

Künstliche Intelligenz und Urheberrecht

Urheberrechtliche Rahmenbedingungen für Anbieter und Nutzer
generativer KI-Systeme

Tagung für Informatik und Recht, 29. August 2023, Bern
Referent: Matthias Städeli, Rechtsanwalt, Partner, RENTSCH PARTNER,
Zürich (staedeli@rentschpartner.ch)

Präsentation

Aufbau

1. Über was man heute spricht?
2. Was versteht man unter generativer KI
3. Urheberrechtliche Grundlagen
4. Aktuelle Beispiele generativer KI
5. Rechtliche Einordnung – Use Cases
6. Ausblick und Regulierung von KI?

Künstliche Intelligenz und Urheberrecht

Fragen und worüber man heute spricht: Einordnung

Headlines

- NZZ vom 10.8.2023: Wer darf mit KI-Stimmen von Musikstars Geld verdienen? Google und Universal verhandeln über eine Lösung
- SRF vom 11.7.2023: Happiger Vorwurf an ChatGPT: Datenklau zu «Trainingszwecken»?
- NZZ vom 8.12.2022: Künstliche Intelligenz kreiert Bilder auf Knopfdruck: Digitalkünstler fürchten um ihre Arbeit
- NZZ vom 9.12.2022: Diese künstliche Intelligenz kann Verse dichten und Programmcode schreiben: Was steckt hinter Chat GPT?

Was versteht man unter generativer KI

Generative KI (General Purpose AI, GPAI)

Unter generativer KI versteht man ein System, das einen Input («prompt») entgegennimmt in ein trainiertes und allenfalls optimiertes Modell (auf der Basis von Trainingsdaten) das durch Software orchestriert wird, die mit dem Vorstehenden einen Output (Text, Bilder, Musik, Videos, Games, 3D, Code etc.) generiert und nicht nur Daten anzeigt (Christian Laux, Entwicklungen im IT-Recht, INGRES Tagung vom 04.07.2023)



Technische Einordnung – Input – Maschine - Output

Mustererkennung und Maschinelles Lernen zur Erzeugung von artifiziellen Erzeugnissen

Maschinelles Lernen wird zum Beispiel durch künstliche neuronale Netzwerke ermöglicht, die vom menschlichen Hirn inspiriert wurden. Ein **neuronales Netzwerk umfasst eine Vielzahl miteinander verbundenen Schichten («layer»)**, wobei jeder Knoten Parameter aufweist, die während des Trainings automatisch angepasst werden.

Beim Training kommen Trainingsdaten und ein Trainingsalgorithmus zum Einsatz. Die Parameter werden während des Trainings iterativ angepasst, so dass eine erzeugte Ausgabe des Netzes immer besser mit einer gewünschten Ausgabe übereinstimmt.

Im Rahmen der KI dienen **Algorithmen** (eindeutige Handlungsvorschrift zur Lösung eines formal beschriebenen Problems, das von einer Maschine selber ausgeführt werden kann) dazu, **KI-Systeme (neuronale Netzwerke) zum Lernen zu bringen**. Die künstlichen, neuronalen Netzwerke werden mit einer grossen Menge von Daten trainiert.

Für Funktionsweise eines Künstlichen Neuronalen Netzwerkes (AI-Systems) braucht es eine

- Schicht mit Neuronen zur Dateneingabe (Input Layer)
- eine Schicht mit Neuronen zur Datenausgabe (Output Layer)
- Dazwischen liegen unterschiedlich viele verdeckte Schichten (Hidden Layers).

Existierende Modelle (z.B. Generative Adversarial Networks (GAN) und Diffusionsmodelle unterscheiden sich hinsichtlich ihrer Architektur und ihres Trainings.

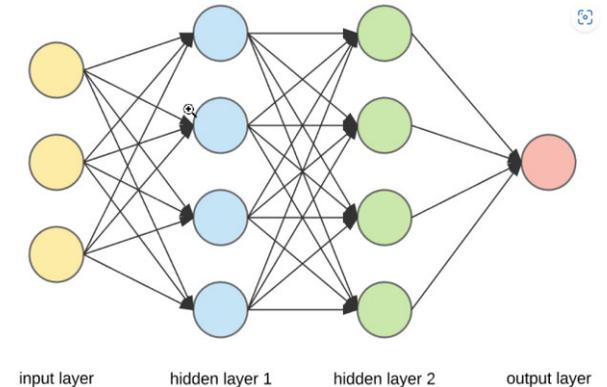


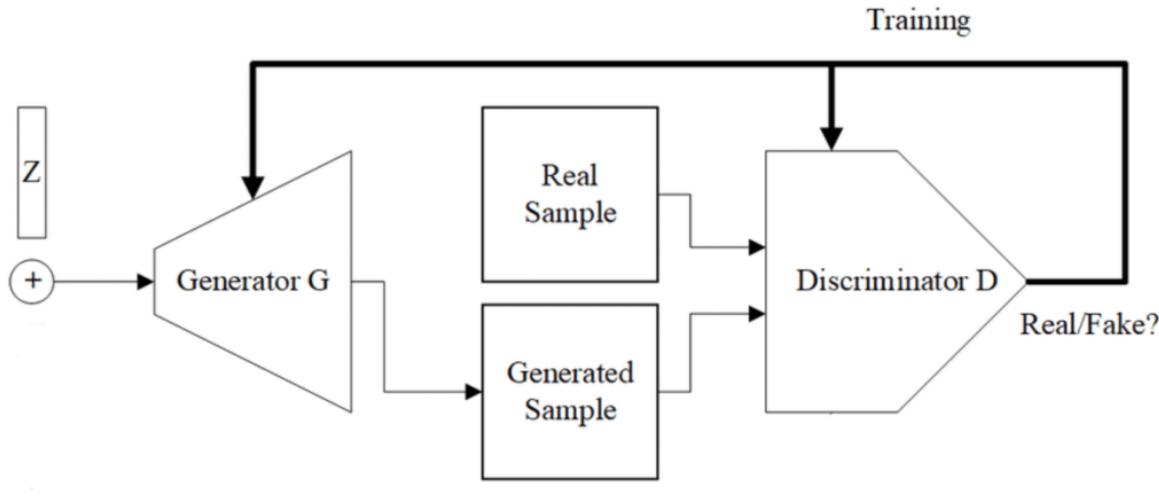
Image 1: Neural Network Architecture

Quelle: Gavril Ognjanovski, Towards Data Science

Technische Einordnung

Generative KI und Training der Algorithmen

Beispiel Generative Adversarial Networks (GAN): Oft werden beim Training eines sog. «Generators» auch zusätzliche neuronale Netzwerke eingesetzt, sog. «Discriminators», um die Parameter des Generators durch Feedback anzupassen. Der Generator versucht, den Diskriminator zu täuschen. Das generative Modell lernt indirekt, indem es als Zielfunktion hat mit seinen synthetischen Ausgaben das Diskriminator Modell zu überlisten, d.h. zu Fehlentscheidungen zu verleiten. Beide Netzwerke werden nach und nach in ihrer Aufgabe besser. Der Generator versucht, den Diskriminator zu täuschen. Beide Netzwerke werden nach und nach in ihrer Aufgabe besser. Der Generator kann anschliessend in der Lage sein, täuschend echte Ergebnisse (Bilder, Texte, Tonaufnahmen, Code, etc.) zu erzeugen.



Urheberrechtliche Grundlagen

Rechtliche Einordnung

Werkbegriff gemäss Art. 2 URG

1. Werke der **Literatur und Kunst** (sehr weiter Begriff), **Computerprogramme** und **fotografische Wiedergaben ohne individuellen Charakter**.
2. Das Werk beruht auf einer **geistigen Schöpfung** (*d.h. Äusserung gedanklicher Tätigkeit eines Menschen, BGE 130 III 172*). Mitumfasst sind Werke unter *bewusster Verwendung von Zufallsprinzipien (zufallsgenerierte Computerkunst, bewusste Auswahl und Präsentation)*
3. Das Werk verfügt über Individualität (**Werkindividualität, d.h. individueller Charakter muss im Werk selber erkennbar sein, Wertungsfrage**).
 - a. Handelt es sich um ein Original oder um eine Kopie eines bestehenden Werks?
 - b. Im Rahmen eines Gedankenexperiments ist zu prüfen, ob das Werk so individuell ist, dass es als sehr unwahrscheinlich erscheint, dass jemand ein identisches Werk unabhängig davon schaffen könnte (statistische Einmaligkeit). Gestalterische Entscheidungen sind zu berücksichtigen, welche im Werk selbst wahrnehmbar sind
 - c. An den Grad der Individualität sind geringere Ansprüche zu stellen, wo der Gestaltungsspielraum klein ist (BGE 125 III 328)

Rechtliche Einordnung

Ausschliesslichkeitsrechte des Urhebers am Werk (Art. 10 URG)

- Der Urheber oder die Urheberin hat das **ausschliessliche Recht zu bestimmen, ob, wann und wie das Werk verwendet wird**. Diese umfassen insbesondere das Herstellungs-, Verbreitungs-, Aufführungs- und **Vervielfältigungsrecht** sowie das Recht zu ändern oder zu bearbeiten.
- Der Urheber kann jedermann verbieten das Werk ohne seine Zustimmung zu verwenden. Eine teilweise Vervielfältigung bei **Bearbeitungen oder Änderungen** des Werks bedarf .
- Für gewisse Nutzungen hat der Gesetzgeber **Schranken** festgelegt (Verwendung für Eigengebrauch etc.)

Rechtliche Einordnung

Schutzdauer (Art. 10 URG)

- Ein Werk ist urheberrechtlich geschützt, sobald es geschaffen ist, unabhängig davon, ob es auf einem Träger festgehalten ist oder nicht.
- Der Schutz erlischt [...] **70 Jahre nach dem Tod** des Urhebers oder der Urheberin (Ausnahme Computerprogrammen und Fotografien ohne individuellen Charakter)

Schranken bei Data Mining

Wissenschaftsschranke in CH – Schranke in DE gemäss § 44 UrhG (Datamining- Analyse der Mustererkennung)

- In CH gelten allgemeine Regeln des Urheberrechts Vervielfältigung nur mit Zustimmung Urheber. Ausnahme in **Art. 24d URG** wird im **Interesse der Forschung sichergestellt, dass Urheberinnen und Urheber das für das Text- und Data-Mining notwendige Kopieren nicht verbieten dürfen** (Wissenschaftsschranke).
- **Weitergehende Schranke in EU/DE:** Gesetzlich ist der § 44b DE- UrhG (**Data Mining**) die Erlaubnis — die eingeführt wurde, weil Big Data Analysen von urheberrechtlich geschützten Inhalten sonst faktisch nicht möglich sind. Vorbehalten **Opt-out**.
- **Damit können in der EU urheberrechtlich geschützte Trainingsdaten (wie Bilder aus dem Internet) lizenzfrei für das Machine Learning verwendet werden. Die Trainingsdaten müssen nach Abschluss des Trainings lediglich gelöscht werden.** Die Urheber haben jedoch die Möglichkeit eines Opt-out, wenn sie ihre Werke nicht für ein Training zur Verfügung stellen möchten.

Anwendungs-Beispiele von durch KI-Systemen generierten Erzeugnisse

Anwendung von Generativer KI zur Erstellung von Text, Bildern, Videos und anderer artifizieller Erzeugnisse

Beispiele für KI-generierte Gemälde

Input: Text Prompt «self portrait by Vincent van Gogh, intense colors»

Output:



Anwendung von Generativer KI bei Erstellung artifiziieller Erzeugnisse

Beispiele für KI-generierte fotorealistische Bilder



Midjourney Version Comparison



Siegerbild von Boris Eldagsen: Sony World Photography Awards 2023

Anwendung von Generativer KI bei Erstellung von Artifiziieller Erzeugnisse

Beispiel Kurzfilme

Exclusive: Watch the world premiere of the AI-generated short film The Frost. | MIT Technology Review



- KI-generierte Videos (Musikvideos, Werbevideos, Assistenten)
- Ausschliesslich erstellt mit nachfolgenden KI-Generatoren

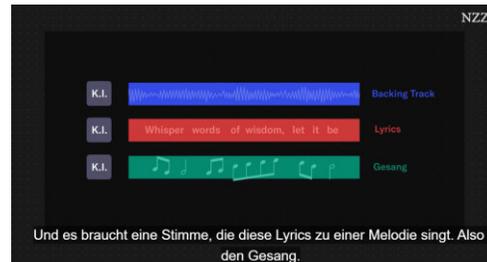


Anwendung von Generativer KI bei Erstellung von Artifiziieller Erzeugnisse

Beispiel Musiksong (voice cloning)

- Mit Hilfe von AI kann ein **neuer Beatles Song** erstellt werden:

<https://www.nzz.ch/video/with-a-little-help-from-ai-kann-kuenstliche-intelligenz-wirklich-die-beatles-wiederaufleben-lassen-Id.1751603>



- **"Heart on My Sleeve"** ist ein Lied, das vom TikTok-Nutzer ghostwriter977 wurde von einer künstlichen Intelligenz (KI) generiert, um **wie die kanadischen Musiker Drake und The Weeknd zu klingen**. Der Song hatte auf TikTok Aufmerksamkeit erregt und Millionen von Aufrufen auf einer Vielzahl von Plattformen erhalten.



Urheberrechtliche Themen – «Use cases»

Künstliche Intelligenz und Urheberrecht

Was ist aus urheberrechtlicher Sicht zu beachten?

Themen

- Was muss bei der **Eingabe (prompts)** von Texten, Bildern, Code, Stimmen, Musik **in KI-Systemen** aus rechtlicher Sicht beachtet werden (**Input**)?
- Darf der Systembetreiber von Generativer KI urheberrechtlich geschützte Werke für das **Training** seiner **Algorithmen** nutzen (Input)?
- Stellen mit **KI generierte Erzeugnisse urheberrechtlich geschützte Werke** dar (Output)?
- **Wem gehören die artifiziellen Erzeugnisse**, die durch KI-Systeme «geschaffen» werden? Wer ist **Urheber bzw. Erfinder**?
- Können **KI-Systeme rechtlich geschützt** werden (Maschinen) ?

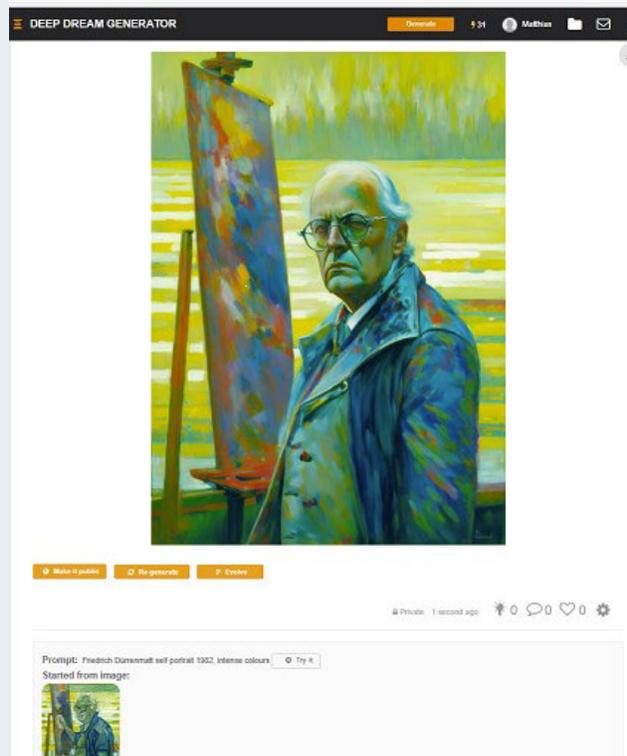
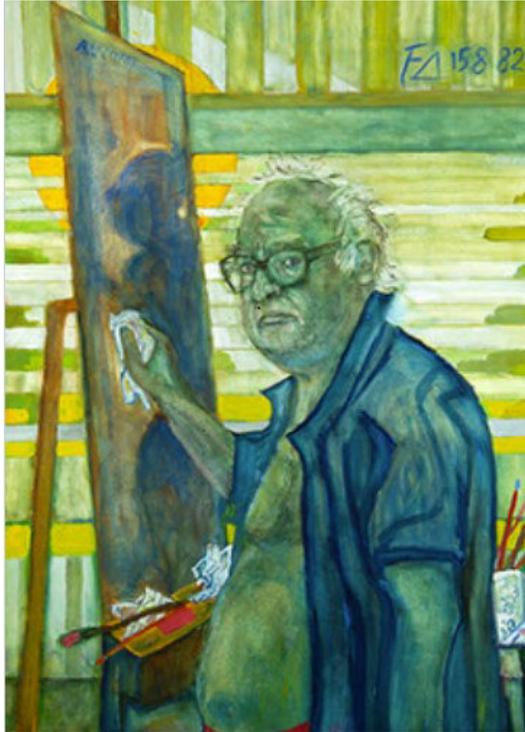
Eigene Prompts als urheberrechtlich geschützte Werke?

Input

- **Prompts: Eigene Texteingaben, die die Erstellung eines Bildes oder Musikstücks (z.B. Metas AI Music Generator) steuern sollen.** Z.B: «Pop dance track with catchy melodies, tropical percussion, and upbeat rhythms, perfect for the beach». Prompt für sich kann, soweit er über Individualität verfügt und eine menschliche Schöpfung ist, als Text urheberrechtlich geschützt sein
- **Prompts: Verwendung fremder Inhalte, Rechte Dritter beachten**

Prompts- Verwendung Werke Dritter

Input – Recht Dritter beachten



Erzeugt durch Deep Dream Generator
Prompt: Bildeingabe mit Text «Friedrich Dürrenmatt, self portrait 1982»

F. Dürrenmatt, † 14. Dezember 1990

Verwendung eines urheberrechtlich geschützten Gemäldes als Input ist eine unzulässige Vervielfältigungshandlung und Bearbeitung.

Vorbehalten: Schranke des Eigengebrauchs?
Speicherung auf kommerzieller Plattform und Veröffentlichung in Web-Galerie (public space) fällt nicht darunter.

Verwendung von Bildern für Trainingsdaten

Input bei KI-Modellen und Training der KI

Beim Training eines generativen KI-Modells werden dem Netzwerk iterativ die Trainingsdaten präsentiert und seine Parameter vom Trainingsalgorithmus so angepasst, dass die Differenz zwischen der generierten Ausgabe und der gewünschten Ausgabe minimiert wird. *In jeder Iteration wird die Güte des Modells anhand einer vorgegebenen Zielfunktion gemessen. Dann wird für jeden Parameter einzeln ermittelt, ob er zur Verbesserung der Zielerreichung erhöht oder verringert werden muss und die Parameter werden angepasst.*

Existierende Modelle unterscheiden sich hinsichtlich ihrer Architektur und ihres Trainings:

- a. **Generative Adversarial Networks (GAN)**: Oft werden beim Training eines sog. «Generators» auch zusätzliche neuronale Netzwerke eingesetzt, sog. «Discriminators», um die Parameter des Generators durch Feedback anzupassen. **Der Generator versucht, den Diskriminator zu täuschen. Das generative Modell lernt indirekt, indem es als Zielfunktion hat mit seinen synthetischen Ausgaben das Diskriminator Modell zu überlisten**, d.h. zu Fehlentscheidungen zu verleiten. Beide Netzwerke werden nach und nach in ihrer Aufgabe besser. Der Generator kann anschliessend in der Lage sein, täuschend echte Ergebnisse (Bilder, Texte, Tonaufnahmen, Code, etc.) zu erzeugen. **Trainingsbilder werden benötigt um sie mit synthetischen Bildern zu mischen. Informationen aus den Trainingsdaten diffundieren indirekt in die Parameter des Diskriminator Modells, um anschliessend in die Parameter des Generator überzugehen.**
- b. **Diffusionsmodelle**: Es werden keine kompletten Bilder synthetisiert, sondern granulare Arbeit an Bildteilen bis hin zu Pixeln. **Arbeit erfolgt, indem aus einem gestörten Eingabebild, ein optisch weniger beschädigtes Ausgabebild erzeugt wird. Z.B. durch Überlagerung von Rauschmustern.** Zielfunktion ist dann die möglichst exakte Reproduktion der gestörten bzw. weggelassenen Bildteile. Diffusionmodelle erlernen ihre Parameter direkt aus den Trainingsbildern.
- c. **Bei keinem Modell werden die Trainingsbilder systematisch in die Parameter der Modelle übertragen**, weder direkt noch indirekt bzw. gekürzt (anders als bei Videodatenkompression wie JPEG). Nur in einzelnen Fällen ist eine Rekonstruktion der Trainingsbilder möglich.

Getty Images/ Stability AI - Klage

Input: Verwendung von Trainingsdaten und Marken Dritter



Klage wegen Urheberrechts-, Marken und Wettbewerbsverletzung und macht Schadenersatz und Gewinnherausgabe

1. Training mit riesigen Datenmengen (12 Millionen urheberrechtlich geschützte Bilder, Bild-/Textkombinationen).
2. Trainingsdatensätze setzen sich aus öffentlich im Internet verfügbaren Daten zusammen, die durch **Web-Scraping** automatisiert heruntergeladen werden.
3. Stability AI hat angeblich die **Wasserzeichen** von Getty Images so gut wie möglich von den Trainingsdaten entfernt. Gleichzeitig erzeugt das Stable-Diffusion-Modell jedoch, häufig eine modifizierte Version des Wasserzeichens von Getty Images



Getty Images/ Stability AI – Klage vom 03.02.2023

Trainingsdaten heruntergeladen durch Web-Scraping

- Urheberrechtlich **massgebend, ob die verkörperten Bilder menschlichen Sinnen wahrnehmbar gemacht** werden können.
- Trainingsbilder werden **nicht systematisch in die Parameter der KI-System übertragen**, auch nicht in kompromierter Form
- **Mangels Verkörperung, keine Vervielfältigung**
- Zentral ist die Frage, ob das KI-Modell von Stability AI Vervielfältigungsstücke der Bilder aus den Trainingsdaten enthalten oder erzeugen können. **Aktuell Gegenstand des Verfahrens.**
- In EU: Rechtfertigung durch Data-Mining Schranke?

Urheberrecht an artifiziellen KI-Erzeugnissen?

Output: Menschliche Schöpfung aufgrund der Gestaltung der einzelnen Schritte fällt u.U. unter Urheberrecht

- Ist Input individuell und im Output Individualität erkennbar (Werkindividualität) kann Urheberrecht beim Benutzer entstehenn
- Schaffung eines Werkes mit KI-Tool als technischem Hilfsmittel für Output (wie Fotoapparat)
 - Auswahl der Vorgaben durch den Text (Prompt)
 - Konfigurationseinstellungen: So können per Prompt für fotorealistische Ergebnisse etwa Vorgaben wie ein Kameramodell, das zu verwendende Objektiv, Lichtverhältnisse und die Tiefenschärfe gemacht werden, die das Ergebnis erheblich beeinflussen.
 - Auch die Reihenfolge der in dem Prompt verwendeten Begriffe spielt für die Gewichtung und damit das Ergebnis eine Rolle.
 - Zudem können mithilfe „negativer Prompts“ unerwünschte Effekte oder Gestaltungen unterbunden werden. Entsprechend stark oder schwach finden sich die Vorgaben aus dem Prompt in dem Bild wieder.

Urheberrecht an artifiziellen KI-Erzeugnissen

Auswahl verschiedener individueller Schritte, welche zu einem Urheberrechtsschutz des Menschen führen kann

Exclusive: Watch the world premiere of the AI-generated short film The Frost. | MIT Technology Review

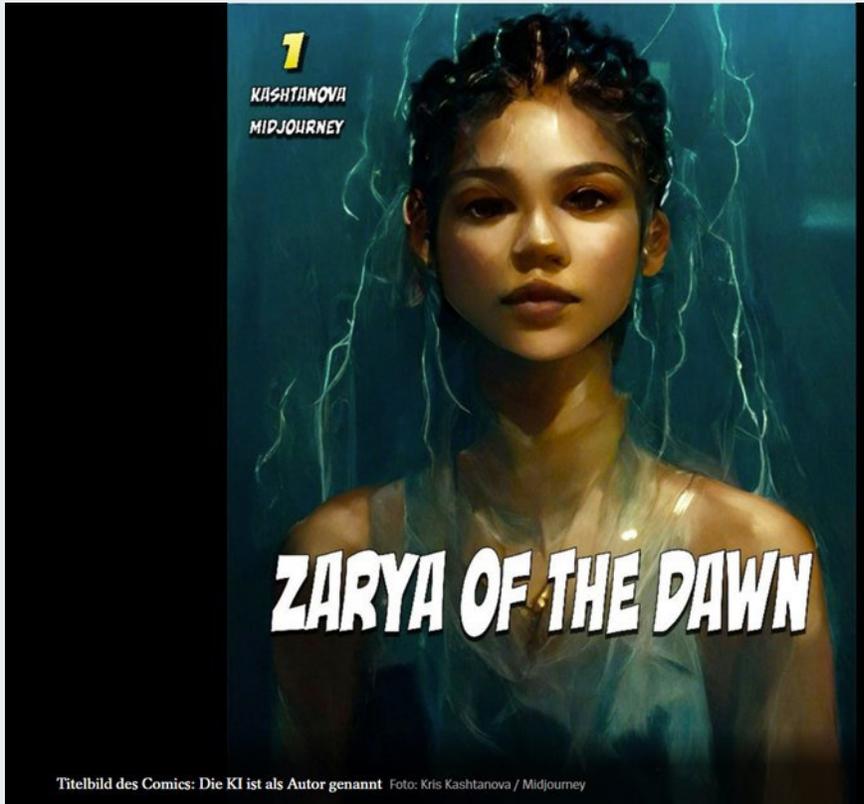


- KI-generierte Videos (Musikvideos, Werbevideos, Assistenten)
- Ausschliesslich erstellt mit nachfolgenden KI-Generatoren



Urheberrecht an artifiziellen KI-Erzeugnissen

US-Copyright Office / Kashtanova: ZARYA OF THE DAWN



- Kashtanova hat Texte der Graphic Novel vollständig selbst verfasst und **Mid-journey als Werkzeug für das Erstellen der Bilder** verwendet
- Ähnlich wie ein Fotograf eine Szene gestalten könne und die Kamera als ein Werkzeug zur Abbildung verwendet werde, sei die Bildgenerierung durch Midjourney von der Künstlerin in einem aufwendigen und iterativen Prozess gestaltet worden.
- **Am 21. Februar 2023 hat das U.S. Copyright Office den Urheberrechtsschutz für die Bildergeschichte bestätigt und dabei deutlich gemacht, dass sowohl der Text als auch die Anordnung von Bild- und Textelementen schutzfähig sind, nicht aber die einzelnen, mithilfe von Midjourney erzeugten Bilder.**

Die Bildergeschichte ist damit der erste öffentlich bekannt gewordene Fall, in dem KI-Kunst explizit Urheberrechtsschutz zugestanden wurde.

Maschine – rechtliche Fragen

Schutz von KI-Systeme

Mögliche Schutzstrategien

Im Rahmen des bestehenden Immaterialgüterrechtsschutz bestehen folgende Schutzmöglichkeiten:

- Teilweiser urheberrechtlicher Schutz von Software, die bei der Architektur benutzt wird
- Patentschutz von computerimplementierten Erfindungen
- Schutz durch Verträge (Nutzungsbedingungen, Lizenzverträge)
- Datenbankschutz

Zusammenfassung

Rechte Dritter und Verträge zu beachten

- **Beachten von Rechten Dritter an Input-Daten und Output-Daten: niemanden Schaden zufügen.**
 - Urheberrecht, Markenrecht, Patentrecht, Designrecht
 - Geheimnisse der Arbeitgeberin, Know-how
 - Datenschutz (Personendaten wie z.B. Stimmen)
 - Vertragliche Bestimmungen (Lizenzbestimmungen des KI-Anbieters)
- **Sich selber keinen Schaden zufügen**
 - Vertragliche Geheimhaltung, Weisungen Arbeitgeberin

Ausblick und Regulierung von KI?

Ausblick und Regulierung von KI

Urheberrecht

- **Aktuell aufgrund nationaler Gesetze uneinheitliche Rechtslage (z.B. Verwendung von Trainingsdaten)**
- **Das geltende Recht enthält flächendeckende Haftungsvorschriften für Schäden, die bei Verletzung von Urheberrechten und beim Einsatz technischer Geräte verursacht werden (Technologieneutralität). Regelungsbedarf?**
- **Internationale Vereinheitlichung für immaterialgüterrechtliche Einbettung artifizieller Erzeugnisse sinnvoll?**
 - Interdisziplinärer Diskurs (Industrie, Kunstschaffen, Politik) nötig!
 - Heutige (schwache) KI-Systeme können immer noch (weitestgehend) als technische Hilfsmittel erfasst werden. Anwendung des bestehenden Urheberrechts möglich
 - Ist die Einführung eines neuen «verwandten Schutzrechts» für von starker KI erzeugten Produkten in der Zukunft nötig (z.B.: «Promptografie»). Einführung verwandter Schutzrechte?

KI- und Urheberrecht

Take-home messages

- **Input/Prompts:** Rechte Dritter wie Urheberrechte, Markenrechte, Persönlichkeitsrechte und vertragliche Rechte (z.B. Geschäftsgeheimnisse zu beachten)
- **Input/Trainingsdaten:** Soweit nicht in Parametern der KI-Systeme verkörpert, liegt keine urheberrechtliche Vervielfältigung vor.
- **Output:** Nur Menschen können Urheberrechte erwerben. Fraglich ob bei KI generierten Erzeugnissen. Beachten von Geschäftsgeheimnissen und Urheberrechten Dritter. Urheberrecht aufgrund verschiedener Gestaltungsschritte von Menschen möglich.
- **KI-Systeme:** Können teilweise über Patentschutz verfügen oder vertraglich abgesichert werden
- **Rechtslage international uneinheitlich - Internationaler Regelungsbedarf?**
 - sinnvoll, neue Vorschriften zu schaffen, die nur für Anwendungen mit Künstliche Intelligenz gelten? Z.B. Haftung?
 - Interessenabwägung Urheberrechtsschutz v.s. Ermöglichen von Innovation durch neue Technologien
 - KI-Urheberschaft im Urheberrecht?
 - Schaffung von «sui generis-Rechten» zum Schutz innovativer Erzeugnisse (z.B. Promptografie)

Fragen?

Matthias Städeli
RENTSCH PARTNER AG
Kirchenweg 8
8008 Zürich
staedeli@rentschpartner.ch