



**Staatsbibliothek  
zu Berlin**  
Preußischer Kulturbesitz

STAATSBIBLIOTHEK ZU BERLIN – PK

# TEXTGENERIERENDE KI UND GUTE WISSENSCHAFTLICHE PRAXIS

Dr. Katrin Frisch, Ombudsman für die Wissenschaft | Dr. Felix Hagenström, Staatsbibliothek zu Berlin  
6. Mai 2024

Kontakt: [katrin.frisch@ofdw.de](mailto:katrin.frisch@ofdw.de) | [felix.hagenstroem@sbb.spk-berlin.de](mailto:felix.hagenstroem@sbb.spk-berlin.de)

## Teil I – Einführung

- Was ist textgenerierende künstliche Intelligenz (TKI)?
- Was ist gute wissenschaftliche Praxis (GWP)?

## Teil II – Wissenschaftliche Autorschaft

- Zustimmung und Verantwortung
- TKI als Assistentin

## Teil III – GWP-divergente und GWP-konforme Nutzung von TKI

- Täuschung und Fehlverhalten
- Transparenz: Dokumentation und Zitation – Aber wie?

## Teil IV - Schluss

- Fragen und Diskussion

Die Folien werden  
im Anschluss auf  
der Webseite zur  
Verfügung gestellt.

Folien basieren teilweise auf: Hahn, Carolin / Janke, Andreas / Jehle, Anna / Schmitz, Christina (2023, June 2). Online-Workshop: ChatGPT & Co. – Künstliche Intelligenz im Schreibprozess. <https://doi.org/10.5281/zenodo.7865178> (Insbesondere folgende Folien sind (fast) 1-zu-1 übernommen aus Hahn et al. 2023: 5, 16, 18, 19, 21, 22, 26, 27)

Inhalte mit Bezug zur guten wissenschaftlichen Praxis entstammen teilweise: Frisch, Katrin, Felix Hagenström, und Nele Reeg. „Textgenerierende KI und gute wissenschaftliche Praxis“. *Zeitschrift für Bibliothekswesen und Bibliographie* 70, Nr. 6 (15. Dezember 2023): 326–36. <https://doi.org/10.3196/186429502370667>.

# Was ist textgenerierende künstliche Intelligenz (TKI)?

## Textgenerierende künstliche Intelligenz (TKI)

- **Terminologisches:**
  - **Large Language Models:** auf großen Datenmengen trainierte Modelle, die natürlichsprachliche Eingaben verarbeiten können
  - **Generative KI:** Themenoffenheit, nicht für einen spezifischen Einsatzbereich entwickelt
  - **TKI:** kann themenoffen und auf Befehl (per ‚Prompt‘) Texte generieren
- **Stochastische Funktionsweise:** Gemini, ChatGPT, Perplexity etc. erstellen Texte aufgrund probabilistischer Kriterien
- Laufende Entwicklung hin zur **Multimodalität:** keine Beschränkung auf Text oder Bild, sondern Kombination unterschiedlicher Formate wie Text, Bild, Audio und Video (sowohl bei Inputs als auch Outputs)
- **Fokus** des heutigen Workshops: **Textproduktion** = KI im Schreibprozess
  - Welche Regeln gibt es? Welche Probleme und Konflikte können auftreten?
  - Was ist gute wissenschaftliche Praxis?

- Abhängigkeit von der Qualität der verwendeten Daten
- Ein Großteil der Quellen kommt aus dem englischsprachigen Raum
- Kontext wird nicht ‚verstanden‘, sondern Text wird nach Wahrscheinlichkeitswerten produziert
- Häufig keine Angabe von Quellen (Ausnahme: z.B. Perplexity)
- Produktion von nicht-objektiven, verzerrten Antworten („Bias“) und „Halluzinationen“
- Meist keine Reproduzierbarkeit des Outputs
- Strenggenommen sind KI-generierte Texte weder plagiat- noch zitierfähig
  - Gleichwohl: die Texte sind verwend- und kennzeichenbar

# Was ist gute wissenschaftliche Praxis (GWP)?

- Verbindliche Standards für alle Hochschulen und Forschungseinrichtungen in Deutschland, die DFG-Fördergelder erhalten (möchten)
- Erste Denkschrift erscheint 1998 (ausgelöst durch den Fall Brach/Herrmann)
- Überarbeitet in 2013 (nach dem Fall Guttenberg); 17 Empfehlungen
- Erneute Überarbeitung 2019; Denkschrift → Kodex
- Jetzt 19 Leitlinien zur guten wissenschaftlichen Praxis
- 3 Ebenen
  - ▶ Leitlinien
  - ▶ Erläuterungen
  - ▶ Onlineportal mit Ergänzungen ([www.wissenschaftliche-integritaet.de](http://www.wissenschaftliche-integritaet.de))



(DFG-Kodex 2019)



## Erläuterung zu Leitlinie 1:

### Verpflichtung auf die allgemeinen Prinzipien

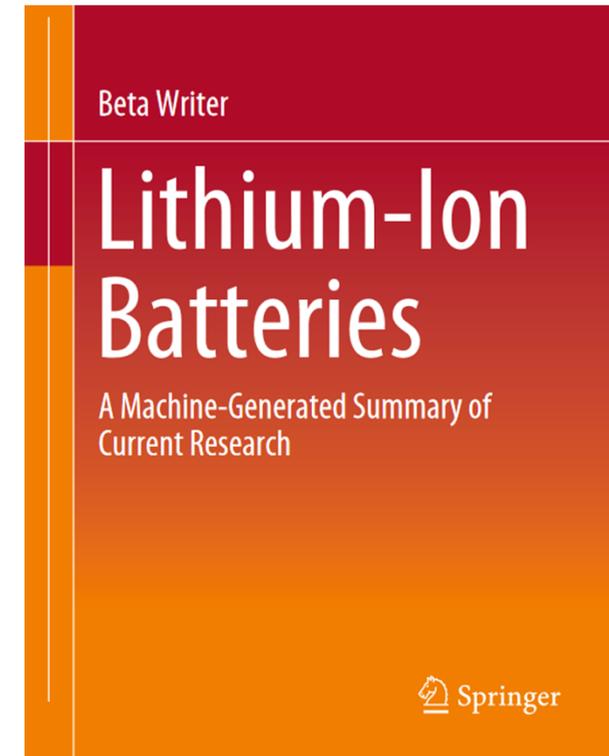
Zu den Prinzipien gehört es insbesondere,

- *lege artis* zu arbeiten,
- strikte Ehrlichkeit im Hinblick auf die eigenen und die Beiträge Dritter zu wahren
- alle Ergebnisse konsequent selbst anzuzweifeln sowie einen kritischen Diskurs in der wissenschaftlichen Gemeinschaft zuzulassen und zu fördern (DFG-Kodex 2019: 9)

- **(Schwerwiegendes) Fehlverhalten:** verbreitetste Definition umfasst die *Erfindung* und *Fälschung* von Daten sowie das *Plagiat* (**fabrication, falsification, plagiarism = FFP**; siehe [NAS/NAE/IoM 1992](#))
- **Fragwürdige Forschungspraktiken (Questionable Research Practices = QRP):** fallen häufig nicht unter die Kategorie Fehlverhalten, verstoßen aber trotzdem gegen die GWP
- In Deutschland ist Fehlverhalten **nicht so eng definiert** (siehe etwa [DFG-Verfahrensordnung zum Umgang mit wissenschaftlichem Fehlverhalten \(VerfOwF\)](#))
- Wissenschaftliches Fehlverhalten erfordert laut der VerfOwF der DFG **Vorsatz oder grobe Fahrlässigkeit**

# Wissenschaftliche Autorschaft

- Genereller Konsens, dass KI keine Autorin sein kann
- Doch: Publikationen mit KI als Autorin existieren (teilweise älter als die Empfehlungen), z.B. Beta Writer (2019): *Lithium-Ion Batteries. A Machine-Generated Summary of Current Research*. New York: Springer.
- Umfrage unter Forschenden zeigte außerdem: 25 % stimmten zu, dass es möglich sein sollte, KI als Autorin zu listen (Fecher et al. 2023)
- KI erfüllt gängige Kriterien von Autorschaft nicht, weil
  - sie keine Verantwortung für den Text übernehmen kann
  - sie keine Zustimmung zum Manuskript geben kann
- Die Verantwortung für den Einsatz von KI und deren „Beitrag“ liegt bei den (menschlichen) Autor:innen einer Publikation
- Möglicherweise muss die Definition wissenschaftlicher Autorschaft in Zukunft angepasst werden



### DFG-Leitlinie 14: Autorschaft

„Autorin oder Autor ist, wer einen genuinen, nachvollziehbaren Beitrag zu dem Inhalt einer wissenschaftlichen Text-, Daten- oder Softwarepublikation geleistet hat. [...]. Sie tragen für die Publikation die gemeinsame Verantwortung, es sei denn, es wird explizit anders ausgewiesen.“ ([DFG-Kodex 2019](#))

### TKI als Assistentin

- Einsatz von KI ≠ GWP-Verstoß
- TKI kann als Assistentin oder Werkzeug begriffen werden, hat bekanntermaßen Risiken und Schwächen, und ihre Nutzung im wiss. Kontext unterliegt u.a. den Regeln der GWP
- Nutzung von KI muss transparent dargelegt werden
- Entstehen durch den KI-Einsatz Verstöße gegen die GWP, tragen die Autor:innen dafür die Verantwortung

# GWP-divergente und GWP- konforme Nutzung von TKI

- Wer die Nutzung von TKI im wissenschaftlichen Kontext nicht offenlegt, d.h. nicht hinreichend kennzeichnet, handelt entgegen der GWP
  - Denn: Die nicht hinreichend offengelegte Nutzung von TKI *täuscht* über die wiss. Leistung und verschleiert die Nachvollziehbarkeit (vgl. DFG-Leitlinien 12 und 13)
- Nutzung von KI muss transparent und sachgerecht angegeben werden
- Einheitliche und klare Regelungen dafür existieren nicht überall, werden z.T. noch ausgehandelt
- Impulsgeber:innen dafür sind maßgeblich
  - Verlage und Journals
  - Fachcommunitys

### Wichtig

Derzeit erfolgt an vielen Hochschulen, bei Verlagen und in den Fachcommunitys eine intensive Auseinandersetzung mit der Frage, wie KI-Tools regelkonform eingesetzt werden können. Es ist zu erwarten, dass auch diejenigen Akteure, die sich bislang nicht äußerten, Empfehlungen veröffentlichen.

- Wer die Nutzung von TKI im wissenschaftlichen Kontext nicht offenlegt, d.h. **nicht hinreichend kennzeichnet**, handelt entgegen der GWP
  - Denn: Die **nicht hinreichend offengelegte** Nutzung von TKI *täuscht* über die wiss. Leistung und verschleiern die Nachvollziehbarkeit (vgl. DFG-Leitlinien 12 und 13)
- Nutzung von KI muss **transparent und sachgerecht** angegeben werden
- Einheitliche und klare Regelungen dafür existieren noch nicht, sondern sind noch im Aushandlungsprozess
- Impulsgeber:innen dafür sind maßgeblich
  - Verlage und Journals
  - Fachcommunitys

### Wichtig

Derzeit erfolgt an vielen Hochschulen, bei Verlagen und in den Fachcommunitys eine intensive Auseinandersetzung mit der Frage, wie KI-Tools regelkonform eingesetzt werden können. Es ist zu erwarten, dass auch diejenigen Akteure, die sich bislang nicht äußerten, Empfehlungen veröffentlichen.

Einschätzungen von Thomas Hoeren bzgl. des grundsätzlichen – gekennzeichneten – Einsatzes von Textgeneratoren im akademischen Kontext:

- „Der markierte Einsatz von KI-Schreibwerkzeugen in wissenschaftlichen Arbeiten wird [...] formal wohl nicht gegen die Regeln guter wissenschaftlicher Praxis verstoßen [...].“ Quelle: Hoeren (März 2023): Rechtsgutachten zum Umgang mit KI-Software im Hochschulkontext, S. 32.
- Maßgeblich für die Kennzeichnung ist: „Für Dritte muss erkennbar sein, welche Textteile in welchem Ausmaß von einer KI generiert wurden.“ Quelle: Hoeren (März 2023, S. 29).

ABER:

- Regeln guter wissenschaftlicher Praxis sind **nicht einheitlich**.
- Für unterschiedliche Fächer & Fachbereiche gelten **unterschiedliche Anforderungen**.
- Entscheidend für die Frage, ob und in welchem Rahmen KI-Tools eingesetzt werden dürfen, sind die **Rahmenvorschriften, Prüfungsordnungen, Selbstständigkeitserklärungen, Publikationsvereinbarungen** usw. der DFG, der Hochschulen, Fachbereiche, Studiengänge und Verlage!

- Dokumentation von KI-Nutzung sollte
  - einerseits der tatsächlichen Arbeitsweise mit der KI gerecht werden;
  - andererseits den Anforderungen der Rezipient:innen
  - Offene Frage: Wie detailliert sollte dokumentiert werden?
- Wo sollte die Nutzung von KI angegeben werden? Verschiedene Vorschläge:
  - Methodenteil
  - Danksagung
  - Bibliographie
  - Contributor Statement
  - Im Text
- Sollen Prompts dokumentiert werden?
  - Interne Dokumentation
  - Supplement/Appendix
  - „Clearly indicate which parts of the manuscript used the output of the language bot, and provide the prompts and questions, and/or transcript in the Supporting Information“ (ACS Nano 2023)

## Wie kennzeichne ich die Übernahme KI-generierter Text(-teil)e?

Allgemeine Empfehlung (angelehnt an die Handreichung der Modern Language Association):  
Folgende **Elemente** sollten angegeben werden:



Literaturhinweis: "How do I cite generative AI in MLA style?" MLA Style Center (17.03.2023),  
<https://style.mla.org/citing-generative-ai/>

Beispiel-Prompt: **In 50 Wörtern: Was ist ein Greif?**

Durch ChatGPT generierter Text: Ein Greif ist eine mythologische Kreatur mit dem Körper eines Löwen und den Flügeln und dem Kopf eines Adlers. Diese Mischung aus Raubtier und Vogel wird oft als Symbol für Stärke und Macht betrachtet. Greife sind in verschiedenen Kulturen und Geschichten zu finden und haben eine lange Tradition in der Mythologie.

→ Wortwörtliche Übernahme des KI-generierten Textteils: [...] „Ein Greif ist eine mythologische Kreatur mit dem Körper eines Löwen und den Flügeln und dem Kopf eines Adlers.“ [...]

→ Paraphrase: [...] Ein Greif ist ein fantastisches Wesen, das den Körper eines Löwen sowie die Flügel und den Kopf eines Adlers hat. [...]

Vorschlag für die Zitation (in einer Fußnote):

**„In 50 Wörtern: Was ist ein Greif?“, Prompt, *ChatGPT*, Version: 24.5.2023, OpenAI, 7.6.2023, <https://chat.openai.com/share/a91fe463-e8a8-4c68-97a5-c2fc73ed2638>**

### Weitere Vorschläge für Verweise auf KI-Programme

- **Hosseini et al. 2023:** Open AI chatbot. 2022c. Response to Query Made by David B Resnik, December 11, 2022.9:54pm EST
- **APA (American Psychological Association):** OpenAI. (2023). ChatGPT (Mar 14 version) [Large language model]. <https://chat.openai.com/chat>
- **Leitfaden «Aus KI zitieren», Uni Basel:** ChatGPT version 3.5, OpenAi: [openai.com/chat](https://openai.com/chat),
  - o «Was ist Geologie», Ausgabe vom 23.03.2023
  - o Hilfe bei der Erstellung der Textstruktur

## Wie kennzeichne ich den Einsatz von Textgeneratoren als Inspirationsquelle?

**Hilfsmittelangabe** zu Beginn oder am Ende der Arbeit einfügen, in dem der **Textgenerator** [Produktname, verwendete Version, Name des Herausgebers/ der Herstellerfirma, allgemeine URL zum Textgenerator] sowie die **Art der Nutzung** [zum Beispiel als Denkanstoß, zur Themenfindung, für erste Strukturierungsvorschläge oder zur stilistischen Überarbeitung des eigenen Texts] spezifiziert wird.

Vorgeschlagene Zitation für detaillierte Prompts/Chained Prompts:  
Eigenleistung soll zum Ausdruck kommen (Veränderung gegenüber der Zitierweise auf der vorhergehenden Folie sind in Orange markiert):

[Allgemein beschreibender Titel des Themas/der Frage], [**detaillierter Prompt** bzw. **Chained Prompt erstellt durch [Name Nutzer:in]**], [*Textgenerator* (kursiv gesetzt)], [Version], [Herausgeber/Hersteller], [Datum der Textgenerierung], [**eindeutige URL mit der Dokumentation der Interaktion**]

### **Wichtig**

Eine Möglichkeit, die Eigenleistung im Umgang mit der KI zu belegen, führt über die Dokumentation der Interaktion mit der Software. So ist es beispielsweise möglich, den kompletten Chat-Verlauf über ein Forschungsrepositorium mit einem Persistent Identifier (z.B. DOI) zu verlinken.

- Kein prinzipieller Ausschluss des Einsatzes generativer KI-Modelle
- Forderung nach „transparentem Umgang mit der Erzeugung von Text- und Bildinhalten“
- Betonung der Bedeutung von „verbindlichen Rahmenbedingungen“
- „Transparenz und Nachvollziehbarkeit des Forschungsprozesses und der gewonnenen Erkenntnisse für Dritte“ als „wesentliche Grundprinzipien wissenschaftlicher Integrität“ – der Einsatz von KI-Modellen entbindet die Wissenschaftler:innen nicht von diesen Grundprinzipien
- Forderung nach Offenlegung, „ob und welche generativen Modelle [...] zu welchem Zweck und in welchem Umfang eingesetzt“ wurden
- Nur natürliche Personen können verantwortliche Autor:innen wissenschaftlicher Publikationen sein
- Bei der Förderantragstellung dürfen KI-Modelle verwendet werden, bei der Erstellung von Gutachten nicht

Siehe:

Präsidium der Deutschen Forschungsgemeinschaft (DFG), „Stellungnahme des Präsidiums der Deutschen Forschungsgemeinschaft (DFG) zum Einfluss generativer Modelle für die Text- und Bilderstellung auf die Wissenschaften und das Förderhandeln der DFG“, September 2023, <https://www.dfg.de/resource/blob/289674/ff57cf46c5ca109cb18533b21fba49bd/230921-stellungnahme-praesidium-ki-ai-data.pdf>.



## Gute wissenschaftliche Praxis und künstliche Intelligenz

🏠 BEITRÄGE > GUTE WISSENSCHAFTLICHE PRAXIS UND KÜNSTLICHE INTELLIGENZ

Hier finden Sie Literatur und weitere Beiträge zur guten wissenschaftlichen Praxis (GWP) und künstlicher Intelligenz (KI). Der Fokus liegt dabei auf Themen, die Forschungspraxis und Wissenschaft als Ganzes betreffen, und weniger auf Lehre und Prüfungsangelegenheiten. Viele Stellungnahmen von Universitäten, die sich hauptsächlich mit Letzterem befassen, sind daher nicht gelistet. Da die Entwicklungen im Bereich KI sehr fluide sind, besteht für den Umgang mit KI zur Zeit zwar schon ein gewisser Konsens, aber es existieren dennoch teils heterogene Empfehlungen in Bezug auf Details.

### Inhaltsverzeichnis

- Allgemein
- Autorschaft und künstliche Intelligenz
- KI-Nutzung in Publikationen transparent machen
- Verstöße und Detektion im Zusammenhang mit KI
- Allgemeine Empfehlungen zum Umgang mit KI
- Einsatz von künstlicher Intelligenz bei Peer Review



Wo finde ich meine  
lokale  
Ombudsperson?



<https://ombudsman-fuer-die-wissenschaft.de/12365/gute-wissenschaftliche-praxis-und-kuenstliche-intelligenz/>

## KI-RESSOURCEN

Hier geben wir einen Überblick über KI-Tools im Kontext von akademischen Lese- und Schreibprozessen. Diese Liste erfüllt weder Werbezwecke noch sind damit Aussagen über die Qualität der jeweiligen Tools verbunden. Im Zuge der gegenwärtig rasanten Entwicklung KI-gestützter Software hat diese Übersicht auch keinen Anspruch auf Vollständigkeit, versteht sich aber als Living Document. Wir freuen uns daher über jede Art von [Ergänzungsvorschlägen](#).

### LITERATURRECHERCHE

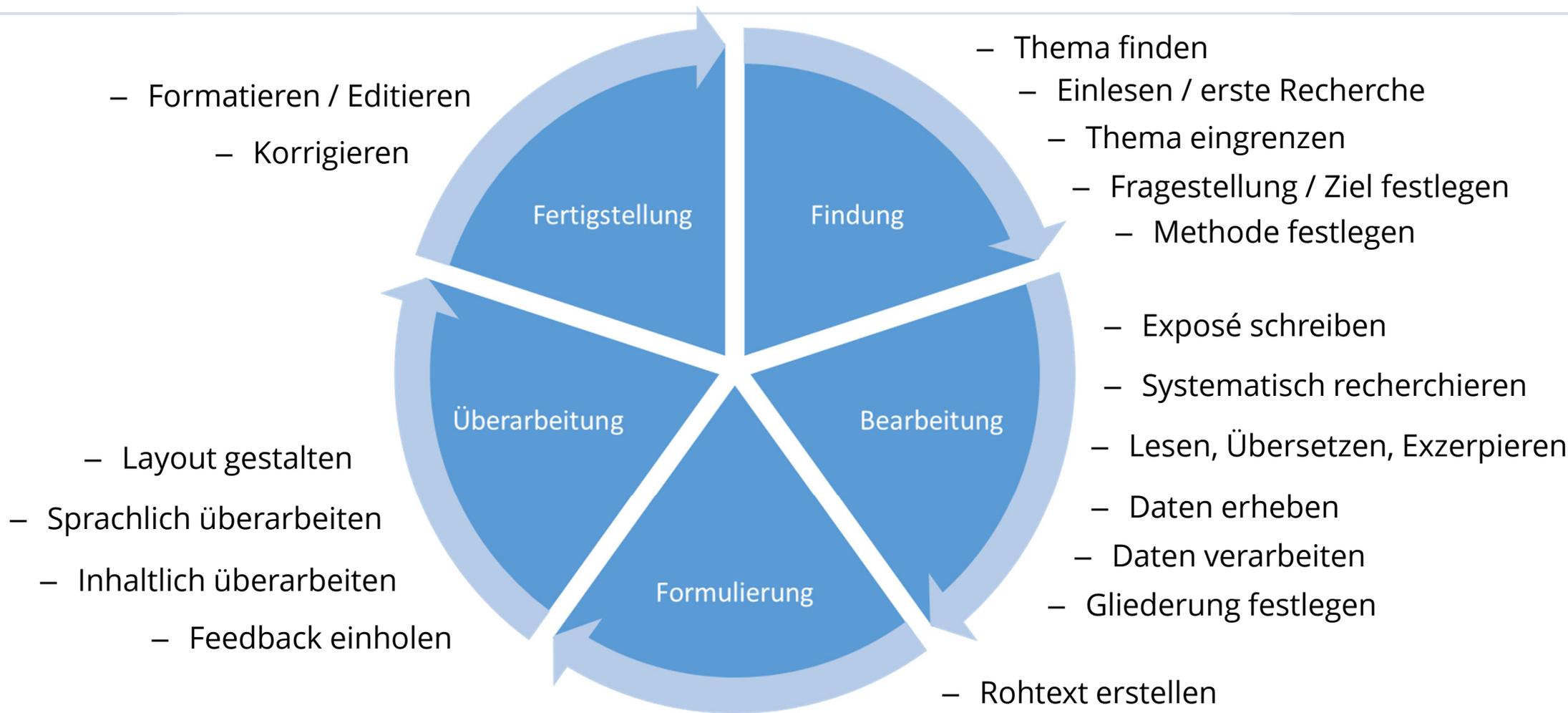
- [connected papers](#)
- [Consensus](#)
- [Elicit](#)
- [Litmaps](#)
- [Open Knowledge Maps](#)
- [Perplexity AI](#)
- [R Discovery](#)
- [Research Rabbit](#)
- [scinapse](#)
- [Semantic Scholar](#)

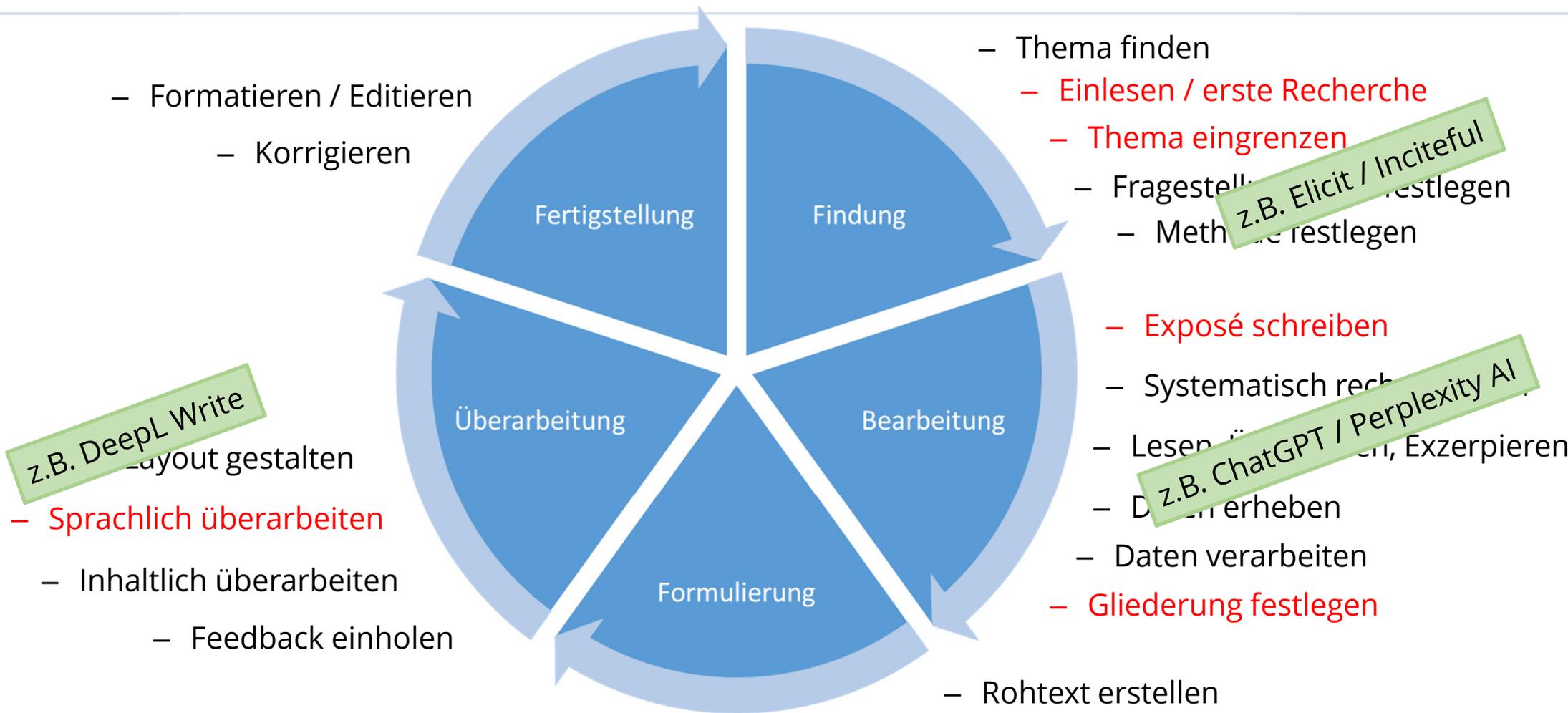
### TEXTVERSTEHEN

- [ChatPDE](#)
- [Explainpaper](#)
- [Humata](#)
- [KISOFT](#) (noch in der Entwicklung)
- [PaperBrain](#)
- [Scispace](#)

### TEXTPRODUKTION

- [Aleph Alpha](#)
- [Alpa](#)
- [CopyCockpit](#)
- [ChatGPT](#)





Es lassen sich mindestens fünf Implikationen unterscheiden:

- (1) the need for accountability in relation to the outcomes produced by LLMs,
- (2) the question of originality with regards to human creativity (e.g., concerns of plagiarism arise)
- (3) the sustainability issue regarding the environmental effects of LLMs,
- (4) the potential exclusion of researchers who lack access to LLMs, raising concerns about universalism,
- (5) the issue of autonomy, in which researchers may become overly dependent on (commercial) AI tools.

Benedikt Fecher u. a.,  
„Delphi Study: Exploring the  
Implications of Large  
Language Models on the  
Science System“ (Zenodo, 6.  
Juni 2023),  
<https://doi.org/10.5281/ZENODO.8009429>.

- *ACS Nano* 2023, 17, 5, 4091–4093. Publication Date: February 27, 2023. <https://doi.org/10.1021/acsnano.3c01544>
- Deutsche Forschungsgemeinschaft. *Leitlinien zur Sicherung guter wissenschaftlicher Praxis*. Bonn, 2019.
- Deutsche Forschungsgemeinschaft: Verfahrensordnung zum Umgang mit wissenschaftlichem Fehlverhalten (VerfOwF). Bonn, 2019.
- Deutsche Forschungsgemeinschaft: „Stellungnahme des Präsidiums der Deutschen Forschungsgemeinschaft (DFG) zum Einfluss generativer Modelle für die Text- und Bilderstellung auf die Wissenschaften und das Förderhandeln der DFG“, September 2023, <https://www.dfg.de/resource/blob/289674/ff57cf46c5ca109cb18533b21fba49bd/230921-stellungnahme-praesidium-ki-ai-data.pdf>.
- Fecher, B., Hebing, M., Laufer, M., Pohle, J., & Sofsky, F. (2023). Delphi Study: Exploring the Implications of Large Language Models on the Science System (1.0) [Data set]. Zenodo. <https://doi.org/10.5281/zenodo.8009430>
- Frisch, Katrin, Felix Hagenström, und Nele Reeg. „Textgenerierende KI und gute wissenschaftliche Praxis“. *Zeitschrift für Bibliothekswesen und Bibliographie* 70, Nr. 6 (15. Dezember 2023): 326–36. <https://doi.org/10.3196/186429502370667>.
- Hahn, Carolin / Janke, Andreas / Jehle, Anna / Schmitz, Christina (2023, June 2). Online-Workshop: ChatGPT & Co. – Künstliche Intelligenz im Schreibprozess. <https://doi.org/10.5281/zenodo.7865178>
- Hoeren, T. Rechtsgutachten zum Umgang mit KI-Software im Hochschulkontext. In: *Didaktische und rechtliche Perspektiven auf KI-gestütztes Schreiben in der Hochschulbildung*. Ruhr-Universität Bochum.
- Hosseini, M., D. B. Resnik und K. Holmes. The ethics of disclosing the use of artificial intelligence tools in writing scholarly manuscripts. *Research Ethics*, 2023. <https://doi:10.1177/17470161231180449>
- “How do I cite generative AI in MLA style?” MLA Style Center (17.03.2023), <https://style.mla.org/citing-generative-ai/>
- National Academy of Sciences, National Academy of Engineering, and Institute of Medicine. 1992. *Responsible Science: Ensuring the Integrity of the Research Process: Volume I*. Washington, DC: The National Academies Press. <https://doi.org/10.17226/1864>