

(Eine) Geschichte der Generativen Modelle

Erarbeitet von
Dr. Jacqueline Klusik-Eckert

Lernziele	1
Inhalt	2
Einstieg	2
Zeitstrahl zur Geschichte	2
Take-Home Message	7
Quellen	8
Disclaimer	8

Lernziele

- Du kannst Meilensteine der Generativen Künstlichen Intelligenz benennen

			erstmalig führende Expert*innen des Forschungsgebietes zusammen. Auf der Agenda stehen viele Bereiche, die für die moderne KI von grundlegender Bedeutung sind. Ein Jahr zuvor hatte McCarthy in der Überschrift seines Konferenzbeitrages erstmals den Begriff „Artificial Intelligence“ (Künstliche Intelligenz) gebraucht, der seitdem für diese Technologie verwendet wird.
1957	1957	Noam Chomsky	Der Linguist Noam Chomsky publiziert „Syntactic Structures“. Hier beschreibt er die grammatischen Regeln für das Parsing und Generieren von natürlicher Sprache.
1958		Perceptron	Frank Rosenblatt stellt das Modell „Perceptron“ vor und legt damit die Basis für die Entwicklung neuronaler Netzwerke. Genau genommen handelt es sich aus heutiger Perspektive bei dem „Perceptron“ um ein „single layer“ neuronales Netzwerk. Quelle: Rosenblatt, F. (1958). The perceptron: a probabilistic model for information storage and organization in the brain. <i>Psychological review</i>, 65(6), 386.
1964	1966	ELIZA	Intelligent? Einfühlsam? Diese Eigenschaften dichteten die Forscherinnen und Forscher ELIZA an, dem ersten Chatbot und damit Vorfahre von ChatGPT. Der Informatiker Joseph Weizenbaum entwickelte am Massachusetts Institute of Technology (MIT) ein Computerprogramm, das menschliche Sprache und ihre Kommunikationsregeln verarbeiten kann: ELIZA. Das „Sprechen“ mit ELIZA in Textform war für die Testpersonen so überzeugend, dass sogar Weizenbaum davor warnte, dass man diesen Maschinen die „Aura des Magischen“ austreiben müsse. Weizenbaum, J. (1966). ELIZA—a computer program for the study of natural language communication between man and machine. <i>Communications of the ACM</i>, 9(1), 36–45. https://doi.org/10.1145/365153.365168
1971		LOGO	Die Programmiersprache Logo wurde bereits Ende der Sechziger Jahre von Daniel G. Bobrow, Wally Feurzeig, Seymour Papert und Cynthia Solomon am MIT entwickelt. Die Idee dahinter war Programmieren so einfach zu gestalten, dass es bereits von Kindern und in

			Schulen angewandt werden kann. Im Prinzip ist Logo damit der Vorreiter von Anwendungen wie Scratch oder Open Roberta (Nepo).
1985	1985	Bayesian Network	Der Informatiker Judea Pearl stellt das Bayessche Netz vor, das als statische Analyse die Grundlage herkömmlicher neuronaler Netze ist, auf die auch die generativen Modelle aufbauen.
1986	1986	RNNs	Jordan, RNNs werden eingeführt. Die Recurrent Neural Networks können besser mit langen Sätzen umgehen und Kontext bewahren. Kurz danach LSTM (Long Short Term Memory).
1989	1989	Bilderkennung	Yann LeCun, Yoahua Bengio und Patrick Haffner zeigen wie CNNs zur Bilderkennung verwendet werden können.
2006	2012	Image Net	Während sich der Großteil der Forscher*innengemeinde auf die Weiterentwicklung von Modellen konzentrierte, erkannte die Forscherin Fei-Fei Li die mangelnde Verfügbarkeit von qualitativ hochwertigen Bilddaten für Forschungszwecke als eine Barriere für den Fortschritt der Disziplin. Sie begann daher den Bilddatensatz Image Net zu erstellen und stellte diesen für die Forschungsgemeinschaft zur Verfügung. Heute beinhaltet der Datensatz mehr als 14 Millionen annotierter Bilddaten und ist Bestandteil des Werkzeugkastens der meisten Forscher*innen und Entwickler*innen im Bereich der KI. https://en.wikipedia.org/wiki/ImageNet
2011		Siri	Die Firma Apple stellt Siri vor: Der virtuelle Assistent ist serienmäßig auf allen iPhones implementiert und begegnet Nutzerinnen und Nutzern als ruhige, freundliche Frauen- oder Männerstimme. Siri beantwortet Anfragen verschiedenster Art, insbesondere zur Steuerung des Gerätes, recherchiert eigenständig im Internet (Wettervorhersagen, Fahrpläne etc.) und beantwortet sogar allgemeine Anfragen („Siri, erzähl mir einen Witz!“) mehr oder weniger humorvoll. Großkonzerne wie Google oder Amazon ziehen bald mit eigenen Systemen nach.
2011		Watson	Watson ist ein Computerprogramm aus dem Bereich der Künstlichen Intelligenz. Es wird von IBM entwickelt,

			um Antworten auf Fragen zu geben, die in digitaler Form in natürlicher Sprache eingegeben werden. Ziel des Projekts ist es, eine hochwertige semantische Suchmaschine zu schaffen. Kurz: Dass der Computer mit einem Menschen redet und sinnvolle Antworten gibt. Hierzu zählt auch, dass der Computer Ironie, Wortwitz und Zwischentöne begreifen kann, dass er fehlende Puzzlesteine entdeckt, dass er zurückfragen und Assoziationsketten entwickeln kann.
2013		VAEs	Diederik P. Kingma and Max Welling stellen mit dem Variational Autoencoder ein neues generatives Modell vor. https://en.wikipedia.org/wiki/Variational_autoencoder#:~:text=A%20variational%20autoencoder%20is%20a,%26%20slab)%20sparse%20coding).
2014	2014	GANs	Der Forscher Ian Goodfellow stellt die Technologie der Generative Adversarial Networks (GANs) vor. Diese ermöglicht das unüberwachte und selbstlernende Architekturen auch neue Bilder erstellen können. Ian Goodfellow et.al. haben im Dezember 2014 ihr Paper zu dem ersten GAN veröffentlicht, eine CNN-Architektur, die eine Bildsynthese produziert. Quelle: https://proceedings.neurips.cc/paper/2014/file/5ca3e9b122f61f8f06494c97b1afccf3-Paper.pdf  Hello world! Sample of GAN-generated faces from Goodfellow et al.'s 2014 paper. The model was trained

			on the Toronto Faces dataset, which since has been removed from the web.
2015			Veröffentlichung Unsupervised representation learning with deep convolutional generative adversarial networks, GANs. Hier wird auch das erste Mal über die Frage nachgedacht, den Latent Space zu manipulieren Quelle: https://arxiv.org/pdf/1511.06434.pdf
2017	2017	Transformer	Der nächste große Durchbruch kam schon 2017 mit der Einführung der Transformer durch Vaswani et al. [7]. Aufbauend auf Transformermodellen gab es allgemein große Verbesserungen bei generativen KI-Modellen, nicht nur bei textbasierten Modellen. Transformer werden inzwischen auch bei der Bildgenerierung und anderen generativen Anwendungsfällen eingesetzt.
2018		GPT	Die erste Version des GPT-Modells wird von Open AI erstellt. Quelle : https://toloka.ai/blog/history-of-generative-ai/
2018		Generiertes „Gemälde“	Weltweite Beachtung findet die Versteigerung des von einer KI hergestellten „Portrait of Edmond Belamy“. Das etwas verschwommen wirkende Bild kommt im britischen Auktionshaus Christie's für 432.500 Dollar unter den Hammer. „Gemalt“ wurde es auf Basis einer Datenbank von 15.000 Porträtgemälden aus der Zeit vom 14. bis zum 20. Jahrhundert. Die Pariser Künstlergruppe „Obvious“ (Hugo Caselles-Dupré, Pierre Fautrel und Gauthier Vernier) ließ einen Algorithmus auf Grundlage dieser Porträts ein Bild erstellen und so lange anpassen, bis ein weiterer Algorithmus das Bild nicht mehr von einem von Menschen gemachten Gemälde unterscheiden konnte. https://www.theverge.com/2018/10/23/18013190/ai-art-portrait-auction-christies-belamy-obvious-robbie-barrat-gans
2018		Google Duplex	Auf einer Entwickler*innenkonferenz stellt Google „Duplex“ vor, ein Sprachbot, der die menschliche Stimme sehr gut nachahmen und einen Termin bei einem Friseursalon vereinbaren kann.
2019	2019	ChatGPT	2019 wurde dann das inzwischen weit verbreitete Transformer-Modell ChatGPT von der Firma OpenAI veröffentlicht. ChatGPT ist ein Large Language Model

			(LLM) und in der Lage, Texte hoher Qualität auf Basis eines Prompts zu generieren.
2021	2021	DALL-E	DALL-E wurde am 5. Januar 2021 von OpenAI vorgestellt.^[6] 05.01.2021 Veröffentlichung Dall-E als multimodale Implementierung von GPT. Dall-E 1 mit Transformer CLIP (Contrastive Language-Image Pre-training) basiert auf Zero-Shot-Lernen und ist mit 400 Millionen Bildpaaren mit Textbeschriftungen aus dem Internet trainiert. Quelle: https://de.wikipedia.org/wiki/DALL-E
2022	2022	DALL-E	April 2022 DALL-E 2.0 für ausgewählte Nutzer*innen, öffentlich verfügbar, dann ab dem 21. April 2022 mit einer Lite Version und September 2022 als Vollversion. Quelle: DALL-E Now Available Without Waitlist. In: <i>openai.com</i>. 28. September 2022, abgerufen am 30. Oktober 2022 (englisch). DALL-E 2 verwendet ein Diffusionsmodell, das auf CLIP-Bildeinbettungen konditioniert ist, die während der Inferenz aus CLIP-Texteinbettungen durch ein früheres Modell generiert werden. DALL-E Now Available Without Waitlist. In: <i>openai.com</i>. 28. September 2022, abgerufen am 30. Oktober 2022 (englisch). Bildgenerator DALL-E 2 ist jetzt offen für alle. In: Spiegel Online. 29. September 2022, abgerufen am 30. Oktober 2022.
2022	2022	Stable Diffusion	Im November 2022 erschien StableDiffusion als Open Source Lizenz von der Computer Vision and Learning Gruppe an der LMU München https://ommer-lab.com/research/latent-diffusion-models/ unterstützt durch die Firma Stability.AI, das auf Diffusion Models aufbaut.

Quelle [3, 4, 5]

[Take-Home Message](#)

Der Erfolg der aktuellen generativen KI-Modelle ist aber nicht nur auf die verbesserten Algorithmen zurückzuführen, sondern auch auf die, im Vergleich zu vorher, immensen Anzahl der Trainingsdaten und die stark verbesserte Rechenleistung, die die Verarbeitung

dieser möglich macht. Sie sind das Ergebnis einer stetigen technischen Verbesserung und haben ihre eigene Geschichte.

Quellen

- Quelle [1] Ad Quelle: Schneider, O. (2019, November 17). Die berühmte Schreib-KI ist zurück—Und erschreckend gut | Netzwoche. <https://www.netzwoche.ch/news/2019-11-17/die-beruechtigte-schreib-ki-ist-zurueck-und-erschreckend-gut>
- Quelle [2] Kremp, M. (2019, November 7). „Talk to Transformer“: Künstliche Intelligenz schreibt Texte fertig. Der Spiegel. <https://www.spiegel.de/netzwelt/web/talk-to-transformer-kuenstliche-intelligenz-schreibt-texte-fertig-a-1295116.html>
- Quelle [3] Offert, F. (2023, November 4). *Ten Years of Image Synthesis*. Fabian Offert. <https://zentralwerkstatt.org/blog/ten-years-of-image-synthesis>
- Quelle [4] Meineck, S. (2023, Juli 4). Olimpias Augen: Das Hype-Theater um moderne Chatbots. netzpolitik.org. <https://netzpolitik.org/2023/olimpias-augen-das-hype-theater-um-moderne-chatbots/>
- Quelle [5] Catani, S. (Hrsg.). (2023). Handbuch Künstliche Intelligenz und die Künste. De Gruyter.

Disclaimer

Transkript zu dem Video „Generative Modelle: (Eine) Geschichte der Generativen Modelle“, Dr. Jacqueline Klusik-Eckert.

Dieses Transkript wurde im Rahmen des Projekts ai4all des Heine Center for Artificial Intelligence and Data Science (HeiCAD) an der Heinrich-Heine-Universität Düsseldorf unter der Creative Commons Lizenz [CC BY 4.0](https://creativecommons.org/licenses/by/4.0/) veröffentlicht. Ausgenommen von der Lizenz sind die verwendeten Logos, alle in den Quellen ausgewiesenen Fremdmaterialien sowie alle als Quellen gekennzeichneten Elemente.