

Bullshit-Zyklen bei Sprachmodellen erkennen und vermeiden



<https://pixabay.com/de/illustrations/schriftzug-mist-schwachsinn-483016/>

Ziele:

- den Bullshit-Zyklus und dessen Risiken im Alltag verstehen
- Chancen und Risiken beim Einsatz von Sprachmodellen erkennen und diskutieren
- Strategien zur Vermeidung des Bullshit-Zyklus entwickeln

Beschreibung:

Sprachmodelle wie ChatGPT liefern ihren Nutzern eine breite Anzahl an verschiedenen Einsatzszenarien, bei denen sich durch den Einsatz dieser Systeme Arbeitsaufgaben erleichtern und beschleunigen lassen. So könnte eine Lehrkraft das System nutzen, um eine Unterrichtsidee zu generieren, diese Idee direkt in ein Arbeitsblatt umsetzen und die Schülerantworten am Ende durch das System bewerten lassen. Die Schüler wiederum könnten dieses Arbeitsblatt mit einem privaten Account durch ChatGPT komplett lösen lassen und der Lehrkraft zurückgeben. Ein solcher Zyklus aus Aufgabenerstellung, -lösung und -bewertung allein durch ein System wie ChatGPT kann allerdings als „Bullshit-Zyklus“ angesehen werden, da dort allein ChatGPT als „Bullshit“-Maschine (vgl. McQuillen (1)) die Arbeit erledigt hat.

Ziel ist es, dass Schülerinnen und Schüler erkennen, dass es bei der Nutzung von Systemen wie ChatGPT zu potentiell kritischen Prozessen kommen kann, bei denen einzig eine künstliche Intelligenz ohne menschliche Interaktion die Inhalte produziert und über deren Korrektheit und deren Nutzen urteilt. Zur Aufklärung über solche Prozesse hat das IQSH ein kurzes Video produziert, das Schülerinnen und Schüler unter <https://go.iqsh.de/ki-bszyklus> finden und entweder im Klassenraum oder zu Hause ansehen können.

Danach können Schülerinnen und Schüler durch die Lehrkraft aufgefordert werden, andere Beispiele für potentiell kritische Beispiele dieses Zyklus' zu finden. Die Ergebnisse können dann in der Klasse diskutiert werden. Aus den Beispielen können dann weitere Diskussionspunkte über die Risiken beim Einsatz von Sprachmodellen abgeleitet werden. Es gilt jedoch unbedingt darauf zu achten, dass die Diskussion hier auch auf potentielle Chancen beim Einsatz von Sprachmodellen im Alltag eingeht. Außerdem kann über Strategien zur Vermeidung von kritischen Prozessen wie dem Bullshit-Zyklus diskutiert werden.

Quellen: (1) McQuillen: <https://www.transformingsociety.co.uk/2023/02/10/chatgpt-the-worlds-largest-bullshit-machine/>

Ablauf:

1. Die Schülerinnen und Schüler schauen das Aufklärungsvideo über den Bullshit-Zyklus bei ChatGPT
2. Die Schülerinnen und Schüler entwickeln in Gruppen eigene Anwendungsfälle für diesen Zyklus im Alltag
3. Die Gruppen stellen ihre Beispiele in der Klasse vor und diskutieren diese gemeinsam.
4. Die Klasse diskutiert über Chancen und weitere Risiken bei der Verwendung von Sprachmodellen im Alltag
5. Optional können mögliche Strategien beim Umgang mit Sprachmodellen diskutiert werden, um kritische Prozesse wie den Bullshit-Zyklus zu vermeiden.
6. Es gibt ein kurzes Fazit am Ende der Stunde.

	Gruppenarbeit Plenum
	Einstieg Erarbeitung
	Denken Austauschen Besprechen
	Austauschen Argumentieren Transfer
	K2 Kommunizieren und Kooperieren K3 Produzieren und Präsentieren K5 Problemlösen und Handeln

Differenzierung und sonderpädagogische Hinweise

- ✓ Aufgabe
Das Video soll zum Start als Hausaufgabe durch die SuS angeschaut und in der nächsten Stunde erklärt werden (Flipped Classroom).
- ✓ Ausrüstung
Präsentationstechnik zur Vorführung des Videos in der Klasse.
- ✓ Sozialform
Die Aufgaben können in Einzel-, Partner-, oder Gruppenarbeit bearbeitet werden.

Was benötige ich zur Erstellung/Durchführung?

- ✓ Endgerät
- ✓ Präsentationstechnik (z.B. Beamer oder Display)
- ✓ Internetzugang

Was benötigen Schülerinnen und Schüler?

- ✓ ggf. Endgerät
- ✓ ggf. Internetzugang

Benötigte Software

- ✓ Internetbrowser

Ähnliche Methoden

- Rollenbilder und Ansichten von Sprachmodellen
- Suchen mit Sprachmodellen
- KI-Texte im Faktencheck
- Unterstützung von Schreibprozessen