

MEDIEN

KÜNSTLICHE INTELLIGENZ

Der Computer, ein Künstler?

Joël Adami

Sogenannte Künstliche Intelligenzen, die Bilder und Texte produzieren, sorgen zunehmend für Aufregung. Künstler*innen machen sich Sorge um ihre Zukunft, andere sehen eine glorreiche Demokratisierung von Kunst in greifbarer Nähe. Die woxx erklärt, wie viel von dem Hype gerechtfertigt ist.

Ein Foto von Papst Benedikt in einer weißen Steppjacke sorgte in den letzten Tagen für Aufregung in sozialen Netzwerken. Das Foto war massenhaft verbreitet worden und Bestandteil vieler Memes, die sich über die Kleidungswahl des Kirchenoberhauptes lustig machten. Das Foto ist jedoch nicht echt, sondern stammt aus einem Bildgenerator, der mit sogenannter Künstlicher Intelligenz (KI) arbeitet. Wer genauer hinschaut, sieht die typischen Ungenauigkeiten solcher Bildgeneratoren: Die Hände des Papstes sehen merkwürdig aus, die Kette seines großen silbernen Kreuzes ist nur auf einer Seite vorhanden und auch der Reißverschluss könnte so nicht funktionieren.

Viele Menschen haben das Bild auf den ersten Blick für echt gehalten, noch mehr haben sich Sorgen darüber gemacht, dass demnächst niemand mehr Realität und Fiktion auseinanderhalten könnte. Eine Diskussion, die nicht neu ist, einfache Möglichkeiten der Bildmanipulation gibt es schließlich spätestens seit der Erfindung von Photoshop. Neu ist höchstens, dass einige Stichworte oder ein beschreibender Satz reichen, um ein Bild zu erzeugen.

Die Programme, die auf Namen wie Midjourney, DALL-E, Stable Diffusion oder StarryAI hören, werden von ihren Entwickler*innen und Medien gerne „Künstliche Intelligenzen“ genannt. Diese Bezeichnung ist aber irreführend: Diese Programme haben kein Bewusstsein und verstehen keine Konzepte. Sie sind lediglich sehr gut darin, Bilder zu erstellen. Sie sind von einer „Artificial General Intelligence“, also „Künstlichen generellen Intelligenz“ sehr weit entfernt – falls es überhaupt je möglich sein wird, eine solche zu entwickeln.

Wie die Bildgeneratoren erfreut sich genau auch der Textgenerator ChatGPT hoher Aufmerksamkeit und

Nutzer*innenzahlen. Er kann Fragen beantworten, Programmiercode erstellen, Gedichte erfinden – kurzum: Texte schreiben. Obwohl die Antworten bei Weitem nicht immer korrekt sind und das Programm keinerlei Möglichkeit hat, richtige von falschen Antworten zu unterscheiden, wird auf politischem Plan vor allem darüber geredet, dass ChatGPT Schüler*innen bei den Hausaufgaben helfen könnte. Dass reiche oder gebildete Eltern ihren Kindern ganz ohne „künstliche Intelligenz“ Hilfe anbieten oder kaufen können, wird jedoch eher selten angesprochen.

Künstliches Lernen

Alle diese Generatoren basieren auf Machine Learning. Die Programme werden mit großen Datenmengen „trainiert“ und lernen so, ihre Aufgaben zu erfüllen. Bei ChatGPT liegt diese darin, einen möglichst plausiblen Text zu schreiben, der zu einer Fragestellung passt. Bei Bildgeneratoren ist die Aufgabe ähnlich: Es soll ein möglichst plausibles Bild zu dem eingegebenen Text produziert werden. Das kann allerdings nur auf der Basis jener Daten passieren, mit dem ein Generator gefüttert wurde. ChatGPT kann zum Beispiel keine Informationen über die russische Invasion der Ukraine liefern, da seine Trainingsdaten nur bis 2021 reichen. Die Vorstellung, dass Generatoren ständig das Internet abgrasen, um sich selbst weiterzuentwickeln, ist zwar weitverbreitet, jedoch falsch.

Doch mit welchen Daten werden diese Generatoren trainiert? Das ist eine Frage, die nicht einfach zu beantworten ist, da die Firmen, die sie entwickeln, darüber nicht unbedingt gerne Auskunft geben. Die Entwickler*innen hinter ChatGPT nennen sich zwar „OpenAI“, wollen über die Trainingsdaten ihres neuesten Modells jedoch keine Informationen mehr herausgeben. Sicher ist: Die Maschinen wurden mit gigantischen Mengen an Text gefüttert. Der stammte aus Büchern, aus Wikipedia, aber auch von allen möglichen Websites. Dieser Trainingsprozess lief jedoch nicht vollständig automatisch ab, sondern brauchte menschlichen Input: Während des Trainingsprozesses gaben Menschen Feedback über

die Qualität des Outputs von ChatGPT, was dann wiederum in das Modell einfließt. „Reinforcement learning from human feedback“ nennen es die Entwickler*innen.

OpenAI wollte ebenfalls verhindern, dass ChatGPT problematische Inhalte wie etwa Beschreibungen von Gewalt, Rassismus oder Sexismus ausgibt. Das dafür nötige Labeling von Inhalten wurde an kenianische Arbeiter*innen ausgelagert, wie das Time Magazine aufdeckte. Die Bezahlung, um sich durch die scheußlichsten Texte der Menschheit, etwa detaillierte Darstellungen von Suiziden, Kindesmissbrauch und Vergewaltigungen zu klicken: zwei Dollar pro Stunde.

Welche Trainingsdaten für Bildgeneratoren benutzt werden, wird offener kommuniziert: Die größten und wichtigsten Modelle wie DALL-E und Midjourney benutzen alle Datensets von Laion (Large-scale Artificial Intelligence Open Network), einem deutschen gemeinnützigen Verein. Laion stellt Milliarden Bilder gemeinsam mit beschreibendem Text in verschiedenen Datenset-Versionen zur Verfügung. Die Bilder hat Laion allerdings nicht selbst heruntergeladen, sondern sie aus dem „common crawl“-Projekt extrahiert. Wer genau entscheidet, welche Bild-Text-Paare in das Datenset aufgenommen werden, ist nicht bekannt. Mehrere Künstler*innen, aber auch die Bildagentur Getty Images, haben Klage gegen die Hersteller von Bildgeneratoren eingelegt, da ihre Bilder ohne ihr Einverständnis zum Training verwendet wurden.

Pornos zum Training

Untersuchungen der Laion-Datensets zeigen, dass sich zum Beispiel viele pornografische Darstellungen darin finden. Eine Suche nach dem Wort „big“ bringt nackte Frauen mit großen Brüsten hervor. „Sex ist im Web so prävalent, dass es sogar eine unschuldige Suche wie ‚big‘ dominiert“, schreibt KI-Forscher Mathieu Jacomy in seinem Blogbeitrag zu Laion. Die Entwickler*innen von Bildgeneratoren arbeiten oberflächlich gegen diese Trainingsdaten und verbieten ihren Pro-

Ein Bild, das der Bildgenerator Midjourney bei dem Begriff „Midjourney“ ausgespuckt hat. Wer hierin ein Selbstporträt einer künstlichen Intelligenz erkennen will, irrt. Allerdings zeigt das Bild sehr gut den typischen „Stil“ von Midjourney.



ILLUSTRATION: CC-BY-SA CHIKORTA/WIKIMEDIA

dukten zum Beispiel allzu pornografische oder gewaltsame Darstellungen.

Der algorithmische Bias (siehe woxx 1628) scheint dennoch immer wieder durch: Im Dezember 2022 machte die App „Lensa AI“ von sich reden. Nutzer*innen konnten ein Foto von sich machen, die App erstellte dann gezeichnete Comic-Avatare. Doch bei vielen Frauen spuckte der Bildgenerator statt den erhofften „magischen Bildern“ etwas ganz anderes aus: sexualisierte Posen, tiefe Dekolletees, abgeschnittene Köpfe. „Lensa AI macht aus mir eine ‚Wichsvorlage‘“, schrieb die Journalistin Barbara Wimmer vom Onlinemagazin Futurzone darüber.

Auch rassistische Stereotypen werden immer wieder bedient. Ein Nutzer, der beim Generator Midjourney „The Queen as a pimp“ eingab, bekam zwar die englische Königin in dem gewünschten „Zuhälter“-Outfit, doch die Monarchin war nun schwarz. Ein subtileres Phänomen ist die Hegemonisierung: Wer die Bildgeneratoren nach Selfies von verschiedenen Personengruppen fragt, wird mit Bildern konfrontiert, auf denen diese Menschen, von alten Ägypter*innen bis zu französischen Soldaten im ersten Weltkrieg immer ein „typisch amerikanisches“ Lächeln mit weit geöffnetem Mund und vielen Zähnen zeigen. Neben den Trainingsdaten ist aber sicher auch die Feinabstimmung der Generator-Entwickler*innen ein Problem, denn diese haben oft eine Standard-Ästhetik, die beim Zielpublikum besonders gut ankommt.

Über ein anderes Problem wurde bislang wenig gesprochen: Der Energiehunger der Bild- und Textgeneratoren. ChatGPT wurde laut OpenAI mit einem Supercomputer trainiert, der 285.000 Prozessoren und 10.000 Grafikkarten-Chips enthielt. Neben dem Training kostet auch die Erstellung von Bildern Prozessorzeit und damit Energie. Es ist nicht nur eine Frage der Zeit, bis ein Bild des Papstes in Steppjacke produziert wird, das nicht mit etwas Hinsehen als Fälschung entlarvt werden kann – sondern auch, wann die Kosten für die Hersteller von Bildgeneratoren so groß werden, dass sie auf ein rein kommerzielles Geschäftsmodell werden umsatteln müssen.