

Sicherer und verantwortungsvoller Umgang mit Sprachmodellen

- erfordert rechtliche, ethische und technische Rahmenbedingungen sowie die Einhaltung spezifischer Grundregeln.

Rahmenbedingungen

- Nutzer*innen müssen sich der rechtlichen und ethischen Vorgaben bewusst sein.
- Wichtige Gesetze: DSGVO, Urheberrecht, AGG.
- Verantwortung der Nutzer*innen: Überprüfung der Rechtskonformität und Abwägung der Nutzung im Kontext.

Grundregel 1 - Keine personenbezogenen Daten

- Anfragen dürfen keine personenbezogenen Daten enthalten.
- Beispiel für unzulässige Anfrage: E-Mail-Adresse einer Person.
- Angemessene Anfrage: Allgemeine Kontaktinformationen.

Grundregel 2 - Keine Bezugnahme auf natürliche Personen

- Anfragen sollten keinen direkten Bezug zu realen Personen haben.
- Beispiel für unzulässige Anfrage: Informationen über eine bestimmte Person.
- Angemessene Anfrage: Allgemeine Fähigkeiten für ein Studium.

Grundregel 3 - Vermeidung von personenbeziehbaren Daten

- Anfragen dürfen keine Daten enthalten, die Rückschlüsse auf Personen zulassen.
- Beispiel für unzulässige Anfrage: Informationen über eine Person in einer Vorlesung.
- Angemessene Anfrage: Allgemeine Themen in Biologievorlesungen.

Grundregel 4 - Respektierung des Urheberrechts

- Anfragen dürfen keine urheberrechtlich geschützten Inhalte anfordern.
- Beispiel für unzulässige Anfrage: Volltext eines Romans.
- Angemessene Anfrage: Zusammenfassung und Analyse eines Werkes.

Grundregel 5 - Wahrung des Rechts am eigenen Bild

- Bilder oder Informationen, die Personen identifizierbar machen, dürfen nicht ohne Zustimmung verwendet werden.
- Beispiel: Erstellung von Porträts ohne Einwilligung.

Grundregel 6 - Diskriminierung vermeiden und verhindern

- Anfragen und Antworten dürfen nicht diskriminierend sein.
- Beispiel für unzulässige Anfrage: Stereotype Annahmen über soziale Gruppen.

Überprüfung der Antworten auf Richtigkeit

- Nutzer*innen müssen die Richtigkeit der Informationen aktiv überprüfen.
- Wichtigkeit der Überprüfung: Vermeidung von Fehlinformationen und Erkennung von Halluzinationen.

Umgang mit potenziell unangemessenen Inhalten

- Kritische Beurteilung der Angemessenheit jeder Antwort ist erforderlich.
- Bereitschaft, unangemessene Antworten zu hinterfragen und zu verwerfen.

Überwachung der Einhaltung

- Einhaltung der Nutzungsbedingungen wird durch stichprobenartige Kontrollen überwacht.
- Konsequenzen bei Verstößen: Verwarnungen bis hin zu rechtlichen Schritten.

Abschluss und Zusammenfassung

- Appell an Nutzer*innen zur Beachtung der Richtlinien.
- Verantwortungsvoller Umgang mit Technologie fördert Ethik, Recht und Datenschutz.