

DALL-E: Texte per KI in Bilder umwandeln

22. Juni 2022 von Silke Grasreiner

Haben Sie schon von den lustigen, mittels KI aus Text generierten Bildern gehört? Vor allem auf Twitter werden die Kreationen gerne geteilt. Die Software dahinter nennt sich DALL-E bzw. ihre Weiterentwicklung DALL-E 2. Damit lassen sich aus wenigen Textvorgaben realistische, nie da gewesene Bilder erstellen. Hier erfahren Sie mehr über diese faszinierende Technologie.



Was ist DALL-E?

DALL-E ist eine neue KI-Software, die **geschriebene Worte in Bilder und Kunstwerke verwandelt**. Der

Nutzer kann einfach einen Text eingeben und das neuronale Netzwerk übersetzt in ein Bild - sogar nicht nur in eins, sondern mehrere, so dass man auswählen kann, welche Umsetzung einem am besten gefällt.

DALL-E wurde entwickelt von OpenAI, einem von Microsoft unterstützten Startup, auf der Basis der Imagen-Technologie von Google. Im Januar 2021 stellte OpenAI die erste Version der Software vor. Inzwischen gibt es ein noch intelligenteres System: DALL-E 2 erzeugt noch realistischere und präzisere Bilder in vier Mal höherer Auflösung.

Auf der OpenAI-Website heisst es: "DALL-E 2 kann originale, realistische Bilder und Kunstwerke aus einer Text-Beschreibung Sprache erstellen. Es kann Konzepte, Attribute und Stile kombinieren." So kann die KI-Software die unwahrscheinlichsten Bilder erzeugen, wie zum Beispiel "ein Koalabär auf einem Motorrad" oder "Teddybären beim Einkaufen im alten Ägypten":



Nach Angaben der Entwickler kann "DALL-E 2 alle Menschen dazu befähigen, sich kreativ auszudrücken. DALL-E 2 hilft uns auch zu verstehen, wie fortschrittliche KI-Systeme unsere Welt sehen und verstehen, was für unsere Mission, KI zum Nutzen der Menschheit zu entwickeln, entscheidend ist."

Wie funktioniert DALL-E?

DALL-E ist ein **neuronales Netzwerk**, das von OpenAI über maschinelles Lernen trainiert wird, Bilder aus Text zu erzeugen. Es kann nicht verwandte Konzepte kombinieren und anthropomorphisierte Versionen von Tieren und Objekten erstellen. DALL-E arbeitet mit CLIP zusammen, einem System, das OpenAI bereits 2021 Jahr vorgestellt hatte. Das neuronale Netzwerk übersetzt die menschliche Sprache zunächst in seine eigene Sprache und wandelt sie erst danach in Bilder um.

Die Oberfläche von DALL-E ist einfach: Es gibt ein Textfeld, in das der Nutzer Text eingeben kann, eine Schaltfläche zum Starten der Konvertierung und die Bilder direkt darunter. Die Entwickler erklären die Funktionsweise von DALL-E 2 so: "Es verwendet ein Verfahren namens 'Diffusion', das mit einem Muster aus zufälligen Punkten beginnt und das das Muster allmählich in Richtung eines Bildes verändert, wenn es bestimmte Aspekte dieses Bildes erkennt."