



AI: KI@Viadrina: KI in der Hochschullehre – Pädagogisches Werkzeug oder ethisches Risiko?

Janne Mesenhöller (sie/-)

Universität Potsdam, Inklusionspädagogik mit dem Schwerpunkt Sprache

Wer bin ich?



Foto: Roesse

- Klinische Psychologin
- seit 2021:
 - Wissenschaftliche Mitarbeitende Universität Potsdam im Arbeitsbereich Inklusionspädagogik mit dem Schwerpunkt Sprache
 - 2021 – 2024 Drittmittelprojekt AKILAS
 - seit 2023 KISS-Pro Projekt

Rückblick



August 2020: „Janne, willst du in einem Forschungsprojekt arbeiten, das sich mit dem Thema KI in der Schule beschäftigt?“



„Öhm...ok?!“



Februar 2021: Start des AKILAS Projekts



...und dann kam der 30.11.2022!

<https://de.wikipedia.org/wiki/ChatGPT#/media/Datei:ChatGPT-Logo.svg>

Wie verändert ChatGPT die Hochschulbildung?



https://hm.edu/aktuelles/news/news_detailseite_284160.de.html

KI in der Bildung

Regeln zu ChatGPT an Unis oft unklar

Stand: 15.05.2023 16:01 Uhr

Mit KI lassen sich in kurzer Zeit Texte erzeugen. Ob Studierende die Programme einsetzen dürfen, ist nicht einheitlich geregelt. Oft hängt die Entscheidung am Lehrpersonal, wie eine *BR*-Umfrage zeigt.

<https://www.tagesschau.de/wissen/technologie/ki-chatgpt-uni-wissenschaft-101.html>

Trotz fehlender Lizenzen – ChatGPT an Hochschulen längst Alltag

18.07.23

Mit der Befragung „ChatGPT-Lizenzen an Hochschulen“ wollte das Hochschulforum Digitalisierung Ende Juni 2023 von den Hochschulmitarbeitenden aus der Community

erfragen. Wie sieht die Nutzung von ChatGPT an Hochschulen aus? Was sind die größten Herausforderungen?

<https://hochschulforumdigitalisierung.de/news/trotz-fehlender-lizenzen-chatgpt-an-hochschulen-laengst-alltag/>

„Der Umgang mit generativer KI an Hochschulen ist mit Unsicherheiten behaftet, die nicht zuletzt aufgrund der stetigen Weiterentwicklung generativer KI nie vollständig ausgeräumt werden können.“

(Friedrich, 2024 zitiert nach Torbor, 2024)



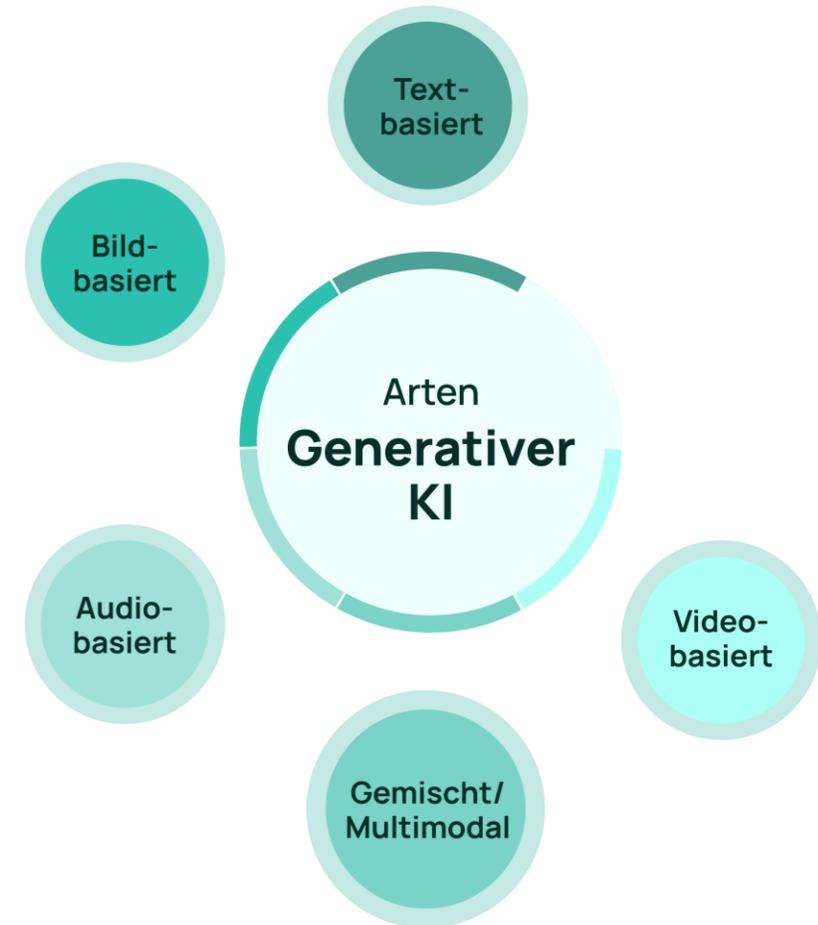
Was ist KI eigentlich?

Künstliche Intelligenz (KI)?

- Es gibt keine einheitliche wissenschaftliche Definition von KI
- Es gibt verschiedene Unterbereiche von KI, die sich darin unterscheiden, wie komplex die Lernverfahren sind, die genutzt werden, um ein System zu trainieren (z.B. Machine/Deep Learning)

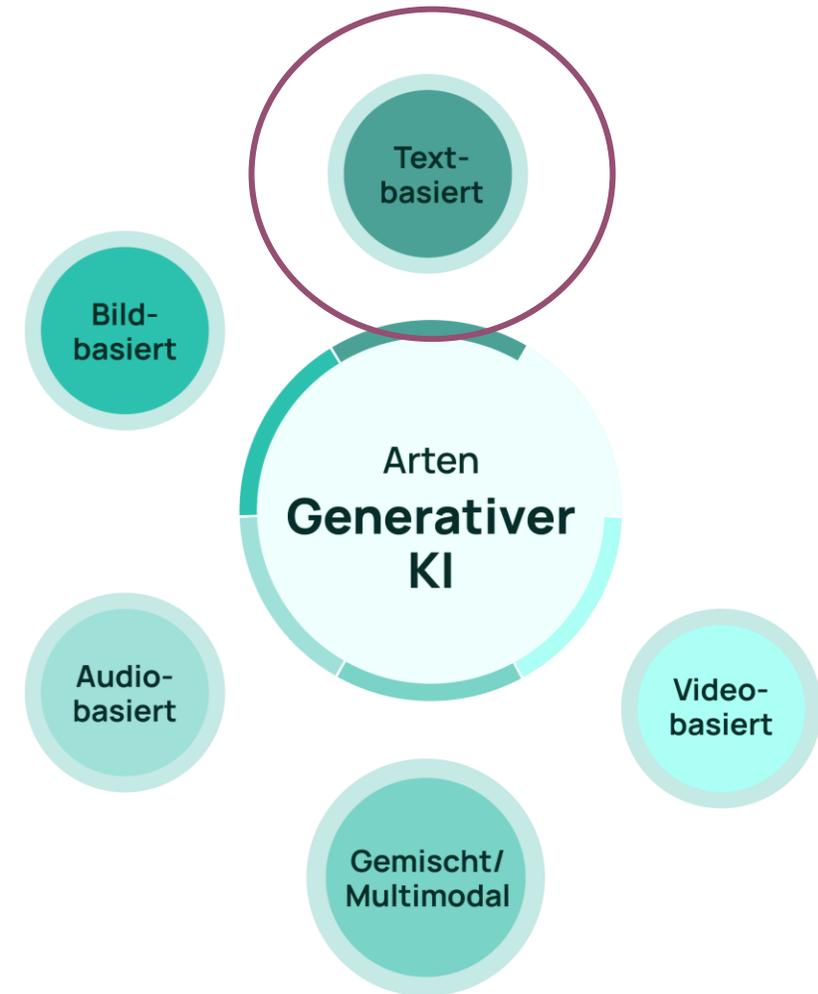
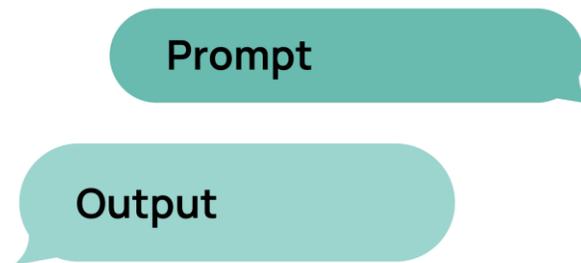
Generative KI

- Beschreibt KI-Modelle, die durch eine Eingabeaufforderung (**Prompt**) scheinbar selbstständig neue Texte, Bilder, Videos oder Audios aus Trainingsdaten generieren können
(Feuerriegel et al., 2024)
- Prompt ist entscheidend für die Ausgabe (**Output**) generativer KI-Systeme



Large-Language-Modelle (LLM)

- Heute in dem Vortrag geht es vor allem um textgenerative KI, die sogenannten **Large-Language-Modelle** (große Sprachmodelle) wie bspw. GPT



Funktionsweise von LLM

- Analyse und Zerlegung des Texts in kleine Teile (Tokens)
- Berechnung der Wahrscheinlichkeit des nächsten Tokens, basierend auf vorherigem Token und Kontext
- Wahl des wahrscheinlichsten Token
- → Output ist sehr wahrscheinlich, aber nicht zwingendermaßen korrekt!
 - *Sog. Halluzinationen* können auftreten!

Default (GPT-3.5)

Hallo ich bin Schülerin der 5. Klasse und schreibe morgen eine Klassenarbeit zum Thema wechselwarme Tiere, kannst du mir bei der Vorbereitung helfen?"

Hallo! Natürlich stehe ich dir gerne zur Verfügung, **um** dich bei der Vorbereitung auf deine Klassenarbeit zu unterstützen.

um = 91.00%
damit = 31.70%
weil = 4.36%
wenn = 2.5%

Total: -0.88 logprob on 1 tokens
(99.97% probability covered in top 5 Logits)



Generiert mit DALL-E

Zum Ablauf heute...

- ~ 1 Stunde interaktiver Vortrag
- ~ 30 Minuten Diskussion
- Möglichkeit für individuelle Beratung:
Morgen (Freitag, 14.02.), 14:30 – 16:00
 - Schreiben Sie mir dazu gerne bis heute Abend 20:30 eine Email an mesenhoeller@uni-potsdam.de

Abkürzungs-/Fachbegriffsverzeichnis

- Link im Chat

mesenhoeller • 16h

Abkürzungs-/Fachbegriffsverzeichnis AI-Ethics in der Hochschule

mesenhoeller
vor 16 Stunden

Verzeichnis wichtiger Begriffe zur KI in der Hochschule (alphabetisch)

- 1. AI Act**
→ Der *Artificial Intelligence Act* ist ein von der EU entwickeltes Gesetzesvorhaben zur Regulierung von KI. Es klassifiziert KI-Systeme nach ihrem Risiko und setzt strenge Anforderungen für Hochrisiko-KI.
- 2. AI Detection Software**
→ Software, die darauf spezialisiert ist, KI-generierte Texte oder Inhalte zu erkennen. Sie wird in Hochschulen zur Überprüfung von Prüfungsleistungen genutzt, ist aber oft unzuverlässig.
- 3. AI Ethics (KI-Ethik)**
→ Bereich, der sich mit den moralischen und gesellschaftlichen Auswirkungen von KI befasst, z. B. Datenschutz, Fairness oder Verantwortung.

Digital Fragen sammeln

- Link im Chat

A rounded rectangular input field with a light gray border and a subtle shadow. On the left side, there is a small gray icon of a person's head and shoulders. To the right of the icon, the text "Type your question" is displayed in a light gray font.

There are no questions asked yet.

Ask the first one!

Blitzlicht

Woher kommen Sie und was bringt
Sie heute her?





Nutzung von KI in der Hochschule

Janne Mesenhöller, KI@Viadrina, AI Ethics in der Hochschullehre – Talk 13.02.2025

Aktueller Stand – KI-Nutzung durch Lehrende

- 54 % Lehrenden nutzt KI im Hochschulkontext „in irgendeiner Form“, vor allem aber zur Vorbereitung der Lehrveranstaltungen (Budde et al., 2024)
- 54% der Lehrenden trauen sich den Umgang mit KI innerhalb einer Lehrveranstaltung zu, aber nur 19% setzen KI-basierte Systeme in konkreten Lehrsituationen ein (Budde et al., 2024)

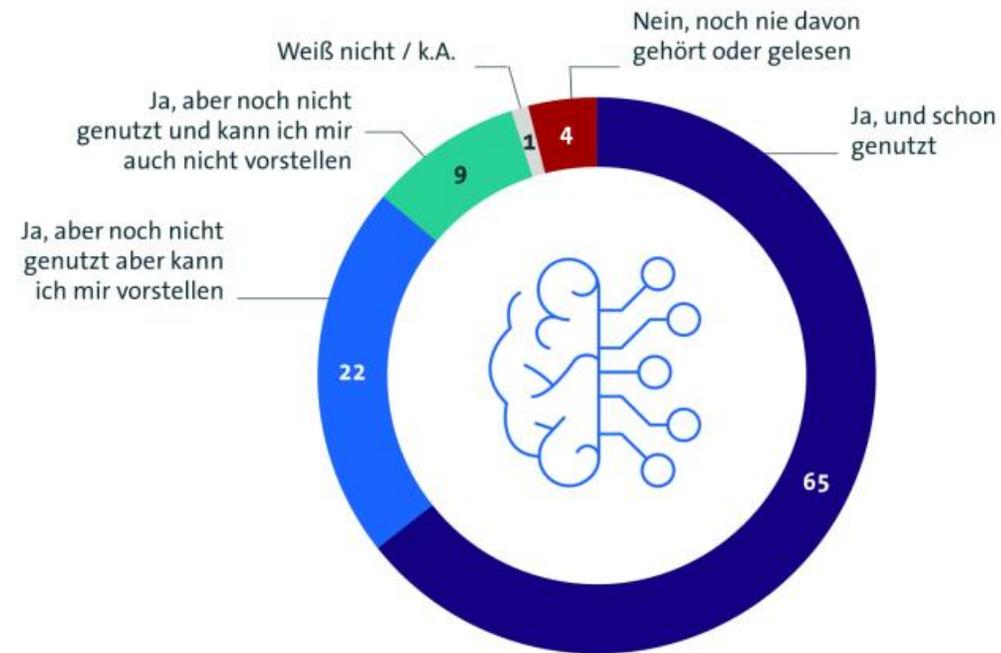
Aktueller Stand – KI-Nutzung durch Studierende

Mehrheit der Studierenden kennt ChatGPT – und nutzt es

Haben Sie schon einmal von der Anwendung ChatGPT gelesen oder gehört?

in Prozent

Basis: Alle Befragten (n=506) | Abweichungen von 100 Prozent sind rundungsbedingt | Quelle: Bitkom Research 2024



bitkom

...aber

- KI wird von Studierenden vor allem zur Recherchezwecken eingesetzt (68%)
- 60% der befragten Studierenden geben an, dass der Einsatz von ChatGPT dazu führt, dass Studierende weniger selbstständig denken und lernen
- Regeln des Einsatzes von KI an Hochschulen in Deutschland sind meist unklar oder Studierende kennen sie nicht

Neben umfassenden Potenzialen von KI gehen auch gewisse Herausforderungen mit dem Einsatz einher!



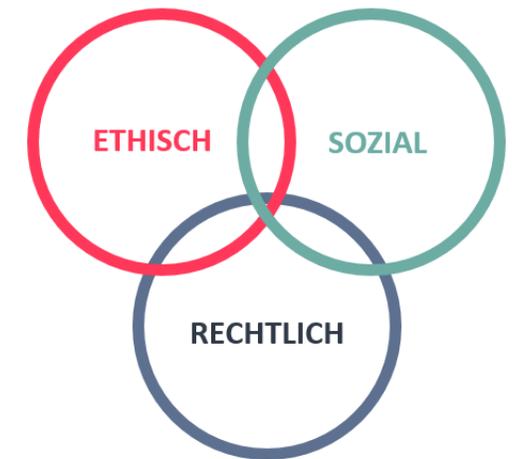


Forschungsfeld ELSI

Janne Mesenhöller, KI@Viadrina, AI Ethics in der Hochschullehre – Talk 13.02.2025

Forschungsfeld ELSI

- **ELSI** - Ethische, rechtliche und soziale Implikationen
- Interdisziplinäres Forschungsfeld, das sich mit potenziellen Auswirkungen wissenschaftlicher und technologischer Fortschritte auf Gesellschaft und Individuum auseinandersetzt
- Ursprung in der Humanmedizin/Biotechnologie (Juengst, 1991)



Eine (sehr bekannte) ethische Herausforderung: Algorithmische Verzerrungen



Verzerrungen in der Entscheidungsfindung von Algorithmen, die durch voreingenommene oder unvollständige Daten entstehen und systematisch bestimmte Gruppen bevorzugen oder benachteiligen.

(vgl. z.B. Akter et al., 2021)

Algorithmische Verzerrungen

Joy Buolamwini
(2017)



<https://mitsloan.mit.edu/ideas-made-to-matter/unmasking-bias-facial-recognition-algorithms>

Relevanz für die Hochschule?

- Studierendenschaft ist sehr heterogen (Kerst et al., 2024)
- Welche Konsequenzen könnten algorithmische Verzerrungen beim Einsatz von KI zur Bewertung von Leistung haben?
 - ...und welche für die eigene Identität?



ELSI der Nutzung von KI im Hochschulkontext

Janne Mesenhöller, KI@Viadrina, AI Ethics in der Hochschullehre – Talk 13.02.2025



Rechtliche Herausforderungen der Nutzung von KI in der Hochschule

Janne Mesenhöller, KI@Viadrina, AI Ethics in der Hochschullehre – Talk 13.02.2025



Rechtliche Herausforderungen der Nutzung von KI in der Hochschule: Datenschutz

Janne Mesenhöller, KI@Viadrina, AI Ethics in der Hochschullehre – Talk 13.02.2025

Rechtliche Herausforderungen: Datenschutz

- In Deutschland gilt Datenschutzgrundverordnung (DSGVO)
- Nutzung von ChatGPT nicht DSGVO-konform (Rojszczak, 2020)
- **Implikation: DSGVO-konforme Nutzung von KI-Systemen**
 - Bspw. KI-fobizz-Assistent (fobizz, 2024) bzw. hochschulinternes LLM
 - Stand März 2024: 30 % der Hochschulen haben Lizenzen für KI-Tools erworben (Budde et al., 2024)
 - Keine Eingabe personenbezogener Daten (SenBJF, 2023)



Möglichkeiten

Kontextverständnis - Merkt sich, was vorab in der Konversation gesagt wurde.

Iteration - Erlaubt nachträgliche Korrekturen generierter Ergebnisse.

Formatierung - Gibt generierte Ergebnisse in gewünschter Form aus.

Anonymisierung - GPT.UP trennt Prompts und Accounts.



Limitationen

Unvollständig - Generiert gelegentlich falsche Informationen.

Vorsicht - Generiert gelegentlich gefährdende oder voreingenommene Informationen.

Limitierung - Das Sprachmodell greift ausschließlich auf Wissen bis zum Dezember 2023 zu.

Datenschutz - Geben Sie keine personenbezogenen Daten ein.

OpenAI GPT-4o (Microsoft) ▾



Hier kannst Du deine Anfrage stellen



Version 1.0

Screenshot GPT.UP

Hochschuleigenes LLM?



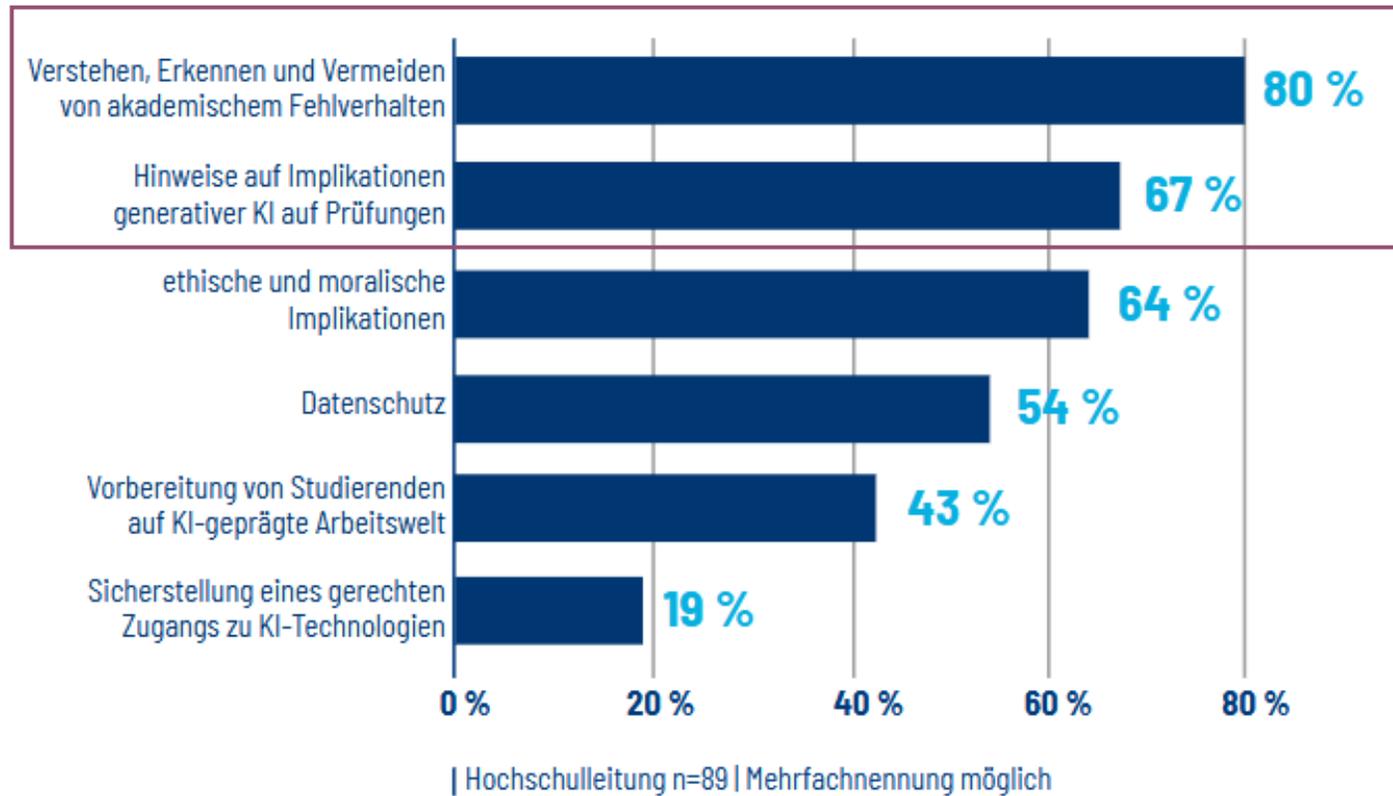
Rechtliche Herausforderungen der Nutzung von KI in der Hochschule: Plagiate/Urheberrecht

Janne Mesenhöller, KI@Viadrina, AI Ethics in der Hochschullehre – Talk 13.02.2025

Herausforderung Urheberrecht/Plagiate

- Hochschulrecht fordert eigenständige Leistungserbringung
- Aber: Was ist eigenständige Leistung im Zeitalter von KI?

Welche Aspekte von KI werden aktuell an den Hochschulen diskutiert?



Budde et al., 2024

Herausforderung Urheberrecht/Plagiate

- AI Erkennungssoftware soll dazu beitragen, Plagiate zu identifizieren
- Performanz wird immer besser (Quidwai et al., 2023)
- ABER:
 - Nach aktuellem Stand ist es nicht zuverlässig möglich, zwischen von Menschen und von einer KI verfassten Texten zu unterscheiden (Weber-Wulff et al., 2023)
 - Hinweise, dass solche Erkennungsprogramme bestimmte Personengruppen benachteiligen (Liang et al., 2023)

Umgang mit Verdachtsfällen

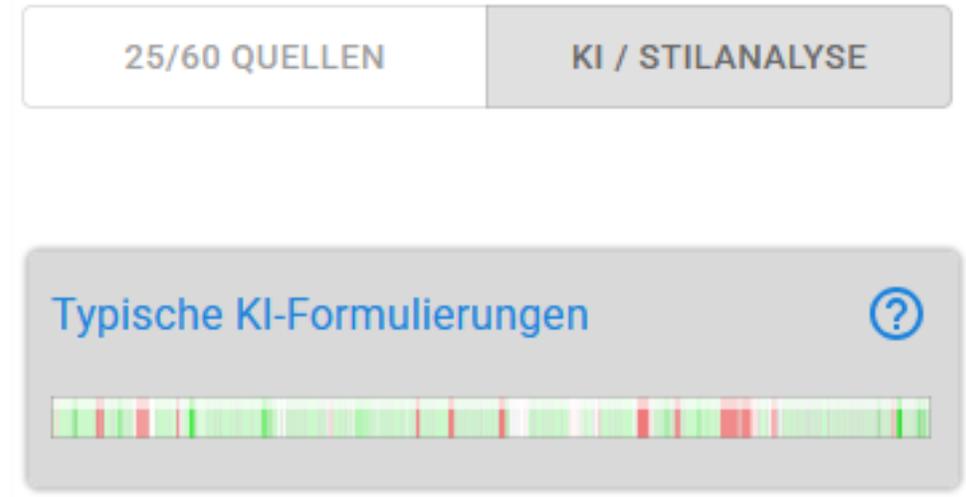
→ Leitlinien der Universität Vechta empfehlen multifaktorielles Prüfverfahren statt reiner KI-Erkennung.

Prüfungsleistungen sollten geprüft werden auf...

- Unregelmäßigkeiten in der Formulierung
- Nicht existierende Quellenangaben
- Ergebnisse von KI-Erkennungstools (z. B. GPT-Zero)

- Klärungsgespräch mit der:dem Studierenden

(Tobor, 2024)



Screenshot PlagAware

Herausforderung Urheberrecht/Plagiate

Was ist ein Plagiat im Zeitalter von generativer KI?

„Im Ergebnis gilt es also, „zwischen KI-gestützten menschlichen Schöpfungen und durch KI erzeugten Schöpfungen“ zu unterscheiden. Wird das KI-Programm von den Nutzenden lediglich als eine Art Hilfsmittel unterstützend herangezogen, so kommt ein Urheberrechtsschutz weiterhin in Betracht. Andernfalls liegen die Voraussetzungen für ein Werk nicht vor, sodass die Nutzenden des KI-Programms keine Urheberschaft an dem KI-generierten Text innehaben.“

Leschke & Salden, 2023, S. 26



Ethisch-soziale Herausforderungen der Nutzung von KI in der Hochschule

Janne Mesenhöller, KI@Viadrina, AI Ethics in der Hochschullehre – Talk 13.02.2025



51% Durch ChatGPT beschäftigen sich Studierende weniger mit den Studieninhalten.

66% Durch den Einsatz von ChatGPT werden sich Prüfungsformate an Hochschulen ändern.

44% Die Nutzung von ChatGPT für Hausarbeiten oder Abschlussarbeiten sollte verboten werden.

Bitkom Research, 2024

Ethisch-soziale Herausforderungen der Nutzung von KI in der Hochschule: Veränderungen der Lehr-, Lern- und Prüfungskultur



Neue (KI-)Kompetenzen: Lernen über KI

Technologisches Verständnis und kritische Reflexion

- Technisches Grundlagenwissen über KI
- Grenzen von KI
- Bewusstsein für Biases und Halluzinationen

Neue (KI-)Kompetenzen: Lernen mit & durch KI

Lernförderliche Interaktion mit KI-Tools

- **Prompt Engineering:** Wie prompte ich so, dass ich einen sinnvollen und hilfreichen Output bekomme?
- **Kritische Reflexion** des KI-Outputs: Wie kann ich einen Output auf seine inhaltliche Korrektheit und Textqualität überprüfen?
- KI als **Sparringpartner:** Wie nutze ich KI-Tools so, dass ich wichtige Lernprozesse nicht an das System delegiere, sondern als Unterstützung nutze?
 - z.B. als personalisierte Schreibunterstützung

Neue (KI-)Kompetenzen: Lernen ohne & trotz KI

Eigenständiges Lernen unabhängig von KI-Tools

- Wissenschaftliche Recherche (ohne KI-Tools)
- Kritisches Denken fördern – mit und ohne KI Tools
 - **Reflexion über den eigenen Lernprozess:** Was habe ich wirklich verstanden – und was nur durch KI bekommen?
 - **Bewusst entscheiden, wann KI sinnvoll ist:** Welche Lernprozesse profitieren von KI, welche nicht?

Akademische Integrität

Integrität als Haltung

- „Das Prinzip der wissenschaftlichen Integrität gilt nicht nur für die Forschung, sondern muss auch in der Lehre aktiv gelebt werden. Generative KI lässt die bisher bekannten Grenzen zwischen Fakten und Fiktion verschwimmen. Deshalb braucht es stets den *"Human-in-the-Loop"*, um Ergebnisse zu validieren, zu hinterfragen, auf Fehler und auf Biases zu überprüfen. Der verantwortungsvolle und transparente Umgang mit GenAI bildet einen Grundpfeiler akademischer Integrität.“ (ETH Zürich, o.D.)
- Verantwortung übernehmen
- Transparenz schaffen
- Fairness einhalten
(ETH Zürich, o.D.)

Um Studierende dabei zu unterstützen,
KI-kompetent zu werden, braucht es KI-
kompetente Lehrende!

Neue Kompetenzen = Neue Prüfungsformate?

- Prüfungen, die Zukunftskompetenzen prüfen
 - Insbesondere die Kompetenz des kritischen Denkens!
- Prüfungsformate, die weg gehen von summativen Assessments hin zu prozessorientierten Prüfungsformaten
 - Aber: Wie? Welche?

Dafür braucht es klare Leitlinien!



**Hochschulforum
Digitalisierung**

BLICKPUNKT

**Leitlinien zum
Umgang mit
generativer KI**

Jens Tobor (HFD/CHE)
07.02.2024

Torbor, 2024



Ethisch-soziale Herausforderungen der Nutzung von KI in der Hochschule: Digital Divide

Janne Mesenhöller, KI@Viadrina, AI Ethics in der Hochschullehre – Talk 13.02.2025

Digital Divide

- Mögliche Vergrößerung der Bildungsschere?
 - *Digital Divide* könnte sich durch den Einsatz von KI weiter vergrößern
 - Zunahme des Matthäus-Effekt?
 - Profitieren Studierende mit ausgeprägten Schriftsprachkompetenzen ggf. mehr von der Nutzung generativer KI als eher leistungsschwächere Studierende (Prompting)?
 - Um Halluzinationen aufzudecken, braucht es fachinhaltliches Wissen



ELSI des flächendeckenden Einsatzes von KI im Hochschulkontext, die oft zu kurz kommen

Janne Mesenhöller, KI@Viadrina, AI Ethics in der Hochschullehre – Talk 13.02.2025

Ausblick:

Belastung der Umwelt

- Training und Nutzung von (insbesondere generativer) KI verbraucht große Mengen von CO₂ und Wasser
- Infrastruktur von Rechenzentren verbraucht ungefähr so viel CO₂ wie Luftfahrtindustrie, ~ 2-4% (Muller & King, 2023)

Ausbeutung von Menschen

- z.B. Clickworker:innen

Deep Fakes

- Desinformationen/Fake News
- (Cyber-)Mobbing
- Vertrauensverlust in audiovisuelle Beweise

Können Sie zwischen einem Bild, das von einer KI bzw. einem Menschen erstellt wurde, unterscheiden?

- Odd One Out ist ein Spiel von Google AI Experiments
- Link im Chat!

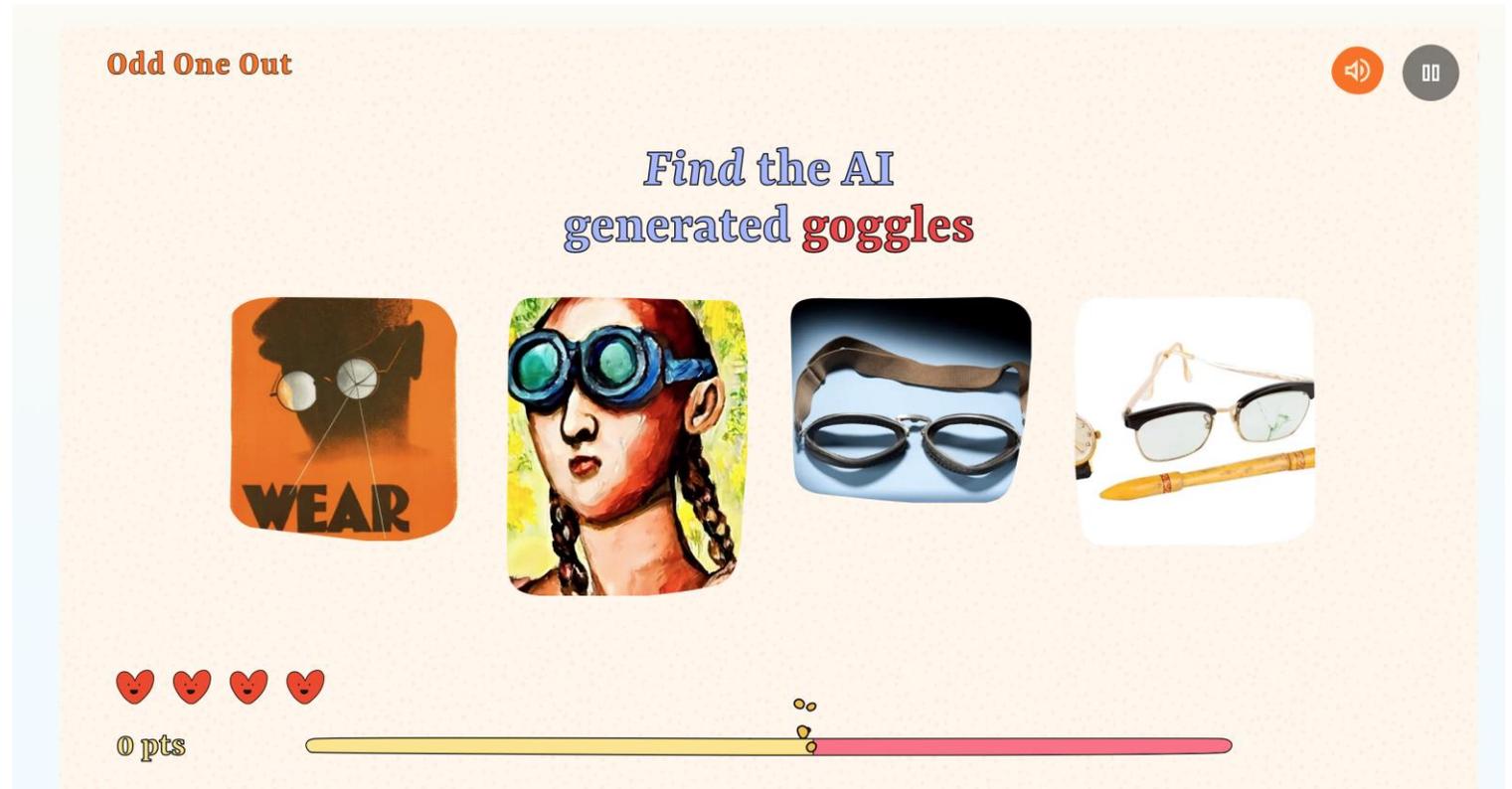


Abb. 9

Ausblick:

Belastung der Umwelt

- Training und Nutzung (insbesondere von generativer) KI verbraucht große Mengen von CO₂ und Wasser
- Infrastruktur von Rechenzentren verbraucht ungefähr so viel CO₂ wie Luftfahrtindustrie, ~ 2-4% (Muller & King, 2023)

Ausbeutung von Menschen

- z.B. Clickworker:innen

Deep Fakes

- Desinformationen/Fake News
- (Cyber-)Mobbing
- Vertrauensverlust in audiovisuelle Beweise

Haftungsfragen

- Wer haftet, wenn KI „versagt“?



KI @Viadrina: KI in der Hochschullehre – Pädagogisches Werkzeug UND ethisches Risiko!

Janne Mesenhöller, KI@Viadrina, AI Ethics in der Hochschullehre – Talk 13.02.2025

Weiterlesen, weiterhören



mesenhoeller
vor 7 Minuten

Weiterhören, weiterlesen



ethz.ch

Akademische Integrität

Akademische Integrität

<https://ethz.ch/de/die-eth-zuerich/lehre/ai-in-education/academic-integrity.html>





**Vielen Dank für Ihre
Aufmerksamkeit!**

Janne Mesenhöller, KI@Viadrina, AI
Ethics in der Hochschullehre – Talk
13.02.2025



Erinnerung:

- Möglichkeit für individuelle Beratung: Morgen (Freitag, 14.02.), 14:30 – 16:00
 - Schreiben Sie mir dazu gerne bis heute Abend 20:30 eine Email an mesenhoeller@uni-potsdam.de
- Keine Rechtsberatung!



Jetzt: Zeit für Fragen!

Janne Mesenhöller, KI@Viadrina, AI
Ethics in der Hochschullehre – Talk
13.02.2025



Literatur

- Akter, S., McCarthy, G., Sajib, S., Michael, K., Dwivedi, Y. K., D'Ambra, J., & Shen, K. N. (2021). Algorithmic bias in data-driven innovation in the age of AI. *International Journal of Information Management*, 60, 102387. <https://doi.org/10.1016/j.ijinfomgt.2021.102387>
- Bitkom Research. (2024). *KI oder Kreide im Hörsaal – so digital sind Deutschlands Hochschulen 2024*. <https://www.bitkom-research.de/node/941>
- Budde, J., Tobor, J., Friedrich J. (2024). Künstliche Intelligenz. Wo stehen die deutschen Hochschulen?. Berlin: Hochschulforum Digitalisierung.
- Buolamwini, J. A. (2017). *Gender shades : intersectional phenotypic and demographic evaluation of face datasets and gender classifiers*. Massachusetts Institute of Technology, School of Architecture and Planning. <https://dspace.mit.edu/handle/1721.1/114068>
- Europäisches Parlament. (2024, 17. April). *Berichtigung: des in erster Lesung am 13. März 2024 festgelegten Standpunkts des Europäischen Parlaments im Hinblick auf den Erlass der Verordnung (EU) 2024/ des Europäischen Parlaments und des Rates zur Festlegung harmonisierter Vorschriften für künstliche Intelligenz und zur Änderung der Verordnungen (EG) Nr. 300/2008, (EU) Nr. 167/2013, (EU) Nr. 168/2013, (EU) 2018/858, (EU) 2018/1139 und (EU) 2019/2144 sowie der Richtlinien 2014/90/EU, (EU) 2016/797 und (EU) 2020/1828 (Verordnung über künstliche Intelligenz)*. https://www.europarl.europa.eu/doceo/document/TA-9-2024-0138-FNL-COR01_DE.pdf
- ETH Zürich. (o. D.). Akademische Integrität- ETH Zürich. <https://ethz.ch/de/die-eth-zuerich/lehre/ai-in-education/academic-integrity.html>
- Falck, J. (2023). *Lernen und Künstliche Intelligenz*. <https://joschafalck.de/lernen-und-ki/>
- Feuerriegel, S., Hartmann, J., Janiesch, C. & Zschech, P. (2024). Generative AI. *Business & Information Systems Engineering*, 66(1), 111–126. <https://doi.org/10.1007/s12599-023-00834-7>

Literatur

- fobizz. (2024). *Digitalen Unterricht einfach und sicher gestalten*. <https://tools.fobizz.com>
- Juengst, E. T. (1991). The Human Genome Project and bioethics. *Kennedy Institute of Ethics journal*, 1(1), 71–74. <https://doi.org/10.1353/ken.0.0139>
- Leschke, J., & Salden, P. (Hrsg.). (2023). *Didaktische und rechtliche Perspektiven auf KI-gestütztes Schreiben in der Hochschulbildung*. Ruhr-Universität Bochum, Universitätsbibliothek. <https://doi.org/10.13154/294-9734>
- Liang, W., Yuksekgonul, M., Mao, Y., Wu, E. & Zou, J. (2023). *GPT detectors are biased against non-native English writers*. <https://doi.org/10.48550/ARXIV.2304.02819>
- Muller, N. & King, N. (7. August 2023). Wie Künstliche Intelligenz der Umwelt schadet. *Deutsche Welle*. <https://www.dw.com/de/wie-k%C3%BCnstliche-intelligenz-der-umwelt-schadet/a-66305844>
- Quidwai, M. A., Li, C. & Dube, P. (2023, 13. Juni). *Beyond Black Box AI-Generated Plagiarism Detection: From Sentence to Document Level*. <http://arxiv.org/pdf/2306.08122>
- Rojszczak, M. (2020). CLOUD act agreements from an EU perspective. *Computer Law & Security Review*, 38, 105442. <https://doi.org/10.1016/j.clsr.2020.105442>
- Senatsverwaltung für Bildung, Jugend und Familie Berlin. (04/2023). *Empfehlung für den Umgang mit KI-Anwendungen am Beispiel von ChatGPT*. <https://www.berlin.de/sen/bjf/service/presse/pressearchiv-2023/pressemitteilung.1316517.php>
- Tobor, J. (2024). *Blickpunkt – Leitlinien zum Umgang mit generativer KI*. Berlin: Hochschulforum Digitalisierung. Version 1.0
- Weber-Wulff, D., Anohina-Naumeca, A., Bjelobaba, S., Foltýnek, T., Guerrero-Dib, J., Popoola, O., Šigut, P. & Waddington, L. (2023). Testing of detection tools for AI-generated text. *International Journal for Educational Integrity*, 19(1). <https://doi.org/10.1007/s40979-023-00146-z>