Staatsbibliothek
zu Berlin
Preußischer Kulturbesitz

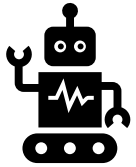
Staatsbibliothek zu Berlin

Literaturrecherche mit KI

Yella Nicklaus
yella.nicklaus@sbb.spk-berlin.de

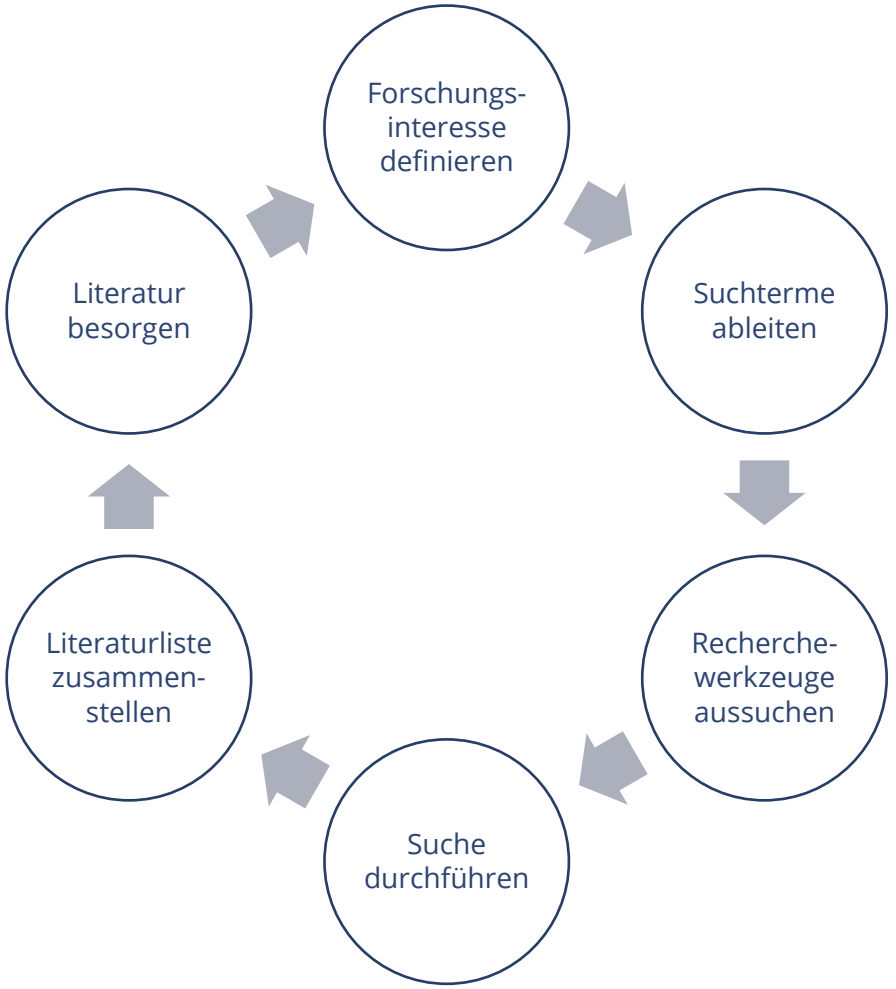
Was erwartet Sie?

- **Einstieg**: Wie recherchieren Sie?
- Der **Literaturrechercheprozess** und wo KI helfen kann
- **LLMs** und Literaturrecherche
- **KI-Rechercheassistenten**
- **Empfehlungen und Fazit**
- Gemeinsames **Ausprobieren**
- **Offene Fragen**

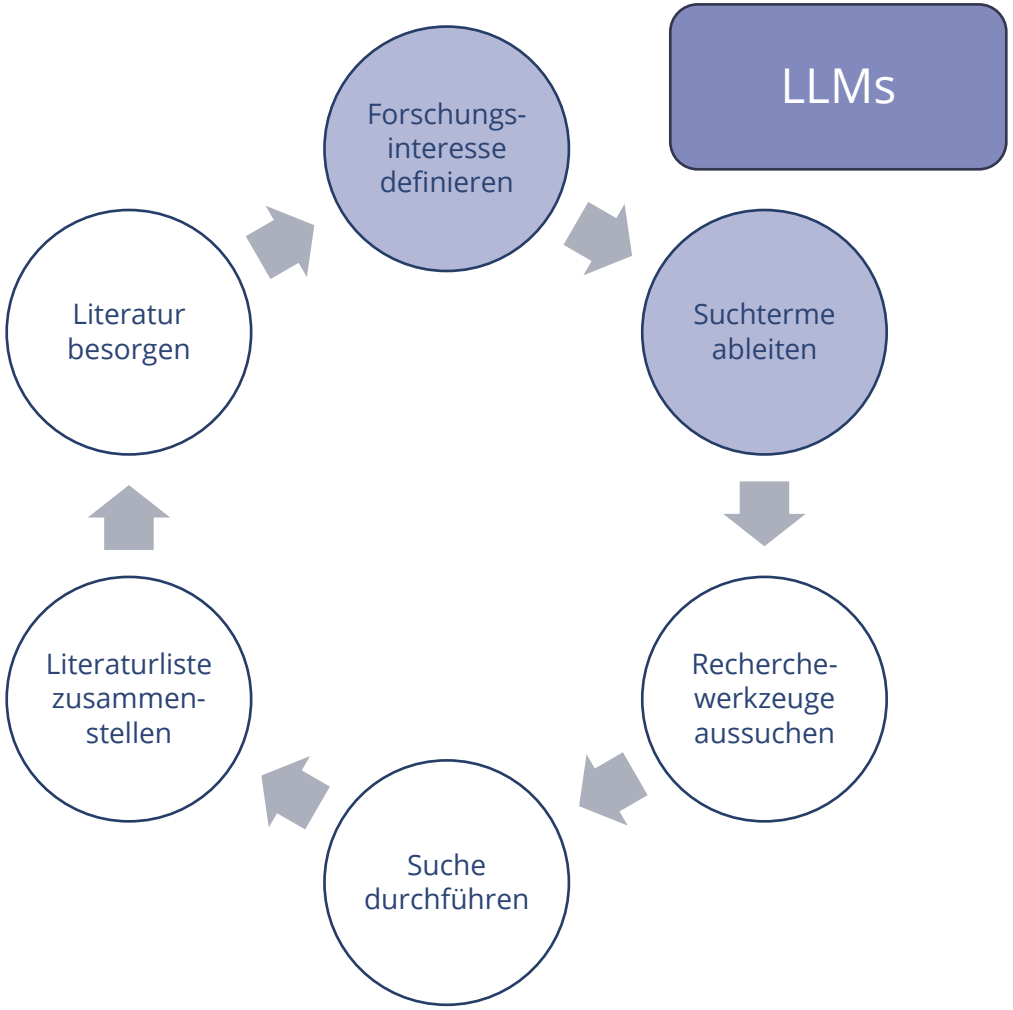


Wie recherchieren Sie?

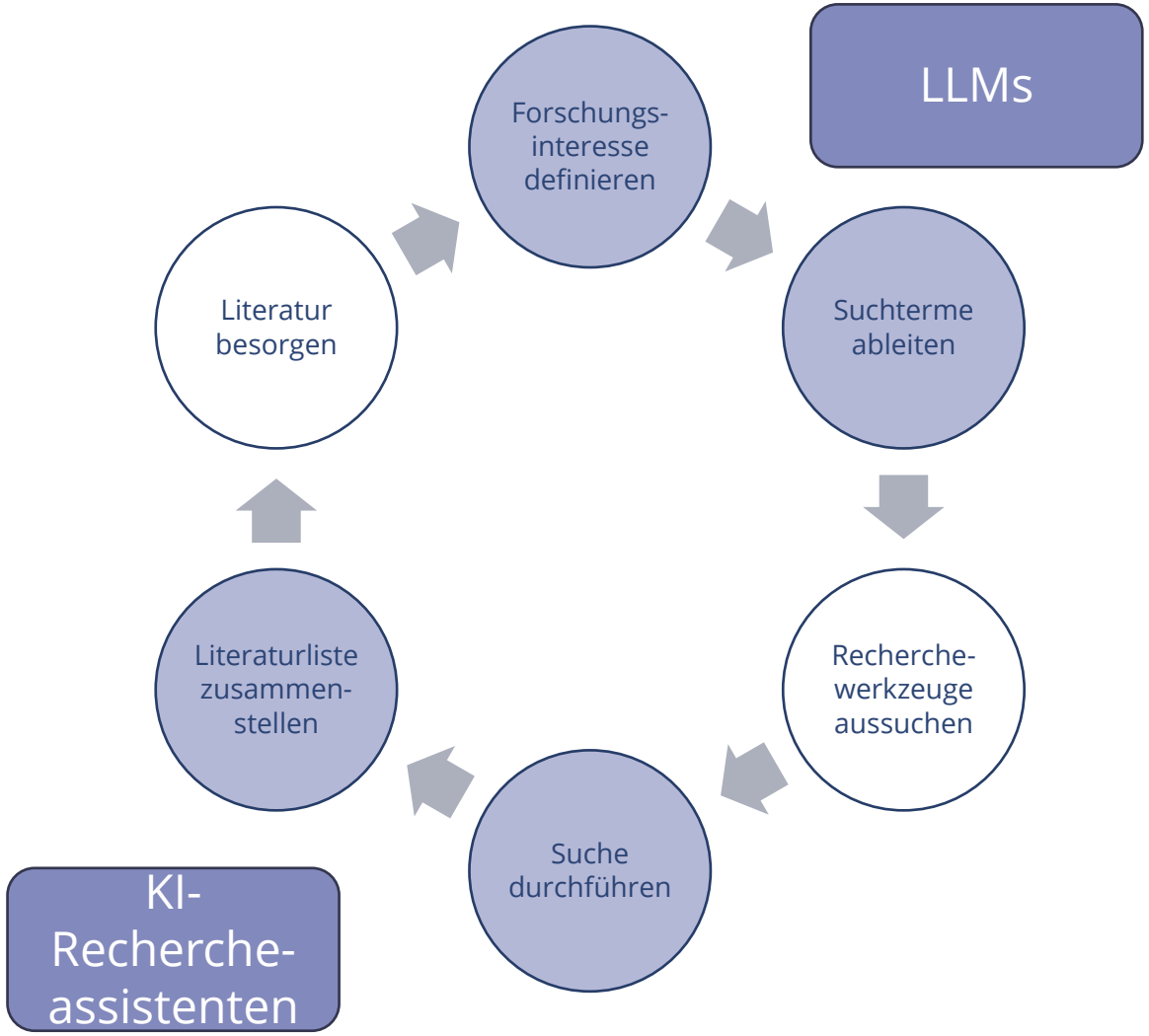
Literaturrecherche klassisch



Literaturrecherche klassisch



Literaturrecherche klassisch



Auf **großen Datenmengen** trainierte KI-Modelle, die **natürlichsprachige** Eingaben verarbeiten und **Text generieren** können.

Interaktive Erklärung der Funktionsweise:
Generative AI exists because of the transformer

- Generierter Output basiert auf **Wahrscheinlichkeiten**, nicht auf gespeicherten Informationen
  → „Halluzinationen“ sind möglich
- **Funktionsweise:** Das Modell setzt Text so fort, dass die Antwort für Menschen plausibel und hilfreich wirkt und weitere Interaktionen generiert
- Nicht für einen spezifischen Use Case entwickelt, sondern **themenoffen**
- Antworten werden immer **neu generiert** → keine Reproduzierbarkeit

Claude

perplexity

ChatGPT

Gemini

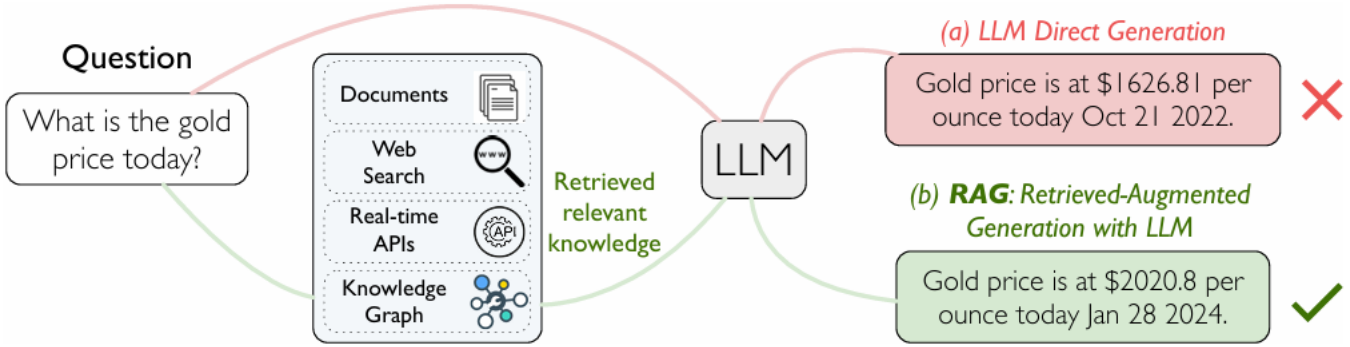


Figure 1: QA using LLMs (a) without RAG vs. (b) with RAG.

Viele LLMs greifen zur Generierung von Antworten nicht nur auf ihre Trainingsdaten zu, sondern nutzen zusätzlich **Web- und Datenbankabfragen**.

Yang, Xiao, et al. ‘CRAG -- Comprehensive RAG Benchmark’. *arXiv [Cs.CL]*, 2024, arxiv.org/abs/2406.04744. arXiv.

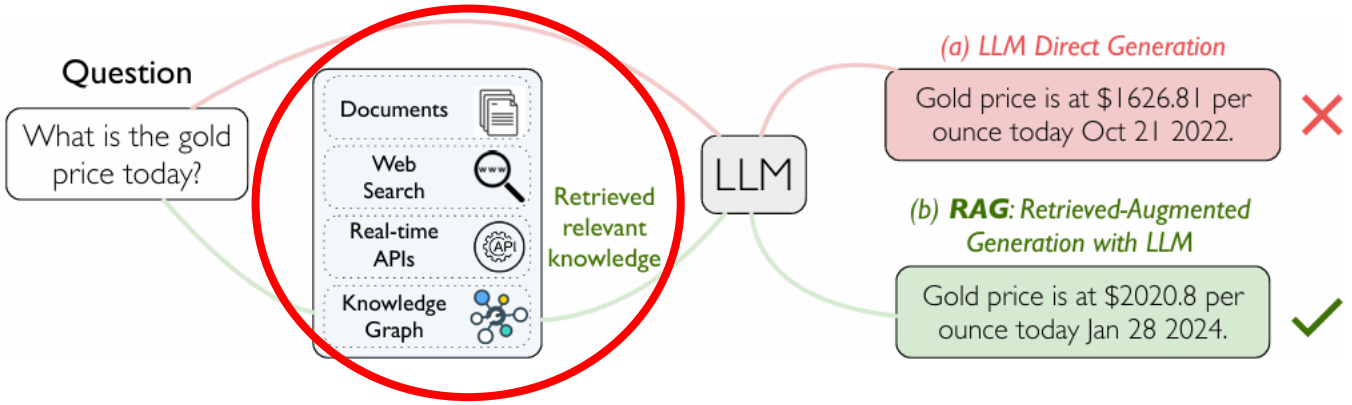


Figure 1: QA using LLMs (a) without RAG vs. (b) with RAG.

Yang, Xiao, et al. 'CRAG -- Comprehensive RAG Benchmark'. *arXiv [Cs.CL]*, 2024, arxiv.org/abs/2406.04744. arXiv.

Viele LLMs greifen zur Generierung von Antworten nicht nur auf ihre Trainingsdaten zu, sondern nutzen zusätzlich **Web- und Datenbankabfragen**.

Welche Datenbestände hier durchsucht werden, ist meist nicht transparent. Im Normalfall handelt es zum großen Teil **nicht um wissenschaftliche Literatur**.

Wofür eignen sich LLMs aktuell?

LLMs können sinnvoll unterstützen bei:

- Brainstorming zur Konzeption eines Forschungsthemas
- Schärfung von Forschungsfragen
- Finden von geeigneten Suchbegriffen
- Unterstützung von Suchstrategien

→ **Vorbereitung der Literaturrecherche**

LLMs eignen sich nicht für:

- Generieren von Literaturlisten
- Beantworten wissenschaftlicher Fragestellungen

Spezialisierte Tools zur **wissenschaftlichen Recherche**, die **LLM-Funktionalitäten** mit einer **wissenschaftlichen Datenbasis** kombinieren.

Zwei Haupttypen: **Finder-** und **Connectortools**

Funktionieren am besten für **Lebens- und Naturwissenschaften**, weil

- Als **Datenbasis** primär **englischsprachige Open-Access-Artikel** und **Preprints** genutzt werden.
- KI-Tools besser mit stark und **einheitlich strukturierten** Texten und **quantitativen Fragestellungen** umgehen können.
- Sich mit diesen Fächern am meisten Geld verdienen lässt.

Können **natürlichsprachige Suchanfragen** verarbeiten und als Output eine **Liste wissenschaftlicher Literaturtitel** ausgeben.

ORKG Ask · Undermind · scite_ · Google Scholar | Labs · Consensus · Elicit · SciSpace · Discovery by Editage

… und teilweise noch mehr:

- Die Treffer in Hinblick auf die Forschungsfrage nach Relevanz sortieren.
- Eine zusammenfassende Antwort auf die Forschungsfrage generieren.
- Einzelne Treffer in Hinblick auf die Forschungsfrage zusammenfassen.
- Weiterführende Fragen auf Grundlage der Ergebnisse beantworten.

Erstellen **grafische Verbindungsnetzwerke** wissenschaftlicher Quellen basierend auf Zitationen oder inhaltlicher Nähe.

OPEN KNOWLEDGE MAPS
Your guide to scientific knowledge

CONNECTED PAPERS

ResearchRabbit™

Inciteful

- Ausgangspunkt: „Seed Paper“ oder vorgeschaltete Stichwortsuche
- Werten die gefundenen Treffer hinsichtlich verschiedener Kriterien aus, z.B. Zitierhäufigkeit, Tags/Keywords, …

Die KI-Tools finden nur Treffer, die in ihrer Datenbasis indexiert sind.

Je nach Tool ist diese Datenbasis unterschiedlich und häufig nicht vollständig transparent. Viele Tools greifen auf den Index der Datenbank **Semantic Scholar** zu. Diese wirbt:

> We index over 200 million academic papers sourced from publisher partnerships, data providers, and web crawls.

Der größte Teil der Suchergebnisse in allen Tools besteht aktuell aus englischsprachigen Open-Access-Artikeln oder Preprints.

Kaum zu finden sind:

- Monografien
- Print-Ressourcen
- Closed-Access-Publikationen

Das derzeit einzige KI-Tool mit einer deutlich größeren Datenbasis ist Google Scholar Labs.

In den meisten Fächern können die KI-Recherchetools klassische Recherchetechniken nicht ersetzen, aber sinnvoll ergänzen.

- **Kombinieren** Sie KI-gestützte und klassische Recherchetechniken.
- KI-generierte Literaturangaben können nach wie vor **fehlerhaft** sein – immer überprüfen, z.B. mit BibFox
- „Alles steht und fällt mit der Datenbasis“ – prüfen Sie, welche Ressourcen überhaupt auffindbar sind.
- Das Feld **entwickelt** sich stetig weiter – überprüfen Sie immer mal wieder, wie sich die Qualität der Funktionalitäten verändert hat.
- KI-Tools haben einen sehr hohen **Energieverbrauch** – prüfen Sie, ob die Tools einen tatsächlichen Mehrwehrt für Ihr Rechercheanliegen bringen.
- KI-Unternehmen verdienen ihr Geld mit **Daten** und verfolgen **kommerzielle Interessen** – gehen Sie sorgsam mit Daten von anderen um und überlegen Sie, welche Daten Sie von sich preisgeben möchten.

Sie arbeiten in Breakout-Räumen von 3-4 Personen. Alle Informationen finden Sie im Pad.

Einigen Sie sich auf eine **Forschungsfrage / ein Forschungsthema** (oder nutzen Sie eins der Beispiele). Probieren Sie mindestens zwei Tools (z.B. ein LLM und ein Finder-Tool) aus und vergleichen Sie:

- Welche Arten von Quellen werden gefunden?
- Ist der generierte Output für Ihr Forschungsthema relevant?
- Falls Zusammenfassungen etc. generiert werden: Sind die Aussagen korrekt und stimmen mit der Quelle überein?

Eine Liste von Tools finden Sie im Chat. Sie können auch weitere Tools, die Sie kennen, ausprobieren.

Am Ende vergleichen wir Ihre Erfahrungen mittels Abfrage im Chat. Bitte notieren Sie, **welche Tools** Sie genutzt haben, welche **Frage** Sie recherchiert haben und wie Ihr **Fazit** ist.

Zeit: 20 Minuten.

Welche Fragen sind offen geblieben?

KI-Findertools:

- ORKG Ask (kostenlos, open Source, datenschutzkonform)
- Consensus
- Undermind
- Elicit
- Scite
- SciSpace
- Google Scholar Labs
- R Discovery

Connector-Tools:

- ResearchRabbit
- OpenKnowledgeMaps (kostenlos, open source, datenschutzkonform)
- Connected papers
- Inciteful

LLMs in sicherer Umgebung:

- ChatAI via Academic Cloud

Verwendete Quellen und Further Reading

- Dreisiebner, S., Gäde, M., Griesbaum, J., Michel, A., Petschenka, A., Tappenbeck, I., Wittich, A. (2026). *Fachspezifische Informationskompetenz-Vermittlung und Künstliche Intelligenz.* https://doi.org/10.18452/35926
- Frisch, Kathrin / Hagenström, Felix (2024). Online-Workshop: Textgenerierende KI und gute wissenschaftliche Praxis. https://blog.sbb.berlin/wp-content/uploads/Stabi_Praesentation_TKI_v20240506.pdf
- 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
- Hochschul- und Landesbibliothek Rhein-Main: Rechercheplattformen mit KI-Unterstützung
- Katalog für Open–Science-Tools – Open Economics Guide der ZBW
- KI-Tools - KI-Ressourcen | Virtuelles Kompetenzzentrum KI und wissenschaftliches Arbeiten
- Kullmann, S., & Hiebl, J. (2023). Artificial Intelligence (AI) Research Assistants in der Praxis. Zenodo. https://doi.org/10.5281/zenodo.10409377
- Lahrsow,M. & Petrovska, L. Universität Tübingen. (2026). Literaturrecherche mit KI – Tipps und Tools. http://hdl.handle.net/10900.3/OER_AXQADCPO
- Lahrsow, M. & Baumgarten, L. Universität Tübingen. (2026). AI Research Tool Overview.
- ULB-LOTSE: KI-Tools für die Literaturrecherche