

Woche 14 Ethik/Recht: Generative Modelle und die ethischen und rechtlichen Herausforderungen

Skript

Erarbeitet von
Dr. Jacqueline Klusik-Eckert

Inhaltsverzeichnis

Lernziele.....	1
Inhalt	2
<i>Generative Modelle und die ethischen und rechtlichen Herausforderungen</i>	<i>2</i>
<i>Take-Home Message.....</i>	<i>4</i>
Quellen	5
Disclaimer	6

Lernziele

- Gespür für die Herausforderungen entwickeln
- Bewusstsein über die aktuellen und zukünftigen Debatten gewinnen

Inhalt

Generative Modelle und die ethischen und rechtlichen Herausforderungen

Mit ihrer enormen Rechenleistung können generative Modelle Abermillionen Varianten von Medien erzeugen. Manches davon macht für uns Sinn, sieht ästhetisch ansprechend aus oder klingt auch ganz gut. Ziemlich viele dieser Ergebnisse landen jedoch auf dem Datenmüll. Man muss nur die Anfrage refreshen.

Die Modelle erwecken aber auch den Schein, Sprache, Texte, Bilder oder Musik wie aus dem Nichts erzeugen zu können. Das stimmt jedoch nicht.

Zu den größten Herausforderungen gehört die Frage zum Urheberrecht. Das betrifft zum einen die Fragen, wer die Urheberschaft auf die neugenerierten Medien innehat.

Portrait of Edmond de Belamy

https://de.wikipedia.org/wiki/Portrait_of_Edmond_de_Belamy#/media/Datei:Edmond_de_Belamy.png

In einem ersten Fall, dem Verkauf des AI-generated Bildes Portrait of Edmond Belamy beim Auktionshaus Christie's, ging es in den anschließenden Rechtsfragen um den Code und das trainierte Neuronale Netz. Gehören die Ergebnisse auch dem Autor des Codes? Hat der Programmierer des Codes ein Recht auf einen Anteil an dem Auktionserlös von 432.500 \$? Oder ist die Person alleiniger Urheber, die das errechnete Bild auf eine Leinwand hat drucken lassen?

Quelle [1]

Bei den Modellen, von denen wir hier sprechen, geht es eher um die Frage, mit was sie trainiert wurden.

Joana:

Hier tauchen ganz ähnliche Probleme auf wie bei ChatGPT, nämlich u. a. die Tatsache, dass GANs oder Stable Diffusion Modelle auf schon existierenden Bildern trainiert wurden. Das eröffnet eine heiße Debatte über die Urheber*innenschaft der generierten Bilder. In diesem Paper wird gezeigt, dass bei Stable Diffusion in der Tat Kunstwerke aus den Trainingsdaten in generierte Bilder kopiert werden.

Hier sieht man in der oberen Zeile von Stable Diffusion generierte Bilder und in der unteren Zeile das Match aus einem Trainingsdatensatz, das am nächsten ist. Auffällig ähnlich, oder? Häufig werden dabei Objekte, an die sich der Algorithmus „erinnert“ aus dem Vordergrund und dem Hintergrund zusammengefügt. Außerdem zeigt das System manchmal ein „rekonstruktives Gedächtnis“, bei dem die „erinnerten“ Objekte semantisch mit dem Ausgangsobjekt übereinstimmen, ohne, dass sie pixelweise identisch sind.

Quelle [2]

Jaron Lanier schrieb bereits 2013 in seinem Buch "Who owns the Future"

"Digital information is really just people in disguise."

"Digitale Informationen sind eigentlich nur verkleidete Menschen".

Hinter den Milliarden von Daten stecken Designer*innen, Musiker*innen, Künstler*innen, die ihren Alltag mit der Produktion von Medien bestreiten. Das ist ihr Job.

Quelle [3]

Video von Oleg Gamulinskii / Pixabay

Und wir als fleißige Datenproduzent*innen schenken auch noch unsere Medien her. Die Debatte um die Rechtsfragen in alle Richtungen ist im Gange. Unklar ist ja auch, wie Medienkünstler*innen ihre mit KI-Unterstützung erstellten Werke sichern sollen. Da kommt auf die Rechtsabteilungen einiges zu. Und auf uns.

Quelle [4 , 5]

Wie sieht es mit den anderen Risiken und Herausforderungen aus?

Generative Modelle versprechen eine enorme Erleichterung und sie werden sich in unterschiedlichen Einsatzszenarien sicherlich etablieren.

Doch im Moment können sie stark voreingenommen sein. Und das liegt an den Biases, den Vorurteilen, die in den Trainingsdaten verborgen liegen. Darüber hinaus ist die Benutzung der Technik nicht gerade fair geregelt. Viele KIs gehören Firmen, die mit ihrer Rechenleistung natürlich auch ein Geschäftsmodell bedienen müssen. Öffentliche Zugänge für alle werden durch nicht-barrierefreie Benutzeroberflächen erschwert.

Quelle [6]

Generell bergen die Modelle auch die Gefahr, dass man sich zu sehr darauf verlässt, dass die Ergebnisse richtig sind. Mit mangelndem Verständnis über ein Thema oder fehlendem Fachwissen lässt sich das auch nur sehr schwer überprüfen. Es besteht die Gefahr, dass gefährliches Halbwissen oder sogar Falschinformationen gestützt und untermauert werden.

Quelle [6]

Es ist auch noch nicht klar, wie in Zukunft Modell-generierte und von Menschen-generierte Medien voneinander unterschieden werden sollen.

So viele Fragen sind noch nicht geklärt. Wie sieht es mit den Kosten für Schulungen und Wartung aus? Wer sichert den Datenschutz und gibt Datensicherheit? Sind diese Systeme überhaupt nachhaltig? Wie sieht es mit dem Ressourcenverbrauch aus? Schließlich fressen die Serverparks für die Hochleistungsprozesse ziemlich viel Strom.

Quelle [7]

Unklar ist auch, wer die Kosten für die Überprüfung von Informationen übernimmt und ob wir an irgendeiner Stelle die Integrität von Inhalten schützen können.
Es kann sein, dass es immer schwieriger wird zwischen echtem Wissen und überzeugend geschriebenem, aber nicht verifiziertem Modell Output zu unterscheiden.

Quelle [8]

Und wie sieht es mit der Mehrsprachigkeit aus? Die großen Sprachmodelle funktionieren sehr gut, wenn es um Englisch geht. Und wie sieht es mit Niederländisch aus? Urdu? Kambodianisch? Wie soll eine Teilhabe an der neuesten Technik ermöglicht werden, wenn die ach so schlaun KIs nur auf wenige, westeuropäische Sprachen trainiert sind?

<https://www.visualcapitalist.com/100-most-spoken-languages/>

Quelle [9]

Alles in allem: Ein fairer Zugang, ein sicherer Gebrauch, sieht anders aus.

Take-Home Message

Generative Modelle haben sich in den letzten zehn Jahren unglaublich schnell verbessert. Das gilt sowohl für die Generierung von Texten, als auch für das Produzieren von Bildern, Animationen und Musik. Die Anwendungsszenarien sind im Moment noch in der Erprobung.

Quelle [10]

Es ist aber jetzt schon abzusehen, dass die Anwendungen Einzug in unseren Arbeits- und Lebensalltag haben werden. Die Technik ist bereit.
Doch sind wir es? Wo werden die Entwicklungen noch hingehen?
Die UNESCO fordert daher, dass die Entwicklungen in der Künstlichen Intelligenz einem auf den Menschen ausgerichteten Ansatz folgen müssen. Humen-Centered

<https://www.unesco.org/en/digital-education/artificial-intelligence>

Quelle [18]

Was die Technik der Künstlichen Intelligenz in den nächsten Jahren noch bringen wird, steht in den Sternen. Klar ist schon heute, dass wir vom Einsatz dieser Verfahren an vielen Stellen profitieren. Wir bestimmen auch die Einsatzszenarien, die rechtlichen Rahmenbedingungen und bauen gewisse Hürden gegen den Missbrauch von Künstlicher Intelligenz auf. Es werden noch viele Debatten auf uns zukommen. Debatten, die wir alle gemeinsam führen müssen. Um das tun zu können, brauchen wir Zugang zum Wissen über diese Techniken und Verfahren. KI für alle.

Quellen

- Quelle [1] Small, Z. (2018, Oktober 29). *Christie's Sells „AI-Generated“ Art for \$432,500 As Controversy Swirls Over Creators' Use of Copied Code*. Hyperallergic.
<http://hyperallergic.com/468060/christies-sells-ai-generated-art-for-432500-as-controversy-swirls-over-creators-use-of-copied-code/>
- Quelle [2] Somepalli, G. et al (2022). *Diffusion Art or Digital Forgery? Investigating Data Replication in Diffusion Models* (arXiv:2212.03860). arXiv.
<https://doi.org/10.48550/arXiv.2212.03860>
- Quelle [3] Lanier, J. (2014). *Who owns the future?* (Simon&Schuster trade paperback edition). Simon & Schuster Paperback.
- Quelle [4] Meineck, S. (2023, Januar 19). Sammelklage: Streit um Bild-Generatoren soll vor Gericht landen. netzpolitik.org. <https://netzpolitik.org/2023/sammelklage-stable-diffusion-bild-generator-rechtsstreit/>
- Quelle [5] Chayka, K. (2023, Februar 10). Is A.I. Art Stealing from Artists? *The New Yorker*.
<https://www.newyorker.com/culture/infinite-scroll/is-ai-art-stealing-from-artists>
- Quelle [6] Mai, K. T., Bray, S. D., Davies, T., & Griffin, L. D. (2023). Warning: Humans Cannot Reliably Detect Speech Deepfakes (arXiv:2301.07829). arXiv.
<http://arxiv.org/abs/2301.07829>
- Quelle [7] *Training a single AI model can emit as much carbon as five cars in their lifetimes*. (2023). MIT Technology Review. Abgerufen 21. Februar 2023, von
<https://www.technologyreview.com/2019/06/06/239031/training-a-single-ai-model-can-emit-as-much-carbon-as-five-cars-in-their-lifetimes/>
- Quelle [8] Vincent, J. (2022, September 15). Anyone can use this AI art generator—That's the risk. The Verge. <https://www.theverge.com/2022/9/15/23340673/ai-image-generation-stable-diffusion-explained-ethics-copyright-data>
- Quelle [9] Soria, C. (2020, Januar 1). Decolonizing Minority Language Technology. State of the Internet's Languages Report. <https://internetlanguages.org/en/stories/decolonizing-minority-language/>
- Quelle [10] Offert, F. (2023). Ten Years of Image Synthesis. Zentralwerkstatt.Org. Abgerufen 21. Februar 2023, von <https://zentralwerkstatt.org/blog/ten-years-of-image-synthesis>

Quelle [11] Quelle 18 UNESCO. Artificial intelligence in education. Abgerufen 20. Februar 2023, von <https://www.unesco.org/en/digital-education/artificial-intelligence>

Disclaimer

Transkript zu dem Video „Woche 14 Ethik/Recht: Generative Modelle und die ethischen und rechtlichen Herausforderungen“, 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](https://creativecommons.org/licenses/by/4.0/) 4.0 veröffentlicht. Ausgenommen von der Lizenz sind die verwendeten Logos, alle in den Quellen ausgewiesenen Fremdmaterialien sowie alle als Quellen gekennzeichneten Elemente.