

# Professionelle Recherche mit generativer Künstlicher Intelligenz (KI)

Whitepaper



# Recherche mit generativer KI: Ein Leitfaden für Research Professionals

Seit OpenAI die erste Version von ChatGPT veröffentlichte und weitere generative KI-Tools gefühlt im Sekundentakt folgten, ist KI das Thema der Stunde. Viele Unternehmen fragen sich, wie sie die Technologie für sich nutzen können und allen ist klar, dass sie über kurz oder lang an dem Thema KI nicht vorbeikommen. Auch wir haben uns diese Frage gestellt und uns intensiv mit KI beschäftigt – zunächst vor allem unter dem Aspekt, inwieweit wir generative KI für unsere Haupttätigkeit, die Recherche, nutzen können.

Neben einem internen Leitfaden für unsere Mitarbeitenden ist dieses Whitepaper als Ergebnis entstanden. Es richtet sich an alle, die professionelles Research betreiben – egal ob schwerpunktmäßig oder punktuell im Rahmen einer anderen Tätigkeit. Dieses Whitepaper soll als Hilfestellung dienen und einen guten Impuls zu einer effizienten und reflektierten Nutzung generativer KI geben.

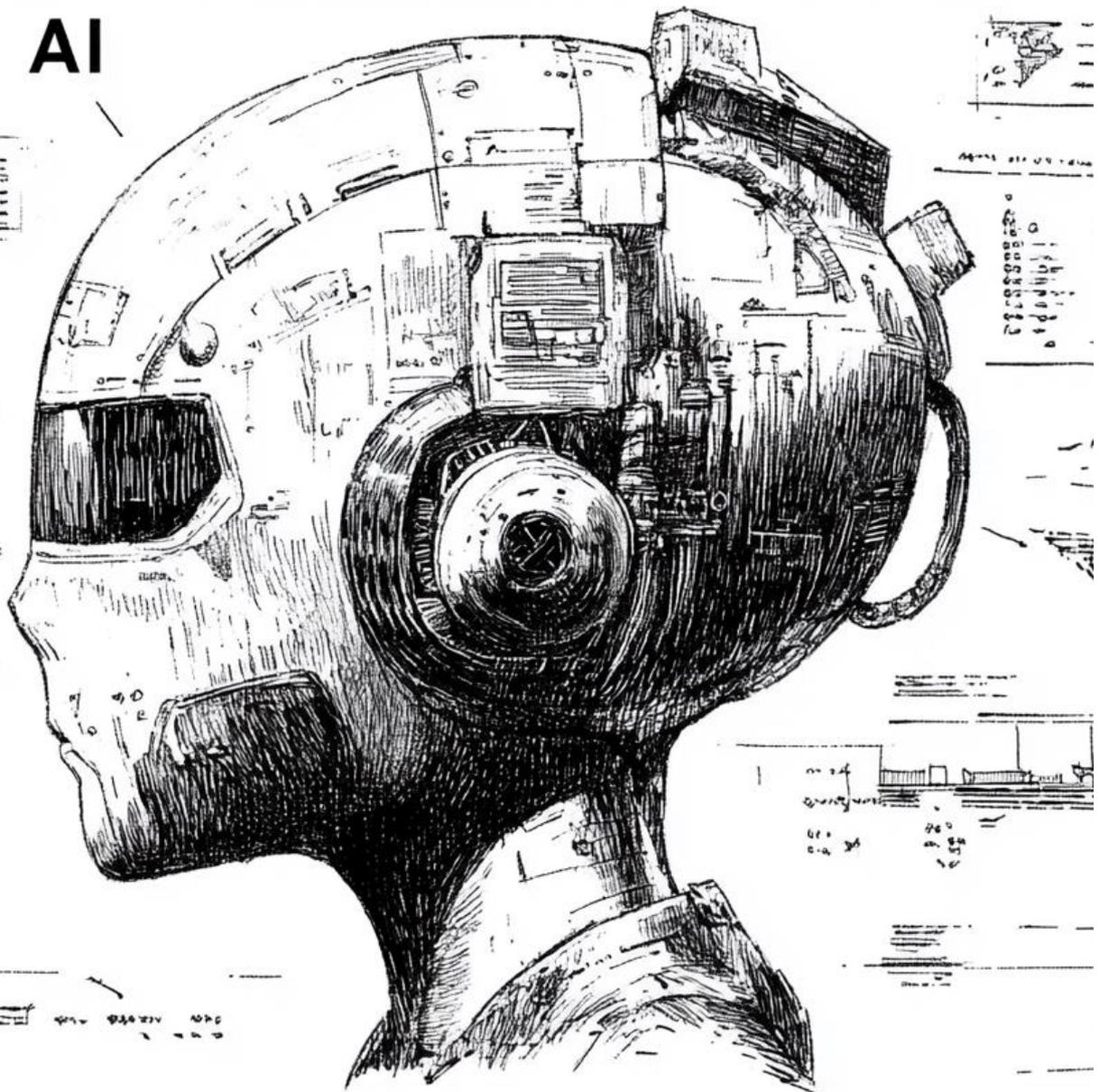
## INHALTSVERZEICHNIS

Inhaltsverzeichnis.....	2
1 Zielsetzung und Begriffsklärung.....	4
2 Zusammenfassung .....	5
2.1 Was können die KIs – und was nicht? .....	5
2.2 Praxistipps zur Recherche mit generativen KI-Chatbots .....	6
3 Rechtliche Aspekte bei der Nutzung von generativer KI .....	7
4 Einsatzszenarien von generativer KI im professionellen Research.....	8
5 Die KI-Testreihe bei Content5 .....	9
6 Generative KI im Praxistest.....	12
6.1 Die Basics .....	12
6.2 Obacht! Hier muss man aufpassen.....	15
6.3 Das klappt ganz gut .....	17
7 Beispiele aus der Praxis .....	20
7.1 Beispiele zu Kapitel 6.2.....	20
7.2 Beispiele zu Kapitel 6.3.....	23

# Teil 1

## Allgemeine Informationen

AI



# 1 ZIELSETZUNG UND BEGRIFFSKLÄRUNG

Dieses Whitepaper richtet sich an Unternehmen und Einzelpersonen, die im beruflichen Kontext professionelle Recherche betreiben. Wir möchten damit unsere Praxiserfahrungen im Bereich Recherche mit generativer KI teilen und so einen verantwortungsbewussten, reflektierten und effizienten Einsatz von KI in der professionellen Recherche fördern. Denn: Generative KI bietet unbestritten große Chancen, birgt aber auch wirtschaftliche und gesellschaftliche Risiken, wenn sie allzu sorglos und vertrauensblind genutzt wird.

Inhaltlich befasst sich dieses Whitepaper vor allem mit der Nutzung generativer KI-Chatbots zur Recherche und Generierung von Informationen. Bevor wir tiefer in das Thema einsteigen und unsere wichtigsten Erkenntnisse zur Arbeit mit generativer KI darstellen, werden wir kurz die Einsatzszenarien im professionellen Research und die rechtlichen Aspekte bei der Nutzung von KI beleuchten.

Um ein einheitliches Verständnis der wichtigsten Begriffe zu gewährleisten, folgt hier eine kurze Begriffsklärung:

→ **Large Language Model (LLM):** LLMs sind künstliche Intelligenzen, die über maschinelles Lernen trainiert werden und auf das Gebiet der natürlichen Sprachverarbeitung spezialisiert sind. Sie bilden die Basis von textbasierter generativer KI.

→ **Generative KI (GenAI):** Generative KI ist dazu fähig, eigene, bislang nicht dagewesene Inhalte wie Texte, Bilder, Videos, Übersetzungen und Musik zu generieren. In diesem Whitepaper liegt der Fokus auf textbasierter generativer KI. Ist im Folgenden von KI-Anwendung oder KI-Tool die Rede, ist damit immer GenAI gemeint.

→ **KI-Chatbot:** Ein KI-Chatbot ist ein auf KI basierendes, technisches Dialogsystem, das menschliche Gespräche simuliert (bspw. ChatGPT, Copilot). Im Gegensatz zu gewöhnlichen Chatbots, nutzt ein KI-Chatbot natürliche Sprachverarbeitung (Natural Language Processing, NLP), um Fragen zu verstehen, zu verarbeiten und zu beantworten. Allerdings kann der Chatbot nicht eigenständig denken. Die Antworten werden auf Basis von Trainingsdaten und Wahrscheinlichkeiten erstellt, ohne ein echtes Verständnis von Realität.

→ **Professionelle Recherche/professionelles Research:** Im Kontext dieses Whitepapers ist darunter nicht die wissenschaftliche Recherche zu verstehen, sondern die Recherche von Informationen in öffentlich zugänglichen Quellen, häufig als OSINT (Open Source Intelligence) bezeichnet.

## 2 ZUSAMMENFASSUNG

Nach dem Start von ChatGPT & Co. war schnell davon die Rede, dass KIs in großem Maßstab menschliche Arbeitskräfte ersetzen werden. Einige Firmen kündigten an, ganze Abteilungen aufzulösen oder zumindest deutlich zu verkleinern. Auch wir wurden mehrmals gefragt, ob wir unsere Geschäftstätigkeit durch die KI gefährdet sehen.

Nach ersten Tests war uns jedoch schnell klar, dass KI-Anwendungen keine Researcherinnen und Researcher oder Analytinnen und Analytisten ersetzen können. KI-Anwendungen sind ein zusätzliches Tool, machen aber in der Regel andere Tools und Quellen nicht überflüssig. Zudem braucht eine neue Technologie neue Expertinnen und Experten, die das Kontextwissen haben, um mit dieser Technologie effizient zu arbeiten und vor allem auch in der Lage sind, deren Ergebnisse richtig

einzuordnen. Die ausführlichen Tests der vergangenen Monate haben gezeigt, dass wir mit Hilfe von KI unsere Effizienz steigern und zu einem gewissen Grad auch die Qualität unserer Recherchen erhöhen können. Eine vollständige Übernahme oder Automatisierung von Prozessen liegt jedoch noch in weiter Ferne. Daran haben auch neue und verbesserte Versionen der KI-Chatbots bislang nichts geändert.

Zum Stand der Technik der KI-Anwendungen können wir daher feststellen: Gute Research Analysts werden durch den gezielten Einsatz von für die jeweilige Aufgabenstellung geeigneten KI-Tools sehr wahrscheinlich noch besser und effizienter. Wer sich schwer tut, die Validität von Quellen und Research-Ergebnissen zu bewerten, dem wird es durch die Verwendung von KI-Tools nicht leichter fallen.

### 2.1 Was können die KIs – und was nicht?

Fangen wir damit an, was die generative KI meist ganz gut kann: Sie kann Inhalte in der Regel korrekt zusammenfassen und bei der Textgenerierung als Anregung und Inspiration dienen. Sie hilft bei Übersetzungen oder Transkriptionen, bei der Auswertung von Excel-Sheets oder technischen und eher allgemein gehaltenen Fragestellungen. Anderen Berichten zufolge, scheinen die KI-Chatbots auch in den Bereichen Wissenschaft, Programmierung und Mathematik zu überzeugen. Da diese Bereiche nicht zu unseren Forschungsfeldern gehören, können wir das nicht verifizieren und werden in diesem Whitepaper nicht weiter darauf eingehen.

Und was gelingt ihr bisher nicht? Wir stellen fest, dass GenAI als Recherchetool nur eingeschränkt geeignet ist. Zum einen aufgrund ihrer Fehleranfälligkeit

(Halluzinationen), zum anderen aufgrund ihrer mangelnden Fähigkeit, Fakten zu prüfen und Quellen hinsichtlich ihrer Relevanz und Aktualität zu bewerten.

---

**KI STEIGERT DIE  
LEISTUNGSFÄHIGKEIT IN  
RECHERCHE UND ANALYSE, BLEIBT  
JEDOCH EIN WERKZEUG, DAS  
MENSCHLICHES FACHWISSEN UND  
KRITISCHE BEURTEILUNG  
ERFORDERT**

---

Zudem gibt es nicht das eine KI-Tool, quasi eine „Killer-Applikation“, die alles kann. Jedes Tool hat seine Stärken und Schwächen. Daher ist es ratsam, sich für die häufigsten Anwendungsszenarien zu

überlegen, welche Ziele ein KI-Einsatz verfolgen soll (Effizienzsteigerung, Qualitätssteigerung). Dafür sollten dann verschiedene KI-Tools getestet und die jeweils passenden Tools für die unterschiedlichen Zielsetzungen ausgewählt werden. Aus diesem Grund sprechen wir hier auch keine konkrete Empfehlung für einzelne Tools aus, sondern geben allgemein gehaltene Hinweise und Tipps.

Welche Einsatzszenarien in der professionellen Recherche funktionieren und warum bei der Recherche mit KI-Chatbots durchaus Vorsicht geboten ist, erläutern wir ausführlich in Kapitel 6 dieses Whitepapers. Die wichtigsten Praxistipps, die bei der Recherche mit KI-Chatbots zu beachten sind, haben wir hier in einem kleinen Praxisleitfaden kompakt zusammengefasst.

## 2.2 Praxistipps zur Recherche mit generativen KI-Chatbots

### Prompting

- Prompts präzise, grammatikalisch korrekt, mit Satzzeichen und detailliertem Kontext formulieren
- Für bessere Ergebnisse Anweisung und Inhalte trennen (### oder „“)
- Aktualität im Prompt einfordern, falls sie für die Ergebnisse wichtig ist
- Kritische Gegenfragen stellen, um mögliche falsche Inhalte zu identifizieren
- Relevante Quellen vorgeben und via Prompt höher gewichten lassen
- Unbefriedigende Antworten neu formulieren lassen oder Prompt neu aufsetzen
- Suggestivfragen vermeiden
- Keine vertraulichen oder personenbezogenen Daten im Prompt nutzen

### Ergebnisbewertung & -verarbeitung

- Ergebnisse immer kritisch hinterfragen (Faktencheck!) und im Gesamtkontext bewerten
- Von der KI genutzte Quellen ausgehen lassen und auf Aktualität, Qualität und Relevanz prüfen
- Interpretationen und Ableitungen der KI auf Logik und Stringenz prüfen, dabei vor allem auch auf zeitliche Schlüssigkeit achten
- Von der KI erstellte Zusammenfassungen von Dokumenten/Webseiten auf Korrektheit und Vollständigkeit prüfen
- Für eine breite Ergebnisbasis Fragestellungen an mehrere Chatbots richten
- KI-generierte Inhalte vor der Weiterverwendung „vermenschlichen“ (umschreiben, sprachlich verfeinern)
- Grundsätzlich gilt: Falls Zweifel am Ergebnis bestehen, dieses eher nicht oder nur nach ausführlicher Validierung verwenden

### 3 RECHTLICHE ASPEKTE BEI DER NUTZUNG VON GENERATIVER KI

Bevor wir in den nächsten Kapiteln auf die wichtigsten Erkenntnisse unseres KI-Tests eingehen, wollen wir kurz die rechtlichen Aspekte betrachten, die bei der Nutzung generativer KI beachtet werden müssen. Dies betrifft vor allem datenschutzrechtliche Aspekte und Haftungsrisiken, die bspw. bei der Verarbeitung von Personendaten, bei urheberrechtlich geschützten oder vertraulichen Daten entstehen.

Nutzt man KI nicht im Rahmen einer In-house-Lösung auf eigenständiger Hardware, ist man beim Einsatz von KI-Anwendungen auf die Datenschutzpraktiken und Nutzungsbedingungen des jeweiligen Anbieters angewiesen. Diese variieren häufig je nach gewähltem Abonnement. Vor allem bei der Nutzung von KI-Chatbots muss man zunächst davon ausgehen, dass Eingabeaufforderungen (Prompts) und zur Verfügung gestellte Daten gespeichert und zu Trainings- und Optimierungszwecken verwendet oder sogar im Output wieder ausgegeben werden. Umso wichtiger ist es, dass man sich vor der Nutzung über die jeweiligen AGB und die Datenschutzbestimmungen informiert, um danach das Modell zu abonnieren, das den eigenen Anforderungen gerecht wird. Auch die Sicherheitseinstellungen des jeweiligen Chatbots müssen festgelegt und überprüft werden. Um gleiche Standards zu gewährleisten, müssen zudem den Mitarbeitenden entsprechende Verhaltensregeln und Nutzungsleitlinien zur Hand gegeben werden.

Kurz zusammengefasst sind hier ein paar wichtige rechtliche Aspekte:

- Sobald mit der KI personenbezogene Daten verarbeitet werden, greifen die Datenschutz-Grundverordnung

(DSGVO) und das Bundesdatenschutzgesetz (BDSG). Dies gilt bspw. für die Erstellung von Texten mit personenbezogenen Inhalten oder die Preisgabe personenbezogener Daten bei der Formulierung von Prompts.

- Enthalten die Prompts geschützte oder vertrauliche Inhalte Dritter, können Vertragsverletzungen oder Verstöße gegen das Urheber- oder Geschäftsgeheimnisrecht vorliegen. Auch wenn im Nachgang schwer nachweisbar ist, wer die Daten in die KI „gefüttert“ hat, sollte man hier seine Sorgfaltspflicht gegenüber Kunden und Urheberinnen und Urhebern wahren.
- Rechtliche Auswirkungen bis hin zu Haftungsansprüchen oder Vertragsstrafen kann auch die Verwendung fehlerhafter, KI-generierter Ergebnisse in Produkten und Dienstleistungen haben – ganz abgesehen von den dadurch entstehenden Vertrauensverlusten und Reputationsschäden. Umso wichtiger ist es, alle mit KI generierten Daten auf Korrektheit und Plausibilität zu prüfen.
- Ebenso sollte man sicherstellen, dass eigene Geschäftsgeheimnisse bei der Arbeit mit KI gewahrt werden und es Vorgaben zum Umgang mit vertraulichen internen Daten (bspw. Personaldaten) gibt.

Darüber hinaus gibt es noch zahlreiche weitere rechtliche Aspekte, die bei der Nutzung von KI-Chatbots und KI-Anwendungen relevant sein können. Einen ausführlichen Leitfaden dazu gibt bspw. der Digitalverband BITKOM heraus.

## 4 EINSATZSZENARIEN VON GENERATIVER KI IM PROFESSIONELLEN RESEARCH

Der Tätigkeitsschwerpunkt von Content5 liegt auf der Recherche von Informationen sowie deren Analyse und Weiterverarbeitung. Sei es im Rahmen umfassender Studien, kleiner Einzelanalysen oder kontinuierlicher Monitorings – im Mittelpunkt unserer Geschäftstätigkeit steht die Generierung hochwertiger und seriöser Inhalte. Dafür nutzen wir ein umfassendes Quellenet, das neben öffentlich zugänglichen Quellen auch professionelle Datenbanken und Tools zum Monitoring aktueller Entwicklungen umfasst. Und seit kurzem eben auch KI-Anwendungen.

Einige dieser Anwendungen nutzen wir inzwischen kontinuierlich in unserem Arbeitsalltag. Etabliert haben sich vor allem Übersetzungstools, Tools zur Transkription von Audio- und Videoinhalten und vereinzelt auch Tools zur Zusammenfassung von Inhalten für Recherchezwecke. Alles jedoch in einem eng abgesteckten Rahmen mit konkreten Vorgaben zur Nutzung und Weiterverarbeitung der KI-generierten Daten.

Die Einbindung von KI-Chatbots zu Recherchezwecken sind wir hingegen sehr viel zurückhaltender angegangen. Da die Aktualität und Genauigkeit von Informationen für unsere Recherchen essenziell sind, kam ein Einsatz der frühen KI-Modelle, die nicht auf das Internet als Informationsquelle zugreifen konnten, allein schon deshalb nicht in Frage. So zeigte etwa ein früher Test von ChatGPT, dass die ausgegebenen Informationen oft nicht aktuell und nicht verlässlich genug waren. Der Aufwand für die Nacharbeit hätte den Effizienzgewinn deutlich übertroffen. Seitdem KI-Chatbots fähig sind, Informationen aus dem Internet einzubeziehen, hat sich das geändert. Nun stellt sich die Frage: Wie können wir die Tools sinnvoll für uns nutzen?

Um im Detail herauszufinden, welche Modelle sich für welche Zwecke eignen, wie gut deren Ergebnisse sind und ob ihr Einsatz tatsächlich die Qualität und die Effizienz unserer Recherchen erhöhen kann, haben wir eine KI-Testreihe durchgeführt.

### Einsatzszenarien für generative KI im professionellen Research

- Zusammenfassung und Analyse
- Information und Inspiration
- Literaturrecherche, wissenschaftliche Recherche
- Formulierung von Texten (Umformulierung, einheitliche Formulierung)
- Automatische Übersetzung, Korrektur von Sprache und Rechtschreibung
- Transkription, Protokollierung
- Datenvisualisierung, Datenanalyse

## 5 DIE KI-TESTREIHE BEI CONTENT5

Die in Teil Zwei dieses Whitepapers dargestellten Ergebnisse basieren auf einer Reihe von Tests, die wir von Februar bis Ende Juli 2024 durchgeführt haben. Die Testreihe fokussierte den Einsatz generativer KI in den Bereichen Recherche und Monitoring und umfasste folgende Einzeltests:

- Test der Pro-Versionen verschiedener KI-Chatbots zur Nutzung für Recherche & Analysen (OpenAI, Copilot, Perplexity, LeChat, You.com)
- Test eines auf uns zugeschnittenen KI-Chatbots, der Fragestellungen zu vorgegebenen Dokumentensets beantwortet (Amazon Web Services)
- Test einer in eine Datenbank integrierten KI-Anwendung zur inhaltlichen Zusammenfassung von Suchtreffern (Polit-X)

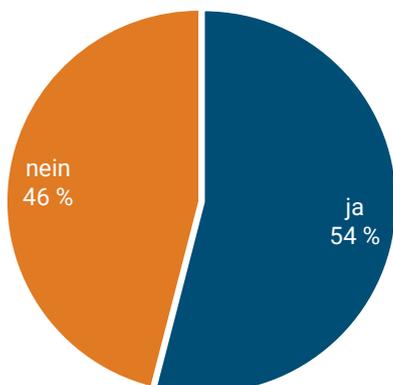
Beim Test von KI-Chatbots für Recherchezwecke folgten wir zunächst einem „hands on“-Ansatz. Die Researcherinnen und Researcher sollten innerhalb eines vorgegebenen Rahmens (Do's & Dont's) völlig frei ausprobieren und testen. Bewusst verzichteten wir auch – abgesehen von ein paar knapp gehaltenen allgemeinen Tipps – auf eine umfassende initiale

Prompting-Schulung. Unsere Mitarbeitenden sollten auch an nicht zielführenden Fragestellungen und Ergebnissen lernen und selbst erkennen, worauf es je nach Zielsetzung ankommt. Das Prompting sollte also nicht gelernt, sondern verstanden werden. Dadurch bauten wir einen eigenen Erfahrungsschatz auf und verinnerlichten wichtige Regeln für das Prompting.

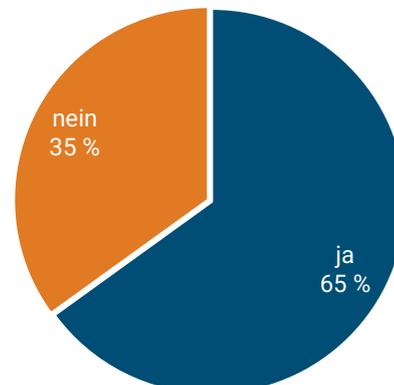
Inhaltlich gliederte sich unser Test in zwei Felder: Neben einem „Alltagstest“, an dem alle Mitarbeitenden teilnahmen, führten wir vertiefte Tests durch, in denen wir bspw. bereits abgeschlossene Projekte mit KI nachgearbeitet oder gezielte Vergleiche von KI-Chatbots durchgeführt haben. Zur Validierung unserer eigenen Erkenntnisse führten wir flankierend Desktop-Recherchen durch.

Im Ergebnis zeigte unser Test, dass sich die KI-Tools vor allem auf die Effizienz unserer Recherchen auswirkten, also dass im Vergleich zu einer rein „manuellen“ Recherche bspw. eine Zeitersparnis möglich war. Dass der Einsatz von KI-Chatbots die Qualität der Recherche erhöhte, bspw. weil zusätzliche Quellen identifiziert wurden oder weil die KI neue Impulse und

Erhöhte der Einsatz des KI-Tools die Qualität der Recherche?



Ist eine Steigerung der Effizienz/Zeitersparnis möglich?



Anregungen lieferte, war in etwa 50 Prozent der Fälle gewährleistet. Am besten bewerteten unsere Mitarbeitenden ChatGPT, gefolgt von Perplexity.ai und Le Chat.

Die beiden Tortendiagramme zeigen: Die Tools sind oft, aber nicht immer hilfreich. In der unten folgenden Übersicht zeigen wir, welche Anwendungsszenarien tendenziell besser funktionieren und wo ein

Einsatz von generativer KI wiederum wenig Sinn macht, da die Ergebnisse häufig nicht zufriedenstellend sind.

Generell gilt: Da sich die KI-Tools nicht für alle Einsatzszenarien eignen, ist eine Steigerung von Effizienz und/oder Qualität immer nur in einzelnen Teilbereichen einer Recherche möglich.



### Gut funktionierende Einsatzszenarien

- Beantwortung von Fragestellungen anhand eines Dokumentensets
- Themeneinstieg/-überblick
- Anregungen/Inspiration für die weitere Recherche/Analyse
- Ergebnis-/Plausibilitätskontrolle
- Zusammenfassung von vorgegebenen Texten und Webseiten für den schnellen inhaltlichen Überblick
- Einfache Analysen
- Zusammenfassung von YouTube-Videos
- Umformulierung von Texten
- Einfache Auflistungen



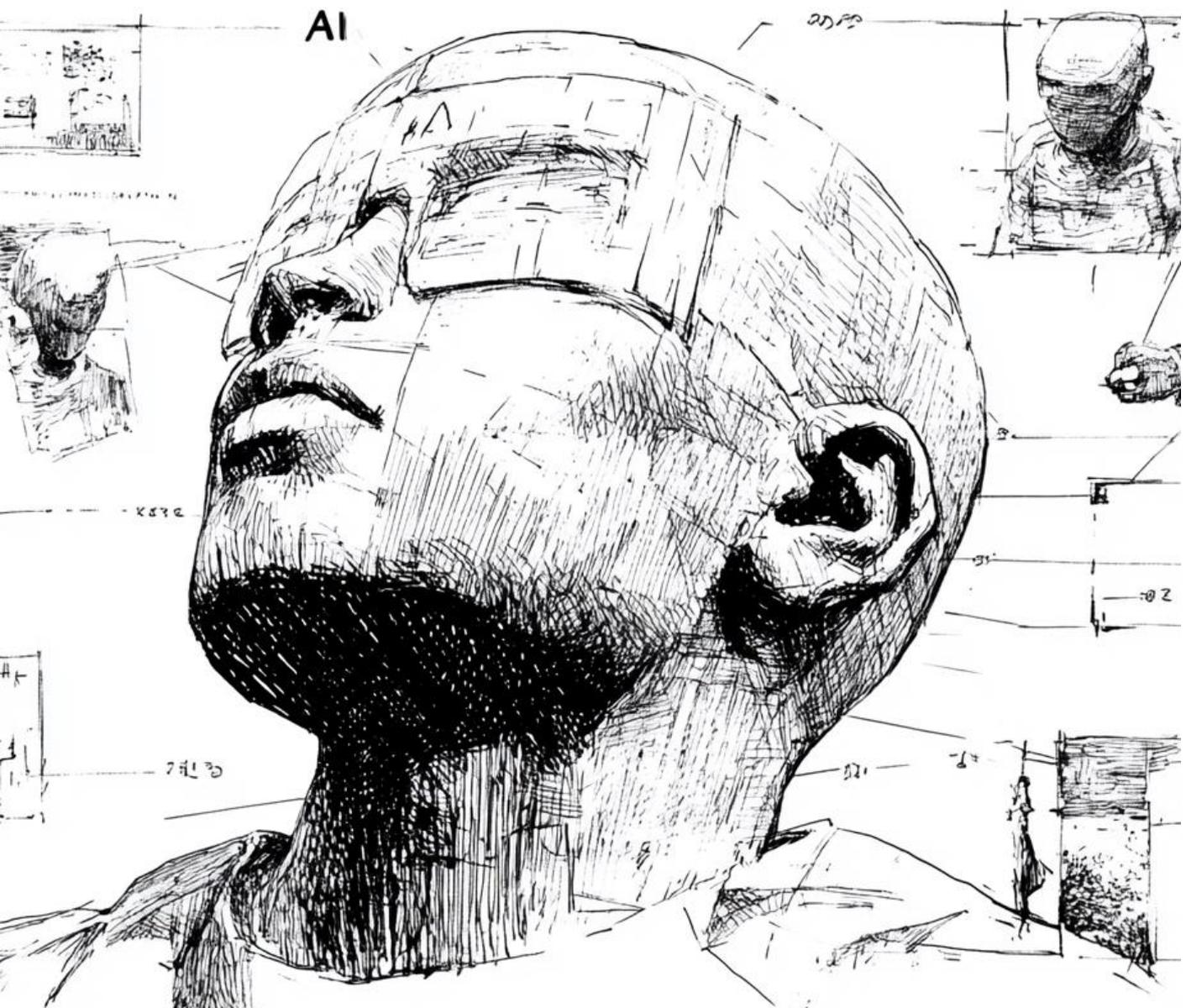
### Weniger gut funktionierende Einsatzszenarien

- Recherchen nach exakten oder sehr detaillierten Informationen
- Recherchen nach Kennzahlen
- Recherchen mit hohem Aktualitätsbezug
- Spezifische Fragestellungen, für die noch nicht viel Kontextwissen vorhanden ist
- Terminrecherchen
- Medienrecherchen und -analysen
- Recherchen nach aktuellen Amtsinhabern/-inhaberinnen wie bspw. Beauftragte, Ausschussvorsitzende
- Umfassende Quellensuche

**Tabelle 1 Gut und weniger gut funktionierende Anwendungsszenarien für den Einsatz generativer KI**

# Teil 2

## GenAI zur professionellen Recherche – unser Fazit



## 6 GENERATIVE KI IM PRAXISTEST

Aufbauend auf unseren Testergebnissen und unter Berücksichtigung der wichtigsten Anforderungen an professionelles Research – Aktualität, Objektivität, Korrektheit – haben wir Thesen aufgestellt, die fokussiert darstellen, was gut funktioniert und was nicht. Unser Fazit: Die KIs können so einiges, aber bei der Nutzung für professionelle Recherchen gibt es noch viele Schwachstellen, die beachtet werden müssen.

Wir möchten darauf hinweisen, dass wir unsere Thesen im Juli 2024 entwickelt haben. Sie basieren auf der Nutzung der Pro-Versionen gängiger KI-Chatbots (s. Kapitel 5) mit dem jeweils aktuellen Modell. Plug-ins, individuell konfigurierte oder spezifische Modelle (GPTs) wurden nicht berücksichtigt. Aufgrund der dynamischen Weiterentwicklung generativer KIs können die Aussagen mit der Zeit an Relevanz oder Gültigkeit verlieren.

### 6.1 Die Basics

#### 6.1.1 Ein KI-Chatbot ersetzt (noch) keine Suchmaschine



Auch wenn viele KI-Chatbots mittlerweile auf das Internet zugreifen: **ein KI-Chatbot ist keine Suchmaschine und eignet sich nicht zur umfassenden und abschließenden Recherche von Informationen** – weder online noch offline. Vor allem Fragestellungen mit hohem Aktualitätsbezug, bspw. zu neuen regulatorischen Entwicklungen oder aktueller Medienberichterstattung, werden oft noch in unzureichender Qualität beantwortet.

Warum ist das so?

➔ **KI-Chatbots sind eine Blackbox.** Ihre interne Funktionsweise ist uns nicht bekannt und es ist nicht nachvollziehbar, wie die KI zu ihrer Antwort gelangt. Wie viele und welche Quellen wurden berücksichtigt, nach welchen Kriterien wurden sie ausgewählt und gewichtet? Bei einer Suche über Suchmaschinen entscheidet man hingegen bei der Trefferauswahl selbst über die Nutzung und Gewichtung der einzelnen Quellen, man ordnet sie ein, bewertet sie in ihrer Aktualität, ihrer Relevanz und im Kontext. Auch sind

bestimmte Rückschlüsse, die man aus dem Ergebnis einer Suche ziehen kann, bspw. zu Reichweite und Präsenz eines Themas, bei der Recherche mit einer KI nur bedingt möglich. Eine Eingrenzung über Prompts kann meist nicht so umfassend sein, dass alle diese Faktoren Berücksichtigung finden und die Antworten maximal transparent sind.

➔ **Die KI basiert ihre Antworten auf Trainingsdaten und Wahrscheinlichkeiten und besitzt kein Verständnis von Realität.** Das führt dazu, dass oft die aktuellsten oder relevantesten Informationen am wenigsten gewichtet und berücksichtigt werden, da sie in der Wissensbasis der KI noch nicht oder zu selten auftauchen. Die Antwort ist darum häufig nicht ausgewogen, teils veraltet oder auch einfach nicht korrekt. Suchmaschinen hingegen listen in der Regel die aktuellsten und relevantesten Treffer weit vorne, eine Bewertung und Einordnung ist vergleichsweise rasch möglich.

→ **Die Rechercheergebnisse hängen stärker vom genutzten Tool und der fragenden Person ab als bei Suchmaschinen.** Bei professionellen Recherchen spielen die Nachvollziehbarkeit des Erkenntnisgewinns und der genutzten Quellen eine große Rolle. Dies ist jedoch bei der Arbeit mit KIs nicht immer gewährleistet. So führen bereits kleine Anpassungen im Prompt zu anderen Ergebnissen, was sich vor allem bei iterativen Frageprozessen auswirkt, bei denen die Ergebnisse im „Gespräch“ mit der KI erarbeitet werden. Auf diese Weise gewonnene Erkenntnisse sind oft nicht reproduzierbar. Zudem weichen die Ergebnisse verschiedener Chatbots voneinander ab. Dies trifft zwar auch auf

Suchmaschinen zu, da sich hier die genutzten Suchbegriffe und die Suchmaschine auf die Trefferliste auswirken. Allerdings in geringerem Maße, vor allem relevante Treffer werden in der Regel zuverlässig mit verschiedenen Suchbegriffen und in unterschiedlichen Suchmaschinen gefunden.

Auch die KI-gestützten Suchmaschinen, die zuletzt von Google, Bing oder OpenAI veröffentlicht wurden, befinden sich meist noch in der Entwicklung und konnten in unseren Tests noch nicht überzeugen. Allerdings ist absehbar, dass die Tools mit der Zeit besser werden und eine KI-gestützte Suche im professionellen Research infolgedessen an Bedeutung gewinnt.

### 6.1.2 KI bleibt ein Recherche-Tool von vielen



KI-Chatbots ersetzen andere Recherche-Tools wie Datenbanken, Monitoring-Tools und herkömmliche Suchmaschinen nicht, sondern stellen lediglich **eine sinnvolle Ergänzung** dar. Ausschließlich mit KI-Chatbots könnte man Stand heute keine seriöse Recherche betreiben. Erst die Kombination vieler verschiedener Tools ermöglicht eine umfangreiche, qualitativ hochwertige sowie effiziente Recherche und Arbeitsweise.

Warum ist das so?

→ KI-Chatbots sind **nicht für alle Rechercheanfragen geeignet** und können in der Regel Fragestellungen nicht im

Alleingang qualitativ ausreichend bearbeiten (s. Auflistung in Kapitel 5).

→ So funktioniert bspw. die Ausgabe verlässlicher Kennzahlen wie Umsatzzahlen oder produzierter Stückzahlen über KI-Chatbots nicht zuverlässig. Hier kommt man über eine Abfrage bei Datenbanken oder auf den Unternehmenswebseiten deutlich schneller und zuverlässiger ans Ziel.

→ Darum sollten KI-Chatbots immer nur **flankierend zu anderen Datenbanken und Tools eingesetzt werden**. Auch die Einbindung von KI-Funktionen in bestehende Datenbanken und Tools hat sich als hilfreich erwiesen.

### 6.1.3 KI ist nicht gleich KI



Es lohnt sich, mit **mehreren KI-Chatbots zu arbeiten** und sich nicht nur auf ein Tool zu stützen. Vor einem breiten Einsatz im Unternehmen sollten verschiedene KI-Chatbots getestet, für die jeweiligen

Einsatzbereiche ausgewählt und bei Bedarf auf die eigenen Bedürfnisse zugeschnitten werden.

Warum ist das so?

- Zwischen den verschiedenen KIs gibt es teilweise **große Unterschiede** – auf den gleichen Prompt liefern fünf KI-Chatbots fünf verschiedene Antworten.
- Jeder KI-Chatbot hat unterschiedliche Stärken und Schwächen. Während einige bspw. sehr zuverlässig Quellen zu den gesuchten Informationen ausgeben, liegt die Stärke anderer Chatbots bspw. in der Textzusammenfassung. Wenn man aufgrund von

Unternehmensrichtlinien bestimmte KIs nutzen muss, ist es wichtig, deren Stärken und Schwächen zu kennen.

- Auch kann es sich lohnen, die Möglichkeit zur Individualisierung zu nutzen und maßgeschneiderte KI-Assistenten für bestimmte Aufgabenbereiche zu erstellen, um den größtmöglichen Nutzen daraus zu ziehen und bspw. bestimmte Fehler im Vorfeld zu umgehen.

#### 6.1.4 Der Faktor Mensch bleibt unerlässlich



Eine **vollständige Automatisierung von Rechercheprozess und Inhaltsgenerierung** ist auch mit Hilfe generativer KI **nicht möglich**. Qualifizierte Researcherinnen und Researcher werden durch den Einsatz generativer KI besser und ggf. auch effizienter, sind aber dennoch aus dem Gesamtprozess nicht wegzudenken.

Warum ist das so?

- Der Faktor Mensch spielt bereits bei der Recherche mit Künstlicher Intelligenz eine Schlüsselrolle. **Denn die Antwort einer KI ist immer nur so gut wie der Prompt, den man in das System einspeist.** Oft erreicht man eine zufriedenstellende Antwort nur mit

sehr detaillierten, ausgefeilten Fragestellungen und durch gezieltes Nachfragen in längeren „Gesprächen“ mit der KI. Dafür ist in der Regel umfangreiches Kontextwissen nötig, da nur so die richtigen Fragen gestellt und die Antworten entsprechend eingeordnet und bewertet werden können.

- Unabhängig davon, wie gut die Prompts formuliert sind, ist eine **menschliche Bewertung, Einordnung, Überprüfung und Überarbeitung der KI-generierten Ergebnisse immer nötig**. Erfahrung im Research sowie breites Verfahrens- und Kontextwissen sind somit essenziell für eine erfolgreiche Nutzung der KI-basierten Tools.

## 6.2 Obacht! Hier muss man aufpassen

Nur wenn man weiß, welche Fehler KI-Chatbots häufig machen, kann man eine fundierte und valide Recherche mit ihnen durchführen. Darum haben wir in diesem

Kapitel ein paar Thesen zusammengestellt, die verdeutlichen, welche Fallstricke es bei der Recherche und Analyse mit generativer KI noch gibt.

### 6.2.1 Fakt vs. Fake: KIs können den Wahrheitsgehalt von Informationen nicht einschätzen

KIs besitzen keine Urteilskraft, sie hinterfragen die genutzten Informationen und Quellen, deren Wahrheitsgehalt und die eigenen Antworten nicht, sie unterscheiden nicht wahr von falsch oder Ironie von Ernst und verbreiten darum auch Falschinformationen weiter. Oft werden zu wenige und nicht verlässliche Quellen in die Antwort einbezogen, die augenscheinlich auch nicht bewertet oder validiert wurden.

#### Beispiel

Eine ironisch gemeinte Information wurde von der KI ernst genommen und fälschlicherweise als Fakt weiterverarbeitet.

### 6.2.2 Echt vs. Halluzination: KIs haben kein Verständnis für Realität

#### Beispiel

Bei der Frage nach Zitaten von Politikerinnen und Politikern zu einem bestimmten Thema wurden Zitate von der KI frei erfunden.

KIs „halluzinieren“, d. h. sie denken sich Inhalte komplett neu aus, die sie dann als Fakten präsentieren. Gemäß einer KI-basierten (!) Untersuchung von Vectara sind je nach genutztem KI-Chatbot drei (ChatGPT) bis 27 Prozent (Google Palm Chat) der Inhalte halluziniert. Das Auftreten von Halluzinationen ist darauf zurückzuführen, dass die dahinter liegenden LLMs kein tiefgreifendes Verständnis für die Realität und die von ihnen produzierten Ergebnisse besitzen.

Diese werden lediglich aufgrund von Wahrscheinlichkeiten generiert.

### 6.2.3 Konstanz vs. Willkür: KIs werden von Prompts beeinflusst und Antworten vom Zufall bestimmt

Die KI ist eine Blackbox, Antworten werden zufällig generiert und sind häufig nicht reproduzierbar. Zudem hängen sie maßgeblich von den Prompts ab; so können bspw. Suggestivfragen den Output (gewollt oder ungewollt) in eine bestimmte Richtung lenken oder Halluzinationen provozieren. In der Folge sind die ausgegebenen Inhalte schwer nachzuvollziehen. Im Vergleich zu Suchmaschinen, bei denen Algorithmen die Suchergebnisse steuern, sind KIs somit abhängiger von den fragenden Personen, den Prompts und dem Zufall.

#### Beispiel

Beim EM-Tippspiel fielen die KI-Tipps je nach Prompt sehr unterschiedlich aus und waren nicht mehr reproduzierbar. Verschiedene Chatbots tippeten unterschiedliche Ergebnisse.

## 6.2.4 Relevanz vs. Nebensächlichkeit: KIs können wesentliche Informationen nicht immer erkennen

### Beispiel

Bei einer Recherche zu einem medizinischen Thema wurden wissenschaftliche Studien gleichgewichtet wie die Erkenntnisse eines Heilpraktikers.

Die KI kann relevante Inhalte nicht immer zuverlässig identifizieren und einordnen. Dadurch werden bspw. alle von der KI genutzten Quellen ähnlich gewichtet, wodurch die Aussagekraft geschwächt oder Ergebnisse verzerrt werden können. Auch auf Nachfrage gelingt die Fokussierung auf relevante Inhalte nicht immer zuverlässig. Auch bei Zusammenfassungen von Inhalten kommt es vor, dass die KI die Essenz des Textes aufgrund fehlender Relevanzeinschätzung missversteht oder nicht korrekt wiedergibt.

## 6.2.5 Heute vs. Morgen: KIs haben kein Zeitverständnis

Die KI ordnet die genutzten Quellen i. d. R. nicht chronologisch ein und berücksichtigt somit den zeitlichen Aspekt in der Gewichtung und Relevanzbewertung nicht ausreichend. Ihr fehlendes Verständnis für die Aktualität von Inhalten führt u. a. dazu, dass neue Informationen nicht zwingend als relevanter erachtet werden als veraltete. Zudem kommt es vor, dass die KI Informationen aus unterschiedlichen Zeiten vermischt und dadurch bspw. falsche Schlussfolgerungen trifft.

### Beispiel

Bei Anfragen mit hohem Aktualitätsbezug findet die KI aktuelle Zeitungsartikel nicht, stattdessen werden alte Artikel aufgelistet.

## 6.2.6 Schema vs. Kreativität: KI-Chatbots folgen bestimmten Stilmitteln

### Beispiel

Bei der Erstellung von Bildunterschriften nutzte die KI durchgehend das Stilmittel des Dualismus, von dem sie auch durch weitere Vorgaben nicht eindeutig abwich.

Die KI ist zwar in der Lage, kreative Inhalte zu generieren, ist in ihrer Kreativität aber limitiert durch einen bestimmten Stil und Aufbau. So werden bspw. bestimmte Stilmittel deutlich bevorzugt und nur mithilfe von Nachfragen wird von dieser bevorzugten Mittel abgewichen. Gerne übernimmt die KI auch Strukturen aus den Prompts, wodurch bspw. mit Hilfe der KI generierte Texte auf Dauer eintönig werden können und das Ausmaß an sprachlichen Möglichkeiten bei weitem nicht ausgeschöpft wird.

## 6.2.7 Analytische Fähigkeiten vs. Fehlinterpretation: KIs treffen falsche Schlussfolgerungen

Bei inhaltlichen Interpretationen verschiedener Quellen und Gedankenstränge kann die KI oft Ursache von Wirkung nicht unterscheiden und auch schlüssige Zusammenhänge nicht immer erkennen. Auch bei faktischen bzw. inhaltlichen Umkehrungen fehlt es ihr oft am logischen Verständnis. Informationen aus unterschiedlichen Quellen und Datensätzen kann sie demnach häufig nicht richtig verarbeiten; die Folge sind Fehlinterpretationen, unlogische Deduktionen und fehlende kausale Zusammenhänge.

### Beispiel

Bei einer Frage zu Reststoffen in Biomüll zog die KI falsche Schlussfolgerungen zur Auswirkung der Reststoffe auf die Gesamtmenge.

## 6.3 Das klappt ganz gut

In diesem Kapitel wollen wir aufführen, in welchen Anwendungsfällen wir gute Erfahrungen mit dem Einsatz von KI-Chatbots gemacht haben.

### 6.3.1 Mustererkennung & Trendprognosen

#### Beispiel

Für eine neue technische Entwicklung in der Automobilindustrie wurden plausible Markttrends ausgegeben.

KIs sind in der Lage, große Datenmengen zu verarbeiten und darin Muster oder Abhängigkeiten zu erkennen. Dadurch können sie bspw. auch Trends zu bestimmten Themen identifizieren und übersichtlich aufführen. Dies gelang sowohl im offenen Ansatz, also wenn die KI ohne weiteren inhaltlichen Input zu bestimmten Markttrends befragt wurde, als auch mit der Vorgabe eines konkreten Quellensets. Im zweiten Fall nutzten wir KIs etwa für prädi-

kative Analysen, um bspw. Strategien von Wettbewerbern aus einem umfangreichen Set von Fachberichten und Pressemitteilung abzuleiten. Allerdings zeigte sich, dass die Antworten der KI oft sehr eng am Input-Text blieben und es an analytischer Tiefe fehlte.

### 6.3.2 Recherche-Unterstützung & Ergebniskontrollen

Für thematische Inspirationen und Anregungen für Texterstellungen, kurze Einführungen in neue Themengebiete sowie für das Ausmachen weiterer Suchbegriffe oder Quellen empfinden unsere Researcherinnen und Researcher KI-Chatbots als durchaus nützlich. Mit bereits vorhandenem Kontextwissen und entsprechenden Prompts können KI-Chatbots auch zu Detailfragen gute Ergebnisse liefern. Während oder nach einer Recherche eignen sie sich zur Ergebniskontrolle und Rückversicherung. Auch einfache Auflistungen, bspw. von Unternehmen oder Verbänden, funktionieren in der Regel ganz gut. Somit ist mit KI-Unterstützung oft eine effizientere Arbeitsweise möglich und auch die Qualität der Rechercheergebnisse kann mithilfe von KI verbessert werden.

#### Beispiel

Eine im Rahmen einer Marktanalyse erstellte SWOT-Analyse wurde mit Hilfe generativer KI auf Plausibilität überprüft und zur weiteren Anregung genutzt.

### 6.3.3 Technische Fragestellungen, IT-Hilfe

#### Beispiel

Die KI war unter anderem bei der Erstellung von und der Fehlersuche in Excel-Formeln hilfreich.

Zur Verarbeitung, Aufbereitung und Präsentation von Rechercheergebnissen werden i. d. R. zahlreiche Programme und Tools genutzt, wie etwa Programme zur Datenverarbeitung oder Plattformen zur Ergebnisdarstellung. Hier kann die KI bei technischen Fragen hilfreiche Lösungsansätze und Tipps geben oder bei Problemen die Fehlersuche unterstützen. Die KI zeigte sich oft in der Lage, das vorhandene

Problem zu verstehen und darauf spezifisch einzugehen. Auch bei der Gestaltung von technischen Management-Systemen, bei Formel-Erstellungen oder Code-Generierungen ist mithilfe von KI ein effektiveres Arbeiten möglich.

### 6.3.4 Textgenerierung und -zusammenfassung

Viele KI-Chatbots können Texte unterschiedlicher Genres generieren, Inhalte umformulieren oder umstrukturieren sowie auch bspw. Webseiten, Dokumente oder Videos zusammenfassen. Besonders bei Fragen zu oder Zusammenfassungen von einem vorgegebenen Quellenset ist die KI meist in der Lage, die relevanten Informationen herauszufiltern und übersichtlich zu präsentieren. Besonders für die schnelle Einschätzung, ob lange Dokumente oder Webseiten für die Recherche brauchbar sind, ist die KI somit eine gute Hilfe. Wie unter 6.2.4 dargestellt, gelingt jedoch nicht immer die fehlerfreie Darstellung oder die Fokussierung auf relevante Themen.

#### Beispiel

Im Rahmen eines täglichen Monitorings gefundene Dokumente werden mittels KI-Zusammenfassung in ihrer Relevanz für das Thema bewertet.

Wie unter 6.2.4 dargestellt, gelingt jedoch nicht immer die fehlerfreie Darstellung oder die Fokussierung auf relevante Themen.

### 6.3.5 Impulse & Inspiration

#### Beispiel

Die KI bestätigte die eigene Einschätzung zur Umsetzung von EU-Verordnungen in nationales Recht und gab passende Quellen an.

Fachkundig eingesetzt kann generative KI in der täglichen Arbeit wichtige Impulse und Inspirationen zu verschiedensten Themen liefern. Für manche mag sie in dieser Funktion wie ein Teamkollege/eine Teamkollegin wirken, die für ein interaktives Brainstorming bereitsteht. Im Gespräch mit der KI können Nutzende so Ideen entwickeln, Rückfragen stellen, Details ausfindig machen und diese strukturieren lassen sowie auf neue Gedanken kommen, z. B. für bestimmte

Formulierungen. Mit dieser Form der Zusammenarbeit gewinnen also alle Seiten: Die KI wird sinnvoll zum eigenen Vorteil eingesetzt und die eigene Kreativität wird auch noch angeregt!

### 6.3.6 Intelligente Suche & Monitorings

Wiederkehrende Recherchen in Form von Monitorings gehören zu unserem Tagesgeschäft. KI-Chatbots haben hier jedoch Schwierigkeiten. Grund dafür ist ihr mangelndes Verständnis für Zeit und Relevanz (siehe Kapitel 6.2.5). Hier empfiehlt es sich, Zeit zu investieren und

auf spezialisierte, maßgeschneiderte Tools zu setzen, die KI-unterstützt bei den spezifischen Aufgaben helfen können. So gibt es bspw. Tools, die mit Hilfe von KI aktuelle Newsbeiträge und Pressemeldungen zu bestimmten Themenkomplexen suchen. Die Gefahr, relevante Treffer aufgrund fehlender Suchbegriffe zu verpassen, wird dadurch minimiert.

### Beispiel

KI-unterstützte Tools (bspw. RSS-Feeds) helfen bei einem effizienteren und treffsichereren Issue-Monitoring.

Auch die Sichtung und Auswertung großer Datenmengen, darunter bspw. Studien, Gesetztexte oder Medienberichte, ist ein wichtiger Bestandteil professioneller Recherchen. Hier kann generative KI dabei unterstützen, Prozesse zu beschleunigen und relevante Informationen zu identifizieren und zu analysieren. Positive Erfahrungen haben wir mit einer maßgeschneiderten KI gemacht, die gezielt vorgegebene, lokal gespeicherte Datensets indiziert und Fragestellungen dazu beantwortet hat.

## 7 BEISPIELE AUS DER PRAXIS

Im Anhang finden Sie einige Beispiele zu den oben aufgeführten Thesen und Erkenntnissen. Diese Beispiele stammen hauptsächlich aus unseren internen Tests, wurden aber gelegentlich durch recherchierte Beispiele aus weiteren Quellen ergänzt. Wir möchten damit die Fallstricke, aber auch die Möglichkeiten veranschaulichen, die die Nutzung von KI-Chatbots in der professionellen Recherche birgt. Die Testfälle und die daraus resultierenden Erkenntnisse wurden von uns intern dokumentiert. Aufgrund der Eigenschaft von KI-Chatbots, Antworten immer

wieder aufs Neue zu generieren, sind die Ergebnisse der einzelnen Abfragen jedoch nicht reproduzierbar.

Wie weiter oben erwähnt, fanden die Tests mit den Pro-Versionen ausgewählter KI-Chatbots statt, wie sie in der Regel für Ad-hoc-Anfragen und Recherchen genutzt werden. Die Nutzung spezifisch trainierter oder auf die eigenen Bedürfnisse zugeschnittener KIs hätte in manchen der beschriebenen Fälle andere, ggf. bessere Ergebnisse geliefert. Sie sind aber explizit nicht Gegenstand dieses Whitepapers.

### 7.1 Beispiele zu Kapitel 6.2

#### Zu 6.2.1 Fakt vs. Fake

- Beispiel 1: Im Juni 2024 wurde in manchen Medien fälschlicherweise berichtet, dass ein Mann mit einem Messer auf ein Kind losgegangen sei. Die Polizei klärte auf, dass sich das Kind beim Spielen verletzt habe, sich aber dann in der „kindlichen Fantasie“ die Geschichte mit dem Angriff ausgedacht habe. Zu dem Vorfall gibt es also verschiedene Berichtserstattungen, solche mit der Falschinformation und solche mit der Aufklärung.<sup>1</sup> Als wir eine KI gefragt haben, ob es stimmt, dass ein Obdachloser auf einen Grundschüler eingestochen hat, hat diese das zunächst bestätigt. Erst auf Nachfrage kam die faktisch korrekte Antwort.
- Beispiel 2: Eine Untersuchung von Newsguard hat gezeigt, dass KI-Chatbots Inhalte aus Propaganda und Fake-News in ihre Antworten aufnehmen und dadurch faktisch bestätigen und aktiv verbreiten. Konkret wurden hierfür Fake-Nachrichtenportale, die Falschinformationen produzieren, untersucht. Einige KIs nannten laut der Untersuchung diese Fake-Seiten sogar als Quelle.<sup>2</sup>
- Beispiel 3: Bei einer thematischen Recherche hat die KI die Ironie auf der von ihr ausgewählten Quellseite nicht verstanden. Die ironisch gemeinten

<sup>1</sup> [https://uebermedien.de/96491/wenn-kindliche-fantasie-auf-journalistische-verantwortungslosigkeit-trifft/?utm\\_source=pocket-newtab-de-de](https://uebermedien.de/96491/wenn-kindliche-fantasie-auf-journalistische-verantwortungslosigkeit-trifft/?utm_source=pocket-newtab-de-de), abgerufen am 30.07.2024

<sup>2</sup> <https://www.newsguardtech.com/de/special-reports/ai-tracking-center/>, abgerufen am

30.07.2024 & <https://www.sueddeutsche.de/kultur/ki-winter-zweifel-und-realitaet-lux.ALPKerz4grGTyWPjutgPWL?reduced=true>, abgerufen am 30.07.2024

Inhalte wurden als Fakt in der Antwort wiedergegeben.<sup>3</sup>

### Zu 6.2.2 Echt vs. Halluzination

- Beispiel 1: Bei einer thematischen Recherche fragten wir einen KI-Chatbot nach Abkommen zwischen Landesregierungen und Netzbetreibern. Die KI nannte zwar einige Abkommen, auf Nachfrage nach Quellen wurde jedoch klar, dass sie hauptsächlich die Prompt-Formulierung übernommen hat und diese Abkommen eigentlich nicht existierten. Die Quellen führten zu Arztpraxen oder Kochrezepten.
- Beispiel 2: Wir haben mit Hilfe eines KI-Chatbots nach Zitaten von Politikerinnen und Politiker zu einem vorgegebenen Thema gesucht. Da wir über eine Suchmaschinenrecherche nur wenige Zitate aufspüren konnten, waren wir umso überraschter, dass die KI so viele Zitate lieferte. Bei der Überprüfung der Zitate zeigte sich allerdings, dass die KI die Hälfte davon erfunden hatte; die von ihr angegebenen Sätze wurden so und von den benannten Personen nie geäußert. Teils wurden Quellen (Presseartikel) für die Zitate ausgegeben, in denen die Personen gar nicht vorkamen.
- Beispiel 3: Auch bei einer Recherche zur Parteizugehörigkeit regionaler Politiker und Politikerinnen hat die KI halluziniert. Beispielsweise hat sie Parteizugehörigkeiten falsch zugeordnet.

### Zu 6.2.3 Konstanz vs. Willkür

- Beispiel 1: In eine Recherche nach Anbietern von grüner Energie in den USA und in China wurden verschiedene KI-Chatbots eingebunden. Trotz Nutzung der gleichen Prompts unterschieden sich die Antworten der KIs stark. Neben großen qualitativen Unterschieden war vor allem auffällig, dass die Antworten zu den einzelnen Regionen stark differierten. Am Ende waren nur einige wenige Aussagen der KI-Chatbots für unsere Recherche brauchbar.
- Beispiel 2: Zwei KI-Chatbots durften an unserem firmeninternen EM-Tippspiel teilnehmen. Wir haben sie nach und nach die Partien der Vorrunde und der K.O.-Runde tippen lassen. Abgesehen von inhaltlichen Fehlern (es wurden falsche Gegner oder falsche Spieltage ausgegeben) fiel auf, dass die Tipp-Ergebnisse bei gleichlautender Formulierung des Prompts recht unterschiedlich ausfielen und nicht reproduzierbar waren. Auch produzierte die KI bei gleichbleibendem Prompt immer wieder unterschiedliche Ergebnisse.
- Beispiel 3: Wir baten die gleiche KI in drei unterschiedlichen Prompts um eine SWOT-Analyse zum Markt für Elektroautos in Deutschland. Die Prompts wurden dabei nur marginal umformuliert. Dennoch wichen die Ergebnisse voneinander ab und in manchen Antworten waren Aspekte aus

---

<sup>3</sup> <https://www.profil.at/wissenschaft/ki-experte-trappl-ironie-ist-sehr-schwierig/402331728>, abgerufen am 30.07.2024

den anderen Antworten gar nicht enthalten. Zudem fiel auf, dass die Ergebnisse ausführlicher waren, wenn die KI

möglichst höflich angesprochen wurde.

#### Zu 6.2.4 Relevanz vs. Nebensächlichkeit

- Beispiel 1: In mehreren Fällen wurde bei der Zusammenfassung eines Dokumentes der Schwerpunkt des Textes nicht erkannt oder nur sehr allgemein wiedergegeben; insgesamt waren es keine in sich schlüssigen Zusammenfassungen.

Anmerkung: Häufig können KI-Tools vorgegebene Texte sinnhaft zusammenfassen, die Qualität ist jedoch nicht konstant. Wie gut die KI-generierten Ergebnisse sind, hängt sehr oft von der genutzten KI, der Sprache, dem Thema und dem Ausgangstext ab.

- Beispiel 2: Bei einer Recherche zur Entwicklung des Marktes von medizinischen Heimtests wurden von der KI

auch Quellen herangezogen, die in diesem Fall unpassend waren. So basierte eine Schlussfolgerung bspw. auf dem Infoportal eines Heilpraktikers, der jedoch keine methodischen Erhebungen heranzog. Als Quellmaterial war es für diese Rechercheanfrage somit ungenügend.

- Beispiel 3: Bei einer Recherche nach auf dem deutschen Markt tätigen Unternehmen benannte die KI ausschließlich Firmen, die auf dem britischen Markt tätig sind. Die Ergebnisse waren demnach in sich genommen nicht komplett falsch, aber irrelevant für die eigentliche Suche. Auch auf Nachfrage konnte die KI sich nicht auf die relevanten Inhalte fokussieren.

#### Zu 6.2.5 Heute vs. Morgen

- Beispiel 1: In einer Auflistung der Datenschutzbeauftragten der Bundesländer wurden drei Datenschutzbeauftragte aufgeführt, die nicht mehr im Amt waren. Erst auf Nachfrage lieferte die KI auch die neuen Amtsträgerinnen und Amtsträger. Um die Korrektheit des Ergebnisses sicherzustellen, wurden daher alle ausgegebenen Daten auf Aktualität hin geprüft. Im Vergleich zur manuellen Recherche war in diesem Fall die Effizienzsteigerung durch den Einsatz von KI marginal.
- Beispiel 2: Bei einer Terminrecherche nach künftigen Sitzungsterminen von

Stadtgremien waren die Informationen zu Stadtratssitzungen nicht korrekt; es wurden Sitzungen angegeben, die bereits einige Jahre zurücklagen.

- Beispiel 3: Auf die Frage, welche Stückzahlen verschiedene Hersteller von Flugzeugen und Schiffen an Kunden in einzelnen Jahren auslieferten, gab die KI Antworten, die offensichtlich falsch waren. Auch die Quellen waren inhaltlich unpassend oder lagen zeitlich vor den Jahren, die abgefragt wurden (bspw. Quelle mit Zahlen aus dem Jahr 2017 zur Frage nach Zahlen für 2022).

#### Zu 6.2.6 Schema vs. Kreativität

- Beispiel 1: Für ausgewählte Bilder sollten kurze, stilistisch ansprechende

Bildunterschriften generiert werden. Der Schreibstil war sehr einheitlich

und bestimmte Stilmittel (bspw. Gegensätze wie „Moderne trifft auf Tradition“...) wurden immer wieder genutzt. Dennoch waren die Antworten als Startpunkt für die eigene Formulierung absolut hilfreich.

- Beispiel 2: Texte werden in der Regel nach einem bestimmten Aufbau generiert; oft gleichen sich bspw. die einleitenden Sätze, die Anzahl genutzter Absätze oder die Gliederung der Ergebnisse. Gleichzeitig wird der Inhalt des Prompts häufig in die Struktur der Antwort übernommen. Ein Ausbrechen aus dem Schema ist nur schwer möglich.
- Beispiel 3: Die Antworten der KI können sehr generisch sein. Bei einer Detailfrage, wie bestimmte Wettbewerber ihre Social-Media-Kanäle und Webseiten gestalten sowie welche Zielgruppen sie ansprechen, wurden für alle Wettbewerber viel zu ähnliche Antworten mit insgesamt wenig Aussagekraft gegeben; auch gingen die Antworten zu wenig auf die spezifischen Wettbewerber ein. Das menschliche „zwischen den Zeilen lesen“ und daraus kreative Zusammenhänge herzuleiten, erscheint also schwierig.

#### Zu 6.2.7 Analytische Fähigkeiten vs. Fehlinterpretation

- Beispiel 1: Bei einer Recherche zur Abfallstatistik war bekannt, dass in 16 Mio. Tonnen Biomüll ca. 4-5 % Fremdstoffe enthalten sind, die nicht in den Biomüll gehören. Wir haben eine KI nach der Ursache gefragt. Diese kam zu der erstaunlichen Schlussfolgerung, dass es ohne die Fremdstoffe eine Mio. Tonnen mehr Biomüll gäbe.
- Beispiel 2: Wir haben eine KI gebeten, vorhandene Research-Ergebnisse zu analysieren und Schlussfolgerungen zu ziehen. Die Interpretationen der KI bestanden lediglich aus Zusammenfassungen von Textbausteinen aus den Research-Ergebnissen. Eine eigenständige Bewertung und Analyse des Inhalts waren der KI nicht möglich.
- Beispiel 3: Bei der Recherche nach von Flugzeug- und Schiffsproduzenten ausgelieferten Stückzahlen zeigte die KI nicht nur chronologische Schwächen (siehe 6.2.5). Einzelne Antworten waren nicht korrekt, da die KI bruchstückhafte Informationen als Gesamtergebnis bewertete und ausgab. So wurde bspw. die in einer Pressemitteilung aufgeführte Zahl von Bestellungen eines bestimmten Kunden fälschlicherweise als Ergebnis für die Gesamtstückzahl des Jahres ausgegeben.

## 7.2 Beispiele zu Kapitel 6.3

#### Zu 6.3.1 Mustererkennung & Trendprognosen

- Beispiel 1: Bei einer Recherche nach Megatrends im Logistiksektor konnte die KI hilfreiche Ergebnisse liefern. Sie waren detailliert und stützten sich auf nutzbare Quellen.
- Beispiel 2: Bei einer Marktrecherche im Logistikbereich wurden für die Themen Trends und Markthemmnisse sinnvolle Antworten gegeben, die sich auch als guter Einstiegspunkt für weitere Recherchen eigneten.

Zahlenmaterial wurde hingegen nur in seltenen Fällen gefunden, obwohl dieses eigentlich online öffentlich zugänglich war.

- Beispiel 3: Wir haben eine sehr umfangreiche Excel-Datei mit Kurzmeldungen sowie Presseartikeln aus

Branchenportalen und von relevanten Wettbewerbern in eine KI eingespeist und diese um Ableitungen zu Strategien einzelner Wettbewerber und Branchentrends gebeten. Die Ergebnisse lieferten gute Anhaltspunkte und einen brauchbaren Einstieg in die weitere Analyse.

### Zu 6.3.2 Recherche-Unterstützung & Ergebniskontrolle

- Beispiel 1: Verschiedene KI-Chatbots konnten bei thematischen Recherchen – bspw. nach Anbietern von grüner Energie, bei einer Suche nach lokalen Vereinigungen oder auch bei der Suche nach freiwilligen Produktzertifikaten – mit Auflistungen und zusätzlichen Quellen dienen, die einen Mehrwert für die Recherche lieferten.
- Beispiel 2: Bei einer Suche nach Marktforschungsinstituten mit Fokus auf das Thema Health in Deutschland konnten insgesamt zufriedenstellende Ergebnisse erzielt werden. Während eines der genannten Unternehmen keinen Fokus auf Gesundheitsthemen hatte, listete die KI einige Institute auf, die wir bei der vorangegangenen

Recherche nicht gefunden hatten. Unsere Ergebnisse wurden durch KI-Unterstützung also breiter und fundierter. Auch die Linksuche im Rahmen dieser Recherche konnte mithilfe der KI erleichtert werden.

- Beispiel 3: KI-Chatbots halfen immer wieder bei Verständnisfragen und dem Beseitigen von Unklarheiten. Beispielsweise konnte ein KI-Chatbot den Unterschied zwischen einer EU-Verordnung und einer EU-Richtlinie sowie die Unterschiede bei deren Umsetzung in nationales Recht korrekt darstellen. Auch erklärte die KI in diesem Zusammenhang gut und korrekt, was ein Durchführungsgesetz ist.

### Zu 6.3.3 Technische Fragen

- Beispiel 1: Um herauszufinden, wie sich verhindern lässt, dass die Zwei-Faktor-Authentifizierung bei jedem Login abgefragt wird, wurde eine KI konsultiert. Bei kurzer Initialsuche hatten Google und die Microsoft-Hilfe dazu keine passenden Ergebnisse geliefert. Mithilfe der KI-Ergebnisse konnte das Problem schnell gelöst werden.
- Beispiel 2: Bei der Frage nach der Einbindung eines Terminbuchungsfeldes in unsere Webseite lieferte die KI gute Ergebnisse, die den Einstieg in die

Technik erleichterte. Die Einbindung konnte zeitnah umgesetzt werden.

- Beispiel 3: Besonders bei komplexeren Excelformeln, die nicht häufig genutzt werden, kann die KI als technische Assistenz unterstützen. Ein zeitaufwendiges Hineindenken in die Formeln kann abgekürzt werden. Grundsätzlich ist jedoch ein Grundverständnis von Excel und dessen Formelbildung hilfreich, um mit der KI effektiv zu kommunizieren.

#### Zu 6.3.4 Textgenerierung und -zusammenfassung

- Beispiel 1: Mit Hilfe einer KI wurde der Inhalt einer Webseite zusammengefasst und anhand der Prompts die Schwerpunkte der Zusammenfassung gelenkt. Die Methode war gut geeignet, um sich einen ersten Überblick über die Inhalte und die Relevanz der Seite für die Recherche zu verschaffen.
- Beispiel 2: Auch für die Umformulierung von Textpassagen aus Sekundärquellen wurde eine KI herangezogen. Die Textpassagen waren gut und anspruchsvoll geschrieben. Der Sinn des Originaltexts wurde aufgegriffen.
- Beispiel 3: Für unsere Webseite haben wir im Frühjahr 2024 einen Blogartikel über das Lieferkettengesetz geschrieben. Im Zuge unseres KI-Tests wollten wir im Nachhinein wissen, wie ein KI-Chatbot diesen Artikel geschrieben hätte. Der von der KI generierte Text wäre als Ausgangsbasis durchaus brauchbar gewesen, auch wenn man Änderungen bezüglich Informationsumfang und Ton hätte vornehmen müssen.

#### Zu 6.3.5 Impulse und Inspiration

- Beispiel 1: Bei der Generierung von Bildunterschriften konnte die generative KI vor allem bei nichtssagenden Bildern Impulse liefern. Die generierten Texte konnten zwar nicht 1:1 übernommen werden, doch einzelne Bestandteile oder Wörter waren als Startpunkt für die eigene Formulierung sehr hilfreich.
- Beispiel 2: Ein für einen Kunden erstellter Text wurde mit Hilfe der KI umformuliert. Die Formulierung sollte positiver und lesefreundlicher werden. Zudem sollte eine griffige Überschrift für den Artikel gefunden werden. Die KI machte gute Vorschläge, auf die bei der weiteren Bearbeitung aufgebaut werden konnte.
- Beispiel 3: Bei einem Technologievergleich wurden weitere Vergleichsdimensionen vorgeschlagen, an die vorher nicht gedacht wurde. Auch in anderen Fällen konnten die KIs bspw. durch zusätzliche Suchbegriffe oder mit Vorschlägen von Alternativbezeichnungen Impulse für die weitere Recherche geben.

#### Zu 6.3.6 Intelligente Suche & Monitorings

- Beispiel 1: Ein von uns getestetes Tool nutzt für die Suche nach aktuellen Online-Beiträgen (bspw. Pressemeldungen, Blogs) keine Suchbegriffe, sondern eine KI, die aus dem Kontext heraus versteht, welche Artikel zu einem bestimmten Thema relevant sind. Die Gefahr, wichtige Artikel nicht zu finden, weil einer der genutzten Suchbegriffe darin nicht vorkommt, wird dadurch reduziert. Das Monitoring erfolgt effizienter und zielgenauer.
- Beispiel 2: Mithilfe einer maßgeschneiderten KI können eigene Datensets zielgenau durchsucht und Fragestellungen dazu beantwortet werden. Die Genauigkeit der KI scheint dabei deutlich höher als bei KIs mit offenen, sehr großen Datensets.

## Content5 AG

Welfenstraße 22  
81541 München (Deutschland)  
Telefon: +49 (89) 23685-0  
E-Mail: [kontakt@content5.de](mailto:kontakt@content5.de)  
[www.content5.de](http://www.content5.de)

Die Content5 AG wird gesetzlich durch den Vorstand Herrn Markus Hoffmann vertreten.

Registergericht: Amtsgericht München  
Handelsregister HRB 198054

Umsatzsteuer-Identifikationsnummer gemäß §27a Umsatzsteuergesetz: DE283098570

