

KI im Hochschulkontext: Chancen und Herausforderungen

Martin Drossos

Weiterbildungsangebote zu KI in der Hochschullehre

19.01.2024

■ **Inputphase:**

- Was ist eine KI?
 - Begrifflichkeit, Gegenstand, Arten
- Betrachtung in verschiedenen Sachverhalten
 - Entwicklung und Training
 - Prompting
 - Ergebnisnutzung
- Folgen für Prüfungen und Lehre

■ **Diskussion:**

- Schritt 1.: Sammeln und Strukturieren von relevanten Fragen und Problemen
- Schritt 2. Überlegen von Antworten und Lösungen

■ **Zusammenfassung**

Gliederung

The image features a complex, three-dimensional network of glowing lines and nodes. The nodes are small, bright spheres, and the connections are thin, glowing lines. The overall color palette is dominated by deep blues and purples, with a gradient of reds and pinks appearing towards the right side. The network has a layered, funnel-like structure, suggesting a neural network architecture. A dark blue rectangular box with a thin cyan border is positioned on the left side of the image, containing the text "Inputphase" in white. The background is dark, making the glowing network stand out prominently.

Inputphase



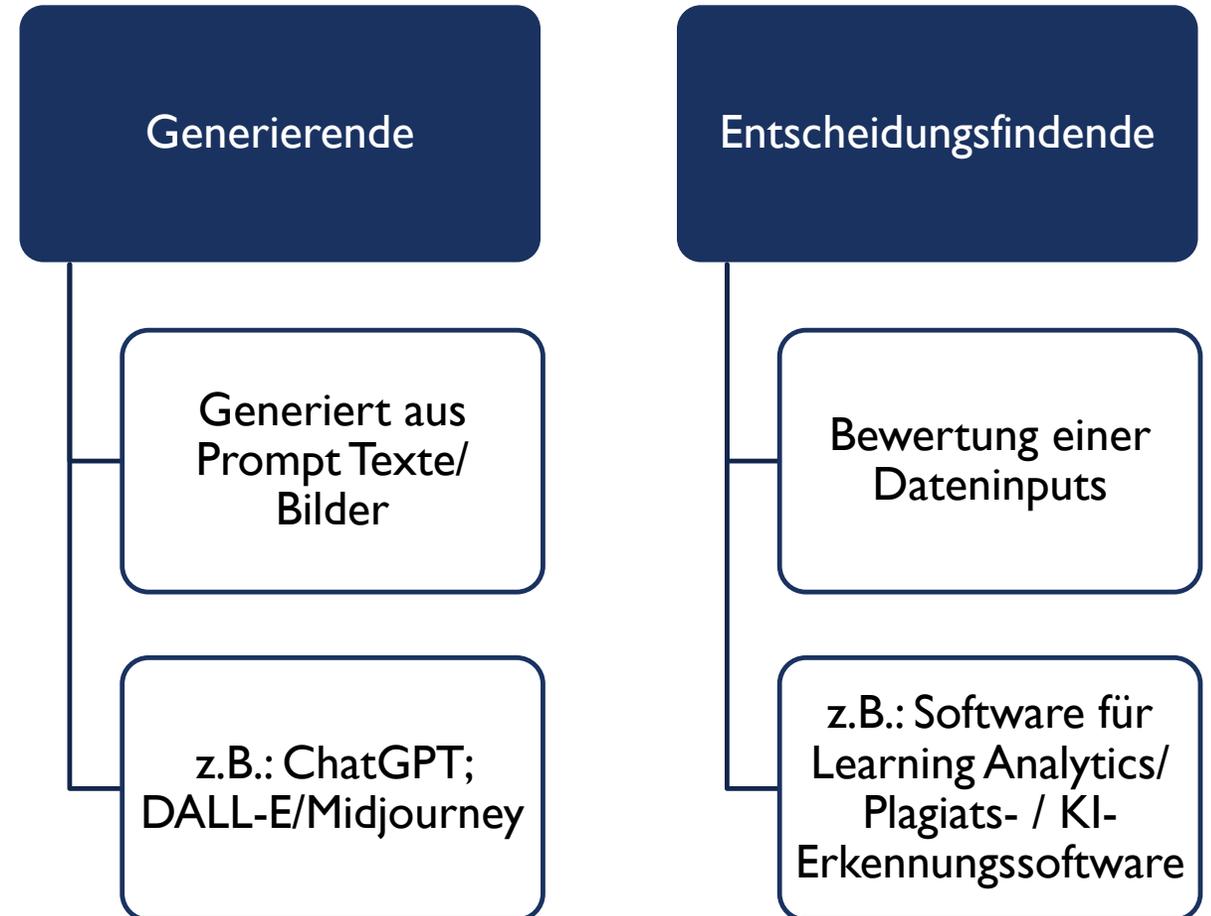
Was ist eine KI?
Begrifflichkeit, Gegenstand, Arten

KI? Worüber genau reden wir?

- KI als Zusammensetzung unterschiedlicher Methoden, um ein Problem möglichst effizient zu lösen.¹
- Software, die mit einer oder mehreren (...) Techniken und Konzepte entwickelt worden ist und (...) Ergebnisse (...) hervorbringen kann, die das Umfeld beeinflussen, mit dem die interagieren.²
- Arten von KI

¹Stuart Russel/Peter Norvig, Artificial Intelligence, Global Edition, A Modern Approach (2021), S. 21; 670.

²Art 3 Abs. 1 Nr. 1 KI- EUVO; <https://eur-lex.europa.eu/legal-content/DE/TXT/?uri=CELEX%3A52021PC0206>



Rechtliche Klassifizierung von KI

- KI- **Verordnung** der EU
 - Einstufung von KI in vier Risikoklassen ausgehen von der Gefahr für best. Rechtsgüter
 - Verbotene KI bzw. unannehmbares Risiko (Art 5 KI- EUVO)
 - Hochrisiko KI (Art. 6f. KI-EUVO)
 - Geringes Risiko
 - Minimales Risiko
 - Auferlegung von Sorgfalsmaßstäbe für die versch. Risikoeinstufungen
- Zu unterscheiden von der EU **Richtlinie** über die KI- Haftung!
 - Soll Haftungslücken beheben und Beweislasten regeln

Generative Pre-trained Transformers

- Künstliche neuronale Netzwerke
- Basieren auf der Transformer-Architektur
- Vorabtraining mit großen Datensätzen
- Fähig zur Generierung neuartiger, menschenähnlicher Inhalte
- „Stochastische Papageie“¹
- <https://www.soekia.ch/GPT/>
 - Veranschaulichung der Funktionsweise von GPT
 - Ist Innovation beim „wahrscheinlichsten Ergebnis“ möglich?

¹ Emily M. Bender, Timnit Gebru, On the Dangers of Stochastic Parrots, S. 6 | 6f.: <https://dl.acm.org/doi/pdf/10.1145/3442188.3445922> ; Charlotte Schubert, Eine Unterhaltung mit ChatGPT: Dataistisches Paradigma, stochastischer Papagei oder einfach nur ein Mißverständnis?, <https://journals.ub.uni-heidelberg.de/index.php/dco/article/view/95246>



Quelle: <https://ideogram.ai/t/trending>
Prompt: GPT stochastic parrot

KI Detektoren

- Unterschied zur Plagiat- Erkennung
- Fundierte wissenschaftliche Untersuchung von KI- Detektoren durch die AG Technologie und Wissenschaftliche Integrität der **European Network for Academic Integrity (ENAI)**: <https://arxiv.org/abs/2306.15666>
 - Können KI-Detektoren zuverlässig von einem Menschen geschriebenen Text erkennen?
 - Können KI-Detektoren zuverlässig einen KI-generierten Text erkennen?
 - Beeinflusst die maschinelle Übersetzung die Erkennung eines vom Menschen geschriebenen Textes?
 - Beeinflusst manuelles Editieren oder maschinelles Paraphrasieren die Erkennung von KI-generierten Texten?
 - Wie konsistent sind die Ergebnisse der verschiedenen KI-Detektoren?
- Ergebnis: große Anzahl an falsch positive und falsch negative Ergebnisse sowie ungenaue Aussagen bzgl. aller Angaben von KI- Detektoren

KI Detektoren

- Schlussfolgerungen und Empfehlungen¹
- Beweis von Täuschungen²
 - Die Prüfungsbehörde trägt die materielle Beweislast
 - Verschiebung der Beweislast bei begründetem Verdacht
 - Regel des Anscheinsbeweises
 - Heranziehung von Indizien problematisch; Spannungsverhältnis zwischen exzellenter Leistung und Täuschung
- Mögliche betroffene Rechtsgüter beim Einsatz von KI- Detektoren
 - Bei falsch negativen Ergebnissen ca. 19%: Keine eigenständige Leistung
 - Bei falsch positiven Ergebnissen ca. 2 %: Falsche Verdächtigung

¹ Constanze Kurz, Debora Weber-Wulff, Maschinelles Lernen: Nicht so brillant wie von manchen erhofft, 2023, <https://netzpolitik.org/2023/maschinelles-lernen-nicht-so-brillant-wie-von-manchen-erhofft/>

² Niehues, Fischer, Jeremias, Dieterich, Prüfungsrecht. 2022, Rn. 236f., 869ff.

Automatisierte Prüfungsaufsicht

- KI schätzt anhand von Bild- und Tonaufnahmen ein, ob ein Täuschungsversuch vorliegen könnte.
- Die Zulässigkeit ist sehr umstritten!¹
- Eine Rechtsgrundlage für eine automatisierte Videoaufsicht wird unter bes.Voraussetzungen in manchen LHG vorgesehen.
- Die Landesdatenschutzbeauftragte einiger Länder halten die Videoaufsicht für unzulässig.
- Mögliche datenschutzrechtliche Hindernisse: Art 22, 9 DSGVO

¹ Jonas Botta, Grundrechtseingriffe durch Online-Proctoring, 21.12.2020, <https://verfassungsblog.de/grundrechtseingriffe-durch-online-proctoring/>
Dagegen: Zulässigkeit der Aufzeichnung: OVG NRW, Beschl. v. 04.03.2021 – 14 B 278/21.

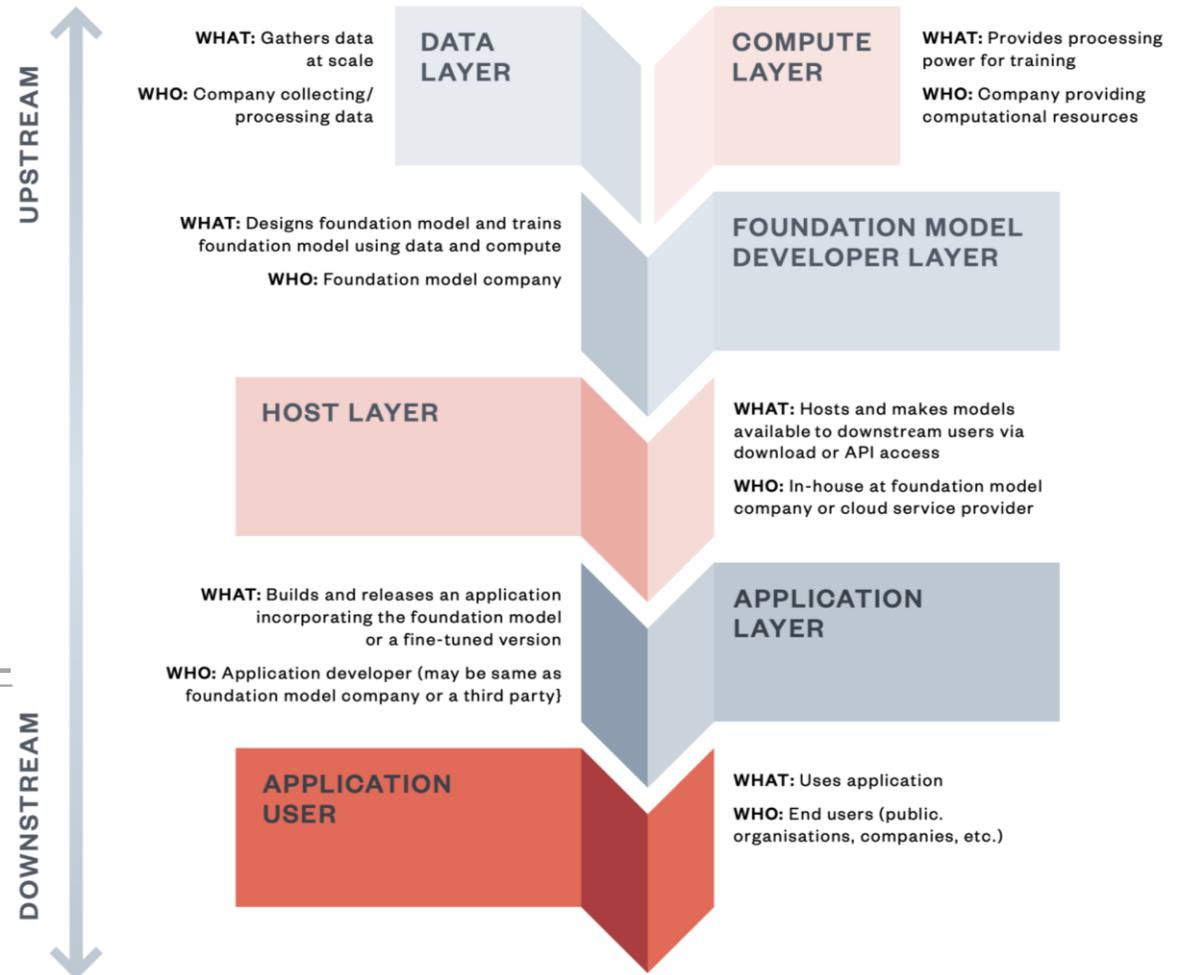


Umgang mit KI in verschiedenen Phasen

Entwicklung und Training der KI

- Komplexes, mehrstufiges Verfahren
- Sehr hohe Kosten
- Viele Beteiligte mit unterschiedlichen Aufgaben
 - Verantwortung und Haftung (KI- EUVO und EURL)
- <https://www.adalovelaceinstitute.org/resource/foundation-models-explainer/>
 - Weiterführender Artikel von Elliot Jones
 - Grafikquelle

Foundation model supply chain



Note: This is one possible model (there will not always be a separate or single company at each layer)

Entwicklung und Training der KI

Urheberrecht

- Werke i.S.v. § 2 UrhG in den Trainingsdaten?
- Nutzungsrechte
 - Erlaubnis
 - Gesetzliche Schranken
 - § 44a UrhG (vorübergehende Vervielfältigung)
 - § 44b UrhG (Text und Data Mining)
 - § 60d UrhG (wissenschaftliche Forschung)

Ethische Fragen

- Förderung von Stereotypen, Ungleichheit und Bias
- Trainingsdatenhygiene
- Datenqualität vs. Repräsentanz

Relevanz in Hochschulen

- Notwendige Sensibilität für die Urheberrechtsverletzung soll bei den Studierenden geschaffen werden.
- Problematik der Nutzung von Prompteingaben zum Training

Prompting

Geistiges Eigentum

- Eingabe urheberrechtlich geschützter Werke
- Prüfer*in lädt Prüfungsleistung hoch, um diese durch die KI überprüfen zu lassen
- Studierende bittet KI Aufsätze zu verarbeiten, um daraus eine Hausarbeit zu erstellen.

Datenschutz

- Verbot der Verarbeitung nach Art. 58 II f) DSGVO
 - Fehlende Rechtsgrundlage
 - Art. 6 I a) DSGVO
 - Art. 6 I I f) DSGVO
- Weitere mögl. Verstöße
 - Art. 5 I DSGVO
 - Art. 12, 13, 14, 17 DSGVO
- Lösung?

Prüfungsrecht

- Prompt-Design als neue Kompetenz?
- Sinnvoller Einbezug in Lernphasen, um vor Schwächen und Gefahren der KI zu unterrichten

Aufgabenerstellung mit Hilfe von KI

Übung: KI-gestützte Aufgabenerstellung

- **Ziel:** Verbesserung von Aufgaben und Tests mit KI-Unterstützung
- **Schritte:**
 - Auswahl eines Themengebiets
 - Festlegung der Schlüsselkonzepte und -kompetenzen
 - KI nutzt vorhandene Aufgaben zur Erstellung neuer
 - Überprüfung und Anpassung
- **Tipps:**
 - Kontext bereitstellen
 - Variation in den Instruktionen ausprobieren.

¹Die Idee für diese Übung ist angelehnt an die Methoden von Lothar Sebastian Krapp und Konstantin Strömel, <https://ki-macht-schule.de/plattform.html>

Prüfungsbewertung mit Hilfe von KI

Übung: Chain-of-Thought (CoT) Prompting: Bewertung von Leistungen¹

- Texte von Studierenden durch Sprachmodell bewerten lassen. Nutzen Sie einen eigenen Kriterienkatalog.
- Vergleich: Naiver Zero-shot-Ansatz ohne Beispiele und Zero-shot-CoT mit letztem Satz "Unterteile die Aufgabe in Zwischenschritte und löse diese Schritt für Schritt".
- Eigenen Ansatz aufschlüsseln: Welche Schritte und Beurteilungskriterien nutzen Sie? Erstellen Sie ein oder zwei Beispiele und testen Sie die Few-shot-CoT Methode.
- **Tipps:**
- Für passenden Kontext sorgen, z.B. "Hintergrund: Dies ist ein Text aus einem Seminar für Geschichtswissenschaft" oder "Hintergrund: Dies ist ein wissenschaftlicher Fachtext".
- Unterschiedliche Operatoren in Prompts testen, wie "Beurteile", "Bewerte", "Nimm kritisch Stellung".

¹Die Idee für diese Übung stammt von Lothar Sebastian Krapp und Konstantin Strömel,
<https://ki-macht-schule.de/plattform.html>

Prüfungsbewertung mit Hilfe von KI

- Die Verwendung des Prompts zum Trainingszwecken durch eine öffentliche KI sollte ausgeschlossen werden!
- Datenschutzrecht
 - Prüfungsleistungen können auch personenbezogene Daten enthalten (EuGH, Urteil vom 20.12.2017 – C 434/16).
 - Erforderlich: Anonymisierung (Erwägungsgrund 26 DSGVO) oder Verarbeitungsgrundlage (Art. 6 I lit. a), c), e) DSGVO) sowie gesetzliche Erlaubnis im LHG.
- Urheberrecht
 - Die in der KI als Prompt eingegebenen Prüfungsleistungen können urheberrechtlich geschützt sein.
 - Erforderlich: Nutzungserlaubnis nach § 31 UrhG oder durch eine Regelung in der Prüfungsordnung.

Ergebnisnutzung

Geistiges Eigentum

- möglicherweise Übernahmen geschützter Werkteile in Ergebnissen
- Rechtsinhaberschaft unklar
 - KI ≠ Person
 - „KI generiert“ vs. „KI gestützt“

Persönlichkeitsrechte

- Ggf. problematisch bei Unwahrheiten in generierten Texten

Prüfungsrecht

- (Un)zulässiges Hilfsmittel?
 - Im Ermessen der Lehrenden, Art. 5 III GG¹
- Problem der Kontrolle & Beweisbarkeit
- KI-robuste und KI-integrative Szenarien kombinieren
 - Die Zukunft der Hausarbeit
 - Neue Prüfungsszenarien?

¹Die Freiheit der Lehre schützt die methodische und inhaltliche Prüfungsgestaltung, OVG Bautzen, Beschl. v. 04.02.2021 – 2 B 27/21.

Ergebnisnutzung durch Studierende

- Grundsatz der eigenständigen Leistungserbringung nur mit Hilfe zugelassener Hilfsmittel.
- Die ungekennzeichnete Übernahme fremder Ausführungen ist keine eigenständige Leistung und ein Plagiat.¹
- Schwerpunkt der fachlichen Bearbeitung muss im Einzelfall festgestellt werden. (Aber Kontrollproblem!)

¹ Auslegung der Rspr. zur Übernahme fremder Texte: BVerwG, Urt. v. 21.06.2017 – 6 C 3/16.

The background features a complex network of glowing lines in shades of blue and red, creating a sense of depth and connectivity. The lines are interconnected, forming a mesh-like structure that recedes into the distance. A central text box is overlaid on the left side of the image.

Diskussion



Vielen Dank!

Martin Drossos

E-Mail: martin.drossos@uni-heidelberg.de