

Vielfalt und Fairness in KI-generierten Bildern

Wie können Anwender von KI-Bildgenerierung zu mehr Vielfalt und Fairness beitragen? In diesem Handout erfährst du Tipps für eine inklusive und bias-bewusste Formulierung von Prompts.

Was ist Bias und warum ist das relevant?

Begriffsverständnis

Einige zentrale Begriffe für unsere Diskussion:

- Bias: Unbewusste Vorverurteilung bestimmter Gruppen.
- Stereotyp: Vereinfachte Annahme über eine Gruppe (z.B. Asiaten sind gut in Mathe).
- Diskriminierung: Ungerechtfertigte Benachteiligung von Gruppen.

Diese Definitionen sind vereinfacht und können je nach Kontext variieren.

Vielfalt und Repräsentation sind relevant für unsere Gesellschaft und Demokratie. KI sollte bestehende Diskriminierung und Benachteiligung nicht reproduzieren oder verstärken, sondern abbauen. Bias (Vorurteile) kann dazu führen, dass KI einzelne Gruppen unfair darstellt oder ausschließt. Generierte Inhalte können stereotype Rollenbilder und Klischees aus den Datensätzen reproduzieren.

Beispiele:

- Beschreibung von Ärzten als männlich manifestiert klassische Rollenbilder.
- Formulierung "der Nutzer" schließt nicht alle Geschlechter ein; Alternativ: "die Person".
- Ethnische Klischees wie "asiatische Mathematiker" werden manifestiert.

Auch im Bildoutput äußert sich Bias durch diskriminierende Darstellungen bestimmter Gruppen.

Beispiele:

- Sexualisierende Darstellung von Frauen bei vielen Anfragen.
- Übertriebene ethnische Stereotype bei bestimmten Berufen oder Kontexten.
- Abwertende Sprache bei marginalisierten Gruppen.

Was können Anwender tun?

- Sensibel bei der Formulierung von Prompts sein
- Geschlecht, Herkunft und andere Aspekte bewusst vielfältig darstellen
- Stereotype vermeiden, individuelle Beschreibungen nutzen
- Problematische Ergebnisse kritisch analysieren und melden
- Feedback an Entwickler geben, um Modelle weiterzuentwickeln

Durch bewusste Anpassungen in der Sprachwahl können wir schrittweise fairere und inklusivere Darstellungen fördern. Es ist jedoch wichtig zu betonen, dass dies ein kontinuierlicher Prozess ist, und solche Anpassungen, obwohl essenziell, nur ein Aspekt einer umfassenden Lösung darstellen.

Beispiele

Textgenerierung

Vorsichtige, bias-vermeidende Formulierung:

Prompt: "Kannst du mir einen Text über eine kleine Gruppe von Lern-Experten schreiben, die innovative Ansätze für die Wissensgesellschaft entwickeln, ohne stereotype Annahmen über Geschlecht oder ethnische Zugehörigkeit zu machen?"

Erklärung: Die Formulierung vermeidet stereotype Annahmen und fokussiert sich auf den Themenkomplex der Wissensgesellschaft.

Explizite Diversitätsformulierung:

Prompt: "Beschreibe eine Gruppe junger Lern-Experten, die sich zusammenschließen, um die Herausforderungen interdisziplinärer Lernnetzwerke zu meistern."

Erklärung: Diese Formulierung gibt spezifische Details und Kontexte an, die helfen, stereotype Annahmen zu vermeiden.

Formulierung für unterschiedliche Lebenssituationen:

Prompt: "Erzähl mir von einer Gruppe von Lern-Experten, die nach verschiedenen Pausen (z.B. Elternzeit, Sabbatical) in die Welt des lebenslangen Lernens zurückkehren und sich den dynamischen Anforderungen der lernenden Organisationen stellen."

Erklärung: Hier wird ein breiteres Spektrum an Erfahrungen durch die Integration unterschiedlicher Lebenssituationen abgedeckt.

Bildgenerierung

Geschlechtsneutrale Formulierung:

Prompt: "Ein Bild von einer Gruppe von Fachleuten, die um einen Tisch sitzen, mit Büchern, Laptops und Notizbüchern vor sich. Im Hintergrund sind Diagramme und Mind-Maps sichtbar, die den Prozess des lebenslangen Lernens darstellen."

Erklärung: Diese Beschreibung legt kein spezifisches Geschlecht fest.

Diversität in Altersgruppen:

Prompt: "Ein Bild von einer Gruppe von Lern-Experten unterschiedlichen Alters und globaler Herkunft, die gemeinsam an einem Workshop zum Thema 'Lernende Organisationen' teilnehmen. Sie sind in einem hellen Raum und diskutieren lebhaft anhand eines Flipcharts."

Erklärung: Dieser Prompt stellt sicher, dass das Bild verschiedene Altersgruppen und Herkünfte zeigt.

Integration von Vielfalt:

Prompt: "Ein Bild einer Gruppe von Fachleuten, die an einem interaktiven Workshop über Lernkulturen teilnehmen. Sie repräsentieren eine Vielfalt von Geschlechtern, Altersgruppen und kulturellen Hintergründen und nutzen verschiedene Lernwerkzeuge und -technologien."

Erklärung: Dieser Prompt sorgt für Vielfalt in Geschlecht, Alter, Hautfarbe und Fähigkeiten.

An die Arbeit! Experimentiere mit den genannten Prompts:

- Teste verschiedene Kombinationen von Geschlecht, Alter, Ethnie und Fähigkeiten.
- Fokus auf positive Attribute und vermeide negative oder fehlende Beschreibungen.
- Experimentiere mit unterschiedlichen Szenarien und Aktivitäten, um Stereotypen herauszufordern.