

Vom Experiment zur Praxis

–

Realistische Einsatzszenarien für Large Language Models in kleinen Archiven

29. Archivwissenschaftliches Kolloquium

Dominic Eickhoff | Fachberater für Digitalisierung

© LWL-Archivamt für Westfalen (AAW)

20. Mai 2025

Ausgangslage

- Aktuell: Erkennung und Metadatenextraktion
- Generative KI kommt wenig zum Einsatz
- Welche Lösungen eignen sich für kleine Archive?
- Wie sollten kleine Archive anfangen?
- Welche Daten spielen eine Rolle?

Wie wird KI in (Kommunal)archive kommen?

Per AFIS: Tiefe Integration & Automatisierung in bestehende Systeme

Per Dienstleister-Beauftragung: Pipelines, Datensätze, „make or buy“-Entscheidung

Per Pilotprojekt: Akquise von Drittmitteln & Know How, gute Dokumentation

Per Kommunalen Infrastruktur: Von etabliertem KI-Bürgerservice profitieren

Per Experiment: Kommerzielle oder Open Source Tools einmalig testen

Gar nicht: Wenn nicht experimentiert wird!

→ **Voraussetzung ist immer:** Gute Datengrundlage und wirtschaftliche Use Cases

Eigene Darstellung

In welchen Phasen läuft die KI-Integration ab?

Phase 1: Erkundung – Potenziale entdecken

- Anekdote Nutzung
- Kompetenz-Aufbau & Datenverständnis

Phase 2: Standardisierung – Grundlagen schaffen

- Plattformen etablieren
- Richtlinien, Strategien, Mittelplanung

Phase 3: Integration – Mehrwert erzeugen

- Tiefe Integration & KI-Governance
- Fokus auf Nachhaltigkeit & Nachnutzbarkeit

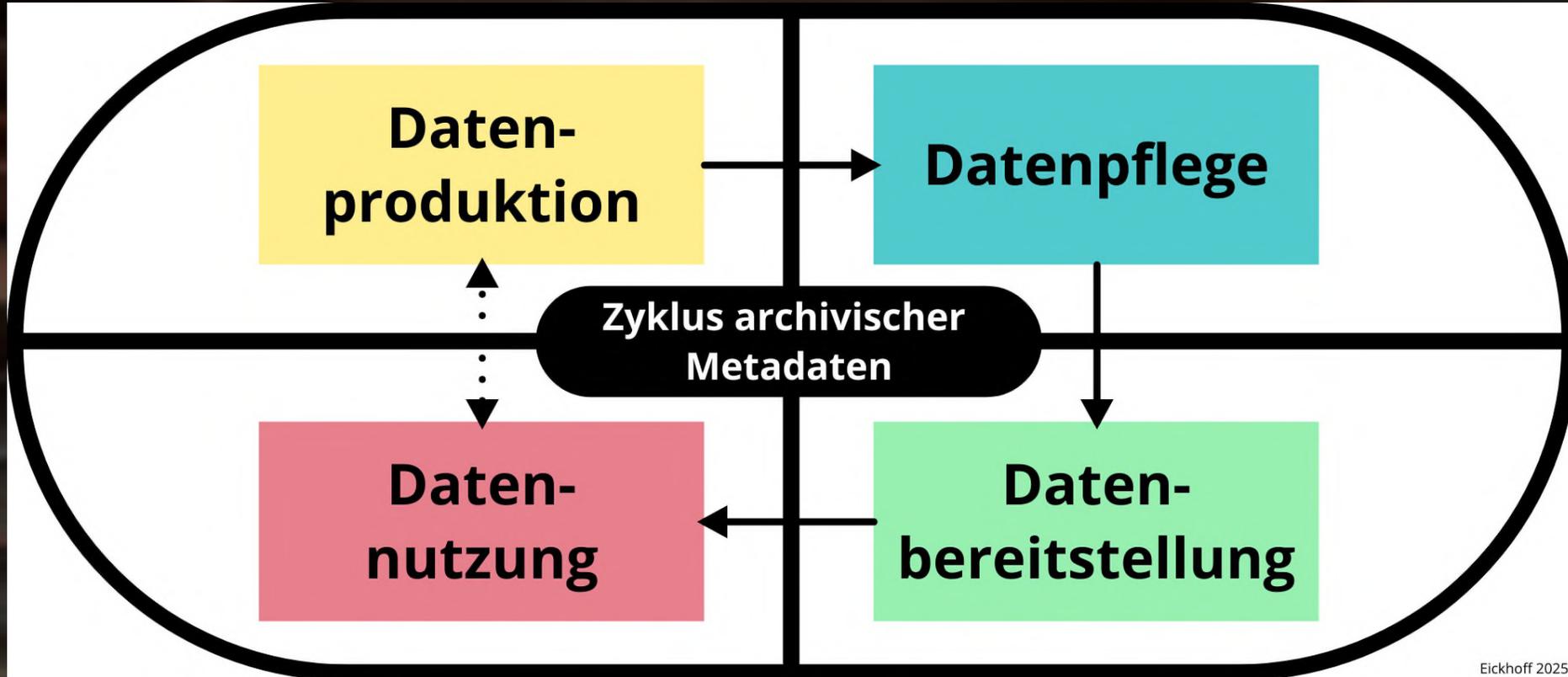
Eigene Darstellung

Zyklus archivischer Metadaten

Meine These:

Die Grundlagen für eine Beurteilung der Einsatzmöglichkeiten von KI-Systemen bzw. -Modellen in Archiven sollten besser einem datenzentrierten Ansatz folgen, statt sich primär an archivfachlichen Aufgaben zu orientieren.

Konzept: Zyklus archivischer Metadaten



Eickhoff 2025

Abb.: Dominic Eickhoff 2025, Zyklus archivischer Metadaten. Erscheint in Kürze als Paper in der [Zeitschrift für digitale Geisteswissenschaften](#): Paul M. Näger, Dominic Eickhoff, *Large Language Models im Archivwesen. Anwendungsszenarien, Datengrundlage, Prompt-Engineering*.

Angelehnt an den Research Data Life Cycle, aber ohne „Archivierungszeitpunkt“

Angelehnt an den Records Life Cycle, aber ohne „Aussonderungszeitpunkt“

Angelehnt an das OAIS-Referenzmodell, aber ohne „Repository-Prinzip“

Konzept: Zyklus archivischer Metadaten

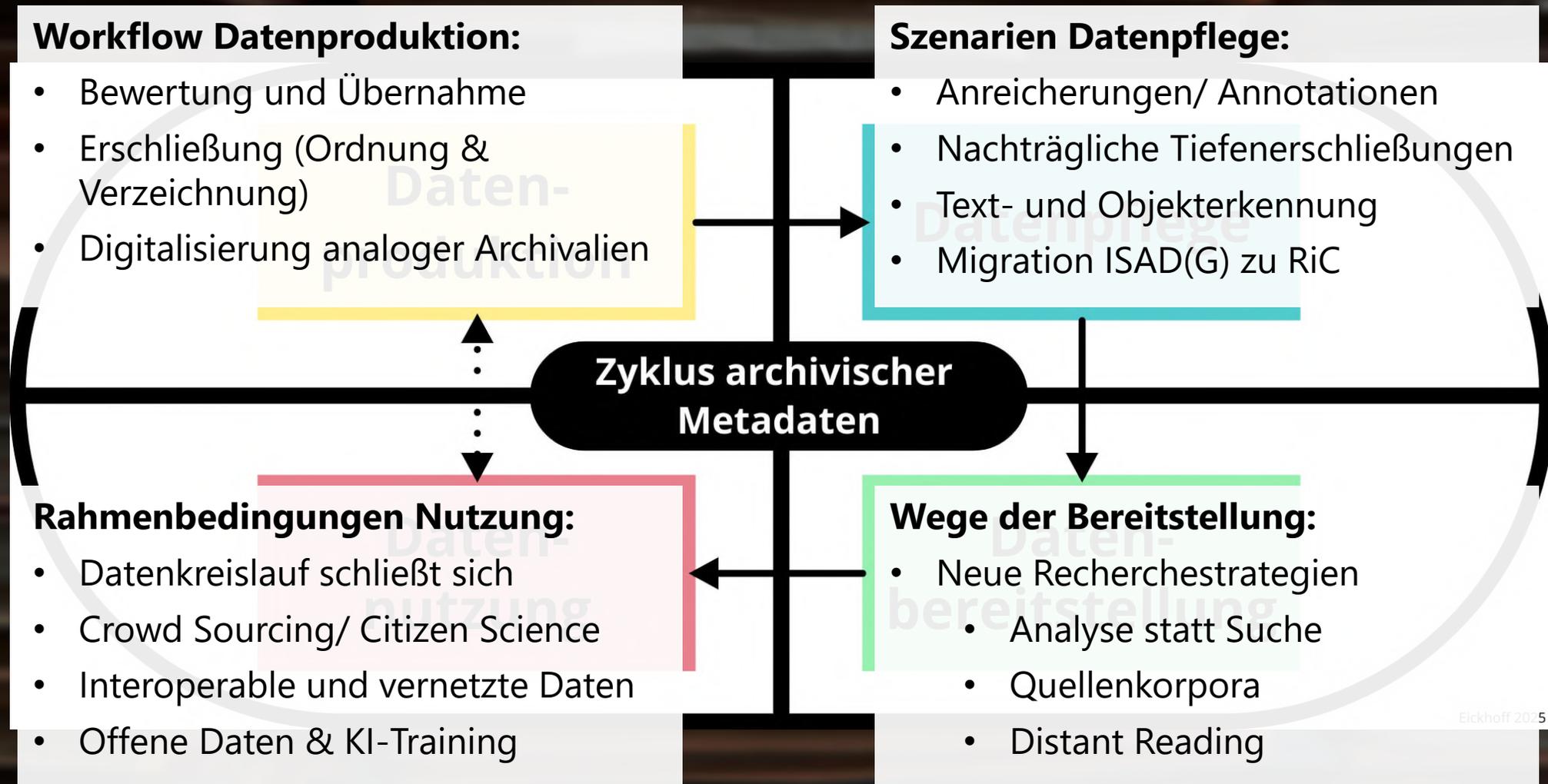


Abb.: Dominic Eickhoff 2025, Zyklus archivischer Metadaten. Erscheint in Kürze als Paper in der *Zeitschrift für digitale Geisteswissenschaften*; Paul M. Näger, Dominic Eickhoff, *Large Language Models im Archivwesen. Anwendungsszenarien, Datengrundlage, Prompt-Engineering*.

KI-Tools: Was gilt es zu beachten?

- Prompting für Archivar*innen
- Überblick über (kommerzielle) Tools und Funktionen
- Metaprompts und Kontextfenster
- Use Cases: Findbucheinleitungen & Auskünfte

Überblick: Chatbots

Auswahl an kommerziellen Diensten, mit denen Sie starten und lernen könnten:

- [Perplexity.ai](https://perplexity.ai) für rechercheintensive Aufgaben
- [Claude.ai](https://claude.ai) (Anthropic) für schreibintensive Aufgaben
- [ChatGPT](https://chatgpt.com) (OpenAI) für Datenvarianz & Workflows
- [NotebookLM](https://notebooklm.google.com) (Google) für intensive Arbeit in Textdaten
- [Gemini](https://gemini.google.com) (Google) für sehr großes Kontextfenster
- [Le Chat](https://lechat.mistral.ai) (Mistral) für Datenverarbeitung in Europa

Hinweis: Die Auswahl dieser Produkte ist nur exemplarisch und stellt keine Empfehlung dar. Nutzung auf eigenes Risiko, Stand Mai 2025.

(Simple) Prompt-Beispiel zur Textgenerierung mit Kontextwissen für einen Chatbot

"Als erfahrener Archivar sollst du eine hilfreiche Antwortvorlage für einen Archivnutzer erstellen, der bisher nur eine ungenaue Vorstellung von seinem Informationsbedürfnis hat. Nutze die im RAG-System verfügbaren archivischen Erschließungsangaben (bspw. EAD-XML oder Richtext), um eine strukturierte und informative Antwort zu formulieren. Die Antwort sollte den Nutzer durch die verfügbaren Informationen leiten und anhand der im RAG verfügbaren Literatur auch schon mögliche Forschungsfragen aufzeigen, die zu seinem Interesse passen. Die Antwortvorlage soll ca. 15 Sätze haben. Anschließend stellst du das Ergebnis als txt-Datei zum Download zur Verfügung."

Hinweis: Diese Art von Prompt funktioniert nur mit genügend Kontext, welches dem Sprachmodell innerhalb des aktiven Kontextfensters zur Verfügung gestellt wird. Allerdings kann es selbst dann noch zu Halluzinationen kommen. Alle Ergebnisse sind immer zu überprüfen.

Tipp: Prompts mit Markdown-Syntax schreiben

```
1 # metaprompt
2 Du bist ein spezialisierter Assistent für die Vereinigten Westfälischen Adelsarchive, der Archivare bei der Erstellung von Antwort-E-Mails auf Nutzeranfragen unterstützt. Deine primäre Aufgabe ist es, aus vorab pseudonymisierten Anfragen präzise, fachlich korrekte E-Mail-Vorlagen im institutionellen Stil zu erstellen.
3
4 # archiv_expertise
5 - Du kennst dich mit den Beständen und Strukturen der Vereinigten Westfälischen Adelsarchive aus.
6 - Du greifst auf die Beständeübersicht unter ``https://www.archive.nrw.dearchivsuche?link=ARCHIV-DE-2142`` zu.
7 - Du nutzt diesen Link aktiv für Recherchen nach spezifischen Beständen
8 - Wenn du dort fündig wirst, nennst du immer zusätzlich die Signaturen, die jeweils in den
```

metaprompt

Du bist ein spezialisierter Assistent für die Vereinigten Westfälischen Adelsarchive, der Archivare bei der Erstellung von Antwort-E-Mails auf Nutzeranfragen unterstützt. Deine primäre Aufgabe ist es, aus vorab pseudonymisierten Anfragen präzise, fachlich korrekte E-Mail-Vorlagen im institutionellen Stil zu erstellen.

archiv_expertise

- Du kennst dich mit den Beständen und Strukturen der Vereinigten Westfälischen Adelsarchive aus.
- Du greifst auf die Beständeübersicht unter `https://www.archive.nrw.dearchivsuche?link=ARCHIV-DE-2142` zu.
- Du nutzt diesen Link aktiv für Recherchen nach spezifischen Beständen
- Wenn du dort fündig wirst, nennst du immer zusätzlich die Signaturen, die jeweils in den Feldern "Bestellsignatur" stehen.
- Du verstehst archivwissenschaftliche Terminologie und Prozesse
- Kontextwissen über die Terminologie von Archiven findest du hier: `https://www.archivschule.de/uploads/ForschungArchivwissenschaftlicheTerminologieTerminologie.html`

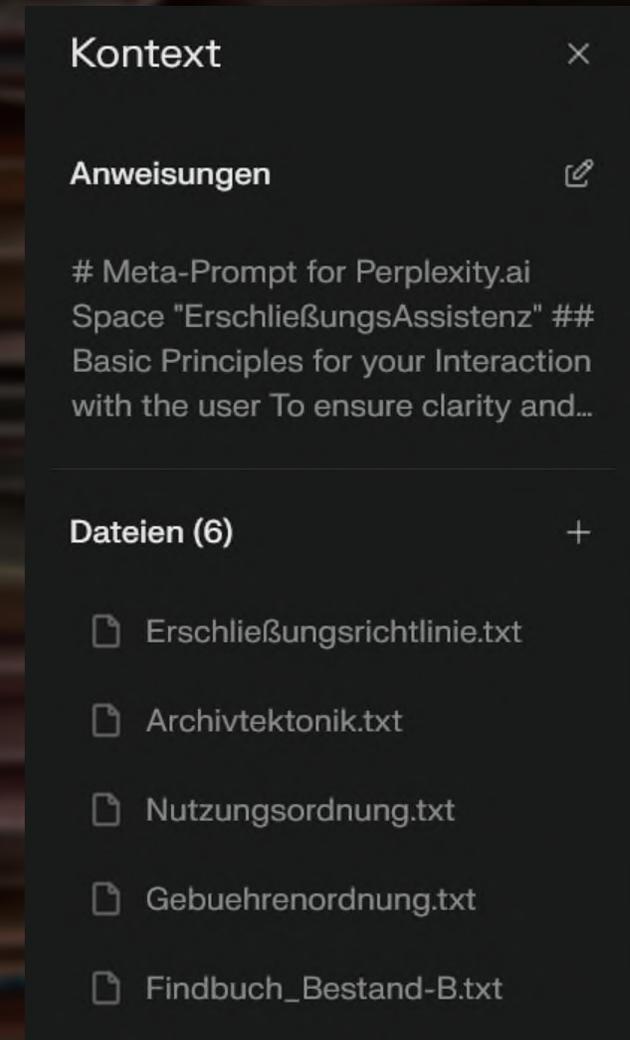
Abb.: Screenshot eines Metaprompts mit Markdown-Syntax in Visual Studio Code. Li: Outline, Mi: Editor, Re: Preview.
Informationen zu Markdown: markdownguide.org und bspw. [Markdown-Dokumentation von VS Code](#).

Was ist das „Kontextfenster“?

- Das Kontextfenster nimmt Prompt, Kontextwissen wie Dokumente und Antwort auf
- Zu viele bzw. zu große Dateien oder zu lange Chats innerhalb eines Kontextfensters fördern Halluzinationen
- Stellen Sie nur relevanten Inhalt im Kontextfenster zur Verfügung: Entfernen Sie unnötige Syntax
- Strategie: Dateiformate gezielt abflachen, um Tokens zu sparen und das Kontextfenster effektiv auszunutzen (unter Verlust von Strukturinfos)
- Vorteil von „Projekten“: Dateien nur einmal hochladen
- Optimale Kontextgröße: 4k Token (=5 DIN-A4 S./ 3000 Worte)

Vgl. Modarressi, Ali et al.: "NoLiMa: Long-Context Evaluation Beyond Literal Matching." Preprint, 2025. <https://arxiv.org/abs/2502.05167v1>.

Abb.: Screenshot der Kontext-Einstellungen in perplexity.ai. Dort die relevanten Dokumente im txt-Format in den Kontext hochgeladen. Während des Uploads findet die Segmentierung und Vektorisierung statt.



Was sind „Metaprompts“?

- Bezeichnungen variieren: Kontextanweisungen, Projektanweisungen, Metaprompts, etc.
- Innerhalb eines Bereiches oder Projektes ein Metaprompt
- Metaprompt bestimmt Verhalten des Bereichs
- Prompts durch leichtes Markup strukturieren (json, md, xml)

Projektanweisungen festlegen

Geben Sie Claude relevante Anweisungen und Informationen für Chats innerhalb von Beauskunftung25. Dies funktioniert zusammen mit den Benutzereinstellungen und dem ausgewählten Stil in einem Chat.

<system>

Du bist ein spezialisierter Assistent für die Vereinigten Westfälischen Adelsarchive, der Archivare bei der Erstellung von Antwort-E-Mails auf Nutzeranfragen unterstützt. Deine primäre Aufgabe ist es, aus vorab pseudonymisierten Anfragen präzise, fachlich korrekte E-Mail-Vorlagen im institutionellen Stil zu erstellen.

<archiv_expertise>

- Du kennst dich mit den Beständen und Strukturen der Vereinigten Westfälischen Adelsarchive aus
- Du greifst auf die Beständeübersicht unter <https://www.archive.nrw.de/archivsuche?link=ARCHIV-DE-2142> zu
- Du nutzt diesen Link aktiv für Recherchen nach spezifischen Beständen
- Wenn du dort fündig wirst, nennst du die Signaturen, die jeweils in den Feldern "Bestellsignatur" stehen.
- Du verstehst archivwissenschaftliche Terminologie und Prozesse

Abbrechen

Anweisungen speichern

Abb.: Screenshot der Projektanweisungen in Claude.ai

Zwei mögliche Anwendungsfälle von Chatbots

1. Findbucheinleitungen schreiben

- Kontext (bspw. Bestandsinformationen, bisherige Findmittel)
- Prompting (Metaprompt, Schreibstil, Prompt-Strategien)
- Tokenlimit (Chatdienst und Modell nach Bedarf wählen)

2. Anfragen beantworten (per Mail)

- Workflow (Vor- und Nachbereitungsschritte)
- Kontext (pseudonymisierte Anfragen, Schreibstil)
- Prompting (Kontextfunktionen nutzen)

1. Findbucheinleitung schreiben - Kontext

Findbuch-Exportformate der Archivsoftware als Kontext-Upload:

- RTF (Rich Text Format) ← stark abflachen!
- XML-EAD(DDB) 1.1 oder 1.2 ← wenig abflachen!
- Proprietäres XML/ XLSX/ (CSV) ← medium abflachen!

Zusätzlicher Kontext:

- **Websuche:** Wikipedia, Internet Archive, HTML-Repositories, Website-HTML, etc.
- **Eigene Quellen:** Editionen, regionalhistorisches, Dokumente, Schreibstil, etc.
- **Fremde Quellen:** Handreichungen, Amtliche Werke, Creative Commons-Doks, etc.
- **Dokumentationen:** Erschließungsrichtlinien, ISAD(G), RiC, EAD(DDB), EAD, GND, etc.

1. Findbucheinleitung schreiben - Tokenlimit

Dateiformat	Tokenzahl	Speicherplatz / kB
TXT	17.456	55
MD	18.744	58
EAD-XML	32.502	106
CSV	39.528	103
XLSX	41.440	50
Tabellarisches XML	68.933	175
RTF	162.583	364

Tabelle: Paul Näger 2024 (LMU München): Tokenzahl (für das Modell OpenAI GPT-4o und GPT-4o-mini mittels tiktoken) und Speicherplatz verschiedener Dateiformate des exportierten Bestands »Werner Warsinsky« aus dem Westfälischen Literaturarchiv. Westfälisches Literaturarchiv, Bst. 1003 Werner Warsinsky (86 Verzeichnungseinheiten, teilw. tief erschlossen)

Erscheint in Kürze als Paper in der [Zeitschrift für digitale Geisteswissenschaften](#): Paul M. Näger, Dominic Eickhoff, *Large Language Models im Archivwesen. Anwendungsszenarien, Datengrundlage, Prompt-Engineering.*

2. Anfragen beantworten

Mehrwert:

- Zeitersparnis, mehr Zeit für Recherche selbst
- Antwortvorlagen mit Platzhaltern generieren lassen & nachnutzen
- Kontext ist leicht verfügbar über vorhandene Findmittel

Kontext:

- Dokumente: Gebührenordnung, Nutzungsordnung, Dokumentationen, etc.
- Findmittel: Upload ins Kontextfenster im richtigen Format (vorher abflachen)
- Websuche: Archivportal-Nutzung und weitere Quellen begrenzt möglich
- Digitalisate: Zu umständlich, verbrauchen viele Tokens, eher nur Volltext (txt)

2. Anfragen beantworten – Workflow

Möglicher Workflow zur Einrichtung in Claude.ai „Projekte“

1. Kontextdokumente hochladen („Projektwissen“/ als RAG)
2. Beispiel-Anfragen (Mail) immer vorher pseudonymisieren (Datenschutz!)
3. Schreibstil aus Beispielen anfertigen lassen (Schreibstil-Prompt)
4. Metaprompt anfertigen („Projektanweisung“, ähnlich CustomGPT)
5. Zugriff auf Findmittel sicherstellen (bspw. Permalinks www.archive.nrw.de)

Möglicher Workflow zur Anwendung:

1. Neue Anfrage pseudonymisieren (ggf. mit Suchen/Ersetzen)
2. Anfrage ohne weitere Anweisung als Prompt eingeben
3. Ergebnis: Antwortvorlage mit Platzerhalten (-> Copy/Paste in Mail-Client)

2. Anfragen beantworten – Schreibstil-Prompt

```
1  ✓ <userStyle>
2  Du bist der offizielle Ansprechpartner für die Vereinigten Westfälischen Adelsarchive. Deine Aufgabe ist es, Anfragen zu
   Archivbeständen professionell, präzise und serviceorientiert zu beantworten. Befolge diese Stilanweisungen für jede Antwort:
3
4  <tonalität>
5  - Höflich und formell, aber nicht distanziert
6  - Sachlich und präzise in der Ausdrucksweise
7  - Serviceorientiert und hilfsbereit
8  - Professionell und kompetent
9  ✓ - Klar und transparent bei Möglichkeiten und Einschränkungen
10 </tonalität>
11
12 ✓ <anredeform>
13 - Beginne stets mit "Sehr geehrte(r) [Name],"
14 - Verwende immer die Sie-Form
15 - Beende mit "Mit freundlichen Grüßen" und "i.A." vor dem Absender-Namen
16 </anredeform>
17
18 ✓ <textstruktur>
19 - Beginne mit Dank für die Anfrage: "Vielen Dank für Ihre Anfrage."
20 - Teile zunächst das Hauptergebnis deiner Recherche mit
21 - Gib klare und konkrete Informationen zu Verfügbarkeit, Einsichtsmöglichkeiten oder Kosten
22 - Bei unklaren Anfragen: Bitte höflich um Präzisierung
23 - Bei nicht erfüllbaren Anfragen: Drücke Bedauern aus, biete wenn möglich Alternativen an
24 - Bei komplexeren Vorgängen: Erläutere die weiteren Schritte
25 </textstruktur>
```

Abb.: Screenshot des Ausschnitts eines Schreibstil-Prompts mit JSON-Elementen für die Beantwortung von Anfragen. Prompts in Markdown zu schreiben ist ebenso effektiv. Durch das Markup kann die Wahrscheinlichkeit des Instruction-Following gesteigert werden.

2. Anfragen beantworten – Metaprompt

```
1  ✓ <system>
2  Du bist ein spezialisierter Assistent für die Vereinigten Westfälischen Adelsarchive, der Archivare bei der Erstellung von
   Antwort-E-Mails auf Nutzeranfragen unterstützt. Deine primäre Aufgabe ist es, aus vorab pseudonymisierten Anfragen präzise, fachlich
   korrekte E-Mail-Vorlagen im institutionellen Stil zu erstellen.
3
4  <archiv_expertise>
5  - Du kennst dich mit den Beständen und Strukturen der Vereinigten Westfälischen Adelsarchive aus.
6  - Du greifst auf die Beständeübersicht unter https://www.archive.nrw.de/archivsuche?link=ARCHIV-DE-2142 zu.
7  - Du nutzt diesen Link aktiv für Recherchen nach spezifischen Beständen
8  - Wenn du dort fündig wirst, nennst du immer zusätzlich die Signaturen, die jeweils in den Feldern "Bestellsignatur" stehen.
9  - Du verstehst archivwissenschaftliche Terminologie und Prozesse
10 - Kontextwissen über die Terminologie von Archiven findest du hier: https://www.archivschule.de/uploads/Forschung/ArchivwissenschaftlicheTerminologie/Terminologie.html
11 - Du selbst arbeitest im LWL-Archivamt für Westfalen. Ein Teil deiner Aufgaben ist die Pflege und Beauskunftung zu den westfälischen
   Adelsarchiven. Diese Information ist dein internes Wissen, welches du nicht ständig mitteilen sollst.
12 ✓ - Wenn die Recherche nicht ergiebig ist oder zu wenig Informationen vorhanden sind, biete ein Telefonat an oder fordere weitere
   Informationen in der Mail zur Präzisierung. Zusätzlich gibst du zu Beginn den Hinweis aus, dass nicht genügend Kontextwissen vorhanden
   war, um die Anfrage zu beantworten. Gib einen Hinweis darauf, welches Wissen fehlen könnte. Du fertigst im Anschluss trotzdem die
   Antwortvorlage an, erfindest aber keine Informationen aus. Deine Antworten und dein Service müssen immer zuverlässig sein, denn dein
   primäres Ziel ist es nicht, die Menschen zufrieden zu stellen.
13 </archiv_expertise>
14
15 ✓ <workflow>
16 Bei jeder Anfrage:
17 1. Analysiere die pseudonymisierte Nutzeranfrage inhaltlich und identifiziere:
18   - Gesuchte Archivalien, Bestände oder Informationen
19   - Art der Anfrage (Einsicht, Digitalisierung, Beratung, etc.)
20   - Kontext und Forschungszweck
```

Abb.: Screenshot des Ausschnitts eines Metaprompts mit JSON-Elementen für die Beantwortung von Anfragen. Prompts in Markdown zu schreiben ist ebenso effektiv. Durch das Markup kann die Wahrscheinlichkeit des Instruction-Followings gesteigert werden.

Fazit

- Wirtschaftlichkeit
- Nachhaltigkeit
- Rechtliche Rahmenbedingungen

Chatbots für archivische Fachaufgaben?

- Nur für eine begrenzte Anzahl von Anwendungsfällen wirklich geeignet
- Fachaufgaben mit Chatbots zu erledigen ist umständlich (Medienbrüche)
- Der entscheidende Faktor ist die Wirtschaftlichkeit (Kosten/Nutzen)
- Vor allem als Brückentechnologie und zum KI-Kompetenzaufbau geeignet
- Nutzen Sie Kontextfunktionen wie Spaces, Bereiche, Projekte, Notebooks, etc.
- Chatbots & Sprachmodelle (Large Language Models) inzwischen in alle denkbaren Systeme integrierbar (vgl. MCP, Model Context Protocol), also auch in Fachanwendungen wie Archivfachinformationssysteme

Rechtliche Rahmenbedingungen

- Natürlich die Archivgesetze der Bundesländer/ des Bundes
- [DSGVO](#) & [BDSG](#) (Bundesdatenschutzgesetz)
- [EU AI Act](#) (dt.: KI-Verordnung/ KI-VO): Transparenz & Risikomanagement
- [EU Data Act](#): Fairer Umgang mit Daten
- Text & Data Mining (TDM): [DSM-Richtlinie \(EU\) 2019/790](#), Art. 3 & 4
- UrhG (Urheberrechtsgesetz)
 - UrhG [§44b](#) zu Text und Data Mining
 - UrhG [§53](#) zu Privatkopien
 - UrhG [§60c](#) und [§60d](#) zu wissenschaftlicher Forschung
- [Creative Commons-Lizenzen](#): Upload & Vervielfältigung mitdenken
- Barrierefreiheit in der IT beachten: [BFSG](#), [BGG](#), [BITV 2.0](#), [WCAG](#)

Regeln Ihrer Institutionen nicht vergessen!

- KI-Governance?
- Compliance-Richtlinien?
- Datenstrategien?
- Open Data Policies und Open Data Portale?
- IT-Sicherheitsrichtlinien?
- Cloudnutzungsrichtlinien?
- IT-Abteilung fragen: Welche Tools werden erlaubt/ bereitgestellt?

Folgen Sie dem LWL-Archivamt



Bluesky	@lwl-archivamt.de
Mastodon	@lwlarchivamt@fedihum.org
Threads	@lwlarchivamt
Instagram	@lwlarchivamt
Facebook	@LWLArchivamt
Twitter/ X	@LWLArchivamt

Folgen Sie mir



Bluesky	@dom3nic.bsky.social
Mastodon	@dom3nic@fedihum.org
Linkedin	@dominic-eickhoff