



lernen:digital
Kompetenzzentrum
Sprachen/Gesellschaft/Wirtschaft

KI Connect 2024

Ethische, rechtliche und soziale Herausforderungen des Einsatzes von generativer KI in der (Hoch-)Schule

Janne Mesenhöller (sie/-)

Universität Potsdam, Inklusionspädagogik im Bereich Sprache, KISS-Pro Projekt



KISS-PRO



Finanziert von der
Europäischen Union
NextGenerationEU

GEFÖRDERT VOM

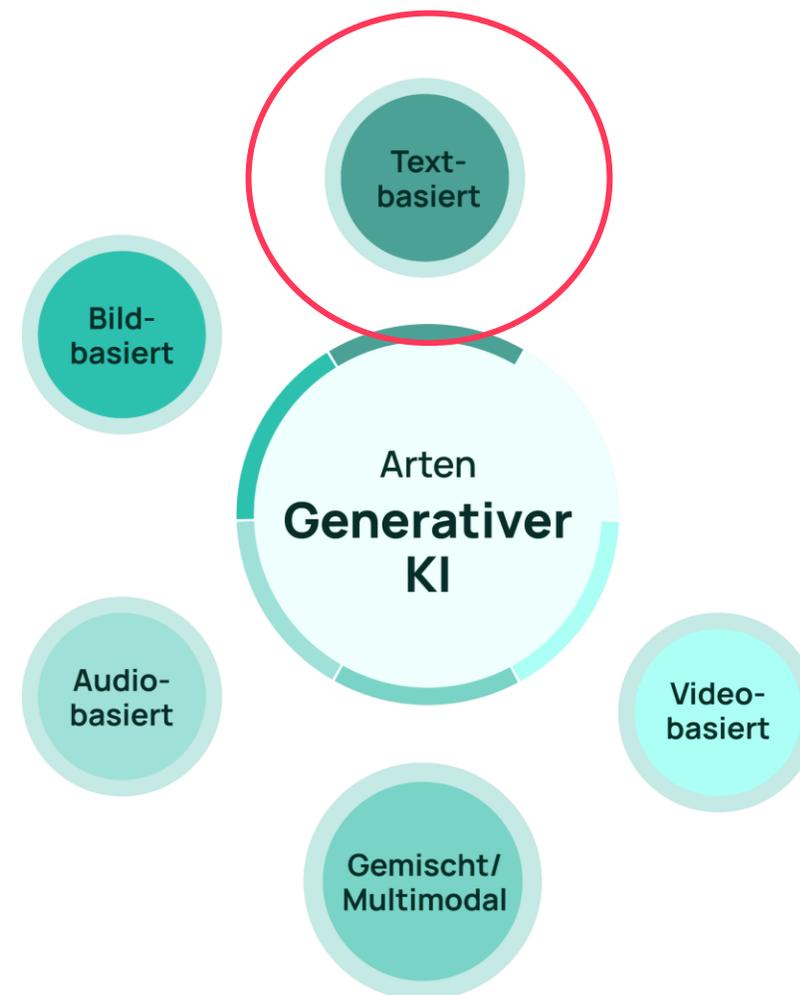
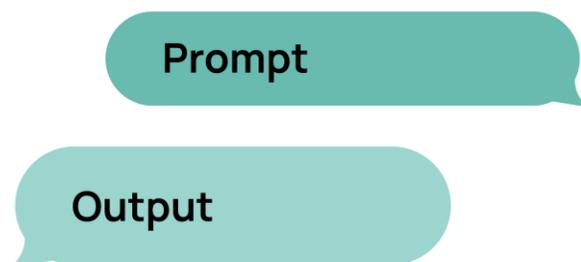


Bundesministerium
für Bildung
und Forschung

Large-Language-Modelle (LLM)



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



Funktionsweise von LLM



- Analyse und Zerlegung des Texts in kleine Teile (Tokens)
- Berechnung der Wahrscheinlichkeit des nächsten Tokens, basierend auf vorherigem Token
- 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)





lernen:digital
Kompetenzzentrum
Sprachen/Gesellschaft/Wirtschaft

Nutzung von KI in der Hochschule



KISS-PRO



**Finanziert von der
Europäischen Union**
NextGenerationEU



GEFÖRDERT VOM

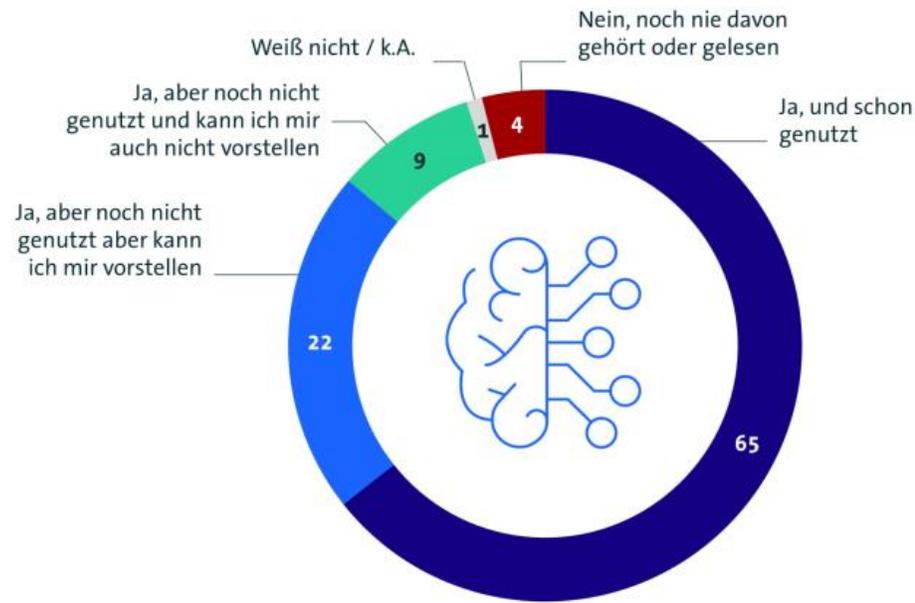
**Bundesministerium
für Bildung
und Forschung**

Aktueller Stand



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

- 54 % Lehrenden nutzt KI im Hochschulkontext „in irgendeiner Form“, vor allem aber zur Vorbereitung der Lehrveranstaltungen

(Bitkom Research, 2024)

...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 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!



(Bitkom Research, 2024)





lernen:digital
Kompetenzzentrum
Sprachen/Gesellschaft/Wirtschaft

Forschungsfeld ELSI



KISS-PRO



**Finanziert von der
Europäischen Union**
NextGenerationEU



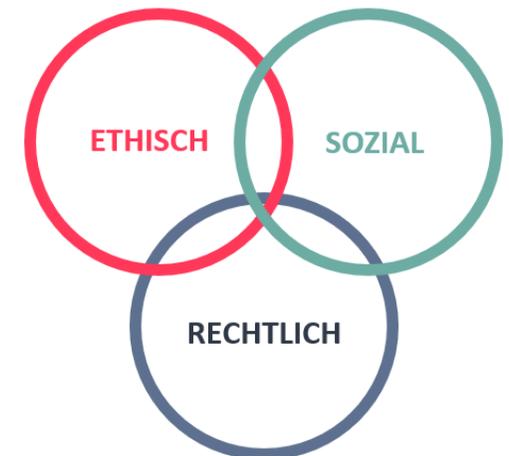
GEFÖRDERT VOM

**Bundesministerium
für Bildung
und Forschung**

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>

Algorithmische Verzerrungen



Google Photos Tags Two African-Americans As Gorillas Through Facial Recognition Software

Maggie Zhang Forbes Staff

I write about technology, innovation, and startups.

Follow



Jul 1, 2015, 01:42pm EDT

Updated Jul 1, 2015, 02:35pm EDT

 **This article is more than 9 years old.**

<https://www.forbes.com/sites/mzhang/2015/07/01/google-photos-tags-two-african-americans-as-gorillas-through-facial-recognition-software/>

KI-Systeme können diskriminierend sein – Wie kann das sein?



Qualität des Output und Funktion eines KI-Systems ist unterschiedlich gut. Das kann unterschiedliche Gründe haben. Dazu gehören:

- welche und wie viele Daten als Trainingsdaten für ein Modell genutzt werden
- wer das Modell wie zu welchem Zweck gestaltet

Relevanz für die Hochschule?



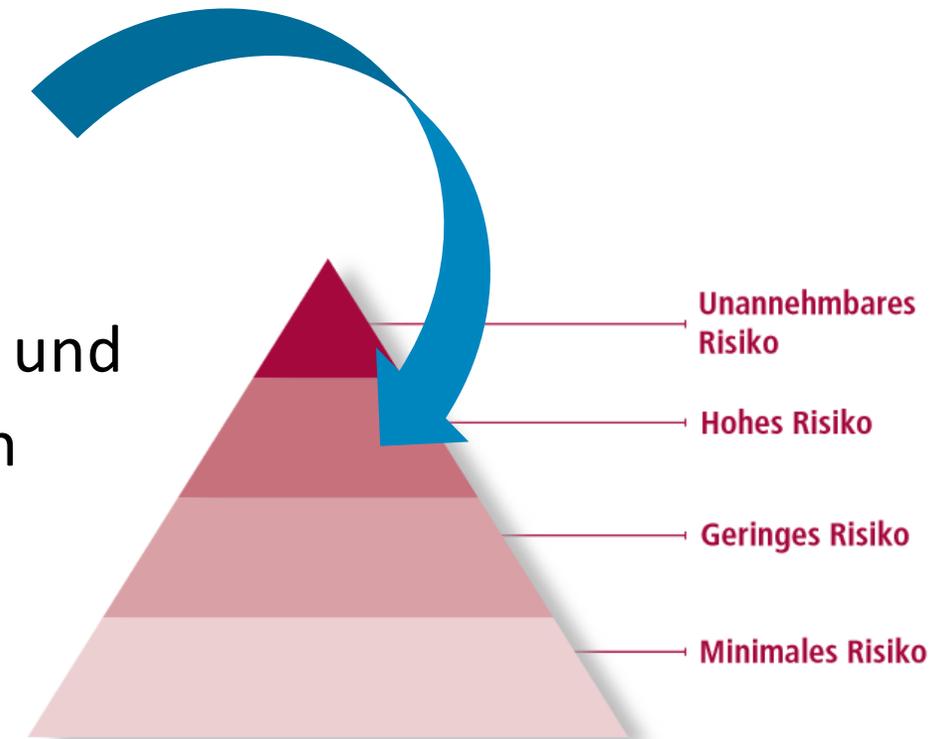
- Studierendenschaft ist sehr heterogen (Kerst et al., 2024)
- Welche Konsequenzen könnten algorithmische Verzerrungen beim Einsatz von KI zur Bewertung von Leistung haben?
 - ...und welche für die eigene Identität?

Der AI Act



- Der kürzlich erschiene AI-Act thematisiert einen Teil dieser Herausforderungen
- Klassifizierung von KI im Bildungskontext als hochrisiko-KI
- Vor allem in Bezug auf Leistungsbewertung und Entscheidung hinsichtlich Bildungszugängen

Europäisches Parlament und Rat der Europäischen Union, 2024





lernen:digital
Kompetenzzentrum
Sprachen/Gesellschaft/Wirtschaft

Rechtliche Herausforderungen der Nutzung von KI in der Hochschule



KISS-PRO



Finanziert von der
Europäischen Union
NextGenerationEU



GEFÖRDERT VOM

Bundesministerium
für Bildung
und Forschung

Rechtliche Herausforderungen



- **Herausforderung: Datenschutz**
 - In Deutschland gilt Datenschutzgrundverordnung (DSGVO)
 - Nutzung von ChatGPT nicht DSGVO-konform (Rojszczak, 2020)

- **Implikation: DSGVO-konforme Nutzung von KI-Systemen**
 - Nutzung von DSGVO-konformen ChatGPT-Alternativen wie bspw. KI-fobizz-Assistent (fobizz, 2024) oder Hochschulinternen/Bundeslandeigenen LLMs
 - Stand März 2024: 30 % der Hochschulen haben Lizenzen für KI-Tools erworben (Budde et al., 2024)
 - Keine Eingabe personenbezogener Daten (SenBJF, 2023)

Rechtliche Herausforderungen



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

Rechtliche Herausforderungen



- **Herausforderung: Plagiate**

- AI Erkennungssoftware soll dazu beitragen, Plagiate zu identifizieren
- Performanz wird immer besser (Quidwai et al., 2023)



Paraphraser



Grammar
Checker



AI Detector

Rechtliche Herausforderungen



- **Herausforderung: Plagiate**

- Nach aktuellem Stand ist es nicht zuverlässig möglich, zwischen von Studierenden und von einer KI verfassten Texten zu unterscheiden (SWK, 2024)
- Hinweise, dass solche Erkennungsprogramme bestimmte Personengruppen benachteiligen (Liang et al., 2023)

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



<https://t.ly/nQ2Ti>

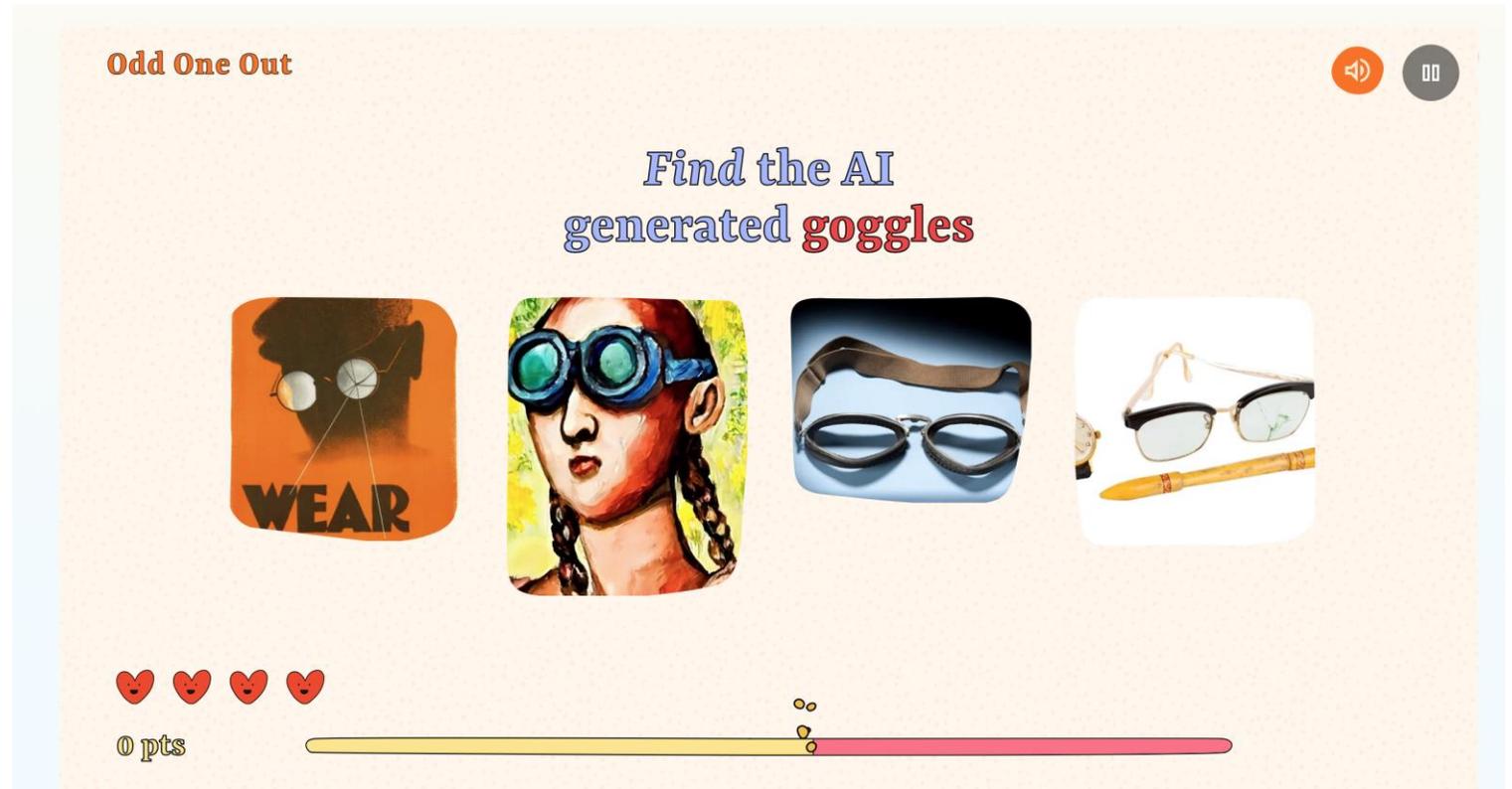


Abb. 9



lernen:digital
Kompetenzzentrum
Sprachen/Gesellschaft/Wirtschaft

Ethisch-soziale Herausforderungen der Nutzung von KI in der Hochschule



KISS-PRO



Finanziert von der
Europäischen Union
NextGenerationEU



GEFÖRDERT VOM

Bundesministerium
für Bildung
und Forschung

Ethisch-soziale Herausforderungen



- **Herausforderung: Mögliche Veränderung der Lehr-, Lern- und Prüfungskultur**
- KI ermöglicht schnelle Erstellung komplexer Inhalte → *Wozu weiterhin Lernen?*



Ethisch-soziale Herausforderungen



- **Mögliche Implikation: Neues Bildungsziel - KI als Lernpartner zielführend und verantwortungsvoll nutzen**
 - KI-bezogenes Wissen und Reflexion
 - Quellenbewertung, -prüfung und -reflexion
 - Inhaltliches Wissen und Prompt-Tuning
 - Selbstregulationskompetenzen

SWK, 2024

Ethisch-soziale Herausforderungen



- **Mögliche Implikation: Neues Bildungsziel - KI als Lernpartner zielführend und verantwortungsvoll nutzen**

- KI-bezogenes Wissen und Reflexion
- Quellenbewertung, -prüfung und -reflexion
- Inhaltliches Wissen und Prompt-Tuning
- Selbstregulationskompetenzen

SWK, 2024

Was bedeutet das für Prüfungen?

- Neue/zusätzliche Kompetenzen prüfen?
- Weg vom summativen hin zum prozessorientierten Prüfen?
- KI bewusst miteinbinden?

Ethisch-soziale Herausforderungen



- **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
 - Hochschullizenzen + Förderung der KI-Kompetenz aller Studierenden!



Ausblick: Weitere Herausforderungen



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“?



**Trotz allem:
AI for good AND for not bad!**



lernen:digital
Kompetenzzentrum
Sprachen/Gesellschaft/Wirtschaft

Vielen Dank für Ihre Aufmerksamkeit!
mesenhoeller@uni-potsdam.de



KISS-PRO



**Finanziert von der
Europäischen Union**
NextGenerationEU



GEFÖRDERT VOM

**Bundesministerium
für Bildung
und Forschung**



lernen:digital
Kompetenzzentrum
Sprachen/Gesellschaft/Wirtschaft

Zeit für Fragen!



KISS-PRO



**Finanziert von der
Europäischen Union**
NextGenerationEU



GEFÖRDERT VOM

**Bundesministerium
für Bildung
und Forschung**

Quellen



- 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>
- 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>
- fobizz. (2024). *Digitalen Unterricht einfach und sicher gestalten*. <https://tools.fobizz.com>
- Jobin, A., Ienca, M. & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence*, 1(9), 389–399. <https://doi.org/10.1038/s42256-019-0088-2>
- 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>
- 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. (2023, August 7). Wie Künstliche Intelligenz der Umwelt schadet. *Deutsche Welle*. <https://www.dw.com/de/wie-k%C3%BCnstliche-intelligenz-der-umwelt-schadet/a-66305844>
- Österreichischer Bundesverlag (2024): KI im Klassenzimmer. Umfrage des Österreichischen Bundesverlags (öbv). Abgerufen von <https://presse.grayling.at/news-oebv-whitepaper-74-der-lehrkraefte-sehen-ki-in-zukunft-als-selbstverstaendlichen-teil-des-unterrichts-?id=200684&menueid=27292&tab=4&l=deutsch>.
- Philologenverband Nordrhein-Westfalen (2024): Umfrage zu ChatGPT & Co. Endergebnisse. Abgerufen von https://phv-nrw.de/wp-content/uploads/2024/01/20240125_Umfrage2_ChatGPT.pdf
- 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>

Quellen



- 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>
- Roos, J. & Schöler, H. (Hrsg.). (2009). *Entwicklung des Schriftspracherwerbs in der Grundschule: Längsschnittanalyse zweier Kohorten über die Grundschulzeit* (1. Aufl.). VS Verl. für Sozialwiss. <https://doi.org/10.1007/978-3-531-91574-6>
- 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>
- Ständige Wissenschaftliche Kommission der Kultusministerkonferenz. (2024). *Large Language Models und ihre Potenziale im Bildungssystem: Impulspapier der Ständigen Wissenschaftlichen Kommission der Kultusministerkonferenz*. <http://dx.doi.org/10.25656/01:28303>
- van de Werfhorst, H. G., Kessenich, E. & Geven, S. (2022). The digital divide in online education: Inequality in digital readiness of students and schools. *Computers and Education Open*, 3, 100100.
<https://doi.org/10.1016/j.caeo.2022.100100>
- van Dijk, J. (2020). *The digital divide*. Polity Press.
- Verordnung (EU) 2024/1689 des Europäischen Parlaments und des Rates vom 13. Juni 2024 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). Amtsblatt der Europäischen Union, L 2024/1689, 12. Juli 2024. <http://data.europa.eu/eli/reg/2024/1689/oj>
- Vodafone Stiftung Deutschland gGmbH. (2024). *Pioniere des Wandels: Wie Schüler:innen KI im Unterricht nutzen möchten*. <https://www.vodafone-stiftung.de/wp-content/uploads/2024/03/Pioniere-des-Wandels-wie-Schueler-innen-KI-im-Unterricht-nutzen-wollen-Jugendstudie-der-VS-2024.pdf>