



KI-basierte Inhaltsanalyse in R

Austausch zur Nutzung und Entwicklung KI-gestützter Webanwendungen



KIPERWEB



Forschungsinstitut
Betriebliche Bildung

- **Update**
 - News & Leaderboard-Update
- **Input**
 - "KI-basierte Inhaltsanalyse in R"
- **Diskussion**

Leaderboard-Update (18.09.2025)



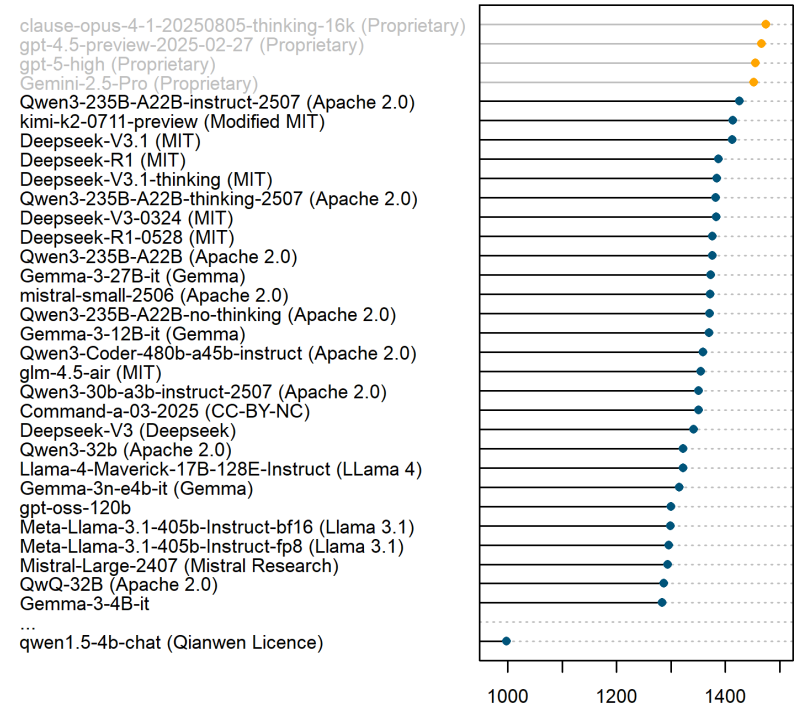
Rechts das aktuelle LMArena-Leaderboard für die Kategorie German mit Blick auf die besten open-weights-Modelle (sowie ausgewählte proprietäre Modelle).

Bemerkenswert:

- **GPT-5-high** wurde als Spitzenreiter abgelöst von **claude-opus-4-1**
- **Deepseek-v3.1** neu eingestiegen, deutlich vor **Deepseek-R1**
- **Mistral-small-2506** (24B) hält seine Position vor dem hervorragenden **Gemma-3-12B-it**, liegt nun aber hinter **Gemma-3-27b-it**.
- Fliegengewicht **Gemma-3n-e4b-it** liegt vor OpenAI's **gpt-oss** aber wieder hinter **Qwen3-32b**
- Ausgewiesenes Schlusslicht ist **qwen1.5-4b-chat** (993)

Arena Score German

based on Imarena.ai on Sep 18, 2025



Fokusthema: KI-basierte Inhaltsanalyse in R



- Prompt „high-tech artificial brain extracts information from a book“

(rechts visualisiert von FLUX.1-schnell, seed 1558494386)

Arbeitsdefinition nach Fischer (2025):

KI-basierte → gestützt auf moderne KI (insb. LLMs)

Qualitative → interpretativ, nicht nur quantifizierend

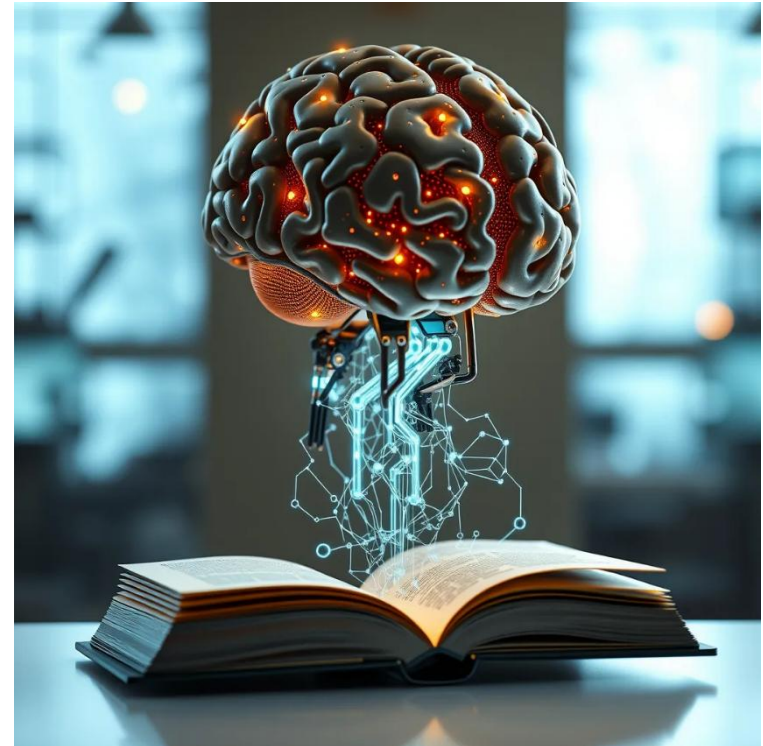
Inhalts- → Kommunikationsinhalte (insb. Texte)

Analyse → hinsichtlich ihrer Bestandteile untersuchen

Source:

Fischer, A. (2025). Die Zukunft der Qualitativen Inhaltsanalyse im Zeitalter Künstlicher Intelligenz. Ein Ablaufmodell KI-basierter Inhaltsanalyse. f-bb-online 03/25.

Fischer, A. (in press). The Future of Qualitative Content Analysis in the Age of Artificial Intelligence. In: J. Dörpinghaus, M. Tiemann (Eds.), Advances in Computational Social Sciences (tba).



Quelle: <https://huggingface.co/spaces/black-forest-labs/FLUX.1-schnell>

Wie kann man sich KI-basierte Inhaltsanalyse vorstellen?



- Allgemeine KI-Systeme (z.B. Huggingface/Chat-ui), MAXQDA's KI ASSIST, KI on premises und/oder CLIs

The screenshot displays three overlapping windows illustrating AI-based content analysis tools:

- HuggingChat** (left): A web interface for interacting with AI models. It shows a sidebar with navigation options like 'Today', 'Qualifikationsvor...', 'AFischer1985', 'Theme', 'Models', 'Assistants', 'Tools', 'Settings', and 'About & Privacy'. The main chat area shows a conversation with 'Interview Thanh'.
- AI Assist** (middle): A window titled 'AI Assist' with a subtitle 'Lassen Sie sich Subcode-Vorlife balance erstellen. AI A Code-Memo ein.' It features a section 'Sprache der Vorschläge' with checkboxes for 'Liste mit Beispielen für', 'Nur aktivierte Dokumente', and 'Hiermit bestätige ich, dass zur Kenntnis genommen'. Below this, there is a text input field and a button.
- LM Studio** (right): A desktop application window titled 'LM Studio'. It shows a chat interface with a 'USER' prompt 'Hallo' and an 'AI' response. Below the chat, there is a code editor displaying a Python script for processing text data. The script includes comments in German and uses libraries like 'uTexts', 'requests', and 'os'.

Below the windows, there are two links:

- <https://www.youtube.com/watch?v=VFft6B2cNj8>
- <https://lmstudio.ai/>

- KI-Kenntnisse/-Kompetenz (i.S.v. Art. 3, EU KI-VO) ebenso so wichtig wie die Wahl geeigneter Tools
- KI-generierte Analysen können nach Fischer (2025)...
 - ... falsch sein (insb. bei suboptimalem Prompting)
 - ... unfair sein (insb. auch bei impliziten Hinweisen wie Dialekten)
 - ... nicht-deterministisch sein (insb. wenn temperature $\neq 0$)
 - ... von jedem Teil des Inputs abhängen (insb. auch bei längeren Dialogen)
 - ... in Zukunft möglicherweise nicht mehr verfügbar sein (insb. proprietäre KI-Systeme)
 - ... ressourcenintensiv generiert werden (insb. bei überdimensionierten KI-Systemen & Prompts)

Vierfelderschema nach Fischer (2025)

Kategorien KI-basierter Inhaltsanalyse auf Basis von Theoriebezug (vertikal) und Verdichtungsintention (horizontal) nach Fischer (2025) – mit *vereinfachten Beispielen* (in blauen Sprechblasen)

Gegeben sei folgender Text:
„[Text]“
Paraphasiere alle angesprochenen Erkenntnisse zur Frage „[Forschungsfrage]“

Gegeben sei folgender Text:
„[Text]“
Liste alle angesprochenen Punkte auf, und gliedere deine Liste nach Kategorien mit Blick auf die Frage „[Forschungsfrage]“

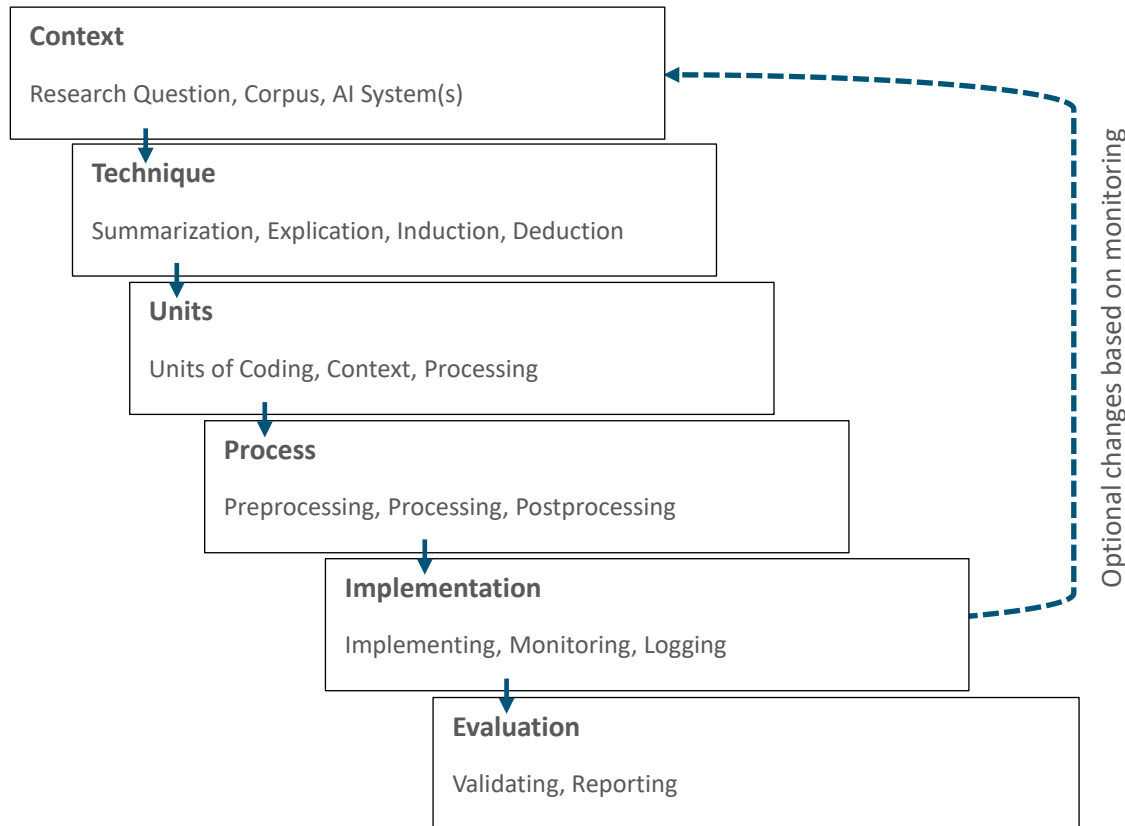
| | |
|--|---|
| KI-basierte Zusammenfassungen: KI wird eingesetzt, um eine Zusammenfassung auf der Grundlage eines Korpus spezifischer Materialien zu erstellen, unabhängig von theoretischen Kategorien. | KI-basierte Explikation: KI wird eingesetzt, um eine Explikation von Materialien innerhalb eines Korpus auf der Grundlage zusätzlicher Kontexte/Materialien/Assoziationen zu erstellen, unabhängig von theoretischen Kategorien. |
| KI-basierte Induktion: KI wird eingesetzt, um auf der Grundlage eines Korpus spezifischer Materialien Schlussfolgerungen zu theoretischen Kategorien und Konzepten zu ziehen. | KI-basierte Deduktion: KI wird eingesetzt, um auf der Grundlage einer Reihe theoretischer Kategorien und Konzepte Schlussfolgerungen zu bestimmten Materialien in einem Korpus zu ziehen. |

Gegeben sei folgender Text:
„[Text]“
Erläutere, inwiefern der Text Hinweise gibt auf die Frage „[Forschungsfrage]“

Gegeben sei folgender Text:
„[Text]“
Ordne diesen Text einer der folgenden Kategorien zu:
„[Kategoriensystem]“
Gib die ID der gewählten Kategorie an („ID“) und begründe deine Antwort („Begründung“).

Anknüpfend insb. an Mayring (2022) unter Berücksichtigung von Literatur wie Döring und Bortz (2016)

Process Model of AI-based Content Analysis (Fischer, 2025)



Beispiel für Anfragen an KI on-premises in R



- Die Funktion `requests` rechts realisiert Anfragen per `httr`-package
- Benötigt wird ein LLM-Server (hier LMStudio im lokalen Netzwerk unter 192.168.2.57)

```
request=function(  
  prompt,  
  system="Du bist ein renommierter Sozialwissenschaftler.",  
  model="gemma-3-12b-it",max=2048, backend="LMStudio"){  
  url="192.168.2.57:1234/v1/chat/completions"  
  body=list(  
    "messages"=list(  
      list("role"="system", "content"=system),  
      list("role"="user", "content"=prompt)),  
    model = model,  
    max_tokens = max,  
    temperature = 0)  
  if(!is.na(stop)) body$stop=c(stop)  
  post=httr::POST(url=url, body=body,  
    httr::add_headers("accept" = "application/json", "Content-Type" = "application/json"),  
    encode="json")  
  post=httr::content(post)  
  if(length(post$choices)>0)post=post$choices[[1]]$message$content  
  post  
}
```

Beispiel für einfache KI-basierte Inhaltsanalyse in R



- *template* ist ein Prompt-Template für die Inhaltsanalyse (anpassen!)
- *texts* enthält Beispieltexpte (z.B. Passagen eines Interviews)
- *sample* realisiert eine Zufallauswahl
- *filename* spezifiziert eine lokale log-Datei (falls was schiefgeht)
- Die *for-Schleife* geht nun alle samples von Texten durch, schickt den Prompt an die KI, und schreibt die Antwort in eine Liste (und in die Log-Datei)

```
template="Gegeben sei folgender Text: \"{{Text}}\".\n\nFormuliere das zentrale Thema des Texts knapp als ein bis maximal drei Worte in deutscher Sprache."
texts=c("Guten Tag, schön dass Sie da sind",
        "Im Folgenden geht es um die Zukunft der Inhaltsanalyse")
sample=sample(length(texts),length(texts))
filename="Logfile.txt"
results=list()
write("",filename,append=F)
range=1:length(texts)
t1a=Sys.time()
for (i in range){
  text=unlist(texts[sample[i]])
  prompt=gsub("\"\\{{text}}\"",text,template)
  response=request(prompt,max=1000)
  response=gsub("[ ]*\n[ ]*", "\n",gsub("(^[ ]+)|[ ]+$)", "",response))
  results[prompt]=response
  output=paste0("\n\n", (i+0), ":\n\n", response)
  write(output,filename,append=T)
  print(paste(i, "/", max(range), ":\n", output))
}
t2a=Sys.time()
print(t2a-t1a)
print((t2a-t1a)/length(range))
```

- Fragen?
- Anregungen?
- Erfahrungen?