

KI-Werkzeuge im Unterricht und in Prüfungen



Digitale Sprechstunde am 27. 1.2025

Klaus Knopper <klaus.knopper@hs-kl.de>
Vizepräsident für Digitalisierung an der Hochschule Kaiserslautern

Zum Kennenlernen: Umfrage 1

Wie oft verwenden Sie KI-Tools im Unterricht?

1) Was sind KI-Tools? 

2) nie

3) selten

4) mehrmals pro Woche

5) oft

6) immer

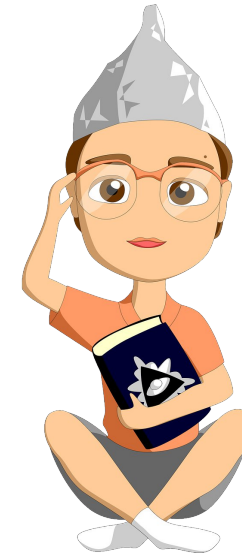
7) Ich bin eine KI. 

Zum Kennenlernen: Umfrage 2

Erlauben Sie die Verwendung von KI-Tools in Studienarbeiten?

- 1) Was sind KI-Tools?**
- 2) Nein, streng verboten**
- 3) Ja, aber nur nach individueller Absprache**
- 4) Generell ja, aber nach festgelegten Kriterien**
- 5) Ich erlaube alles.**

Experimente und Infos zum Kurs



→ <https://play.hs-kl.de/ki>

Künstliche Intelligenz

- ...ist nicht das, was man intuitiv damit verbindet („schlaue Maschinen“)
- Definition des [→ europ. Parlaments](#):
„**Künstliche Intelligenz** ist die **Fähigkeit einer Maschine, menschliche Fähigkeiten** wie logisches Denken, Lernen, Planen und Kreativität zu **imitieren**.“ ^{*)}
- Definition nach Stuart J. Russell und Peter Norvig^{**)} in vier unabhängigen Kategorien:
 - 1) **Nachbildung** von **menschlichem Denken**
 - 2) **Nachbildung** von **rationalem Denken**
 - 3) **Nachbildung** von **menschlichem Verhalten**
 - 4) **Nachbildung** von **rationalem Verhalten**

Während 1. und 2. eher eine „**Simulation menschlichen Verhaltens**“ als Ziel definieren (vergl. frühe erfolgreiche Versuche mit „**Eliza**“ von **Joseph Weizenbaum 1966**^{***)}), fließt in 2. und 3. auch eine menschenunabhängige Art „**Vernunft**“-Simulation ein.

^{*)} Was ist künstliche Intelligenz und wie wird sie genutzt? | Aktuelles | Europäisches Parlament. 14. September 2020, abgerufen am 18. Juli 2021.

^{**)} Stuart J. Russell, Peter Norvig: Künstliche Intelligenz: Ein moderner Ansatz. Pearson Studium, Berkeley 2004, ISBN 3-8273-7089-2 (englisch: Artificial Intelligence: A Modern Approach.).

^{***)} → <https://de.wikipedia.org/wiki/ELIZA>

Was bedeutet „generative KI“ eigentlich?

- Inhalte **erzeugen** (Input → Output, auch „Transformation“)
- I.d.R. mit vortrainiertem („pre-trained“) großem Sprachmodell („**Large Language Model**“)
- GPT: **G**enerative **P**re-trained **T**ransformer
- Funktionsweise (aktuell), eigener Definitionsversuch: **Berechnung** von **Wahrscheinlichkeiten** für auf **eine Frage** „am **besten passende**“ **Text/Grafik/Audio-Bausteine** („Tokens“) - mit Steuerungsmöglichkeit – nebst **Zusammenfügen einer „optimalen“ Antwort** mit einstellbarer Varianz („Zufall“), **ohne „semantisches Verständnis“ des generierten Inhalts** durch den Algorithmus.

KI-Tools sind toll...

- **Erzeugen** von **professionell** bis „perfekt“ **gestalteten, zielgruppenspezifischen Inhalten**.
- **Texte, Bilder, Sprache, Musik, Programmcode...**
- Im **Browser**, per **App**, mit **Sprachein-/ausgabe nutzbar**.
- KI-Tools **großer Firmen: Niederschwellig**, einfach zu bedienen (gelegentlich sogar **barrierearm**), **umfangreiche Datenbasis**, mit oder ohne **Anbindung an Internet-Recherche**.
- KI-Tools können **Lösungsansätze vorschlagen**, wenn man **selbst keine Idee hat**, wie eine Aufgabe zu lösen ist.
- KI-Tools können **sehr schnell Lehrmaterial generieren** (nein, diese Folien sind nicht von ChatGPT erstellt. ;-)



Die Philosophie-Seite: Weiß/versteht der KI-Algorithmus, was er tut?

- Falsche Frage. ;-)
- Je **ähnlicher** der Output einer **menschlichen Antwort** kommt, desto mehr hat man **den Eindruck**, dass die Maschine „denkt“ oder „versteht“ → **bessere Simulation**.
- Auch die Frage, ob eine **KI ein „Bewusstsein“ entwickeln** könnte, ist akademisch (zumindest, so lange man noch nicht einmal verstanden hat, **wie das menschliche Gehirn funktioniert**).
- Wenn es hilft/glücklich macht, sich **vorzustellen**, die KI würde den Menschen **persönlich verstehen** und könnte sogar ein*e „elektronische*r Freund/Freundin“ sein → *go for it*, aber **Awareness für die technischen Zusammenhänge ist wichtig, um Enttäuschungen und Gefahren zu vermeiden**.



Was ist denn das Problem? → Herausforderungen

- **Kommerzielles Interesse** der Anbieter liegt **nicht in der Präzision / universellen Einsetzbarkeit**, „Wahrheit“, sondern eher „**Gefallen**“ (Kundenbindung erwünscht)
→ **Faktencheck wichtiger als je zuvor!**
- Ohne eigenes Hintergrundwissen bzw. „begleitetes Surfen“ **für Nutzer schwierig**, zwischen „**Fake Information**“ und **Fakten zu unterscheiden**. Auch „**Halluzinationen**“ der KIs möglich und bei **unzureichendem eigenen Wissen schwer zu erkennen**.
→ **Kern-Aufgabe der Lehrer*innen kann systembedingt nicht von einer KI übernommen werden** (was in gewisser Weise eine beruhigende Art Jobsicherheit für uns ist, s.a. [→ Job-Futuromat](#))
- **Urheber*innenschaft** der Rohdaten i.d.R. unklar → Problem für Zitation
- **Eigene Leistung der Schüler*innen fair bewerten?**
- **Bias** → Art der **Fragestellung beeinflusst Ergebnis** („stets gefällige Antworten“)

Juristisches

- **DSGVO** (hatten wir schon):
Nutzung von KI-Tools von außereuropäischen Anbietern problematisch, auch wenn Rechenzentren im EU-Raum, da Inhalte ausgeleitet werden.
Integration von entsprechenden „Assistenten“ in Standardsoftware muss ggf. vor der Nutzung deaktiviert werden.
- **KI-Verordnung** (seit Juli 2024 in Kraft, aber noch nicht anwendbar):
→ **Amtsblatt der europäischen Union**
Risikobasierter Ansatz; Definition von sog. „**KI-Hochrisiko-Systemen**“ (→ **Anhang III**), für welche **erhöhte Anforderungen** und **Dokumentations-/Kennzeichnungspflicht** gelten, die aufgelisteten Hochrisiko-Bereiche enthalten u.a. auch Einsatz von KI in der Bildung, v.a. bei **KI-gestützter Bewertung von Leistungen** sowie **Zugangsprüfungen** und **Bewerber*innenauswahl**.
- **Urheberrecht** → S. Nächste Folien.



Ist ein KI-generierter Text ein Plagiat?

- Das **Urheberrecht** ist ein sog. **natürliches Recht**, das **ausschließlich Menschen** zusteht und auch ohne „Vertrag“ alleine **durch Schaffung eines Werks** entsteht bzw. anwendbar ist.
- Daher **können KIs keine Urheber** sein.
- Daher handelt es sich bei einem **KI-generierten Text / Bild etc. NIEMALS um ein Plagiat** („Aneignung von Leistungen anderer Menschen“), und die **„Plagiats-Kriterien“ greifen nicht**.
- **Menschengemachte Rohdaten**, die KIs als Basis verwenden, können sehr wohl **urheberrechtlich geschützt** sein. Da sie aber nur in „Schnipseln“ verwendet werden ist die **Signifikanz einzelner Werke im Output schier unmöglich zu klären**.
- **Unwissenschaftliche Zitation** bzw. **Verwendung von Elementen ohne Quellen-Nachweis** kann dennoch zur **Abwertung bis zum Nichtbestehen der Prüfung** führen.
- **Regeln für KI-generierte Inhalte** notwendig.

Kann man KI-generierte Inhalte, ggf. durch KI -Tools, erkennen?

- Die kurze Antwort: **Nein**, inzwischen (seit GPT-4*) wirklich nicht mehr.
- **Heuristische Ansätze** (Vergleich von KI-generierten Texten mit gleichem Thema mit der abgegebenen Arbeit, z.B. GptZero) **funktionieren nicht mehr**.
- **Paradox**: Durch **KI modifizierte Texte** werden u.U. von **KI-Checkern eher als menschlich bewertet**, als **menschengenerierte Originale**, neues Geschäftsmodell: „KI-based humanifier“ *)
- Da zuverlässiger Täuschungsnachweis bei KI-Verwendung unmöglich
→ **Verbote wirken nicht**.
- Lösungsansatz 1 (Prager Wirtschaftsuni): Abschaffung der schriftlichen Abschlussarbeiten (?) **)
- Lösungsansatz 2: Ausgewählte (oder alle?) **KI-Tools erlauben** und **Regeln festlegen** plus **Aufwertung eines Pflicht-Kolloquiums zur schriftlichen Arbeit**.

*) → <https://www.bloomberg.com/news/features/2024-10-18/do-ai-detectors-work-students-face-false-cheating-accusations>

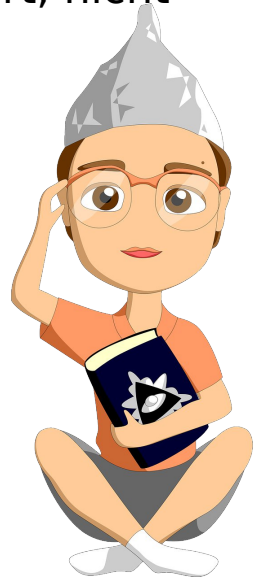
**) → <https://www.zdf.de/nachrichten/politik/ausland/chatgpt-kuenstliche-intelligenz-bachelorarbeit-100.html>

Unser Ansatz zur Regulierung (HS-KL)

- **Art und Umfang** von **KI-generierten** oder **KI-modifizierten** Inhalten zu **erlauben oder auch nicht, obliegt den Prüfenden**.
- Die **Prüflinge** erhalten **vor jeder Prüfung von den Prüfenden** die **hierfür geltenden Vorgaben**, als [→ verbindliche Handreichung](#).
- **Beispiele zu sinnvollen Zitationsformaten** bei KI-Elementen (z.B. **verwendeter Prompt, Version der verwendeten Software, vollständige Inputs/Outputs im elektronischen Anhang**).
- **Ziel: Nachvollziehbarkeit / Plausibilität für die Prüfenden herstellen.**
- Entbindet die **Prüflinge** nicht vom **Fakten-Check** (→ ChatGPT ist keine Primärquelle)
- Entbindet die **Prüfenden** nicht vom **Fakten-Check** (→ Nachvollziehbarkeit der von der KI mitverwendeten Primärquellen).
- **Wenn nicht nachprüfbar**, sind **KI-generierte Abschnitte in Frage zu stellen** (→ unwissenschaftliche Zitation).

Schlimme Fehler

- Einen **Chatbot** nach seiner Meinung fragen.
- „**Bias**“ bereits **in der Frage** („Erkläre mir, warum ... schlecht / gut ist“). → Demo
- **Chatbot** wie eine „**Suchmaschine**“ benutzen (Ergebnisse sind immer generiert, nicht zitiert)
- Aktuellen **Kontext** / Rolle bei **Chatbots** nicht kennen
- ...



Neue Aufgaben

- **Digitalkompetenz stärken** (auch Verständnis fördern, wie KI-Tools arbeiten, was sie können, was nicht). → Nutzer*innen **nicht** nur **Konsumenten**, sondern **aktive Mitgestalter**.
- **Kritische Auseinandersetzung** mit **KI-generierten Inhalten** fördern (→ s.a. „EvilGPT“-Chatbot)
- **Datenschutz-Awareness** („Woher stammen die Daten, wohin gehen unsere Eingaben?“ → wichtig für Forschungsfragen → Gültigkeit von transatlantischen Datenschutzabkommen jetzt und zukünftig fraglich → US „Patriot Act“, „Cloud Act“ ff.)
- **Open Source LLMs** für spezielle Aufgaben, z.B. eigene Inhalte der (Hoch-)Schule und Lehrmaterialien einbinden („Erklär mir mit einfachen Worten ... und zeige mir, wo es in welchem Dokument steht.“, s.a. FAIRD-Projekt der HS-KL) *)

*) Hier geht es NICHT nur um den Kosten-Aspekt, sondern um digitale Souveränität und eigene Gestaltungsmöglichkeiten.

KI-Strategie

- Kickoff zur Mitarbeit an der zukünftigen Digitalisierungs- und KI-Strategie fand bereits am 27.11.2024, multi-Hybrid an drei Standorten und online.
- Offener und transparenter Prozess, Teilnahme immer noch möglich, Mattermost-Team: → <https://mattermost.gitlab.rlp.net/hs-kl/channels/ki>

(oder “digitalisierung” allgemein)



ENDE

Dieser Vortrag steht unter einer Creative Commons Lizenz



<http://creativecommons.org/licenses/by-nd/4.0/>

Klaus.Knopper@hs-kl.de