

Mit Künstlicher Intelligenz lebensgeschichtliche Interviews interpretieren

Ein Werkstattbericht

Dennis Möbus

1. Einleitung

In der qualitativen Sozialforschung ist das Interpretieren in der Gruppe Teil des Forschungsalltags. Ab den 1960er Jahren haben sich zunächst in den USA, dann auch in Deutschland verschiedene Traditionen der gemeinsamen Analyse von Forschungsdaten etabliert (Reichert 2013: 32). Auch in der Oral History gibt es immer wieder Versuche, das Auswerten in der Gruppe populärer zu machen (Arp 2024), doch ist es neben der vorherrschenden Praxis, Quellen individuell auszuwerten, eine Herausforderung, Termine und Räumlichkeiten zu koordinieren und eine konsistente Gruppe zu formieren (vgl. Reichert 2013: 31). Das hängt auch mit strukturellen Unterschieden zwischen sozialwissenschaftlich und geschichtswissenschaftlich geprägten Fächern zusammen: Während in ersteren wesentlich häufiger in großen Projekten gemeinsam Daten erhoben werden, sind Oral-History-Studien oftmals kleiner angelegt und werden nur selten von einer Gruppe Interviewer:innen durchgeführt. Daraus ergibt sich für die sozialwissenschaftlichen Forschungsprojekte viel eher eine natürliche Gruppenstruktur. Die Vorteile einer gemeinsamen Interpretation liegen jedoch auch für Historiker:innen auf der Hand: Die Gruppe trägt zur Validierung einer Interpretation bei und öffnet dabei noch weitere Perspektiven, die dem oder der Einzelforscher:in möglicherweise nicht auffallen (Reichert 2013: 13 ff.). Doch wie können sich Historiker:innen entsprechende Hilfestellung suchen, wenn keine Interpretationsgruppen verfügbar sind? Erste Versuche, mit generativer Künstlicher Intelligenz Interpretationsgehilfen zu bauen, zeigen, dass dies ein interessanter Weg sein könnte, eigene Interpretation barrierearm zu prüfen, über eigene Argumente zu „diskutieren“ und weitere Inspiration oder übersehene Interpretationsansätze zu bekommen (Krähnke et al. 2025). Dieser Werkstattbericht vermittelt einen Eindruck über Möglichkeiten, State-of-the-Art sowie Stärken und Schwächen, die mit dem KI-gestützten Interpretieren für Oral Historians verbunden sind. Im Mittelpunkt werden das Design einer DSGVO-konformen KI-Infrastruktur, der Aufbau und die serielle Verknüpfung von Prompts sowie erste Einblicke in die entstandenen Interpretationsangebote stehen.

2. Versuchsaufbau

In dem Versuch ging es darum, eine ausgewählte, vorab anonymisierte Interviewpassage, bereitgestellt in Form des Transkripts, mittels KI zu interpretieren und aus der

Perspektive verschiedener Personen zu diskutieren. Die „Diskussion“ soll der Berücksichtigung unterschiedlicher Forschungsperspektiven und Schulen dienen und durch Widersprüche in Interpretationsansätzen aufzeigen. Überprüft werden sollte die Qualität der Interpretation und deren Variation in Abhängigkeit von der vorgegebenen Forschungsrichtung. Im Sinne einer ersten Annäherung wurden die Ergebnisse rein qualitativ überprüft. Eine performative Podiumsdiskussion im Rahmen einer Konferenz für digitale Geistes- und Kulturwissenschaften, bei der die KI-Diskussion von realen Personen vorgelesen und anschließend von „realen“ Forscher:innen diskutiert wurde, diente der weiteren interdisziplinären Einordnung.

Zentrale Bestandteile des Versuchsaufbaus sind zunächst Small- (SLM) und Large Language Models (LLM), die aus Gründen des Datenschutzes lokal betrieben werden können, des Weiteren eine Oberfläche zur Bedienung der Sprachmodelle und schließlich eine Entwicklungsumgebung, um mittels einer Programmiersprache komplexere KI-Architekturen bauen zu können. Für einen Erstversuch ist ChatGPT¹ bzw. die für professionelle Anwendungen bereitgestellte OpenAI-Plattform² mit *GPT 5.1* für die KI-Diskussion einer anonymisierten Interviewpassage von zehn Minuten Länge genutzt worden. In einem nächsten Schritt wurde mit derselben anonymisierten Interviewpassage über ein browserbasiertes Programmernotebook und einer Schnittstelle zur KI-Plattform *together.ai*³ das offen verfügbare LLM *Llama 3.3 70B Instruct*⁴ des US-Konzerns Meta getestet. Während der Tests ist eine äußerst komplexe Prompt-Architektur entstanden, zeitgleich nahm die Entwicklung und Optimierung kleiner, für den offlinebetrieb optimierter Sprachmodelle (SLM) zu. In einer letzten Phase konnten kleine, lokal über das ChatGPT nachempfundene Interface *Ollama*⁵ betriebene SLMs mit weniger komplex verschalteten, aber in sich gut ausgearbeiteten Prompts erstmals überzeugende Ergebnisse in einer vollständig abgeschlossenen und sicheren Umgebung liefern.

3. Transkript

Die Interviewpassage stammt aus den Beständen des Archivs „Deutsches Gedächtnis“ im Institut für Geschichte und Biographie der FernUniversität in Hagen.⁶ Das Interview mit Frau A. (im Transkript „IP“) wurde 1981 im Rahmen des Projekts „Lebensgeschichte und Sozialkultur im Ruhrgebiet 1930 bis 1960“ geführt (Niethammer 1983 a und b; Niethammer/Plato 1985). Das Transkript der für die KI-Interpretation und -Diskussion zugrunde gelegten Passage wird im Folgenden vollständig abgedruckt:

I: Können wir anfangen? Also wäre schön, wenn Sie mit Kindheit beginnen würden.

IP: Ich bin gebürtig aus [Ort] im Sauerland und bin am [Tag, Monat] 25 geboren. Und war das erste Enkelkind, einzige Enkelkind lange Zeit und bin sehr

1 <https://chatgpt.com/> (30.11.2025).

2 <https://platform.openai.com/docs/overview> (30.11.2025).

3 <https://www.together.ai/> (30.11.2025).

4 <https://huggingface.co/meta-llama/Llama-3.3-70B-Instruct> (30.11.2025).

5 <https://ollama.com/> (30.11.2025).

6 <https://deutsches-gedaechtnis.fernuni-hagen.de/de/interviews/adg0001> (30.11.2025).