

- openHPI: ChatGPT: Was bedeutet generative KI für unsere Gesellschaft? -

Die Rolle von Zielsetzung, Kontrolle und Tests

Johannes Hötter
Christian Warmuth

Wichtigkeit der Zielbeschreibung

- Akkurate **Zielbeschreibung** essentiell um gute, den eigenen Zielen entsprechende Ergebnisse zu erhalten
- Wandel von Erstellung der Ergebnisse zur **Formulierung** des **Zieles** durch Menschen
- Beispiel: Prompt Engineering



Die Rolle von Zielsetzung, Kontrolle und Tests

Johannes Hötter,
Christian Warmuth

openHPI

Rahmenbedingungen, Grenzen und Regeln



- Formulierung von **Rahmenbedingungen**, Grenzen und Regeln essentiell um **gesamtgesellschaftlich** erstrebenswerte Ergebnisse zu erhalten
- Problem: Schwierigkeit (vergleichbar mit Social Engineering)
- Beispiel: Jailbreaks und anderes ungewolltes Verhalten

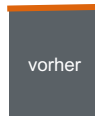
Die Rolle von Zielsetzung, Kontrolle und Tests

Johannes Hötter,
Christian Warmuth

openHPI

Kontrolle und die Versuchung

- Aus eigener Erfahrung:
Waschmaschine im Urlaub mit der Aufschrift
"Der orangefarbene Hellseher brennt"
- Korrekte Übersetzung aus dem
französischen: "Die orangene LED leuchtet"
- Gefahr: "Blindes Vertrauen auf Technologie"



Ergebnis



**Die Rolle von
Zielsetzung,
Kontrolle und Tests**

Johannes Hötter,
Christian Warmuth

openHPI

Folie 4

Kontrolle und die Versuchung

- Ergebnisse bereits so gut, dass die **Versuchung sehr hoch** ist, Inhalte weniger zu kontrollieren
- Kontrolle und Feedback essentiell, um Ergebnisse zu verbessern
- Verändert sich weiter mit höherer Qualität



Die Rolle von Zielsetzung, Kontrolle und Tests

Johannes Hötter,
Christian Warmuth

openHPI



Appell: Die Wichtigkeit von korrekten Zielstellungen, Kontrolle und Tests sind nicht zu unterschätzen in der generativen KI!

- openHPI: ChatGPT: Was bedeutet generative KI für unsere Gesellschaft? -

Die Rolle von Zielsetzung, Kontrolle und Tests

Johannes Hötter
Christian Warmuth