

Prof. Dr. Lasse Scherffig – Geschlechteraspekte in der gestalterischen Forschung zu generativer KI



Prof. Dr. Lasse Scherffig

hat die Professur für Interaction Design an der KISD (Köln International School of Design) der Fakultät für Kulturwissenschaften inne. Aktuelle Schwerpunkte der Forschung bilden die Rolle künstlicher Intelligenz im Design sowie die inklusive und partizipative Gestaltung von Technologie mit Menschen mit Behinderung. Prof. Dr. Lasse Scherffig leitet das Teilprojekt der TH Köln im BMBF-geförderten Projekt KITeGG (KI greifbar machen und begreifen: Technologie und Gesellschaft verbinden durch Gestaltung).

Ausgangslage

Statistische Modelle, mit deren Hilfe sich Bilder produzieren lassen, verändern als „generative Künstliche Intelligenz“ derzeit nachhaltig die Produktion visueller Inhalte. Diese statistischen Modelle beziehen die zur Bildgeneration nötigen Wahrscheinlichkeitsverteilungen aus Trainingsdaten. In diesem Prozess werden die strukturellen Gemeinsamkeiten großer Bilddatensätze für die Erzeugung "neuer" Inhalte nutzbar gemacht.

Die so produzierten Inhalte reproduzieren (Bianchi 2023) und verstärken (Birhane 2023) die in den Daten enthaltenden diskriminierenden Vorurteile. Generative KI steht im Zeichen unterschiedlicher Biases, die sich aus den vorurteilsbehafteten kulturellen Inhalten speisen, auf deren Grundlage das Training erfolgt - auch, weil diese Inhalte nach "ästhetischen" Kriterien gefiltert werden, die auf Normen einer westlichen, „männlich“ geprägten Bildkultur basieren (Buschek und Thorp 2024).

Erste Ideen zu Forschungsfragen

Gemeinsam mit Prof. Thomas Hawranke, PhD von der Kunsthochschule Halle arbeitet Prof. Dr. Lasse Scherffig an einem Antrag für ein Kooperationsprojekt, das Standardisierungstendenzen von Design und visueller Kultur im Sinne einer "template culture" (Lorusso 2023) in den Blick nehmen und generative KI hier in eine Entwicklung einordnen soll, die lange vor den aktuellen Systemen beginnt, sie ermöglicht und durch sie fortgesetzt wird.

Aufbauend auf gemeinsame Vorarbeiten zur Rolle von Klassifikation in generativen Verfahren (Scherffig und Hawranke 2024), sollen diese einerseits in die Geschichte der Digitalisierung des Designs eingeordnet werden, andererseits sollen Methoden entwickelt werden, die mittels Programmierung und Interfaceentwicklungen das generative Potential dieser Verfahren, einschließlich seiner Vorurteile und Grenzen, erfahrbar machen und unterlaufen.

Die Arbeit im Rahmen des Fellowships besteht darin, die Rolle von Geschlechteraspekten hierbei gesondert in den Blick zu nehmen. Einerseits heißt das, das Feld einschlägiger Arbeiten aus dem Bereich der kritischen KI zu kartieren, andererseits einen Standpunkt dazu zu entwickeln, wie Geschlechteraspekte im geplanten Projekt methodisch berücksichtigt werden sollen.

Literatur

Bianchi, Federico; Kalluri, Pratyusha; Durmus, Esin; Ladhak, Faisal; Cheng, Myra; Nozza, Debora; Hashimoto, Tatsunori; Jurafsky, Dan; Zou, James; Caliskan, Aylin (2023). Easily Accessible Text-to-Image Generation Amplifies Demographic Stereotypes at Large Scale. In: 2023 ACM Conference on Fairness, Accountability, and Transparency, New York, NY, USA: ACM, S. 1493-1504.

Birhane, Abeba; Prabhu, Vinay; Han, Sang; Boddeti, Vishnu N. (2023). On Hate Scaling Laws For Data-Swamps, in: arXiv. Online unter: <https://arxiv.org/abs/2306.1314>

Lorusso, Silvio (2023). What design can't do. Essays on design and disillusion, Eindhoven: Set Margins' Publications.

Scherffig, Lasse; Hawranke, Thomas (2024). Hund. Gassi gehen im Latent Space. In: Ina Bolinski; Thomas Hawranke; Stefan Rieger (Hg.), Virtuelle Tiere. Lebewesen zwischen Code und Kreatur, Bielefeld: transcript.

GEFÖRDERT VOM



Bundesministerium
für Bildung
und Forschung

Technology
Arts Sciences
TH Köln