

Unterrichtsidee Deutschunterricht mit KI Interpretation gegen die KI

Fach	Deutsch
Jahrgangsstufe	10–12 (Oberstufe empfohlen)
Zeitbedarf	90 Minuten (Doppelstunde) + 45 Min. Folgestunde (empfohlen)
Denklinien	KI als Analysegegenstand + KI als Anlass für Urteilkraft
ALLit-Domänen	E (Erkunden & Verstehen) + B (Bewerten & Einordnen) + V (Weiterdenken & Verantworten)
Tools	ChatGPT, Claude oder vergleichbar; ggf. zwei verschiedene Modelle zum Vergleich

Kernidee

Schülerinnen und Schüler interpretieren einen literarischen Text zunächst eigenständig nach fachlichen Methoden. Erst danach lassen sie denselben Text von einer KI interpretieren – und vergleichen die Ergebnisse systematisch. Der Vergleich macht sichtbar, worin die spezifische Qualität menschlicher Interpretation besteht (kontextuelle Deutung, persönliche Lesart, Textbezogenheit der Belege) und wo KI-Interpretationen typische Muster zeigen (Paraphrase statt Deutung, nicht belegte Verallgemeinerungen, fehlender historischer Kontext, scheinbare Vollständigkeit).

Lernziele

- (1) SuS vertiefen ihre eigene Interpretationskompetenz durch den Vergleich mit einer KI-Interpretation.
- (2) SuS identifizieren Qualitätsmerkmale einer fachlich tragfähigen Interpretation und grenzen sie von oberflächlicher Zusammenfassung ab.
- (3) SuS erkennen typische Muster von KI-Interpretationen (z. B. Abwesenheit kontextueller Tiefe).
- (4) SuS reflektieren die Grenzen von KI bei hermeneutischen Aufgaben und formulieren begründet, was nur menschliche Interpretation leisten kann – oder ob diese Grenze schwächer ist als erwartet.
- (5) SuS untersuchen, welche kulturellen Perspektiven und Interpretationstraditionen im KI-Output sichtbar (oder unsichtbar) werden, und reflektieren die Implikationen für den Umgang mit KI-generierten Deutungen.
- (6) SuS formulieren eine eigene, begründete Haltung zum Einsatz von KI bei Interpretationsaufgaben.

Fachdidaktische Prüffragen

⚠ Bitte durch Deutschkolleg:innen prüfen

- (1) Ist die Reihenfolge (erst eigene Interpretation, dann KI-Vergleich) fachdidaktisch sinnvoll, oder besteht das Risiko, dass SuS ihre eigene Deutung nachträglich entwerten, wenn die KI flüssiger formuliert?
- (2) Welche Textsorten/Textlängen eignen sich besonders gut für diesen Vergleich? Ist Lyrik zu voraussetzungsvoll, oder ist sie gerade deshalb besonders ergiebig?
- (3) Passt das Vergleichsraster (a)–(f) zu den in der Oberstufe üblichen Interpretationskategorien, oder sollte es an ein etabliertes Raster (z. B. EPA-Operatoren) angepasst werden?
- (4) Besteht das Risiko, dass das Szenario das Interpretieren selbst delegitimiert („Warum soll ich das noch selbst machen?“) – und wie kann die Lehrkraft dem entgegenwirken?
- (5) Ist die Kategorie (f) „Kulturelle Perspektive“ für SuS angemessen operationalisierbar, oder braucht sie einen konkreteren Arbeitsauftrag (z. B. „Wie würde ein:e Literaturwissenschaftler:in aus Japan diesen Text lesen?“)?

Ablauf – Doppelstunde (90 Minuten)

Phase	Zeit	Beschreibung
Vorbereitung	Vor der Stunde	Lehrkraft wählt einen geeigneten literarischen Text (Kurzprosa, Gedicht oder Textauszug), der interpretativ ergiebig ist: mehrere Deutungsebenen, sprachliche Verdichtung, ggf. historischer Kontext. Der Text sollte den SuS nicht bereits bekannt sein. Die Lehrkraft generiert vorab eine KI-Interpretation (1–2 Prompts, Ergebnis drucken oder digital bereithalten) und prüft, ob diese ausreichend substantiell und zugleich typisch fehlerhaft ist.
Einstieg	5'	Lehrkraft teilt den literarischen Text aus, kein Hinweis auf KI. Stiller Leseauftrag.
Erarbeitung I	30'	Einzelarbeit (15 Min.): SuS verfassen eine eigene Deutungshypothese und stützen sie mit mindestens drei konkreten Textbelegen. Keine Hilfsmittel außer dem Text. Partnergespräch (15 Min.): Austausch der Deutungen, Diskussion abweichender Lesarten, ggf. Schärfung der eigenen Hypothese.
Konfrontation	20'	Lehrkraft teilt die KI-Interpretation aus (oder projiziert sie). SuS vergleichen die KI-Interpretation mit ihrer eigenen anhand des erweiterten Vergleichsrasters (siehe unten).
Sicherung, Reflexion & Transfer	25'	Plenum – Befunde sammeln (10 Min.): Wo war die KI überraschend gut? Wo war sie typisch schwach? Was kann menschliche Interpretation, das KI nicht kann – und ist das tatsächlich so, oder sind wir zu schnell zufrieden mit dieser Grenzziehung? Gemeinsame Formulierung: „Drei Dinge, die eine gute Interpretation ausmachen und die wir bei der KI vermisst haben.“ Transfer (5 Min.): Leitfrage: Was folgt aus dem Vergleich für mein eigenes Interpretieren? Was nehme ich für meine nächste Interpretation konkret mit? Haltung formulieren – Domäne V (10 Min.): SuS formulieren in Einzelarbeit eine persönliche Regel für den Umgang mit KI-Interpretationen (1–2 Sätze). Beispiel-Impuls: „Formuliere eine Regel, die du dir selbst geben würdest, wenn du das nächste Mal eine KI-Interpretation liest.“ 2–3 Regeln werden im Plenum geteilt.
Puffer	10'	Zeitreserve für Phasenübergänge, intensivere Plenumsdiskussion oder Vertiefung einzelner Raster-Kategorien. Falls nicht benötigt: Beginn mit der Prompt-Erprobung der Folgestunde.

Erwartbare Schülerprodukte

- (1) Eigene Deutungshypothese mit Textbelegen (Einzelarbeit).
- (2) Ausgefülltes Vergleichsraster mit sechs Kategorien (Partnerarbeit).
- (3) Persönliche KI-Interpretationsregel (Einzelarbeit, 1–2 Sätze).
- (4) Optional (Folgestunde): Verbesselter Prompt und Reflexion über Prompt-Qualität.

Vergleichsraster

SuS vergleichen die KI-Interpretation mit ihrer eigenen anhand der folgenden sechs Kategorien:

	Kategorie	Leitfrage
(a)	Deutungstiefe	Benennt die KI nur Inhalt – oder deutet sie auch? Gibt es eine echte These?
(b)	Textbelege	Zitiert die KI konkret und passgenau – oder bleibt sie allgemein?
(c)	Kontextwissen	Nutzt die KI historischen, biografischen oder literaturgeschichtlichen Kontext?
(d)	Ambiguität	Lässt die KI Mehrdeutigkeit zu – oder vereindeutigt sie vorschnell?
(e)	Eigenständigkeit	Enthält die KI-Interpretation etwas, das die eigene Deutung bereichert?
(f)	Kulturelle Perspektive	Wessen Interpretationstradition spiegelt die KI wider? Welche kulturellen Annahmen stecken im Output – und welche Perspektiven fehlen? Wie würde eine Deutung aus einem anderen kulturellen Kontext aussehen?

Warum Kategorie (f)?

LLMs sind „multilingual but monocultural“: Sie beherrschen viele Sprachen, bilden aber überwiegend westliche, anglophone Interpretationstraditionen ab. Im Deutschunterricht wird das greifbar, wenn die KI z. B. ein Gedicht der Romantik mit psychologisierenden Kategorien deutet, statt die philosophische oder religiöse Dimension zu sehen. Kategorie (f) macht diese Schlagseite zum Lerngegenstand – und erfüllt damit Goldene Regel 4 (Ethik ist eingebaut, nicht angehängt) und Planungsprinzip 4 (Wessen Perspektive fehlt?).

Folgestunde: Prompt-Erprobung (45 Minuten)

Phase	Zeit	Beschreibung
Rückblick	5'	Kurze Rekapitulation der Ergebnisse aus der Doppelstunde: Welche Schwächen hatte die KI-Interpretation? Was fehlte?
Prompt-Erprobung	25'	SuS prompten selbst und versuchen, durch gezielte Prompts eine bessere KI-Interpretation zu erzeugen. Varianten: (a) einzeln am eigenen Gerät, (b) in Partnerarbeit – eine:r promptet, eine:r protokolliert. Diskussionsfrage: Kann man eine KI durch gute Prompts zu einer guten Interpretation bringen? Was müsste man dafür wissen? Erkenntnis: Wer gut prompten will, muss selbst interpretieren können.
Meta-Reflexion	15'	Optional – Vertiefung: Vergleich zweier verschiedener KI-Modelle; oder SuS schreiben einen metareflexiven Text: „Was hat mir der Vergleich mit der KI über mein eigenes Interpretieren gezeigt?“

Quellenrahmen: OECD/EC AILit Framework (2025), Singapur MOE EdTech Masterplan 2030, UNESCO AI CFT (2024), Australian Framework for GenAI in Schools (2023/25), Code.org/TeachAI (2025).