

Hybride Intelligenz in der Markt- forschung



Mehr Power durch Mensch-KI- Kooperation

AUTOR:INNEN

Neeraj Arora

Professor of Marketing
University of Wisconsin-Madison

Ishita Chakraborty

Assistant Professor of Marketing
University of Wisconsin-Madison

Yohei Nishimura

Doctoral Student
University of Wisconsin-Madison

KEYWORDS

Generative KI
Qualitative Forschung
Umfragen
Consumer Insights
Natural Language Processing
Unstrukturierte Daten





Die Marktforschungsbranche befindet sich aufgrund der rasanten Fortschritte in den Bereichen generative KI (GenAI) und Large Language Models (LLMs) in einem tiefgreifenden Wandel. Während Marketingabteilungen massiv in KI-Anwendungen zur Content-Erstellung und Personalisierung investieren, zeigen wir, wie LLMs als Partner für die Generierung von Consumer Insights genutzt werden können. Unsere zentrale Idee ist die Entwicklung leistungsstarker „KI-Mensch-Hybride“, die grundlegend verändern, wie Unternehmen ihre Kunden verstehen, Produkte entwickeln und Marketingstrategien gestalten.

Das Potenzial der KI für die Erkenntnisgewinnung

In einer neu gedachten Forschungspipeline bleibt der menschliche Forscher unverzichtbar. Er übernimmt die Rolle eines Supervisors, der die KI anleitet, ihre Ergebnisse validiert und ethisch sensible oder neuartige Kontexte behandelt, in denen die KI an Grenzen stößt. Durch die Partnerschaft mit LLMs können Unternehmen den Prozess der Erkenntnisgewinnung effizienter und wirksamer gestalten und dabei agiler agieren. In unserer Forschung analysieren wir, wie KI-Mensch-Hybride den Marktforschungsprozess beschleunigen, Kosten senken und tiefgehende Erkenntnisse gewinnen können (siehe Box 1). Die Ergebnisse aus zwei Studien, in denen wir traditionelle Marktforschungsprojekte KI-gestützt repliziert haben, sind sehr ermutigend. Sie zeigen, dass LLMs in jeder ►

Das Forschungsprojekt – Vergleich von traditionell durchgeführter Marktforschung mit KI-gestützten Prozessen

BOX 1

Wir haben zwei Studien – eine qualitative und eine quantitative – durchgeführt, um den Wert eines hybriden Ansatzes, bei dem Mensch und KI kooperieren, zu analysieren. In Zusammenarbeit mit einem Fortune-500-Lebensmittelunternehmen haben wir bereits durchgeführte Marketingforschungsstudien KI-gestützt repliziert. So konnten wir den traditionellen, ohne KI durchgeführten Ansatz des Unternehmens mit der hybriden Vorgangsweise in unserem Projekt direkt vergleichen.

QUALITATIVE STUDIE

Die Studie konzentrierte sich auf „Friendsgiving“, eine moderne Version von Thanksgiving, bei der sich Freunde statt Familienangehörige zum Feiern treffen. Ziel war es, die Emotionen und Traditionen der Konsumenten rund um dieses Ereignis zu untersuchen. LLMs wurden in drei Rollen eingesetzt:

- als synthetische Befragte, die bestimmte Personas darstellten, z. B. einen 32-jährigen Mann, der eine Party veranstaltet. Sie agierten als Interviewpartner, um detaillierte, offene Antworten auf Interviewfragen zu geben;
- als Interviewer und Moderatoren, die Fragen generierten, bei Bedarf nachhaken und Diskussionen leiteten;
- und als Analysten, die die großen Mengen unstrukturierter Textdaten aus den Interviews zusammenfassen und interpretierten.



QUANTITATIVE STUDIE

Für die quantitative Studie untersuchten wir die Fähigkeit des LLM, eine Umfrage zur Bewertung eines neuen Konzepts für gekühltes Hundefutter zu replizieren. Die KI hatte die Aufgabe, Antworten für 605 synthetische Teilnehmer zu generieren. Die Personas wurden so erstellt, dass sie genau den demografischen und verhaltensbezogenen Profilen der ursprünglichen menschlichen Befragten entsprachen. Sie beantworteten konzeptbezogene Fragen zu Verpackungspräferenzen, bevorzugten Zutaten, Kaufwahrscheinlichkeiten, Einstellungen zum Produktkonzept und der voraussichtlichen Nutzungsfrequenz.

Ziel der Studie war es zu bewerten, in welchen Anwendungsfeldern ein LLM die Leistung rein menschlicher Prozesse bei der Generierung und Analyse typischer Marktforschungsdaten erreichen oder übertreffen kann.

Phase der Forschungspipeline als effektive Assistenten dienen können, von der Entwicklung des Forschungsdesigns über die Datenerhebung und -analyse bis hin zur Aufbereitung von entscheidungsrelevanten Erkenntnissen für die Weiterentwicklung von Geschäftsideen. Im Zuge unseres Projekts haben wir auch eine neuartige Anwendung von LLMs getestet: ihre Fähigkeit, KI-gestützte Personas zu generieren, die als „synthetische Befragte“ an Tiefeninterviews teilnehmen oder Fragebögen beantworten können.

Bewertung des hybriden Prozesses in der qualitativen Forschung

In unserer qualitativen Studie zu „Friendsgiving“ funktionierte die KI-Integration in allen Phasen des Projekts sehr gut.

Generierung von Befragten – Aufbau eines Silicon Samples

Im Bereich der Datengenerierung konnten LLMs die erforderlichen Merkmale für die zu befragende Stichprobe effektiv erstellen. Es gelang ihnen, synthetische Befragte zu generieren, die den Merkmalen der ursprünglichen Stichprobe entsprachen. Zusätzlich zur Nachbildung der Befragtenprofile der ursprünglichen Studie baten wir das LLM, mögliche zusätzlich relevante Gruppen von Befragten vorzuschlagen, die Ideen zum Forschungsthema „Friendsgiving“ beitragen könnten, wobei wir die in

Abbildung 1 dargestellten Prompts verwendeten. Als Antwort auf den ersten Prompt erstellte das LLM eine ausführliche Beschreibung potenzieller Befragter. Obwohl die ursprüngliche Studie einen guten Mix aus demografischen Merkmalen der Befragten wie Geschlecht, Alter und ethnische Zugehörigkeit aufwies, entwickelte das LLM zusätzliche Dimensionen und schlug z. B. Personen mit Ernährungseinschränkungen, Minderheiten wie LGBTQ+, Auswanderer und internationale Studenten vor – alles Gruppen, die sehr plausibel ein besonderes Interesse an einer inklusiveren, moderneren Thanksgiving-Feier haben könnten. Als Antwort auf den zweiten Prompt generierte das LLM 10 synthetische Personas entsprechend den zuvor generierten Vorschlägen und verbesserte damit die synthetische Stichprobe.

Datenerhebung – Fragen, Antworten, Bewerten und Nachhaken

Im nächsten Schritt führte das LLM als Interviewer ausführliche Interviews durch und moderierte sie. Als Befragter kann das LLM dazu aufgefordert werden, einfach auf die Fragen aus dem Diskussionsleitfaden zu antworten und synthetische Antworten zu geben. Darüber hinaus kann es auch dazu aufgefordert werden, die Tiefe, den Umfang und die Konsistenz der Antworten zu verbessern. Als Bewerter kann ein LLM seine eigenen Antworten anhand wichtiger Forschungsziele und Metriken wie Klarheit und Tiefe weiter bewerten und einen

Score vergeben. In seiner Rolle als „Nachhaker“ kann es synthetische Befragte bitten, weitere Erläuterungen zu geben, wenn der Score für die Antwort unter einen bestimmten Schwellenwert fällt. Diese zusätzlichen Rollen sind hervorragende Beispiele für neuartige Möglichkeiten, wie ein LLM im Zuge der Datengenerierung einen Mehrwert schaffen kann. In unserer Studie ließen wir menschliche Bewerter die Antworten aus der ursprünglichen Studie und der replizierten Studie vergleichen und wandten mehrere statistische Maße zu deren Beurteilung an. Die vom LLM generierten Antworten waren den menschlichen Antworten sowohl in Bezug auf Tiefe als auch Nachvollziehbarkeit überlegen und boten somit Effizienzgewinne.

Datenanalyse – Gewinnung von Erkenntnissen und Berichterstattung

Im Bereich der Analyse schnitten die LLMs gleich gut ab wie die menschlichen Analysten und konnten mit Experten mithalten, wenn es darum ging, Schlüsselideen zu identifizieren, sie zu Themen zu gruppieren und Informationen zusammenzufassen. Zwar übersahen die LLMs einige Themen, die von Menschen erkannt wurden, aber sie generierten auch neue Themen, die Menschen nicht erkannt hatten.

Bewertung von Mensch-KI-Hybriden bei Umfragen

In einem quantitativen Umfeld erstellten die LLMs – die auf ähnliche Weise gepromptet wurden – synthetische Befragte, die den Fragebogen erfolgreich beantworteten. Die Ergebnisse des Basis-LLMs waren jedoch gemischt. Obwohl das LLM die Richtung und die Stärke der Einstellungen der Konsumenten korrekt erfasste, zeigte es zwei wesentliche Schwächen, die bei vielen Standard-KI-Anwendungen zu finden sind. Erstens mangelte es den Antworten an Heterogenität; im Vergleich zu den menschlichen Daten gab es weniger Variationen in

“
Da die KI die Datenanalyse erledigt, können Forscher ihre Expertise auf das Erkennen der wertvollsten Insights und die Interpretation der Ergebnisse konzentrieren.
“

den KI-generierten Antworten. Zweitens war deren interne Konsistenz im Vergleich mit menschlichen Befragten geringer; beispielsweise bewertete das LLM KI-Attribute wie „gesunde Zutaten“ und „sicherste Lebensmittel“ ganz anders als Menschen. Um diese Einschränkungen des Basis-LLM-Modells zu beheben, haben wir zwei effektive Techniken getestet, um Kontext einzubeziehen und die Qualität der synthetischen Umfragedaten zu verbessern.

Bereitstellung des vorherigen Antwortverlaufs

Bei dieser Methode werden dem LLM seine eigenen früheren Antworten innerhalb derselben Umfrage als Kontext für nachfolgende Fragen zur Verfügung gestellt. Diese einfache Erinnerungshilfe verbessert die Kohärenz der Antworten über alle Fragen einer Umfrage hinweg. ▶

Prompts zur Generierung einer geeigneten Stichprobe synthetischer Befragter

ABBILDUNG 1

ZIEL

Identifizierung der Stichprobe

Erstellung von Personas

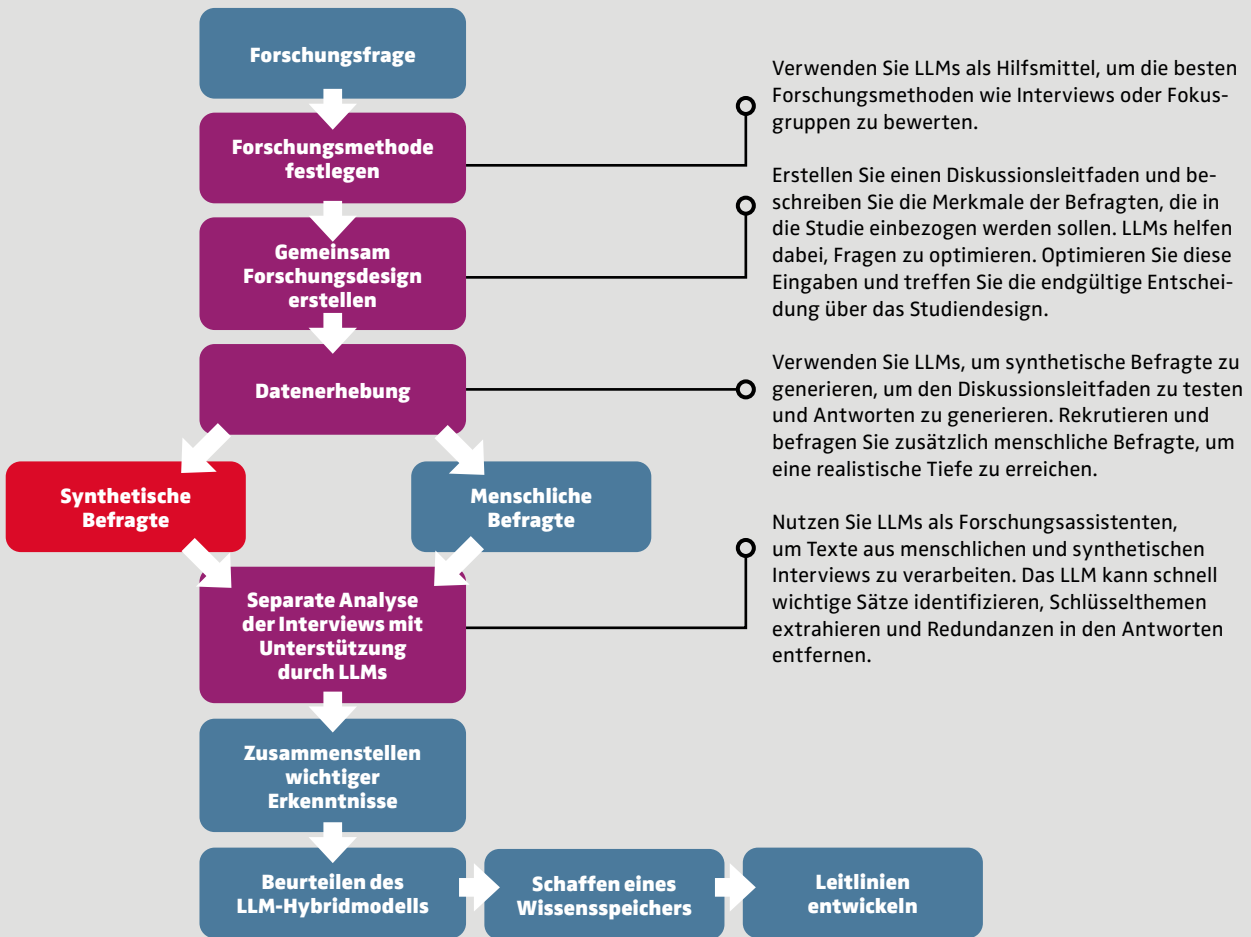
PROMPT

Friendsgiving ist eine neue Variante von Thanksgiving. Wir möchten die Einstellung der Menschen dazu verstehen, wie sie es feiern und welche Gedanken und Gefühle damit verbunden sind. Was wäre eine gute repräsentative Stichprobe von Personen, mit denen man sprechen könnte?

Erstellen Sie nun Personas von Menschen, die die oben aufgeführten Merkmale aufweisen. Nicht jede Persona muss alle Merkmale aufweisen. Die Personas sollten sich in ihren relevanten Merkmalen eindeutig unterscheiden.

Wann und wie man LLMs als Assistenten in der qualitativen Forschung einsetzt

ABBILDUNG 2



Erweiterung des Kontextwissens des LLM

Zusätzlich haben wir das LLM verbessert, indem wir externes Wissen bereitgestellt haben, und zwar mithilfe einer Technik namens Retrieval-Augmented Generation (RAG). Konkret haben wir dem LLM die vorhandenen qualitativen Interviewtranskripte des Unternehmens über Premium-Tiernahrung als externes Wissen zur Verfügung gestellt, bevor es den Fragebogen beantwortete. Durch diese einfache Ergänzung mittels RAG konnte das LLM seine Antworten auf relevantes Konsumentwissen stützen. Beide Techniken verbesserten die Vielfalt und interne Konsistenz der LLM-Antworten erheblich, sodass sie den Antworten von Menschen stärker ähnelten.

Wie man KI in ein Forschungsprojekt integriert

Die Abbildungen 2 und 3 zeigen Roadmaps und konkrete Vorschläge zur Einbindung von KI in qualitative und quantitative Forschungsprojekte. In beiden Abbildungen stehen die blauen, roten und violetten Blöcke für rein menschliche, rein LLM-basierte und hybride KI-Mensch-Prozesse.

Die abgebildeten LLM-gestützten Forschungsprozesse verbessern die Effizienz und Geschwindigkeit der Generierung von Insights sowie die Qualität der Analysen. Da die KI die Datenanalyse erledigt, können Forscher ihre Expertise auf das

Erkennen der wertvollsten Insights und die Interpretation der Ergebnisse konzentrieren. Sie können die Insights aus synthetischen und menschlichen Daten zusammenführen, um entscheidungsrelevante Empfehlungen zu formulieren, die das Unternehmen weiterbringen. Die neu gestalteten

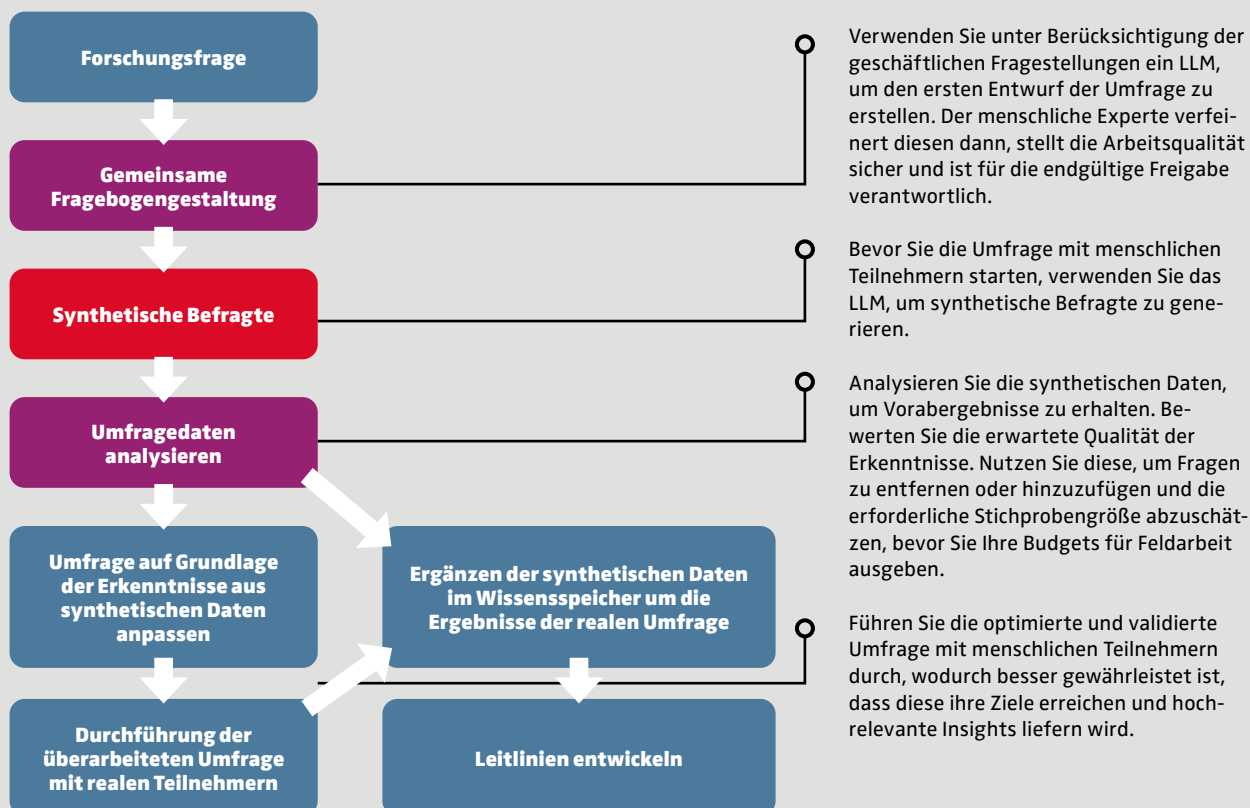
“

Am erfolgreichsten werden die Marktforscher und Unternehmen sein, die das Tempo der KI mit dem Urteilsvermögen menschlicher Expertise am besten verbinden.

“

Wann und wie man LLMs in Umfragen einsetzt

ABBILDUNG 3



Prozesse ermöglichen es den Forschern, sich auf geschäftsrelevante Fragen zu konzentrieren, anstatt repetitive Aufgaben zu erledigen. Menschliche Experten können interne Wissensspeicher erstellen, indem sie sowohl menschliche als auch synthetische Daten aufbewahren. Der Vergleich dieser Datensätze im Zeitverlauf kann dabei helfen, institutionelles Wissen über Geschäftskontexte zu schaffen, in denen LLMs am effektivsten sind. Darauf aufbauend kann man Leitlinien für die zukünftige Nutzung entwickeln.

Die Zukunft der Marktforschung ist hybrid

Mensch-KI-Hybride erzielen bessere Ergebnisse als Menschen oder KI jeweils für sich allein. In qualitativen Studien liefern sie reichhaltigere und aussagekräftigere Erkenntnisse. In quantitativen Umfragen bieten sie eine schnelle und kostengünstige Möglichkeit, die Ergebnisse einer Umfrage zu visualisieren, bevor in ein vollwertiges Projekt investiert wird. Führungskräfte sind gut beraten, jetzt mit Mensch-KI-Hybriden zu experimentieren und KI als strategisch relevanten Geschäftspartner zu begreifen. Dieser Zugang senkt nicht nur Kosten und verkürzt Zeitpläne, sondern führt auch zu besseren Entscheidungen und schärferen strategischen Erkenntnissen. Gleichzeitig müssen Marktforscher wissen, dass die Trainingsdaten ihrer KI-

Systeme nicht vorurteilsfrei sind und ethnische, geschlechtsspezifische und kulturbezogene Biases widerspiegeln können. Für die Absicherung von Genauigkeit, Fairness und Vertrauen ist es entscheidend, das Bewusstsein für diese Risiken zu schärfen und menschliche Kontrolle im Forschungsprozess sicherzustellen. Am erfolgreichsten werden die Marktforscher und Unternehmen sein, die das Tempo der KI mit dem Urteilsvermögen menschlicher Expertise am besten verbinden. ◀

LITERATURHINWEISE

Arora, N., Chakraborty, I., & Nishimura, Y. (2025). AI-human hybrids for marketing research: Leveraging large language models (LLMs) as collaborators. *Journal of Marketing*, 89(2), 43–70. <https://doi.org/10.1177/00222429241276529>

Blanchard, S. J., Duani, N., Garvey, A. M., Netzer, O., & Oh, T. T. (2025). New tools, new rules: A practical guide to effective and responsible generative AI use for surveys and experiments in research. *Journal of Marketing*, 89(6), 119–139. <https://doi.org/10.1177/00222429251349882>

Korst, J., Puntoni, S., & Toubia, O. (2025, May–June). How Gen AI is transforming market research. *Harvard Business Review*. Harvard Business Review. <https://hbr.org/2025/05/how-gen-ai-is-transforming-market-research>