

10 Unterstützung durch künstliche Intelligenz (KI)

In diesem Kapitel erfahren Sie etwas über

- den Einsatz künstlicher Intelligenz bei der qualitativen Inhaltsanalyse,
- die Diskussion um LLMs in der Wissenschaft,
- die Doppelrolle der KI als Werkzeug und Partnerin bei der Datenanalyse,
- praktische Anwendungsmöglichkeiten der KI bei der Datenexploration und Kategorienbildung,
- weitere Anwendungsmöglichkeiten bei der Codierung und Analyse der codierten Daten,
- Probleme des Datenschutzes und ethische Herausforderungen sowie
- die Frage möglicher disruptiver Effekte von generativer KI.

10.1 KI und ihre Rolle in der qualitativen Inhaltsanalyse

Dieses zehnte Kapitel haben wir neu für diese sechste Auflage dieses Lehrbuches geschrieben. Es wurde zu einem Zeitpunkt verfasst, als riesige Fortschritte in der Entwicklung künstlicher Intelligenz mit großem Tempo passierten, es aber gleichzeitig (noch) an verbindlichen, überall gültigen Regelungen des Einsatzes von KI mangelte. Die Rolle, die künstliche Intelligenz (KI) potenziell in der qualitativen Inhaltsanalyse spielen kann, lässt sich nicht unabhängig von der allgemeinen KI-Rezeption im Bereich der Wissenschaft und den dort diskutierten Regelungen der Nutzung bzw. Einschränkung der Nutzung bestimmen. Es sind die gesetzlichen Vorgaben hinsichtlich des Datenschutzes, die ethischen Normen und die Regeln guter wissenschaftlicher Praxis, welche die Rolle determinieren, die KI in der Analyse von Forschungsdaten überhaupt spielen kann bzw. spielen darf. Als das Unternehmen OpenAI im November 2022 den Chatbot ChatGPT vorstellte, fand dies zunächst in der breiten europäischen Öffentlichkeit wenig Beachtung. Durch den kostenfreien Zugang von ChatGPT wurden jedoch generative große Sprachmodelle, sogenannte Large Language Models (LLMs), allgemein zugänglich und damit nahm die Aufmerksamkeit für künstliche Intelligenz sehr schnell exponentiell zu. Die Resonanz auf ChatGPT war nicht nur in den Medien riesig. Fast alle, die den Chatbot selbst ausprobierten, waren von den Leistungen verblüfft. Es stellte sich eine Art Wow-Effekt ein, doch gleichzeitig wurden auch Ängste und dystopische Befürchtungen virulent. Die Hochschulen und Wissenschaftsorganisationen,

wie die DFG (Deutsche Forschungsgemeinschaft) und der DHV (Deutscher Hochschulverband), befassten sich dann sehr schnell mit dem Thema „Künstliche Intelligenz“. Bereits im Heft 1/2023 erschien in der DHV-Zeitschrift „Forschung und Lehre“ ein Artikel unter der Überschrift „Meilenstein der KI-Entwicklung? Der Chatbot ChatGPT“ (Weßels, 2023), in dem die Autorin die Auswirkungen von generativen KI-Modellen für die Zukunft des Lehrens und Lernens an Hochschulen fokussierte. Sie schätzte die Entwicklung der LLMs als bahnbrechend ein:

„Seit dem 30. November 2022 ist meine Welt (...) gefühlt eine andere Welt, die uns in die ‚Neuzeit‘ führt, von der wir noch nicht wissen, ob wir sie lieben oder fürchten sollen.“ (ebd., S. 26)

KI-Modelle bieten, so Weßels, „die Möglichkeit, große Datenmengen schneller und präziser zu analysieren und neue Erkenntnisse und Perspektiven zu gewinnen. Sie können auch dazu beitragen die Qualität und Validität von Forschungsergebnissen zu verbessern“ (ebd., S. 27). Weßels forderte die Hochschulen und die Wissenschaftler und Wissenschaftlerinnen dazu auf, sich bewusst für die verantwortungsvolle und ethische Verwendung von LLMs zu entscheiden und Richtlinien einzuführen. Dies ist allerdings bislang nicht geschehen. Im Bereich von Bildung und Wissenschaft besteht immer noch Unsicherheit darüber, wie mit den neuen Möglichkeiten der generativen KI umgegangen werden kann. Während an der Spitze der Bildungshierarchie, in den Ministerien und Universitätspräsidenten, nach wie vor keine Einigkeit über die Zulässigkeit und die Limits der KI-Nutzung besteht, zeigen empirische Studien, dass zur gleichen Zeit im Bildungssystem von unten eine Generation nachwächst, die ganz selbstverständlich mit künstlicher Intelligenz in Form von LLMs arbeitet. So zeigte schon eine im Juli und August 2023 durchgeführte Studie des Bayerischen Forschungsinstituts für digitale Transformation (bidt) unter mehr als 3 000 Internetnutzenden, dass 73 % der erwachsenen Schülerinnen und Schüler und 78 % der Studierenden generative KI einsetzten. Annähernd die Hälfte der KI-Nutzenden war überzeugt, dadurch bessere Noten erreicht zu haben, und ein ähnlich hoher Prozentsatz der Lehrkräfte war überzeugt, dass KI zu einer Leistungssteigerung führe (Schlude et al., 2024). Die Prozentzahlen der Nutzenden haben sich seither vermutlich stetig erhöht. Fakt ist also, dass LLMs von sehr vielen genutzt werden, es aber nach wie vor keine Richtlinien für die Nutzung gibt. Allgemeine Stellungnahmen pro KI-Nutzung und konkretes Handeln vor Ort passen häufig nicht gut zueinander: Während die DFG in ihrer Stellungnahme betont, „dass der Einsatz von KI-Modellen im wissenschaftlichen Arbeiten nicht ausgeschlossen werden sollte, da er erhebliche Chancen und Potenziale bietet“ (DFG, 2023), lässt die TU München zur gleichen Zeit einen Studienbewerber nicht zum Studium zu, weil sie vermutet, dass der Bewerber

bei der Abfassung seines Bewerbungssessays ChatGPT benutzt hat. Dass es im Wissenschaftsbereich bislang nicht gelungen ist, klare operationale Regeln zu formulieren, lässt sich gut anhand der in der Zeitschrift „Forschung und Lehre“ erschienen Artikel nachzeichnen. Dort finden sich sowohl Beiträge, die KI ablehnen und ihre Leistung geringschätzen (bspw. Cap, 2023), als auch Beiträge, die für eine Integration in Forschung und Lehre plädieren (Hoeren, 2023) und davon überzeugt sind, dass sich durch KI die Qualität und Validität von Forschungsergebnissen verbessern lässt (Weßels, 2023).

Diese unsichere Situation macht es uns nicht leicht, die Funktion zu bestimmen, die generative KI, insbesondere LLMs, in der qualitativen Inhaltsanalyse übernehmen kann. Bis zum Zeitpunkt der Abfassung dieses Kapitels (Juni 2024) ist die Entwicklung generativer Sprachmodelle in rasantem Tempo verlaufen, die Modelle werden immer zahlreicher und immer leistungsfähiger. Gleichzeitig wurden generative Sprachmodelle auch bereits in Software zur qualitativen Datenanalyse wie MAXQDA und ATLAS.ti integriert. Was die Analyse sozialwissenschaftlicher Forschungsdaten betrifft, so sind in diesen QDA-Programmen schon seit längerem Funktionen enthalten, die sich wie die *Sentiment Analysis* oder *Topic Modeling* auf KI-Techniken stützen, ohne dass dies als Problem gesehen worden wäre. Niemand hätte eine Ablehnung der Masterarbeit oder Dissertation riskiert, wenn er oder sie Topic Modeling benutzt hätte, im Gegenteil wäre dies vielleicht als interessante methodische Innovation bewertet worden. Mit KI in Form von generativen Sprachmodellen verhält sich dies nun völlig anders. Warum? Wir sehen hierfür zwei Gründe, erstens die ungleich größere Mächtigkeit von LLMs und zweitens deren Doppelcharakter als *Werkzeug* und als *Assistenz*.

Die mächtigste Funktion *als Werkzeug* ist wohl die Fähigkeit der KI, in sehr kurzer Zeit sehr große Datenmengen zusammenzufassen und dabei die wichtigsten Punkte herauszuarbeiten. Diese Fähigkeit ist weitaus mächtiger als das bisherige Instrumentarium von QDA-Software, welches es „nur“ ermöglicht, aus sehr großen Textmengen alle Textstellen zusammenzustellen, die bestimmte Worte, Wortkombinationen oder thematisch codierte Segmente enthalten.

Gänzlich neu ist aber die Rolle der KI als *Assistenz*, als Co-Analyst bei der Exploration und Auswertung der Daten. Hier ist vor allem die Fähigkeit zum Chat hervorzuheben: Nun ist es möglich, in einen direkten Dialog mit den zu analysierenden Daten zu treten und in mehreren Zyklen Erkenntnisse zu gewinnen. Wie bei einem menschlichen Assistenten bzw. Assistentin kann die Assistenzrolle allerdings unterschiedlich interpretiert und praktiziert werden. Assistenz kann bedeuten, lediglich einfache Arbeiten zu delegieren; Assistenz kann aber auch bedeuten, wesentliche Teile der Arbeit oder sogar die gesamte Arbeit, inklusive der Erstellung des analytischen Berichts, zu delegieren. In diesem Fall würden die Forschenden dann nur noch Politikern ähneln, die das vom Redenschreiber verfasste Manuskript lediglich vorlesen.

Es ist vor allem diese Doppelrolle der generativen KI als Werkzeug und als Assistenz, die es den Universitäten und Forschungseinrichtungen schwer macht, zu operationalen Regeln für den Einsatz von LLMs zu kommen. Die Rolle als Werkzeug würde sich noch relativ einfach regeln lassen. Hier können, wie die DFG in ihrer o. g. Stellungnahme vorschlägt, die etablierten Standards guter wissenschaftlicher Praxis als Grundlage dienen. In Bezug auf die Nutzung von LLMs als Werkzeug würde dies bedeuten, Transparenz herzustellen und offenzulegen, welche Methoden in welchem Umfang eingesetzt wurden. Wie detailliert hierbei zu verfahren wäre, müsste dann noch festgelegt werden. Orientiert man sich daran, wie derzeit der Einsatz von Computersoftware beispielsweise bei der Literaturrecherche, bei statistischen Berechnungen, bei der Analyse qualitativer Daten oder bei der Texterstellung in Forschungsarbeiten offengelegt wird, so erscheinen allzu detaillierte Nachweise nicht erforderlich zu sein, schließlich ist es auch nicht üblich, bei jeder statistischen Tabelle das Hilfsmittel anzugeben, etwa durch die Anmerkung: „Diese Faktorenanalyse wurde mithilfe von SPSS Version 25 erstellt“. Auch bei einer qualitativen Inhaltsanalyse ist es nicht üblich, bei jedem Zitat das Werkzeug zu benennen („Dieses Zitat wurde mit Hilfe der von MAXQDA erstellten Zusammenstellung aller codierten Segmente der Kategorie ‚xy‘ ausgewählt“).

Weitaus schwieriger ist es, präzise Regelungen für die Assistenzrolle zu finden. Am Ende von Dissertationen, Bachelor- und Masterarbeiten muss eine sogenannte Selbstständigkeitserklärung eingefügt sein. In dieser wird erklärt, die Arbeit eigenständig verfasst zu haben und sämtliche fremde Quellen angegeben zu haben. In diesem Rahmen gab es Assistenz schon lange vor der Existenz von generativer KI. Konstruktive Doktorandencolloquien, kritische Lektüre von Kolleginnen und Kollegen, Tipps und Ratschläge von Betreuenden sowie Korrekturlesen und Verbessern durch kompetente Mitmenschen, all das war keineswegs verboten, sondern gängige und erwünschte Praxis – und natürlich wurden auch Tools und Datenbanken für die Online-Recherche über den Stand der Forschung sowie Zusammenfassungen und automatisch erstellte Zusammenfassungen genutzt. Detaillierte Regelungen oder gar Verordnungen zu Art und Umfang solcher Assistenz Tätigkeiten existieren nicht, letzten Endes bildet hier Vertrauen die Grundlage: Wer könnte schon etwas dagegen haben, dass beispielsweise die Masterarbeit durch die Durchsicht des befreundeten Deutschlehrers fehlerfrei und sprachlich wesentlich verbessert wird. Auch die Weiterentwicklung des Kategoriensystems aufgrund von intensiven Diskussionen im Graduiertenkolleg ist mit der Selbstständigkeitserklärung vereinbar.

Was bedeutet all dies nun für die Rolle von generativer KI im Rahmen der qualitativen Inhaltsanalyse? Potenziell kann KI ein sehr mächtiges Werkzeug für die qualitative Inhaltsanalyse sein, das mächtigste Werkzeug, das wir derzeit zur Verfügung haben und das mit großem Tempo weiterentwickelt wird. Aber generative KI ist eben nicht nur ein Werkzeug, sondern kann für die qua-

litative Inhaltsanalyse eine kompetente Assistenzrolle übernehmen: Mit der KI können wir Ideen über das Kategoriensystem und die Relevanz von Kategorien austauschen, wir können direkt mit den Daten kommunizieren, Zusammenhänge explorieren, große Textmengen zusammenfassen und dies in ungeahnter Geschwindigkeit tun. Large Language Models können sogar noch mehr: Sie können interpretieren, verschiedene Lesarten gegenüberstellen, Kategorien für die Analyse entwickeln, Besonderheiten entdecken, ja sogar Zusammenhänge visualisieren. All das ist jetzt bereits möglich und die Möglichkeiten werden kontinuierlich mehr. Im folgenden Kapitel 10.2 werden wir konkret forschungspraktisch in die Materie einsteigen und potenzielle neue, integrative Analysemöglichkeiten vorstellen. Dies tun wir aus der Sicht von aktiven Pionieren, die untersuchen, wie diese neuen Techniken und Möglichkeiten im Rahmen der qualitativen Inhaltsanalyse genutzt werden können. Generative KI stellt uns gleichzeitig auch vor große Herausforderungen, vor allem Ethik und Datenschutz betreffend; diese erörtern wir anschließend in Kapitel 10.3.

10.2 Praktische Anwendungsmöglichkeiten

Auf den folgenden Seiten zeigen wir die vielfältigen Möglichkeiten auf, wie künstliche Intelligenz, insbesondere generative Sprachmodelle, die Durchführung einer qualitativen Inhaltsanalyse bereichern kann. Es ist uns wichtig zu betonen, dass KI dabei nicht als einfacher Automatismus funktioniert, der aus eingegebenen Interviews oder anderen Texten einen fertigen Analysebericht erstellt. Vielmehr begreifen wir KI als Assistenzfunktion, die in den einzelnen Arbeitsschritten der qualitativen Inhaltsanalyse wertvolle Hilfe bietet. Als Forschende tragen wir dabei die uneingeschränkte Verantwortung für die Analyse, was bedeutet, dass wir die von der KI generierten Ausgaben kritisch überprüfen und falls nötig an unsere Forschungsziele anpassen müssen, um die Qualität und Integrität unserer Arbeit sicherzustellen (Rädiker, 2024a).

Der Einsatz von KI ist in allen drei Varianten der qualitativen Inhaltsanalyse möglich. Unsere Ausführungen gliedern sich daher entlang der Schritte des allgemeinen Ablaufmodells (Abb. 11, S. 106), beginnend mit der initiierenden Textarbeit, über die Bildung von Kategorien, das Codieren und die Analyse der codierten Daten, bis hin zur Erstellung des abschließenden Berichts. KI lässt sich natürlich auch in dem vorbereitenden Schritt der Transkription gewinnbringend einsetzen, dies haben wir bereits im Kapitel 8.1 erläutert.

Zur Umsetzung der im Folgenden beschriebenen Möglichkeiten können Sie verschiedene Softwarelösungen verwenden:

- ChatGPT, Microsoft Copilot und andere Online-Tools, die Chat-Anfragen mit Hilfe großer Sprachmodelle beantworten.