



Was ist generative KI – und was kann sie leisten?



KONTOR4 GmbH

Was ist generative KI – und was kann sie leisten?

Generative künstliche Intelligenz ist eine Art von künstlicher Intelligenz (KI), die Bilder erstellen kann, die einer Reihe von vorgegebenen Wörtern entsprechen. Diese Art von KI hat in den letzten Jahren die Aufmerksamkeit von Technologiegiganten und Risikokapitalfirmen im Silicon Valley auf sich gezogen. Analysten erwarten, dass diese Technologie in einer Vielzahl von Branchen zum Einsatz kommen wird.

Und das trotz der Tatsache, dass die von diesen Computerprogrammen erzeugten Bilder nicht perfekt sind: An Händen zum Beispiel erscheinen zusätzliche Finger, unnatürliche Biegungen der Gliedmaßen und so weiter.

Gleichzeitig stoßen die Bildgeneratoren auch auf Probleme bei der Verarbeitung von Text, z. B. bei der Erzeugung nichtssagender Symbole. Aber diese bildgenerierenden Programme könnten der Beginn eines Technologiebooms sein. "Der Begriff 'generative KI' ist in den letzten drei Monaten zu einem Modewort geworden.

Die KI-Technologie hat in den letzten Jahren große Fortschritte gemacht. Dies gilt insbesondere für das Verständnis und die Verarbeitung von Daten. KI-Modelle und die dahinterstehenden Algorithmen können jetzt erkennen, ob auf einem Foto eine Katze zu sehen ist, und sie können auch jeden Tag Milliarden von Suchergebnissen für Google generieren. Generative KI-Modelle gehen jedoch noch einen Schritt weiter, indem sie Dinge erschaffen, die es vorher nicht gab.

Was ist generative KI – und was kann sie leisten?

Wie generative künstliche Intelligenz funktioniert

Die Bildgenerierung verwendet Techniken aus einer Teilmenge des maschinellen Lernens, die als Deep Learning bezeichnet wird. Deep Learning hat die meisten Fortschritte in der KI vorangetrieben - ein Großteil davon begann im Jahr 2012, als ein wegweisendes [Papier aus dem Jahr 2012 über die Bildklassifizierung](#) das Interesse an der Technologie neu entfacht hat. Deep Learning verwendet Modelle, die an großen Datensätzen trainiert werden, bis das Programm die Beziehungen in diesen Daten versteht. Dieses Modell kann dann in Anwendungen verwendet werden, um beispielsweise zu erkennen, ob sich ein Hund auf einem Bild befindet, oder um Text zu übersetzen.

Bildgeneratoren arbeiten, indem sie diesen Prozess umkehren. Anstatt Englisch ins Französische zu übersetzen, übersetzen sie englische Sätze in Bilder. Sie bestehen normalerweise aus zwei Hauptteilen: Einem, der den Anfangssatz verarbeitet, und einem anderen, der die Daten in Bilder umwandelt.

Der erste Teil, Generative künstliche Intelligenz, basiert auf einer Methode namens Generative Adversarial Networks (kurz GAN). Früher wurden diese GANs oft verwendet, um Fotos von Personen zu erstellen, die nicht existierten. Im Wesentlichen arbeiten sie, indem sie zwei KI-Modelle gegeneinander ausspielen, um bessere Bilder zu erstellen, die ein vorgegebenes Ziel erreichen. Einen Überblick über die Funktionsweise einer solchen GAN-Struktur gibt es u.a. bei Google: https://developers.google.com/machine-learning/gan/gan_structure

Was ist generative KI – und was kann sie leisten?

Neuere Methoden verwenden jedoch häufig Transformatoren; ein Konzept, das erstmals 2017 in einem Papier von Google vorgeschlagen wurde. Transformatoren nutzen viel größere Datensätze – die Kosten für das Trainieren dieser Technik belaufen sich allerdings auf Millionen von Dollar. Bekannteste Vertreter dieser Transformatoren sind beispielsweise GPT-3, LaMDA und Wu-Dao. Letzterer ist übrigens der derzeit größte seiner Art: 10x größer als GPT-3. Nicht unwichtig zu erwähnen, dass es sich um eine chinesische Entwicklung handelt – und dass weder Dokumentationen noch Schnittstellen zur Verfügung stehen. Zumindest nicht für die westliche Welt ...

Bekannte Bildgeneratoren

Der erste Bildgenerator, der viel Aufmerksamkeit erregte, war [Dall-E](#), ein 2021-Projekt des Silicon-Valley-Startups OpenAI. OpenAI hat dieses Jahr ein leistungsfähigeres Update veröffentlicht.

Ein weiterer beliebter KI-basierter Bildgenerator ist [Craiyon](#), früher bekannt als Dall-E Mini, der online verfügbar ist. Nachdem ein Benutzer einen Satz eingegeben hat, kann er die resultierende Zeichnung innerhalb von Minuten im Browser sehen. Seit dem Start im Juli 2021 generiert Craiyon nun etwa 10 Millionen Bilder pro Tag, insgesamt 1 Milliarde nie zuvor gesehene Bilder, so Boris Dayma, Schöpfer von [Craiyon Productive AI](#), einer Plattform für KI und maschinelles Lernen. Von ihm stammt die Aussage:

„Das Beeindruckendste ist, dass generative KI auch neue Dinge schaffen kann. Sie schaffen nicht nur ähnliche alte Dinge. Mit Bildern kann man Neues schaffen. Dinge, die völlig anders sind als das, was bisher gemacht wurde.“

Was ist generative KI – und was kann sie leisten?

Nachdem die Nutzung Anfang dieses Jahres sprunghaft angestiegen ist, geht Daima mit Craiyon All-In. Er sagte, er habe sich auf die Verwendung von Anzeigen konzentriert, um die Benutzer wegen der hohen Serverkosten der Website frei zu halten. Der Twitter-Account [“craiyon”](#) der sich dem Posten der verrücktesten und kreativsten Bilder verschrieben hat, hat mehr als 1 Million Follower.

Aber das Projekt, das am meisten Begeisterung auslöste, war [Stable Diffusion](#), das im August dieses Jahres der Öffentlichkeit vorgestellt wurde. Der Quellcode ist auf GitHub verfügbar und kann auf einem Computer, in der Cloud oder über eine Programmierschnittstelle ausgeführt werden. Dadurch können Anwender den Programmcode an ihre eigenen Zwecke anpassen oder neue Programme darauf aufbauen.

Stable Diffusion ist beispielsweise über ein Plug-in in Adobe Photoshop integriert, mit dem Benutzer Hintergründe und andere Teile eines Bildes generieren können, die sie dann direkt in der App mit Ebenen und anderen Photoshop-Tools bearbeiten können.

Der Programmierer und Künstler Christian Cantrell hat das Plugin entwickelt. Das Plugin sei zehntausendfach heruntergeladen worden. Künstler sagten ihm, dass sie es an unzähligen Orten verwendet hätten, die er nicht erwartet hatte, wie zum Beispiel bei der Animation von Godzilla oder beim Erstellen von Bildern von Spider-Man in jeder Pose, die sich der Künstler vorstellen konnte.

Weil es manchmal schwierig ist, die richtigen Worte zu finden, um ein Bild zu erstellen, gibt es die Suchmaschine [Lexica](#). Dort werden Bilder von Stable Diffusion mit der genauen Wortfolge verknüpft, die zu ihrer Generierung verwendet wurden. Auf Plattformen wie Reddit und Discord werden Tricks genannt, wie man Menschen anleitet, Sätze einzugeben, die sie zum Generieren von Bildern verwenden möchten.

Was ist generative KI – und was kann sie leisten?

The next big thing?

Viele Investoren sehen in der **generativen künstlichen Intelligenz** eine potenziell transformative Technologie, wie das Smartphone oder die Anfänge des Internets. Durch diesen Wandel wird der potenzielle Markt, der diese Technologie nutzen könnte, erheblich vergrößert. Start-ups, Cloud-Service-Anbieter und Chiphersteller könnten von dieser neuen Technologie profitieren.

Cantrell argumentiert, dass die **generative künstliche Intelligenz** einer grundlegenden Technologie namens Datenbank ähnelt. „Generative KI ist ein bisschen wie eine Datenbank“, sagte er, „und Datenbanken helfen dabei, das enorme Potenzial von Anwendungen freizusetzen. Fast jede Anwendung, die wir jemals in unserem Leben verwendet haben, baut auf einer Datenbank auf, aber niemanden interessiert es, wie eine Datenbank funktioniert, sie wissen nur, wie man sie benutzt.“

Einige Fachleute sind der Meinung, dass die Start-ups, die diese neuen Technologien vorantreiben, schließlich die Softwaregiganten herausfordern könnten, die derzeit die KI beherrschen. Dazu gehören Google, die Facebook-Muttergesellschaft Meta und Microsoft.

Clement Delangue, CEO von Hugging Face, sagte: „Es wird eine große Anzahl neuer Unternehmen im Wert von Billionen Dollar geben, und diese Startups werden auf dieser neuen Technologie basieren.“ Entwicklerplattformen ähnlich wie GitHub, die vor trainierte KI-Modelle hosten, einschließlich Craiyon und Stable Diffusion. Ziel ist es, Programmierern den Aufbau von KI-Technologie zu erleichtern.

Was ist generative KI – und was kann sie leisten?

Auch Cloud-Dienstleister wie Amazon, Microsoft und Google könnten profitieren, da generative KI eine rechenintensive Technologie ist. Meta und Google haben einige der besten Köpfe auf diesem Gebiet eingestellt, um diese fortschrittliche Technologie in die Produkte des Unternehmens zu integrieren. Im September kündigte Meta eine KI-Initiative namens „[Make-A-Video](#)“ an, die die Technologie auf die nächste Stufe heben wird, indem Videos statt nur Bilder generiert werden.

Meta-CEO Mark Zuckerberg postete auf seiner Facebook-Seite: „Dies ist ein erstaunlicher Fortschritt. Ein Video zu erstellen ist viel schwieriger als ein Foto zu erstellen, da das System nicht nur jedes Pixel richtig machen muss, sondern auch vorhersagen muss, wie es sich im Laufe der Zeit ändern wird.“ Kürzlich veröffentlichte Google auch Code für ein Programm namens [Phenaki](#), das Text in minutenlange Videos umwandeln kann. Um mit [KI Video erstellen](#) zu können – kann man auch schon auf Lösungen zurückgreifen, die komplett online funktionieren. In dem verlinkten Blogbeitrag stellen wir diese Lösung näher vor.

Der Ansturm könnte auch Chipherstellern wie Nvidia, AMD und Intel Auftrieb geben, deren Grafikprozessoren ideal für das Training und den Einsatz von KI-Modellen sind. Jensen Huang, Chief Executive von Nvidia, hob kürzlich die generative KI als Schlüsselanwendung für die neuesten Chips des Unternehmens hervor und sagte, dass solche Technologien die Kommunikation bald revolutionieren könnten.

Was ist generative KI – und was kann sie leisten?

Ethische Fragen

Es gibt einige ethische Probleme mit computergenerierten Bildern. Das erste ist die Frage der Arbeitsplätze. Obwohl viele Programme leistungsstarke Grafikprozessoren erfordern, sind computergenerierte Inhalte immer noch deutlich günstiger als die Lohnkosten professioneller Illustratoren, deren Stundensätze deutlich über den Kosten für eine Software liegen. Generative KI könnte große Probleme für Künstler, Videoproduzenten und andere bedeuten, die ihren Lebensunterhalt damit verdienen.

„Es stellt sich heraus, dass maschinelle Lernmodelle besser, schneller und billiger als Menschen werden können“

Michael Dempsey, geschäftsführender Gesellschafter bei Compound VC.

Generative KI stellt auch komplexere Herausforderungen in Bezug auf Originalität und Eigentum. Das KI-Modell wird mit einer großen Anzahl vorhandener Bilder trainiert, und ob die Urheber der Originalbilder das Urheberrecht an den im Originalstil erzeugten Bildern haben, ist noch umstritten. Ein Künstler gewann kürzlich einen Kunstwettbewerb im US-Bundesstaat Colorado mit Bildern, die hauptsächlich von einer generativen KI namens [MidJourney](#) erstellt wurden. Er sagte in einem Interview nach seinem Gewinn, dass er eines der Hunderte von Bildern, die er erstellt hatte, ausgewählt und dann in Photoshop angepasst und bearbeitet habe.

Was ist generative KI – und was kann sie leisten?

Auf einigen der von Stable Diffusion generierten Werken erscheinen Wasserzeichen, was darauf hindeutet, dass ein Teil des ursprünglichen Datensatzes urheberrechtlich geschützt ist. Einige Anleitungen mit Hinweisen empfehlen, dass Benutzer den Namen eines bestimmten lebenden Künstlers verwenden, um bessere Ergebnisse bei der Nachahmung des Stils dieses Künstlers zu erzielen. Im September 2022 haben [Getty Images und Shutterstock das Hochladen von KI-generierten Bildern in ihre Bilddatenbank verboten](#), weil Bedenken wegen Urheberrechtsverletzungen bestehen.

Die Bildgeneratoren können auch verwendet werden, um neue Bilder von markenrechtlich geschützten Charakteren oder Objekten zu erstellen, wie z. B. Minions, Marvel-Charaktere oder Game of Thrones-Charaktere. Da die Bilderzeugungssoftware immer besser wird, hat sie auch das Potenzial, Benutzer dazu zu bringen, falsche Informationen zu glauben, oder Bilder oder Videos von Ereignissen zu zeigen, die nie stattgefunden haben.

Entwickler müssen sich auch mit der Möglichkeit auseinandersetzen, dass KI-Modelle, die mit großen Datenmengen trainiert wurden, Verzerrungen in Bezug auf Geschlecht, Rasse oder Kultur in den Daten enthalten können, was dazu führen könnte, dass die Modelle solche Verzerrungen in ihrer Ausgabe aufweisen. Hugging Face hat [Material zu ethischen Fragen](#) und Diskussionen zur verantwortungsvollen Entwicklung von KI-Modellen veröffentlicht:

„Wir sehen kurzfristige und unmittelbare Herausforderungen bei diesen Modellen, da es sich um probabilistische Modelle handelt, die auf großen Datensätzen trainiert werden und dazu neigen, eine Menge Vorurteile zu absorbieren“, sagte Clement DeLange, CEO von Hugging Face.

Generative KI wurde einmal gebeten, zu zeichnen ein Bild eines "Software-Ingenieurs", und es erzeugte ein Bild eines weißen Mannes, so der CEO.

Was ist generative KI – und was kann sie leisten?

Die so genannte Deepfake-Technologie ist in der Lage, mit künstlicher Intelligenz realistische künstliche Videos einer Person zu erstellen. In einem Deepfake-Video wird das Gesicht einer Person nahtlos mit dem einer anderen ausgetauscht, wodurch die Illusion entsteht, dass die Person in dem Video etwas sagt oder tut, was sie in Wirklichkeit nie gesagt oder getan hat. Deepfakes können zwar für harmlose Späße verwendet werden – sie haben aber auch das Potenzial, für böartige Zwecke eingesetzt zu werden. So könnte beispielsweise einem Politiker ein schädliches Geständnis vorgetäuscht werden, oder einer Führungskraft aus der Wirtschaft eine rassistische oder sexistische Äußerung vorgetäuscht werden. Mit der Weiterentwicklung der künstlichen Intelligenz werden die Fälschungen wahrscheinlich immer realistischer und schwieriger zu erkennen sein. Dies könnte ernsthafte Auswirkungen auf unsere Gesellschaft und unsere Demokratie haben.

Ausblick

Die Lösungen sind bereits da und können genutzt werden! Und ja: Sie sind an einigen Stellen noch ausbaufähig. Doch gemessen am aktuellen Stand der Entwicklung ... ist das schon beunruhigend gut, was da geliefert wird. Zugegeben: Die Ausgabe/Anzeige der generierten Bilder ist schwer zu beeinflussen. Das wird sich vermutlich auch in den nächsten Versionen noch nicht grundlegend ändern. Auch die Bildauflösungen müssen noch besser werden – aber auch dafür gibt es schon ergänzende Lösungen, mit denen die Ergebnisse “hochskaliert” werden können. Und für kreativ arbeitende Menschen können diese Werkzeuge als Ergänzung für Ideen oder Prototypen verwendet werden. Glücklicherweise sieht es nicht so aus, als würde ein ganzer Berufszweig durch diese Software überflüssig werden. Aber es ist sicherlich hilfreich, sich mit den Möglichkeiten und Tools rund um [Künstliche Intelligenz im Marketing](#) zu beschäftigen. Denn aufzuhalten – ist deren Nutzung schon lange nicht mehr.

So kontaktieren Sie uns

KONTOR4 GmbH

Agentur für neue Medien

Georgstraße 48

30159 Hannover

Tel.: 0511-99995890

E-Mail: info@kontor4.de

Internet: www.kontor4.de

Telefonisch erreichen Sie uns wochentags am besten in der Zeit von 9:00-17:30 Uhr. Sollten wir nicht persönlich zu erreichen sein, sprechen Sie Ihr Anliegen bitte auf unseren Anrufbeantworter.