. Die Ergebnisse werden vielmehr durch natürlichsprachliche Eingabeaufforderungen erzeugt.



Abbildung 2.1 Midjourney bietet deiner Kreativität unbegrenzte Möglichkeiten.

Bevor du dein erstes Bild mit Midjourney kreieren kannst, musst du ein Konto erstellen und ein Abonnementmodell auswählen. Dann musst du dich mit dem Chatbot verbinden. Danach kannst du dich schließlich an die ersten einfachen Motive wagen. Wenn du etwas sicherer geworden bist, ist es an der Zeit, dich mit der Oberfläche zu beschäftigen. Mit der Zeit wird die Anzahl deiner Bildkreationen zunehmen. Dann ist es sinnvoll, dir eine Strategie zur Organisation deiner Werke zu überlegen. Wie das funktioniert, erfährst du in diesem Kapitel.

2.1 Erste Schritte

Es gibt prinzipiell zwei Möglichkeiten, mit Midjourney zu kommunizieren: per Discord-Chat oder direkt über die Midjourney-Website. Die erste Variante ist flexibler, was die Möglichkeit der Steuerung der Bildgenerierung über *Parameter* angeht. Die zweite Variante ist einsteigerfreundlicher. Auf jeden Fall musst du zunächst einen Zugang zu Midjourney auf der Website einrichten. Du erfährst im folgenden Abschnitt, wie du dafür vorgehst.

2.1.1 Den Zugang zu Midjourney einrichten

Um Midjourney nutzen zu können, benötigst du ein Konto. Ich beschreibe nachfolgend zunächst den Weg, wie du Midjourney in Verbindung mit *Discord* nutzt.¹ Diese Variante bietet ein größeres Maß an Flexibilität. Discord war ursprünglich als Kommunikationsdienst für Computerspieler (auch *Gamer* genannt) entwickelt worden.

Solltest du noch keinen Zugang zu Discord besitzen, so begib dich auf <https://discord.com> und klicke dort auf die Schaltfläche EINLOGGEN. Über den Link REGISTRIEREN kannst du nun einen Zugang zu Discord anlegen (siehe Abbildung 2.2).

Du erhältst eine E-Mail mit einem Bestätigungslink, der anzuklicken ist, um das Konto zu aktivieren. Melde dich nun einmal testweise bei Discord an und prüfe, ob der Zugang funktioniert.

Anschließend begibst du dich auf www.midjourney.com. Klicke im Fenster unten rechts auf die Schaltfläche SIGN IN. Melde dich nun in dem sich öffnenden Fenster mit deinen Discord-Zugangsdaten an. Erteile im nachfolgenden Dialog die Berechtigung, dass Midjourney einen Server auf Discord hinzufügen darf (siehe Abbildung 2.3). Daraufhin wird dieser zur Liste der von dir in Discord bereits abonnierten Server hinzugefügt. Schließlich wirst du an die Midjourney-Website im Browser weitergeleitet.

¹ Zum Zeitpunkt der Drucklegung war es noch nicht möglich, ein Konto direkt auf der Midjourney-Homepage über das Webfrontend einzurichten. Dies kann sich in der Zwischenzeit geändert haben.

Account erstellen

E-MAIL *
maria.musterfrau@web.de

ANZEIGENAME
Marie Musterfrau

BENUTZERNAME *
maria_muster_art

PASSWORT *
[Redacted]

GEBURTSDATUM *
5 Mai 1995

Weiter

☒ Ich habe die Nutzungsbedingungen und die Datenschutzerklärung von Discord gelesen und akzeptiert.

[Hast du bereits einen Account?](#)

Abbildung 2.2 Der Weg zu Midjourney führt über ein Discord-Konto.

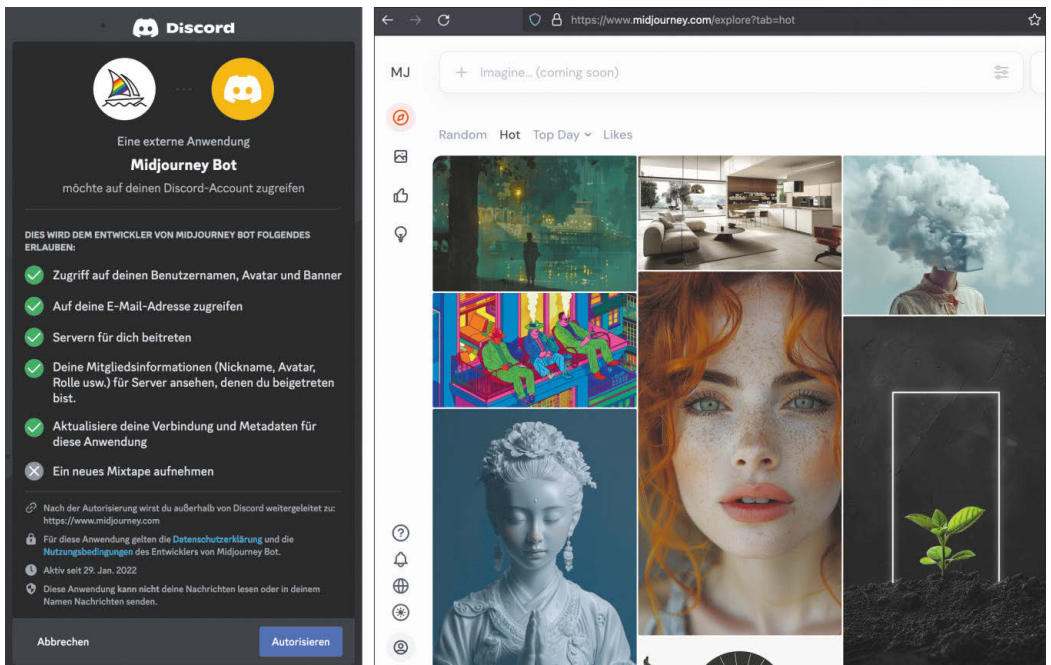


Abbildung 2.3 Nach der Anmeldung mit den Discord-Login-Daten landest du auf der Midjourney-Website. Hier kannst du direkt mit der Bilderstellung loslegen, sofern du ein Abo abgeschlossen hast.

Obwohl man (ein Abonnement vorausgesetzt) Bilder direkt auf der Midjourney-Webseite durch Eingabe von Prompts in die Promptzeile generieren kann, nutzt die Mehrzahl der erfahrenen Midjourney-Nutzer*innen die Übergabe von Prompts per Discord-Chat. Schauen wir uns diese Möglichkeit einmal ausführlich an.

Begib dich auf die Seite <https://discord.com> und melde dich (falls nicht schon geschehen) mit deinen Discord-Zugangsdaten an. Klicke auf der linken Seite in der Rubrik der Server auf das Midjourney-Symbol (das kleine Segelschiff in Abbildung 2.4).

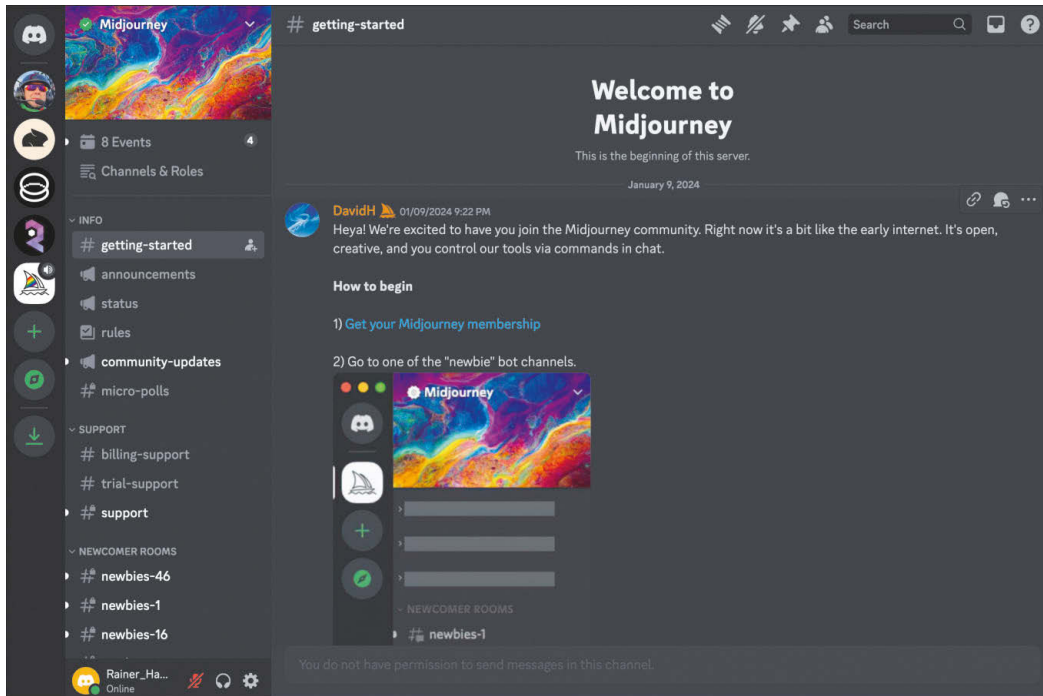


Abbildung 2.4 Der Zugriff auf Midjourney via Discord bietet mehr Möglichkeiten als der Zugriff über die Midjourney-Website.

Im linken Bereich des Fensters erscheinen nun verschiedene Kommunikationskanäle. Als Einsteiger solltest du einmal den Kanal #getting-started anklicken: Hier erfährst du (sofern du der englischen Sprache mächtig bist) Wissenswertes über die ersten Schritte in Midjourney. Aber keine Sorge, nachfolgend gibt es dazu eine ausführliche deutschsprachige Einführung.

Wenn du einen Eindruck davon erhalten möchtest, was andere Anwender*innen aktuell mit Midjourney gestalten, dann klick dich einfach einmal durch die #newbies-Kanäle.

Hier machen Einsteiger ihre ersten Schritte. Die Profis trifft man eher in den #general-Kanälen oder im #showcase.

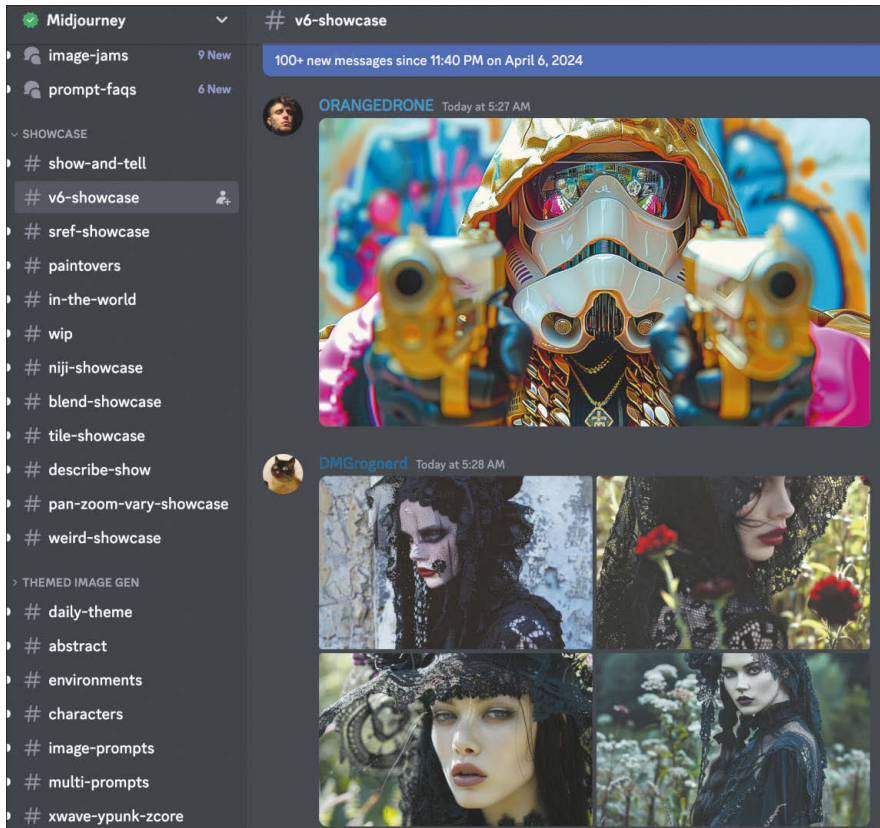


Abbildung 2.5 Die aktuellen Arbeiten der Profis kann man im »Showcase«-Bereich bewundern. Auf der linken Seite des Discord-Fensters lassen sich unterschiedliche Kanäle durch Anklicken auswählen.

2.1.2 Was kostet der Spaß?

Während Craiyon und zum Teil auch Microsoft Copilot mit integriertem Dall-E kostenlos nutzbar sind, musst du für Midjourney zwingend ein Abo abschließen. Dies war in der Startphase von Midjourney einmal anders, als es ein kleines Freikontingent an Rechenzeit gab. Aufgrund des Missbrauchs durch einige schwarze Schafe wurde dieses aber wieder abgeschafft.

Tabelle 2.1 zeigt die verschiedenen Abo-Optionen (Stand August 2024) nebst Preis für monatliche oder jährliche Abrechnung. Letztere fällt stets günstiger aus.

	Basic	Standard	Pro	Mega
Preis (monatlich)	10 US\$	30 US\$	60 US\$	120 US\$
Preis (jährlich)	96 US\$	288 US\$	576 US\$	1152 US\$
»schnelle« Rechenzeit pro Monat	200 min	15 h	30 h	60 h
gleichzeitige schnelle Bildjobs	3	3	12	12
Bildgenerationen im Relaxed Mode	–	unbegrenzt	unbegrenzt	unbegrenzt
Stealth Mode	–	–	✓	✓

Tabelle 2.1 Die Abo-Modelle von Midjourney

Dazu einige Erläuterungen:


- **Preis:** Die Preise sind für das jeweilige Modell in US-Dollar angegeben. Der Preis in Euro variiert aufgrund des aktuellen Wechselkurses. Als Zahlungsmittel werden die gängigen Kreditkarten (Visa, Master) sowie Google Pay und Apple Pay akzeptiert.
- **»Schnelle« Rechenzeit:** Die Generation eines Midjourney-Bildes bzw. 4er-Grids nimmt ca. 1 Minute Rechenzeit auf einem Midjourney-Cloud-Server in Anspruch. Diese Rechenzeit wird »schnelle Rechenzeit« genannt, weil man ohne Wartezeit direkten Zugriff auf die leistungsfähigen GPU-Cluster erhält. Bei den teureren Plänen besteht die Möglichkeit, eine unbegrenzte Anzahl von Bildern im sogenannten *Relaxed Mode* erstellen zu lassen. In diesem Fall muss man zwar deutlich länger auf das Ergebnis warten, es wird dafür aber keine wertvolle »schnelle Rechenzeit« abgerechnet.
- **Gleichzeitige schnelle Bildjobs:** Du kannst die Erstellung mehrerer Bilder in Auftrag geben, indem du Prompts im Chat unmittelbar nacheinander abschickst. Im Basic- und Standardplan sind drei parallele Jobs möglich, im Pro- und Megaplan sind es sogar zwölf.
- **Stealth Mode:** In diesem Modus wird der Prompt, den du für ein bestimmtes Bild verwendet hast, nicht der Allgemeinheit angezeigt. Diese Option bieten nur der Pro- und Megaplan.

Die im Rahmen eines Plans erworbene Rechenzeit wird jeden Monat zum Abrechnungsdatum erneuert. Es ist nicht möglich, verbleibende Rechenzeit in den Folgemonat »hinüberzuretten«. Unabhängig vom speziell gewählten Plan kannst du aber stets Rechenzeit dazukaufen, wenn du merkst, dass du bis zum Ende des Monats nicht mit deinem Kontingent auskommen wirst.

Tipp: Welcher Plan eignet sich für mich?

Wenn du zunächst nur ein wenig in die Bilderstellung mit Midjourney hineinschnuppern möchtest, genügt ein einmaliges Monatsabonnement des Basic-Plans für 10 US\$. Damit kannst du zunächst ca. 200 Bilder erstellen. Wenn du merkst, dass es dir Spaß macht (und davon gehe ich stark aus), dann kannst du auf ein Basic-Jahresabo für $12 \times 8 \text{ US\$} = 96 \text{ US\$}$ wechseln. Wenn du schließlich völlig von Midjourney überzeugt bist, dann kannst du auf einen Jahresplan Standard für $12 \times 24 \text{ US\$} = 288 \text{ US\$}$ upgraden. Damit kannst du in einem Jahr ca. 10.000 Bilder im Schnellmodus erstellen. Das sollte auch die ambitioniertesten »Promptograf*innen« zufriedenstellen.

Um einen Plan zu buchen, hast du zwei Möglichkeiten (siehe Abbildung 2.6):

1. In Discord: Gib in einem beliebigen Kanal des Discord-Servers den Befehl `/subscribe` ein, gefolgt von der -Taste. Klicke anschließend auf den nun erscheinenden Link namens MANAGE ACCOUNT. Du wirst dadurch auf die Midjourney-Abo-Seite weitergeleitet.
2. Direkt über die Midjourney-Website: Klicke, nachdem du dich eingeloggt hast, auf deinen Login-Namen am unteren linken Bildrand. Folge dem Link unter MANAGE SUBSCRIPTION.

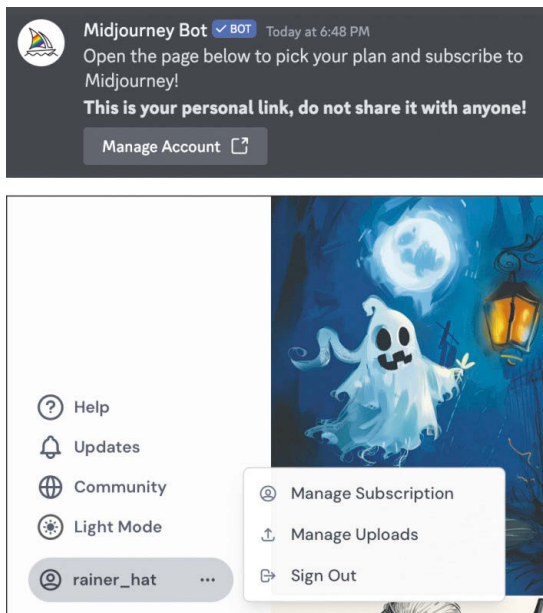


Abbildung 2.6 Der Weg zum Abo kann sowohl aus Discord (oben) oder von der Midjourney-Website aus (unten) erfolgen.

3. Wähle im folgenden Fenster schließlich das passende Abo-Modell aus (siehe Abbildung 2.7) und gib im Dialog deine Rechnungsinformationen sowie das bevorzugte Zahlungsmittel an. Danach kannst du sofort mit der Bilderzeugung loslegen.

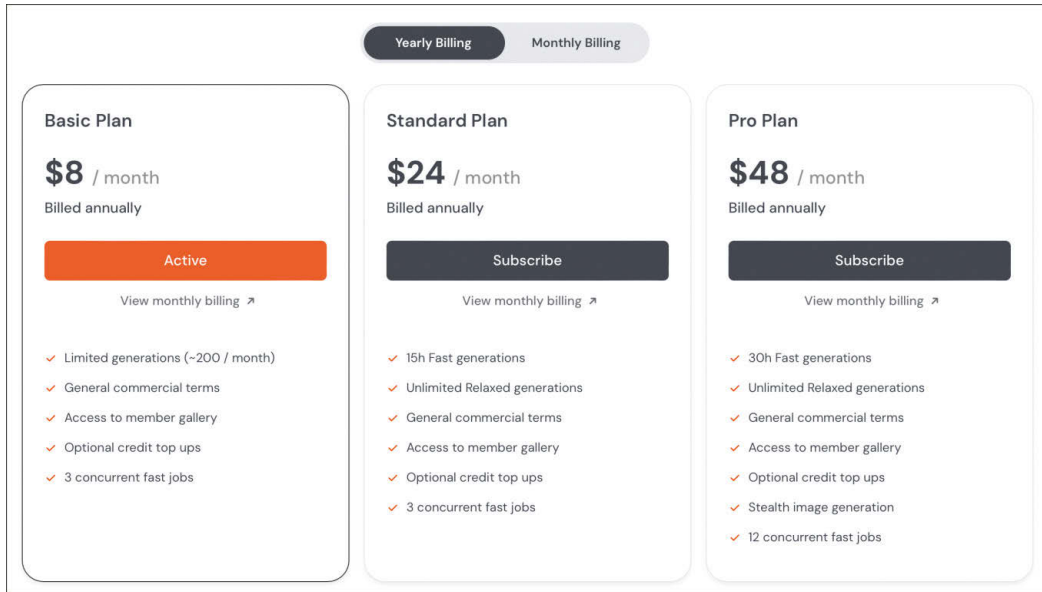


Abbildung 2.7 Im vorliegenden Fall habe ich bereits den Basic-Plan gebucht. Es gibt an dieser Stelle die Möglichkeit, zu »besseren« Plänen zu wechseln.

2.1.3 Midjourney mobil

Aufgrund der Tatsache, dass auf Midjourney über den Discord-Chat zugegriffen werden kann, lässt sich die KI auch auf mobilen Endgeräten nutzen. Abbildung 2.10 zeigt die Bedienung von Midjourney über die Discord-App auf einem iPhone. Am unteren Bildrand findest du das Eingabefeld für Prompts. Mittels der abgebildeten QR-Codes kannst du die für dein Mobilsystem passende App rasch installieren.



Abbildung 2.8 Discord für Android (links) und iOS (rechts)

Midjourney mobil per Browser

Alternativ kannst du auch per Mobilbrowser direkt auf das Midjourney-Webfrontend zugreifen. Es bleibt abzuwarten, ob Midjourney selbst seinen Nutzer*innen eine eigene App spendieren wird.

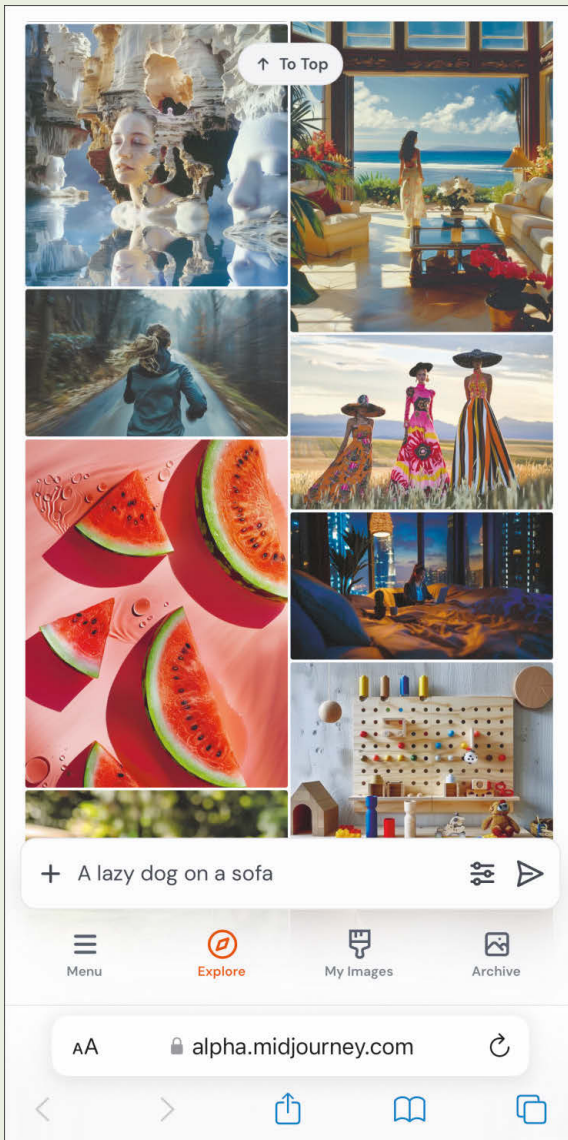


Abbildung 2.9 Das Midjourney-Webfrontend kann von jedem mobilen Browser aus genutzt werden.



Abbildung 2.10 Midjourney lässt sich auch auf dem Smartphone via Discord bedienen.

2.1.4 Dein erstes Bild mit Midjourney

Lass uns ein erstes Bild in Midjourney erstellen! Das kannst du entweder über das Prompteingabefeld auf der Website oder über den Discord-Chat machen. Wenn du dich für die zweite Möglichkeit entscheidest, gibt es einen kleinen Trick: Normalerweise würdest du dich in einen der vielen Newbie-Channels begeben. Da sich dort viele Neulinge tummeln, werden dort sehr viele Bilder generiert. Dadurch wird das Chatfenster permanent aktualisiert, und es besteht die Gefahr, dass du dein eigenes Bild bzw. das Ergebnis deines Prompts aus den Augen verlierst. Um den Überblick zu behalten, solltest du folgendermaßen vorgehen:

Unmittelbar nach Abschluss deines Midjourney-Abonnements erhältst du eine Direktnachricht vom Midjourney Bot mit einer Bestätigung. Um diese zu lesen, klicke einfach auf das Maskensymbol in der linken oberen Ecke des Discord-Fensters. Dadurch gelangst du in den Bereich der Sofortnachrichten. Klicke jetzt in der Liste auf MIDJOURNEY BOT (siehe Abbildung 2.11). In diesem Channel kannst du nun direkt mit dem Bot kommunizieren, ohne dass dich jemand stört. Insbesondere kannst du nun auch damit beginnen, eigene Bilder mit dem `/imagine`-Befehl zu erstellen.

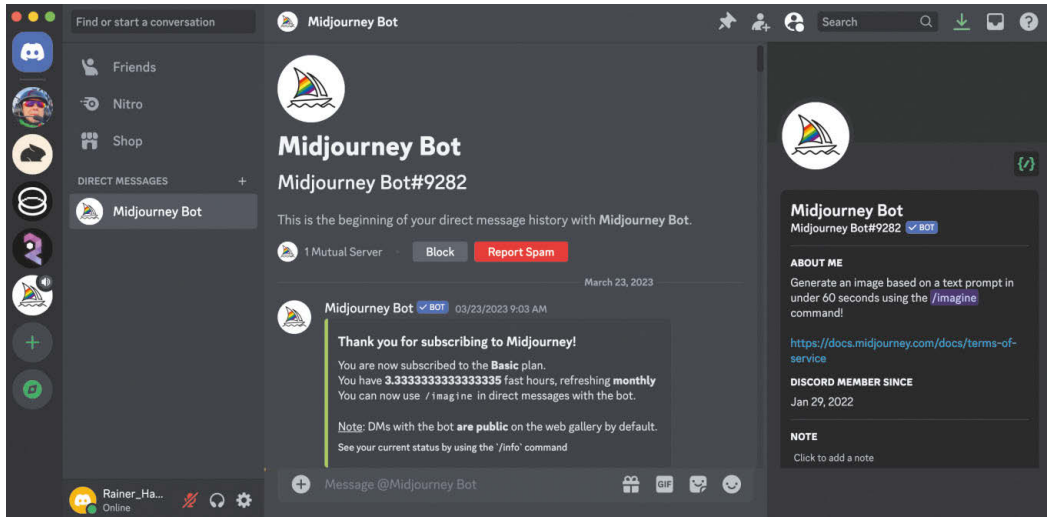


Abbildung 2.11 Im Chatkanal mit dem Midjourney Bot kannst du ungestört arbeiten.


Wie du mit einem eigenen Discord-Server mehr Ordnung in deine Kreationen bringen kannst, erfährst du in Abschnitt 2.3.2. Wenn du Midjourney direkt über die Website bedienst, brauchst du dir solche Gedanken natürlich nicht zu machen.

Für dein erstes Bild solltest du einen sehr einfachen, überschaubaren Prompt wählen, z. B.:

Eine junge Frau mit rotem Hut im Central Park New York

Wie bereits erwähnt, erzeugt man die besten Ergebnisse mit Midjourney durch englische Prompts. Solltest du nicht so fit im Englischen sein, zeige ich dir in Abschnitt 2.1.5, »Verstehst du meine Sprache?«, welche Tools du für die Übersetzung verwenden kannst. Wir übergeben somit folgenden Prompt an Midjourney:

`/imagine a young woman with a red hat in central park new york`

Groß- und Kleinschreibung spielen beim Prompten keine Rolle. Bereits bei der Eingabe des Schrägstrichs bietet dir Midjourney eine Reihe von Befehlen zur Vervollständigung an, unter anderem `/imagine`. Wähle dieses Kommando aus und gib den obigen Prompt ein. Betätige anschließend die -Taste und lass dich überraschen, was sich Midjourney zur beschriebenen Szene ausdenkt. Das Bild erscheint zu Beginn des Erstellungsvorgangs zunächst verwaschen. Diesen Prozess der Bilderzeugung nennt man *Diffusion*.

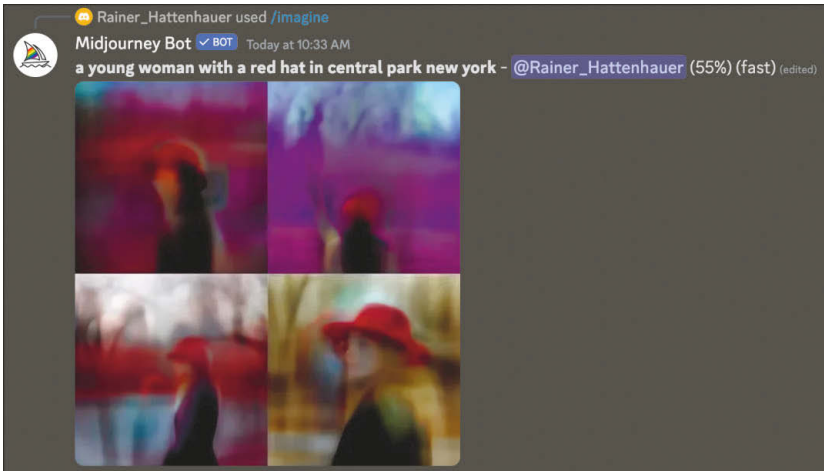


Abbildung 2.12 Das Bild entwickelt sich langsam, der prozentuale Fortschritt wird hinter dem Prompt angezeigt.

Die Konturen des Motivs werden nun immer deutlicher. Nach ca. 1 Minute erscheint schließlich das (vorläufige) Endergebnis in Form eines Viererrasters (im Englischen auch *Grid* genannt, siehe Abbildung 2.13).

Mithilfe der Schaltflächen unter dem Bildraster kannst du entscheiden, wie es weitergehen soll. Die Teilbilder sind intern von 1 bis 4 nummeriert. Links oben befindet sich Bild 1, rechts oben Bild 2, links unten Bild 3 und rechts unten schließlich Bild 4 (siehe Abbildung 2.13).

- **U1–U4:** Damit entscheidest du dich für ein Bild mit der entsprechenden Nummer im Vierer-Grid und skalierst dieses Bild hoch. Der Buchstabe U steht für Upscale.
- **V1–V4:** Sollte dir eines der Bilder gut gefallen, kannst du auf dessen Basis ein neues Grid erzeugen. Dieses zeigt dann leichte Variationen des ausgewählten Bildes. Das V steht in diesem Fall für Variation.
- **Doppelpfeil:** Wenn dich bisher keines der Bilder des Rasters überzeugt, dann kannst du über diese Schaltfläche noch einmal neu beginnen.

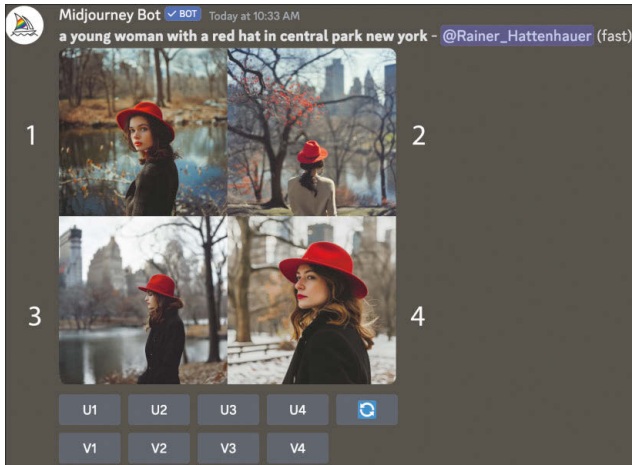


Abbildung 2.13 Nach der Erzeugung stehen dir vier verschiedene Varianten des Motivs zum Weiterverarbeiten zur Verfügung. Die Nummerierung der Bilder wurde zur Orientierung ergänzt.

An dieser Stelle begnügen wir uns zunächst damit, ein Bild aus dem Raster auszuwählen, um dieses als Grundlage für weiterführende Arbeiten zu nutzen. Im vorliegenden Fall wähle ich das vierte Bild und betätige daher die Schaltfläche U4.

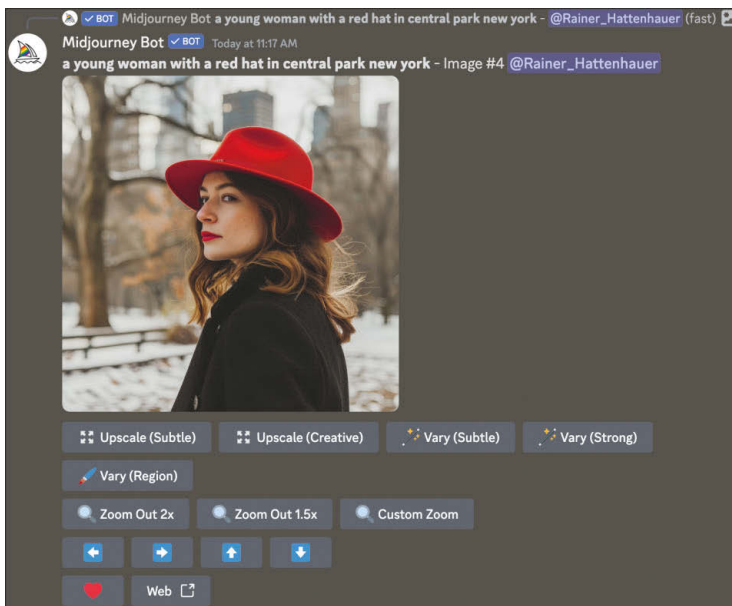


Abbildung 2.14 Ein Bild aus dem Raster wurde ausgewählt. Die Bedeutung aller sichtbaren Schaltflächen wirst du in Abschnitt 2.2.2 noch ausführlich kennenlernen.

An dieser Stelle unterscheiden sich die Midjourney-Versionen. Während bis Version v5 an dieser Stelle bereits das Endergebnis vorliegt, muss ab Version v6 nun noch explizit ein weiterer Vorgang – ein Upscaling – vorgenommen werden, um das Bild detailliert auszugestalten. Dabei wird ein erneuter Diffusionsprozess angeworfen, was durch Anklicken einer der beiden Upscale-Schaltflächen unter dem Bild erreicht wird. Ich entscheide mich an dieser Stelle für UPSCALE (SUBTLE), um die Kreativität bei der Ausschmückung mit Details im Rahmen zu lassen. Wer kreativere Ergebnisse wünscht, klickt auf die Schaltfläche UPSCALE (CREATIVE).

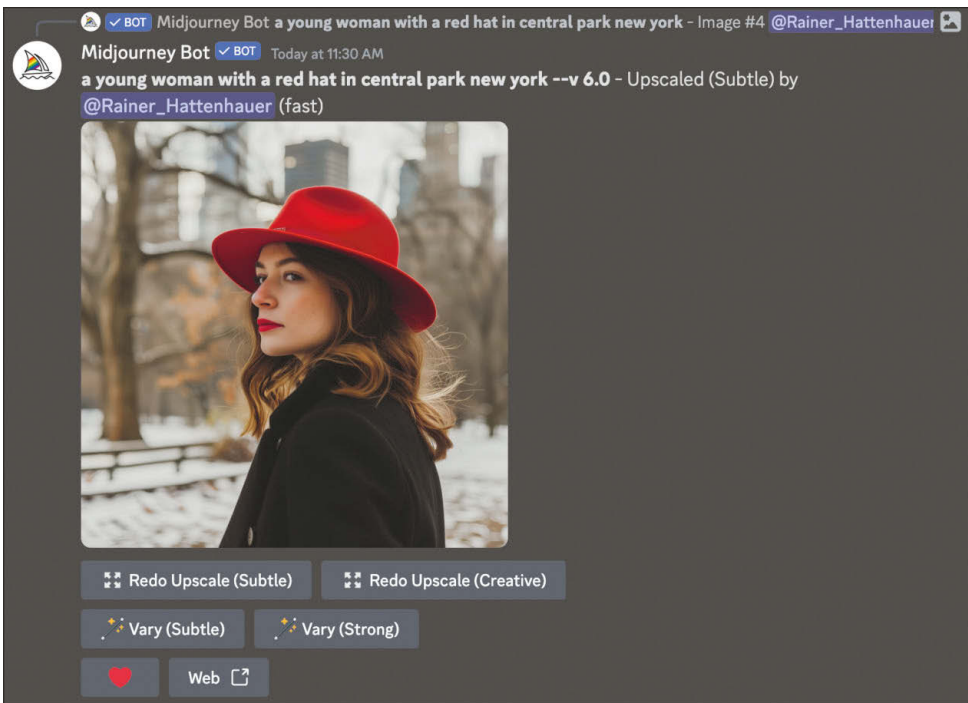


Abbildung 2.15 Nach dem Upscaling steht das Bild in voller Auflösung zur Verfügung. Sollte dir das Ergebnis nicht gefallen, so lässt sich der Prozess über die Schaltfläche »Redo Upscale« wiederholen.

Du kannst dir einen Eindruck von der Qualität des detaillierten Bildes verschaffen, indem du dieses anklickst. Möchtest du schließlich das Bild in voller Auflösung auf deinen Rechner herunterladen, dann klicke auf die Schaltfläche WEB. Dadurch wird ein Browserfenster geöffnet, und du landest direkt auf der Midjourney-Website. Führe einen rechten Mausklick über dem Bild im Browser durch und wähle den entsprechenden Kontextmenüpunkt zum Herunterladen des Bildes. Die Bezeichnung kann dabei

von Betriebssystem zu Betriebssystem variieren. Fertig! Du hast dein erstes Bild mithilfe einer KI erstellt.

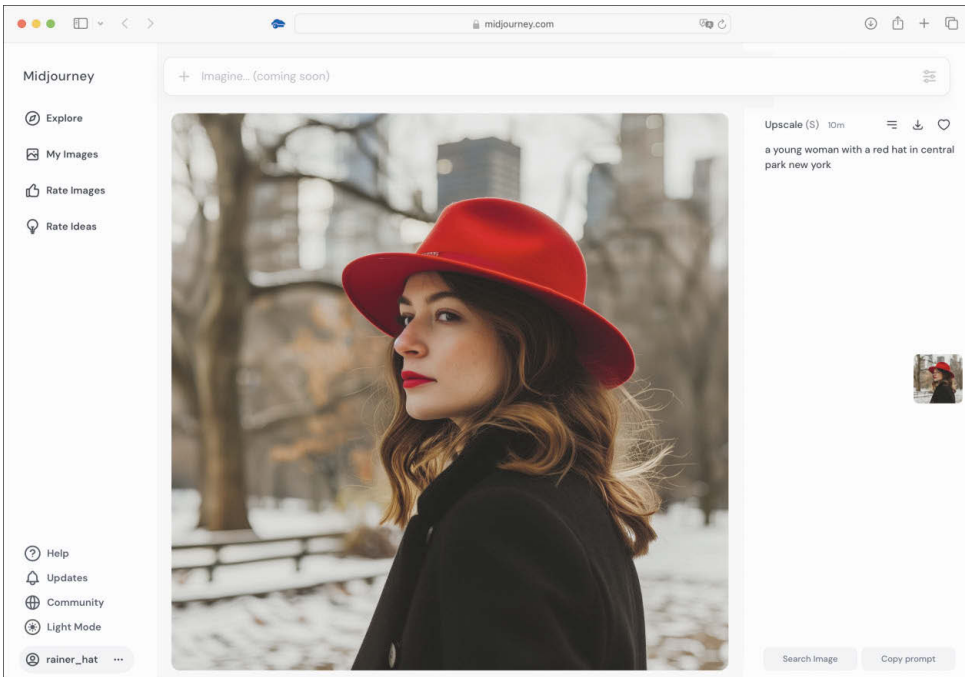


Abbildung 2.16 Das fertige Bild kann von der Midjourney-Seite per Browser heruntergeladen werden.

2.1.5 Verstehst du meine Sprache?

Die besten Ergebnisse erzielt man bei Midjourney, wenn die Prompts auf Englisch verfasst werden. Das mag bei einfachen Szenen bei der Mehrzahl der Leser*innen mit durchschnittlichen Englischkenntnissen noch funktionieren. Spätestens bei der Übersetzung umfangreicher Superprompts stößt man jedoch schnell an Grenzen. Hier bieten sich Übersetzungstools wie Google Translate (<https://translate.google.com>) oder DeepL (www.deepl.com/translator) an. Insbesondere DeepL besticht durch Übersetzungsergebnisse, die dank des Einsatzes von Methoden der künstlichen Intelligenz denen eines Muttersprachlers in nichts nachstehen.

DeepL kann für gelegentliche Übersetzungen kostenlos genutzt werden. Wird ein bestimmtes Kontingent pro Monat überschritten, ist ein Abonnement erforderlich. Dieses beginnt bei 8 €/Monat.

Kapitel 3

Die KI-Leinwand – Workshops für Einsteiger

Mit wenigen Worten viel bewirken, zumindest in künstlerischer Hinsicht, das gelingt dir mit Midjourney & Co. Die Ergebnisse, die selbst Laien mit bildgenerativer KI erzielen können, genügen dabei auch professionellen Ansprüchen.

Nachdem du im letzten Kapitel die Oberfläche und die Steuerungsmöglichkeiten von Midjourney ausführlich kennengelernt hast, geht es nun endlich in die Praxis. Du lernst das Bauprinzip erfolgreicher Prompts kennen, erstellst fotorealistische Szenen und be gibst dich auf die Spuren der großen Maler*innen. Schließlich erfährst du, wie du die vielfältigen Optionen und Parameter nutzen kannst, um faszinierende Bilder zu erzeugen.

3.1 Ein Bild für alle Fälle

Der Weg zu atemberaubenden Bildkompositionen führt oft nicht direkt über einen komplexen Prompt, sondern entwickelt sich aus einfachen Stichworten. Wie du das Prinzip »vom Einfachen zum Komplexen« bei der Nutzung von bildgenerativer KI anwenden kannst, erfährst du in diesem Abschnitt.

3.1.1 Iteratives Prompten

Wie entstehen große Kunstwerke? Ist es wirklich so, dass der Maler vor einer leeren Leinwand sitzt, sich ein Motiv überlegt und dieses schließlich mit Pinsel und Farbe in die Wirklichkeit umsetzt? Nun, diese Legende entspricht wohl kaum der Realität. Vielmehr ist es so, dass Künstler*innen mit Skizzenblöcken arbeiten, auf denen sie Bildideen grob skizzieren und dann in mehreren Anläufen versuchen, ihre Vorstellungen umzusetzen.

Auch bei einer bildgenerativen KI ist es sinnvoll, sich von groben Skizzen zum Endergebnis zu tasten, indem das Bild schrittweise verfeinert wird. Ein solcher Prozess wird *Iteration* genannt. Die meisten Anwender*innen von Midjourney sind nach dem ersten Kontakt oft enttäuscht, wenn ihre komplex und ausführlich beschriebene Bildvorstellung nicht 1 : 1 umgesetzt wird. Der Trick heißt: schrittweise vom Einfachen zum Komplexen.

Nehmen wir das folgende Beispiel: Du möchtest ein Bild erstellen, auf dem ein Paar Hand in Hand auf einer Palmeninsel in den Sonnenuntergang schlendert. Dann beginne mit dem Paar und dem Sonnenuntergang. Wir formulieren den Prompt auf Englisch und schreiben:

a couple in front of the setting sun

Das erstellte Vierer-Grid geht schon in die richtige Richtung (siehe Abbildung 3.1).



Abbildung 3.1 Das Paar vor dem Sonnenuntergang wird gut getroffen.

Was an dieser Stelle auffällt, ist die Vorprägung der KI: Bei der Übergabe des Begriffs »Paar« (engl. *couple*) im Prompt wird in den seltensten Fällen ein gleichgeschlechtliches Paar dargestellt. Dies müsste gegebenenfalls explizit im Prompt definiert werden. Man spricht in diesem Fall von einem *Bias*, also einer Verzerrung, die dadurch entsteht, dass in den vielen Millionen, wenn nicht Milliarden Trainingsbildern überwiegend gemischtgeschlechtliche Paare dargestellt werden.

Kommen wir zur nächsten Eigenschaft, die das Bild haben soll. Das Paar soll Hand in Hand der untergehenden Sonne entgegengehen. Ändern wir die Aufgabenstellung entsprechend ab:

a couple, walking hand in hand into the setting sun

Das Komma ist wichtig, um den Prompt in logische Einheiten zu unterteilen.



Abbildung 3.2 Das Setting wird geändert: Das Paar wandert nun in den Sonnenuntergang.

Nun muss noch der Ort – die Location – angepasst werden. Das Paar soll sich auf einer einsamen Insel befinden. Am linken Bildrand sollen Palmen stehen, im Hintergrund soll das weite Meer zu sehen sein. Das erreichen wir durch Ergänzen weiterer Promptstichwörter:

a couple, walking hand in hand into the setting sun. Lonely island, palm trees on the left edge of the picture, the vast sea in the background.

(Übersetzung: Ein Paar, das Hand in Hand in die untergehende Sonne geht. Einsame Insel, Palmen am linken Bildrand, das weite Meer im Hintergrund.)



Abbildung 3.3 Das Setting wurde noch einmal angepasst. Ich wähle hier das Bild unten rechts. Bei der Bildgeneration kann es durchaus zu Artefakten kommen, das sind Fehler in der Bildlogik. So scheint sich das Paar in Bild 3 auf der Wasseroberfläche zu bewegen.

Durch Drücken der Schaltfläche **UPSCALE** (in meinem Fall **U4**) kannst du nun eines der Bilder vergrößern. Beachte, dass ab Midjourney-Version v6 ein weiterer Upscale-Vorgang (**UPSCALE (SUBTLE)** oder **UPSCALE (CREATIVE)**) notwendig ist, um das Bild in voller Auflösung mit allen Details zu erzeugen. Abbildung 3.4 zeigt ein mögliches Endergebnis.



Abbildung 3.4 Das durch den iterativen Prozess entstandene Ergebnisbild entspricht recht gut der ursprünglichen Idee.

3.1.2 One-Shot-Prompts

Auch sie gibt es, die ultimativen, kurzen Prompts, die mit wenigen Worten eine Szene auf den Punkt bringen. Man spricht in diesem Zusammenhang auch von *One-Shot-Prompts*.

Zum Beispiel kann ein Frosch, der auf einem Blatt sitzt, mit wenigen Worten spektakulär in Szene gesetzt werden:

a frog on a leaf, close-up

(Übersetzung: Ein Frosch auf einem Blatt)

Der Begriff *Close-up* steht hier für eine Nahaufnahme.



Abbildung 3.5 One-Shot-Prompts können sehr effektiv sein.

Weitere Beispiele für One-Shot-Prompts sind:

a dog on a skateboard

(Übersetzung: Ein Hund auf einem Skateboard)



Abbildung 3.6 Wie wär's mit einem sportlichen Hund?

... oder wer's sportlicher mit einem menschlichen Motiv mag:

a surfer riding a big wave, action zoom shot

(Übersetzung: Ein Surfer reitet auf einer großen Welle, Action-Zoomaufnahme)



Abbildung 3.7 ... und »Action«: Ein Surfer reitet eine Welle.

Inhalt

1	Die faszinierende bunte Welt der Bild-KIs	9
1.1	Kunstwerke – von Profis und Laien wie von Zauberhand erschaffen	9
1.1.1	Aufbruch in eine neue Welt	10
1.1.2	Benötigte Skills	11
1.1.3	Dein Fahrplan in die Kreativität	12
1.2	Am Anfang steht der Prompt	15
1.2.1	Prompts und Diffusionsmodelle	15
1.2.2	Das erste eigene Bild	18
1.3	KI-Werkzeuge für Profis	19
1.3.1	DALL-E	19
1.3.2	Midjourney	21
2	Einstieg in Midjourney	23
2.1	Erste Schritte	24
2.1.1	Den Zugang zu Midjourney einrichten	24
2.1.2	Was kostet der Spaß?	27
2.1.3	Midjourney mobil	30
2.1.4	Dein erstes Bild mit Midjourney	32
2.1.5	Versteht du meine Sprache?	37
2.2	Die Oberfläche und grundlegende Befehle	38
2.2.1	Rundgang durch das Webinterface	38
2.2.2	Auf mein Kommando	42
2.2.3	Promptparameter	48
2.2.4	Geschmackssache – das Webinterface	52
2.2.5	Von Version zu Version	55
2.3	Organisatorisches	57
2.3.1	Den Überblick im Midjourney-Frontend behalten	58
2.3.2	In Discord den Überblick behalten – mit einem eigenen Server	62
2.3.3	Tipps und Tricks	64
2.4	Rechtliche Aspekte bezüglich KI-generierter Bilder	66
2.4.1	Das Recht am (eigenen) KI-Bild	67
2.4.2	Kennzeichnungspflicht	69

3 Die KI-Leinwand – Workshops für Einsteiger 73

3.1	Ein Bild für alle Fälle	73
3.1.1	Iteratives Prompten	73
3.1.2	One-Shot-Prompts	77
3.1.3	Der Aufbau komplexerer Prompts – 2-4-2-Takt	81
3.1.4	Fotorealistische Szenen – Werbefotograf*innen ersetzen	83
3.1.5	Auf den Spuren der großen Meister	89
3.1.6	Tempus fugit – Midjourneys Metamorphose	92
3.2	Unendliche Weiten und Charaktere – Landschaften und Porträts	94
3.2.1	Landschaftsfotografie und -malerei – the Sky's the Limit	94
3.2.2	Professionelle Porträts erstellen	100
3.2.3	Porträts mit speziellem Look	104
3.2.4	Du kommst hier nicht rein! Objekte per Prompt aus Szenen entfernen	108
3.3	Fingerspitzengefühl – Funktionen und Parameter gezielt anwenden	111
3.3.1	Von der Schnecke zum Turbo – Ideen kostengünstig oder schnell umsetzen	111
3.3.2	Formatfüllend – mit dem Seitenverhältnis spielen	114
3.3.3	Stilvoll – Stilzuweisungen anwenden	118
3.3.4	Von chaotisch bis verrückt – wie du Leben in deine Bilder bringst	121
3.3.5	Animes und Comics für alle – der Niji-Modus	125
3.3.6	Scharfe Sache – hochskalierte und hochaufgelöste Bilder erstellen	127

4 Midjourney für kreative Köpfe 131

4.1	Grafisch gestalten mit KI – Logos, Flyer und mehr	131
4.1.1	Eigene Logos und Icons – mit KI erstellt	131
4.1.2	Flyer und Einladungen	136
4.1.3	Professionelle Vektorgrafiken im Handumdrehen	139
4.1.4	Geistesblitze materialisieren – mit Sketchnotes	143
4.1.5	Die eigene Tapete und den eigenen Teppich designen	145
4.2	Themenfotografie	147
4.2.1	Naturfotografie	147
4.2.2	Food-Fotografie	150
4.2.3	Nahaufnahmen, Makros, Close-ups	152

4.3	Midjourney – stilvollendet	153
4.3.1	Stile mit Style Reference übertragen	153
4.3.2	Eigene Stile entwerfen – der Style Tuner	157
4.3.3	Genau mein Stil – Bildstile von Skizzen übertragen	159
4.4	Pimp my Picture – die KI mit eigenen Bildern füttern	161
4.4.1	Lieblingsbild reloaded	161
4.4.2	Potpourri – Mash-ups erstellen	168
4.4.3	Doppelgänger – der eigene Avatar aus der KI	171
4.4.4	Von der Skizze zum Bild – KI als Kunst-Booster	175
4.5	Abseits ausgetretener Pfade – Spaß mit Bild-KI	180
4.5.1	Historische Selfies und Zeitdokumente	180
4.5.2	Drama, Baby!	183
5	Midjourney ausreizen – Workshops für Fortgeschrittene	187
5.1	Der KI-Zauberstab – Bildelemente variieren und anpassen	187
5.1.1	Kaninchen aus dem Hut zaubern – Objekte ergänzen und umwandeln	187
5.1.2	Bilder aufmischen – mit dem Remix-Modus experimentieren	193
5.1.3	Hintergrundwissen – wie man den Hintergrund austauscht	198
5.1.4	Schwarz auf weiß – Schrift in Bildern einsetzen	200
5.2	Konsistente Bilder und Charaktere erzeugen	206
5.2.1	Die Harmonie der Ähnlichkeit – den Seed-Parameter nutzen	206
5.2.2	Eineiige Zwillinge – ähnliche Charaktere für das Storytelling generieren	208
5.2.3	Ein Kinderbuch illustrieren	213
5.3	Prompting für Fortgeschrittene	214
5.3.1	Der vollautomatische Promptgenerator	214
5.3.2	Tipps und Tricks zu Midjourney	221
5.3.3	Wissenswertes zum Webfrontend	231
5.3.4	Bildrepliken nahe am Original erstellen	234
5.4	Ein Füllhorn voller Prompts	235
5.4.1	Promptbörsen	235
5.4.2	Treffpunkt der Meister*innen – die Community	236
5.4.3	Tutorials und mehr	238
5.4.4	AI Imagelab	240

6 Masterclass – spezielle Werkzeuge 241

6.1	Als die Bilder laufen lernten – KI für Videoclips	241
6.1.1	Videoclips auf Zuruf	241
6.1.2	Der Ton macht die Musik – Videos per KI vertonen	246
6.2	Bildbearbeitung per KI aufrüsten	249
6.2.1	Firefly und Photoshop, das Dreamteam	249
6.2.2	Intelligente Nacharbeit – KI-Bilder reparieren und ergänzen	252
6.2.3	KI-generierte Texte mit Adobe Firefly	255
6.2.4	Eigene Bilder per KI erweitern	257
6.2.5	Skylum Luminar Neo – die Photoshop-Alternative	257
6.3	Alles in allem – Text- und Bild-Kombis	260
6.3.1	Bilder im Dialog erstellen – die multimodale KI in ChatGPT nutzen	260
6.3.2	ChatGPT – Custom GPTs zur Prompterstellung nutzen	263
6.3.3	Microsoft Copilot – der Konkurrent von ChatGPT und Midjourney	268
6.4	Stable Diffusion – eine Bild-KI lokal betreiben	271
6.4.1	Die eigene KI hosten	272
6.4.2	Stable Diffusion lokal installieren	273
6.4.3	Erste Schritte in Stable Diffusion	275

7 Neue Horizonte – die Zukunft von Midjourney & Co. 279

7.1	Chancen	279
7.2	Risiken und Fazit	282
7.2.1	Urheberrechtsverletzungen	283
7.2.2	Verbreitung von Falschinformationen	284
7.2.3	Auswirkung auf den Arbeitsmarkt, das soziale Umfeld und die Umwelt	287
7.2.4	Ethische und soziale Auswirkungen	288
7.2.5	Fazit und Ausblick	288

Anhang 291

A	Tabelle einiger One-Shot-Prompts	291
B	Glossar	297

Index	305
-------	-----

Midjourney & Co.

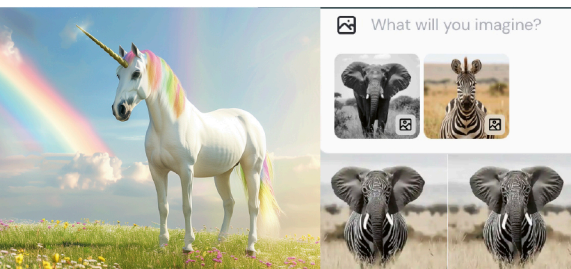
Wie du mit KI beeindruckende Bilder schaffst

Du willst faszinierende Kunstwerke erstellen und dich kreativ ausleben, aber Malen und Zeichnen sind nicht deine Stärken? Dann probier's mal mit KI!

Rainer Hattenhauer zeigt anhand vieler praktischer Beispiele, wie du mithilfe von Midjourney überzeugende Ergebnisse erzielst. Mit geschickten Prompts bearbeitest du (deine) Fotos, erstellst individuelle Grußkarten, humorvolle Bilder oder sogar Logos und Flyer für die Arbeit. Und wenn du einfach nur deiner Kreativität freien Lauf lassen willst, liegst du mit diesem Buch genauso richtig. Hier kannst du ohne Vorkenntnisse mit Midjourney & Co. in die Welt der KI-Bilder eintauchen.

Aus dem Inhalt

- + Am Anfang steht der Prompt
- + Oberfläche und Befehle
- + Workshops für den Einstieg
- + Logos, Flyer und mehr
- + Themenfotografie
- + Eigene Fotos bearbeiten
- + Fortgeschrittene Tipps und Techniken
- + KI für Videoclips
- + Eine Bild-KI lokal betreiben



»Eine faszinierende Reise durch die Welt der KI-Kunst, die deine Kreativität entfesselt und deine Vorstellungskraft neu definiert.«



Rainer Hattenhauer ist promovierter Physiker und unterrichtet Physik, Mathematik und Informatik an einem Gymnasium. Neben seiner schulischen Tätigkeit ist der Technik-Nerd seit mehr als zwei Jahrzehnten Autor aus Leidenschaft. Das zeigt sich auch in seiner Begeisterung für künstliche Intelligenz und deren kreativen Einsatz.

