

## Warum Siri meine Oma nicht versteht ... und andere Ungerechtigkeiten im Zeitalter Künstlicher Intelligenz

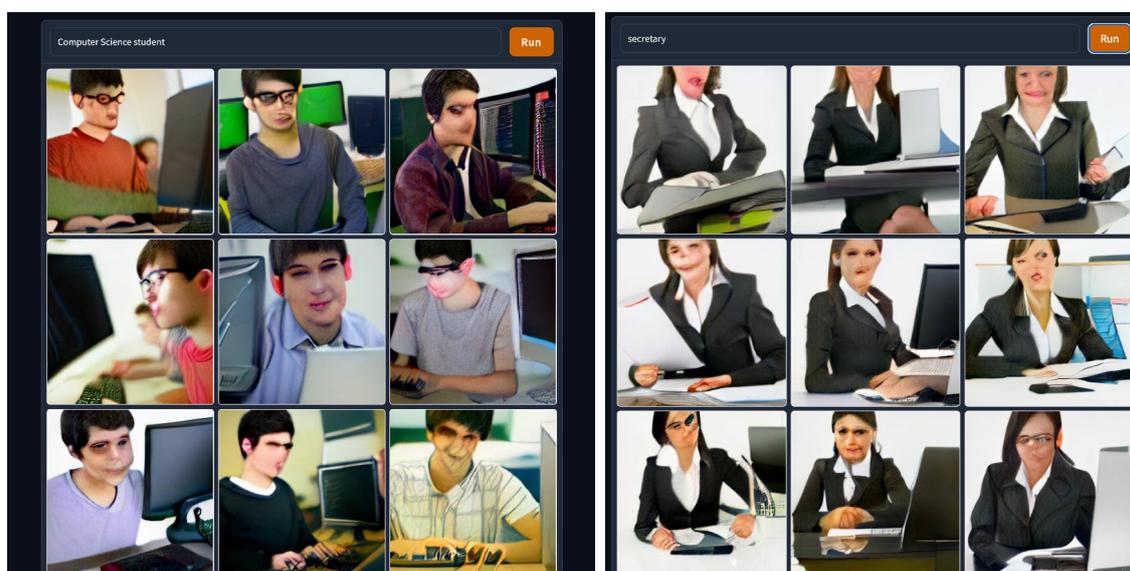
Keynote von Martina Mara

Kurzzusammenfassung

---

### Ausgangslage

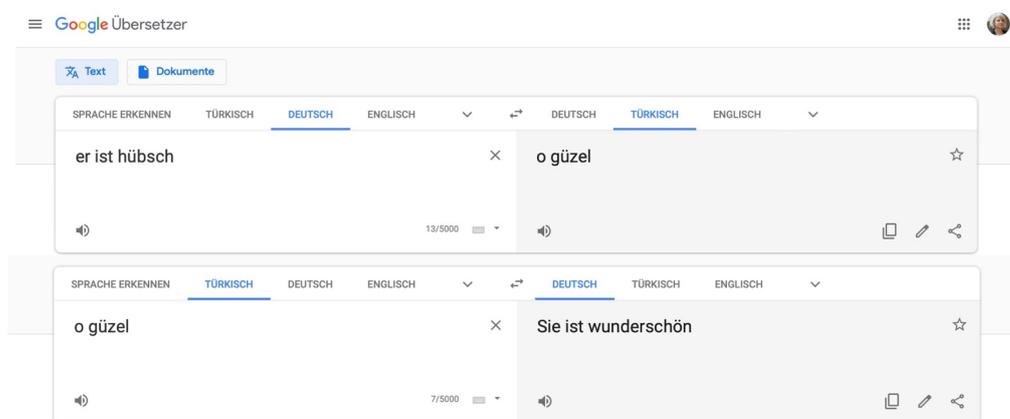
- Künstliche Intelligenz (KI) spielt in unserem Alltag eine immer größere Rolle. Zu Anwendungsbeispielen zählen unter anderem Routenplaner, Chatbots, Musik- oder Film-Empfehlungen, Gesichtserkennung, Online-Übersetzer oder Matching-Algorithmen auf Dating-Plattformen.
- Die Entwicklung von KI und Gestaltung von Technikprodukten liegt nach wie vor fest der Hand meist junger Männer. Laut UNESCO-Bericht zum Digital Gender Gap (<https://en.unesco.org/ld-blush-if-i-could>) geht zum Beispiel nur jedes 14. IT-Patent an eine Frau. Im Bereich Machine Learning – einer Hauptsparte der Künstlichen Intelligenz – sind nur 12 Prozent der Forscher\*innen weiblich.
- Eine neue KI-Anwendung, die sich derzeit rasant verbreitet, ist Bildsynthese in Form automatisierter Text-to-Image-Systeme wie etwa Dall-E 2 (<https://openai.com/dall-e-2/>), Imagen (<https://imagen.research.google>) oder Dall-E Mini (<https://github.com/borisdayma/dalle-mini>). Auf Basis statistischer Analysen riesiger Bild- und Text-Datenbanken und großen Energieaufwänden erlernen diese Systeme, natürlichsprachliche Texteingaben zu interpretieren und dazu passende Bilder zu erstellen. Ein Problem dabei ist unter anderem, dass stereotype Darstellungen von Männern und Frauen, die in den Datenbanken vorhanden sind, durch visuelle Outputs dieser Systeme reproduziert werden können. Siehe dazu etwa folgende von Dall-E Mini generierte Bildbeispiele, in denen „computer science student“ stereotyp männlich, „secretary“ jedoch stereotyp weiblich visualisiert wird:



## Geschlechter-Gerechtigkeit im Bereich KI: Drei aktuelle Challenges

### Challenge 1: Stereotype in Datensätzen

- Datengetriebene KI (Machine Learning, Deep Learning, Neuronale Netze) folgt Zielen, die von Menschen vorgegeben sind, und lernt aus Daten, die ebenfalls menschengemacht sind. KI-basierte Übersetzungsprogramme wie Google Translate lernen zum Beispiel aus Millionen von Texten, die Menschen in der Vergangenheit geschrieben haben. Algorithmen zur Personalauswahl lernen aus Personal-Entscheidungen, die Menschen in der Vergangenheit getroffen haben.
- Es ist daher nachvollziehbar, dass KI-Systeme per se nicht objektiver oder neutraler als wir Menschen selbst sein können, sondern dass sie uns im Gegenteil als Gesellschaft den Spiegel vorhalten. Vorurteile, Fehler oder veraltete Ansichten zu Geschlechterrollen, die in unseren Daten stecken, werden von KI-Systemen wiedergekaut. Siehe dazu folgendes Übersetzungsbeispiel von Google Translate:



### Challenge 2: Mangelnde Repräsentanz in Datensätzen

- Weil es in Datensätzen, mittels denen KI-Systeme trainiert werden, teilweise an Repräsentativität hinsichtlich aller Nutzer\*innengruppen mangelt, kann es passieren, dass nicht alle Personengruppen ausreichend mit ihren jeweiligen Charakteristika und Bedürfnissen darin vertreten sind. Ein Praxisbeispiel, das die potenziell negativen Konsequenzen mangelnder Repräsentanz in Datensätzen für Nutzer\*innen vor Augen führt, sind Sprachassistenzsysteme, die laut zahlreicher anekdotischer Berichte Sprachbefehle älterer Frauen schlechter verstehen als etwa solche von jüngeren Männern.
- Siehe dazu Alters- und Geschlechtsverteilung auf der Plattform Mozilla Common Voice:



### Challenge 3: Überholte Rollenbilder im Tech-Design

- Veraltete Geschlechter-Klischees scheinen sich aktuell über den Umweg neuer Technologie-Designs wieder in unseren Alltag zurückzuschleichen. Sprachassistenzsysteme wie Alexa (Amazon) oder Siri (Apple), die rund um die Uhr per Zuruf für den User bereitstehen, oder soziale Roboter für den Pflegebereich werden häufig stereotyp weiblich konnotiert und von Designer\*innen etwa mit weiblichen Namen oder Stimmen versehen.
- Forschung am Robopsychology Lab der JKU zeigt, dass männliche Nutzer im Durchschnitt KI-Assistenzsysteme bevorzugen, die über eine weibliche Stimme verfügen und gleichzeitig passiv und folgsam auftreten, während weibliche Nutzerinnen solchen stereotypen Designs großteils ablehnend gegenüberstehen (Moradbakhti, Schreibelmayr & Mara, 2022). Der Markt scheint bisher diesen männlichen Präferenzen zu folgen.
- Sowohl auf Seite der Hersteller\*innen wie auch auf Seite der Nutzer\*innen wird es in Zukunft um Bewusstseinsbildung gehen, denn eine Reproduktion von veralteten Geschlechterrollen (hier beispielsweise „die passive, dienende Frau“) durch Technik-Gadgets kann auch für das gesellschaftliche Bild realer Frauen problematische Auswirkungen haben.

### Dringend notwendig: Deutlich mehr Diversität im KI-Bereich

#### Wir brauchen:

#### Bewusstsein über die Herausforderungen ...

... bei gleichzeitiger Anerkennung, dass sie sich nicht rein technisch lösen lassen

#### Mehr Diversität ...

... in den Daten, aus denen wir KI-Systeme lernen lassen

... in der demografischen Zusammensetzung von Entwicklungs-Teams

... in Perspektiven, Expertisen und Mindsets

... in der Ansprache von Zielgruppen für digitale Bildungsangebote

#### Kontakt:

Univ.-Prof.<sup>in</sup> Dr.<sup>in</sup> Martina Mara  
Head of LIT Robopsychology Lab  
Johannes Kepler Universität Linz  
[martina.mara@jku.at](mailto:martina.mara@jku.at)  
[www.jku.at/lit/robopsychology](http://www.jku.at/lit/robopsychology)

Initiative Digitalisierung Chancengerecht:  
[www.idc.vision](http://www.idc.vision)