

Neue Tools, neue Aufgaben

Wie Manager Generative KI für bessere Marketing- Insights nutzen

AUTOR:INNEN

Oded Netzer

Arthur J. Samberg Professor of Business
Columbia Business School

Simon J. Blanchard

Dean's Professor
McDonough School of Business
Georgetown University

Nofar Duani

Assistant Professor of Marketing
Marshall School of Business
University of Southern California

Aaron M. Garvey

Associate Professor of Marketing
Haslam College of Business
University of Tennessee

Travis Tae Oh

Assistant Professor of Marketing
Sy Syms School of Business
Yeshiva University

KEYWORDS

Generative KI
Umfragen
Experimente
Marktforschung
Textanalyse
Validierung



GettyImages/Qi Yang



Stellen Sie sich Folgendes vor: Es ist Montagmorgen, und Ihr Team hat gerade eine neue Produktidee entwickelt. Bereits am Freitag liegen Konzepttests, Kundenreaktionen und ein vollständiges Insights-Deck vor – alles unterstützt durch generative KI (GenAI). Was früher Monate gedauert hat, passiert heute in wenigen Tagen. GenAI-Tools können Daten zusammentragen, adaptive Interviews durchführen und die Stimmung von Konsumenten innerhalb von Minuten zusammenfassen. In schnelllebigen, hart umkämpften Märkten ist das ein echter Game-changer. Doch das enorme Potenzial hat auch

Tücken: Die neuen Tools können gute Forschung zwar beschleunigen – aber auch Fehler potenzieren. Die Kunst liegt im richtigen Einsatz.

Das Funktionsprinzip von GenAI

Das Herzstück vieler GenAI-Systeme sind große Sprachmodelle (LLMs). Diese Modelle werden mit Milliarden von Wörtern trainiert, um das jeweils nächste Wort in einer Sequenz vorherzusagen. Sie sind hervorragend darin, Texte zu erzeugen, die sprachlich stimmig und kontextuell passend sind. Das macht sie zu leistungsfähigen Tools für das Erstellen von Fragebögen, das Zusammenstellen ▶

von Reports oder das Schreiben von Code. LLMs arbeiten jedoch probabilistisch. Sie wählen das wahrscheinlichste nächste Wort – nicht zwangsläufig das faktisch richtige. Deshalb „halluzinieren“ sie gelegentlich Fakten, geben Quellen oder Ergebnisse falsch wieder oder weichen von klar definierten Konzepten ab. Moderne GenAI-Systeme gehen über reine LLMs hinaus und ergänzen diese mit zusätzlichen Funktionen. Retrieval-Augmented Generation (RAG) ermöglicht es, externe Dokumente einzubeziehen und hochgeladene Reports oder Datensätze als Kontext zu nutzen. Code-Ausführungsumgebungen erlauben es GenAI-Systemen nicht nur, Code vorzuschlagen, sondern Programme tatsächlich auszuführen und statistische Analysen anzuwenden. Diese systemischen Erweiterungen machen GenAI für Nutzer zugänglicher und erhöhen ihr Potenzial für Forschung und Unternehmen. Sie verändern jedoch weder den probabilistischen Grundcharakter der Technologie noch die Prinzipien sauberer, rigoroser Forschung.

Scannen der externen Wissenslandschaft

Eine einzige Abfrage kann strukturierte Zusammenfassungen aus Dutzenden von wissenschaftlichen Artikeln, Branchenberichten oder Fachpublikationen liefern und dabei aufkommende Themen in dicht besetzten Forschungsfeldern sichtbar machen. Für eine erste Eingrenzung ist diese Geschwindigkeit von unschätzbarem Wert. Gleichzeitig birgt genau diese Effizienz auch Risiken. Da GenAI Informationen rekonstruiert, statt sie direkt abzurufen, kann es zu Vermischungen von Ergebnissen oder sogar zur Zitierung falscher Quellen kommen. Für Manager ist die Schlussfolgerung eindeutig: GenAI-Zusammenfassungen sind als Ausgangspunkt zu verstehen, nicht als belastbare Wissensquelle. Sie helfen, relevante Themen und Zusammenhänge aufzudecken, müssen jedoch vor strategischen Entscheidungen immer anhand der Originaldokumente überprüft werden. Verantwortungsvoll eingesetzt kann GenAI tagelange Desk Research auf wenige Stunden verkürzen. Unkritisch genutzt entstehen schnell scheinbar überzeugende Argumente, die einer genauen Überprüfung nicht standhalten.

“

GenAI ist ein Generator von Alternativen, aber Manager müssen diese Alternativen filtern und verfeinern.

“

Insights aus internem Wissen gewinnen

Neben externen Quellen verfügen viele Unternehmen über umfangreiches, oft wenig genutztes internes Wissen: frühere Befragungsdaten – insbesondere offene, qualitative Antworten aus großen Umfragen oder Fokusgruppen –, Kundenstudien, Beratungsreports oder interne Wettbewerbsanalysen. GenAI kann helfen, Muster und Erkenntnisse aus dieser fragmentierten Informationsbasis herauszuarbeiten und ältere Insights wieder aktiv nutzbar zu machen. Bei der Arbeit mit internen Quellen hat Vertraulichkeit jedoch oberste Priorität. Das Hochladen proprietärer Daten in öffentliche Chat-Schnittstellen birgt das Risiko, sensible Unternehmensinformationen preiszugeben. Manager sollten stattdessen auf Enterprise-APIs oder lokal betriebene LLM-Deployments setzen, die Datensicherheit gewährleisten. Unter diesen Voraussetzungen wird GenAI zu einem leistungsfähigen Instrument, um jahrelang angesammeltes Wissen zu kombinieren – und isolierte Archive in handlungsrelevante Entscheidungsgrundlagen für aktuelle Fragestellungen zu verwandeln. ▶



Midjourney AI prompt by GCO

Für Manager ist es entscheidend, diese Grundlagen zu verstehen. So wie man Umfrageergebnisse nicht beurteilen sollte, ohne das Stichprobenverfahren zu kennen, sollte man sich auch nicht auf GenAI verlassen, ohne ihre Grenzen zu erkennen. In den folgenden Abschnitten zeigen wir, wie sich GenAI-Fähigkeiten in unterschiedlichen Settings der Insights-Generierung verantwortungsvoll und sorgfältig einsetzen lassen.

Desk Research mit GenAI

Jedes Forschungsprojekt beginnt damit, zu verstehen, was bereits bekannt ist. GenAI beschleunigt diese Phase der Desk Research erheblich.

Die Kunst des präzisen Promptings: Das richtige Konstrukt treffen

BOX 1

Generative KI kann in Sekundenschnelle Fragebogenitems, Bilder und weitere Stimuli entwerfen – allerdings nur, wenn klar ist, was gemessen werden soll. Sind Prompts vage formuliert, weichen Modelle häufig auf verwandte Ideen aus. Das Ergebnis kann zwar ausgefeilt klingen, aber das gewünschte Konstrukt völlig verfehlen.

Beispiel

Ein Einzelhändler wollte Konsumentenreaktionen auf „natürlich gewachsenes Obst“ untersuchen und bat GenAI, Items zur Messung der „visuellen Attraktivität“ zu erstellen. Das System generierte Aussagen wie: „Dieses Obst sieht frisch und hochwertig aus.“ Sprachlich sauber, inhaltlich jedoch falsch: Das Konstrukt verschob sich von der Attraktivität der Form hin zu Frische und Qualität – zwei klar unterschiedliche Dimensionen.

Bei visuellen Stimuli zeigte sich eine ähnliche Problematik. Wenn im Prompt klare Definitionen enthalten waren, kam es zu substanziell anderen Ergebnissen.

Abbildung 1 zeigt, welche Ausgaben GenAI erzeugt hat, als es beauftragt wurde, einen „visuell attraktiven“ bzw. einen „visuell unattraktiven“ Apfel zu erstellen. Präzise Vorgaben im Prompt konnten die Ergebnisse deutlich verbessern – und verhindern, dass am eigentlichen Messziel vorbeigearbeitet wird.

Variationen in den Ergebnissen bei unterschiedlichen Prompts für „Visuelle Attraktivität“ eines Apfels

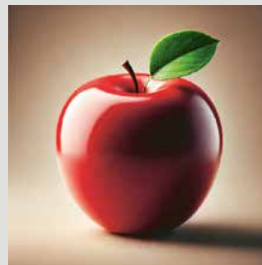
ABBILDUNG 1

VAGE

Prompt-Ansatz

„Erstellen Sie zwei separate Bilder, auf denen jeweils ein optisch attraktiver und ein visuell unattraktiver Apfel zu sehen ist.“

Attraktiv



Unattraktiv



MIT KONTEXT

Ergänzung von Zeile 1 um die Definition und den Kontext, z. B. durch die Spezifizierung von „visueller Attraktivität“ als natürliche Variation in der Form und nicht als Anzeichen von Verderben oder als Qualitätsmangel.



PRÄZISIERT*

Weitere Spezifizierung von Hintergrund, Winkel, Farbvariationen, Oberflächenstruktur, Beleuchtung usw.



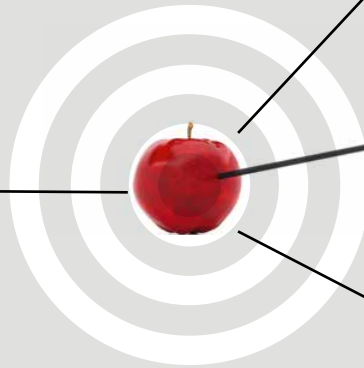
*Präzisiert Prompt: „Eine hochauflösende Nahaufnahme eines einzelnen roten Apfels vor einem weißen Hintergrund, aufgenommen auf Augenhöhe. Der Apfel hat eine asymmetrische und untypische Form. Seine Färbung zeigt eine Mischung aus Rot- und Orange-Gelbtönen. In einem großen Bereich in der Mitte geht die rote Pigmentierung in eine starke natürliche Verfärbung mit weniger Glanz über. Die Schale hat kleine, punktuelle Verfärbungsflecken, die dem Apfel ein organisches, auf natürliche Weise unvollkommenes Aussehen verleihen. Die Oberfläche ist glatt mit sichtbaren Texturvariationen, aber der Apfel wirkt trotzdem frisch und unbeschädigt. Das Licht ist weich und gleichmäßig und betont die natürlichen Unregelmäßigkeiten, ohne den Apfel angeschlagen oder verdorben erscheinen zu lassen.“

So bleiben Prompts auf Kurs

ABBILDUNG 2

Fügen Sie Kontext hinzu

Beschreiben Sie die Bewertungssituation oder den Betrachtungsrahmen, z. B. „ein einzelner Apfel vor weißem Hintergrund“, um die Interpretation zu steuern.



Definieren Sie das Konstrukt klar

Geben Sie explizit an, was eingeschlossen und ausgeschlossen ist. Beispiel: „Visuelle Attraktivität bezieht sich auf natürliche Formvariationen, nicht auf Verderben oder Produktqualität.“

Iterieren und überprüfen Sie

Selbst präzise Anweisungen erfordern menschliches Urteilsvermögen, um Biases, Mehrdeutigkeit oder konzeptionelle Abweichungen zu erkennen.

iStock/Evgeny Ostroushko

Gestaltung von Fragebögen und Messinstrumenten

Eine der größten Stärken von GenAI liegt in der Erstellung von Fragebogenitems, Skalen und experimentellen Stimuli. Was früher Zeit, Kreativität und oft mehrere Iterationsschleifen erforderte, lässt sich heute in Sekunden generieren. Ein einzelner Prompt kann Dutzende potenzieller Items oder Varianten von Werbetexten liefern und verschafft Managern damit eine Breite an Optionen, deren manuelle Erstellung kaum leistbar gewesen wäre.

Die Herausforderung besteht darin, dass diese Outputs nicht immer konzeptionell präzise sind, wie Abbildung 1 zeigt. Eine Lösung besteht darin, das Modell mit klaren Definitionen zu steuern, die festlegen, was das zentrale, zu erforschende Konstrukt beinhaltet – und was nicht (siehe Box 1). Doch selbst dann bleibt eine sorgfältige menschliche Prüfung unverzichtbar, um mögliche Verzerrungen oder Unklarheiten zu identifizieren.

Für Stimuli wie Werbetexte oder Vignetten gilt dasselbe Prinzip: GenAI ist ein Generator von Alternativen, aber Manager müssen diese Alternativen filtern und verfeinern.

Datenerhebung und -analyse

GenAI verändert die Art und Weise, wie Daten erhoben werden, wie man mit Befragten interagiert und wie die gewonnenen Daten anschließend kodiert sowie ausgewertet werden.

Meinungsforschung wird adaptiv und skalierbar
Kundenbefragungen basierten lange auf festen, vorab geskripteten Fragen. GenAI verändert das grundlegend, indem sie adaptive Formate ermöglicht, in denen ein KI-Agent seine Anschlussfragen an die Antworten der Teilnehmenden anpasst. Unternehmen können heute KI-Interviewer in Umfrageplattformen integrieren, um Konsummotive vertieft zu erforschen, fokusgruppenähnliche Diskussionen zu simulieren oder dynamische Treat-

ments in Experimenten einzusetzen. Solche Agenten können zudem mit den Forschenden selbst interagieren – etwa bei der Konzeption oder Analyse der Studie. Der Vorteil liegt in qualitativer Tiefe bei quantitativer Skalierung: Statt Hunderter kurzer, statischer Antworten erhalten Manager Hunderte adaptiver Gespräche, jeweils auf die einzelne Person zugeschnitten.

Dieser Ansatz ist zwar revolutionär und nützlich, bringt aber Risiken mit sich. Da GenAI Antworten probabilistisch generiert, können zwei Teilnehmende trotz identischer Antwort leicht unterschiedliche Folge-Prompts erhalten. Das erzeugt Nebengeräusche in den Daten und schwächt die Vergleichbarkeit. Hinzu kommt das Risiko, dass die KI vollständig vom vorgesehenen Skript abweicht. Um dem entgegenzuwirken, müssen Unternehmen umfangreich vortesten, mit Verstärkungsprompts arbeiten, um die KI fokussiert zu halten, und Gespräche vollständig transkribieren, um Kontrollmöglichkeiten zu schaffen. Mit diesen Absicherungsmaßnahmen kann Meinungsforschung das Kundenverständnis deutlich vertiefen, ohne dabei an Genauigkeit einzubüßen.

“

Ohne konsequente Prüfung von Sorgfalt, Validität und Reproduzierbarkeit kann GenAI ebenso leicht irreführen wie aufklären.

“



Codierung und Analyse werden skalierbar

Während KI-Interviewer den Umfang dessen erweitern, was gefragt werden kann, beschleunigt GenAI-gestütztes Coding die Verarbeitung der Antworten. Über Jahrzehnte hinweg hatten Manager Schwierigkeiten, die riesigen Mengen an Kundenbewertungen, Social-Media-Beiträgen, frei formulierten Antworten und Interviewtranskripten zu klassifizieren. GenAI ermöglicht es nun, Texte, Bilder und Videos in wenigen Minuten zu codieren und unstrukturierte Daten in nutzbare, handlungsrelevante Erkenntnisse zu überführen.

Automatisierung macht Validierung jedoch nicht überflüssig – im Gegenteil, sie wird noch wichtiger. KI-codierte Variablen müssen wie alle anderen geprüft werden: durch Abgleich mit Selbstauskünften bei der Messung von Einstellungen und Präferenzen, durch Benchmarking mit menschlichen Codern bei interpretativen Beurteilungen oder durch Überprüfung anhand realer Outcomes bei der Vorhersage von Verhalten. GenAI kann außerdem Analyse-Code in R, Python oder anderen Statistiksoftware-Anwendungen entwerfen – und teilweise auch ausführen – und senkt damit die Einstiegshürde für Teams mit begrenzter Programmiererfahrung.

Trotzdem sollten Analysten sich nicht mit Ergebnissen zufriedengeben, die vollständig innerhalb eines Chatfensters erzeugt wurden. Jeder generierte Code sollte auch mit spezialisierter Analyse-Software ausgeführt und validiert werden, um Genauigkeit und Reproduzierbarkeit sicherzustellen. Richtig eingesetzt kann GenAI sowohl den verarbeitbaren Datenumfang als auch den Umfang der Analyse erheblich erweitern – vorausgesetzt, Manager verbinden Geschwindigkeit mit methodischer Disziplin bei der Replikation ihrer Workflows.

Schnellere Insights – aber jeder Schritt muss geprüft werden

Erkenntnisse, deren Gewinnung früher Wochen dauerte, lassen sich heute in wenigen Tagen oder

sogar noch schneller erzielen. Doch Geschwindigkeit ohne methodische Sorgfalt ist gefährlich. Irreführende Sekundärforschung kann Strategien in die falsche Richtung lenken, schlecht konzipierte Survey-Items können Segmentierungsstudien untergraben. Inkonsistente KI-Interviews gefährden die Reliabilität von Erkenntnissen, und der ungeprüfte Einsatz von KI-Coding kann Reproduzierbarkeit und Generalisierbarkeit aushöhlen.

Manager müssen daher ein neues Mindset entwickeln: GenAI ist kein Orakel, sondern ein Beschleuniger. Sie erweitert das Machbare, ersetzt aber nicht die Verantwortung, Genauigkeit zu prüfen, Validität sicherzustellen und Datenschutz zu gewährleisten. Generative KI ist nicht nur ein weiteres Produktivitätswerkzeug – sie verändert grundlegend, wie Umfragen und Experimente konzipiert, durchgeführt und ausgewertet werden.

Für Manager liegt der Nutzen in schnelleren Innovationszyklen, einer breiteren Exploration von Ideen und der Möglichkeit, die Insights-Generierung zu skalieren, die früher durch Zeit- und Kostenrestriktionen begrenzt war. Technologie allein garantiert jedoch keine besseren Entscheidungen. Ohne konsequente Prüfung von Sorgfalt, Validität und Reproduzierbarkeit kann GenAI ebenso leicht irreführen wie aufklären. Es liegt in der Verantwortung der Manager, neue Tools mit neuen Regeln zu verbinden. Diejenigen, die dies tun, werden GenAI als echten strategischen Vorteil nutzen: Sie können schneller entscheiden, ohne dabei methodische Sorgfalt zu opfern. ◀

LITERATURHINWEISE

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>