Einfach, weil's wichtig ist.

ERGO

Von SEO zu KI-Suchoptimierung

Erkenntnisse aus der Versicherungswirtschaft mit Empfehlungen für den Versicherungsmarkt und andere Branchen

Wie große Sprachmodelle digitale Sichtbarkeit neu definieren

ERGO

Das **ERGO Innovation Lab** arbeitet an der Spitze neuer Technologien und Services rund um Versicherung, Risiko und Finanzen. Am Merantix AI Campus in Berlin angesiedelt, hinterfragt das Team kontinuierlich den Status quo der Versicherungswirtschaft. Dabei entstehen richtungsweisende Projekte, geprägt von technologischen Entwicklungen und digitalen Trends.

ECODYNAMICS

Die Expert:innen von **ECODYNAMICS** beraten und qualifizieren Unternehmen im Umgang mit generativer KI und digitalen Geschäftsmodellen. Das Unternehmen ist seit 2021 Teil des OpenAI Early Adopter Programms und spezialisiert auf die Einführung generativer KI, den Einsatz KI-gestützter Agentensysteme sowie die Analyse und Optimierung von Webseiten für KI-gestützte Suchumgebungen.

„Bis 2026 wird das Volumen klassischer Suchanfragen um 25 Prozent zurückgehen. Das Suchmaschinenmarketing wird Marktanteile an KI-gestützte Suche und Agenten verlieren."

Gartner Inc.

Die Art, wie Menschen online nach Informationen suchen, verändert sich grundlegend. Die klassische Stichwortsuche ist zwar noch weit verbreitet, wird jedoch zunehmend von fortschrittlicheren, gesprächsbasierten Ansätzen abgelöst. Grundlage dafür sind große Sprachmodelle, die auf künstlicher Intelligenz beruhen.

Diese Systeme liefern nicht einfach Listen mit Links. Sie erkennen die Absicht hinter einer Frage, geben präzise und verständliche Antworten und verfeinern ihre Ergebnisse mit aktuellen Daten. Für die Versicherungswirtschaft ist diese Entwicklung besonders relevant, denn dort spielen Klarheit, Zugänglichkeit und Vertrauen eine zentrale Rolle für die Entscheidung von Kund:innen.

Menschen erwarten heute schnelle, personalisierte und zutreffende Informationen. KI-gestützte Suchsysteme können diese Anforderungen besser erfüllen als herkömmliche Suchmaschinen. Sie ersparen Nutzer:innen, sich durch mehrere Seiten mit Treffern zu klicken, und liefern stattdessen direkt passende, kontextbezogene Antworten. So entsteht ein deutlich vereinfachter Zugang zu Produkten und Dienstleistungen aus dem Versicherungsumfeld. Wie jede technologische Neuerung bringt auch dieser Wandel neue Herausforderungen mit sich. Digitale Inhalte wurden über Jahre hinweg vor allem auf die Logik klassischer Suchmaschinen zugeschnitten. Mit dem Aufkommen KI-gestützter Abrufsysteme stehen Versicherer:innen vor neuen Anforderungen.

Inhalte müssen heute nicht nur für Rankings sichtbar sein, sondern auch von Sprachmodellen referenziert und als vertrauenswürdig eingestuft werden. Das verändert den Maßstab für Sichtbarkeit. Gleichzeitig entstehen Risiken durch verzerrte oder fehlerhafte Antworten. Es braucht daher neue Strategien für die Struktur, Qualität und Auffindbarkeit von Inhalten in einer sich dynamisch entwickelnden Suchumgebung.

Das ERGO Innovation Lab hat gemeinsam mit ECODYNAMICS diese Veränderungen genau untersucht und ihre Auswirkungen auf die Versicherungsbranche analysiert.

Dieses Whitepaper bietet eine umfassende Einordnung des aktuellen Stands KI-gestützter Suche und zeigt, welche Rolle sie künftig für Informationsbeschaffung, Kundenkontakt und Entscheidungsfindung spielt.

Unsere Analyse basiert auf mehr als 33.000 ausgewerteten Suchtreffern und über 600 untersuchten Webseiten. Sie zeigt, wie Sprachmodelle die Suche nach Versicherungsangeboten verändern und wie Nutzer:innen mit diesen Angeboten interagieren. Wir betrachten, welche strukturellen und technologischen Anpassungen als Reaktion auf diese Entwicklung notwendig sind. Darüber hinaus analysieren wir die weiterreichenden Auswirkungen auf die Branche und benennen zentrale Trends, die die Zukunft KI-gestützter Suche prägen. Auf dieser Grundlage geben wir Empfehlungen, wie sich der Wandel strategisch nutzen lässt.

Damit Versicherer in diesem veränderten Umfeld sichtbar bleiben, müssen sie ihr digitales Vorgehen neu ausrichten. Inhalte sollten so gestaltet sein, dass sie mit der Logik KI-gestützter Systeme kompatibel sind. Nur dann werden sie zuverlässig erkannt, verarbeitet und in Antworten eingebunden. Andernfalls gewinnen andere Anbieter:innen an Sichtbarkeit – etwa Makler:innen, deren Inhalte klar strukturiert, maschinell lesbar und gut vernetzt sind.

Diese Entwicklung wird durch den zunehmenden Einsatz agentischer Systeme zusätzlich beschleunigt. Nutzer:innen stellen ihre Fragen direkt in einem System, erhalten personalisierte Antworten, vergleichen Angebote und treffen Entscheidungen im laufenden Dialog. Ganze Recherchen, Produktvergleiche und Vertragsabschlüsse finden in einer durchgehenden Gesprächssituation statt. Klassische Webseiten treten dabei in den Hintergrund. Wer mit seinen Inhalten in diesen neuen Abläufen nicht vorkommt, verliert den Zugang zu Kund:innen.

Inhaltsverzeichnis

Wie KI-gestützte Suche klassische SEO-Strategien verdrängt

KI-Tools wie ChatGPT verändern die Art und Weise, wie Menschen nach Informationen suchen, indem sie Linklisten durch direkte, dialogorientierte Antworten ersetzen. Dieser Wandel verändert das Nutzerverhalten, stellt traditionelle Suchmaschinen vor Herausforderungen und schafft neue Chancen und Risiken für Unternehmen, insbesondere in Branchen wie der Versicherungswirtschaft.

Die Entwicklung KI-gestützter Suche

Die Integration von Suchtechnologie in große Sprachmodelle hat grundlegend verändert, wie Nutzer:innen mit Informationen umgehen. Immer mehr Menschen wenden sich Systemen wie ChatGPT oder Perplexity zu, um gezielte Antworten zu erhalten, statt klassische Listen mit Links zu durchforsten. Klassische, schlüsselwortbasierte Suchmaschinen wie Google haben sich zu KI-gestützten Modellen weiterentwickelt, die Kontexte verstehen, direkte Antworten erzeugen und gesprächsähnliche Sucherfahrungen ermöglichen. Diese Systeme liefern keine einfachen Links und Textausschnitte mehr, sondern formulieren zusammengefasste Antworten und vereinfachen so den Suchprozess. Eine Schlüsseltechnologie hinter dieser Entwicklung ist Retrieval Augmented Generation (RAG). Dabei greifen Sprachmodelle in Echtzeit auf relevante externe Informationen zu, um die Genauigkeit und Relevanz ihrer Antworten zu verbessern.

Laut Gartner, einem führenden Marktforschungsunternehmen, wird das Suchvolumen klassischer Suchmaschinen bis 2026 um 25 Prozent zurückgehen, da KI-gestützte Dialogsysteme verstärkt genutzt werden. Google bleibt mit einem Marktanteil von 89 Prozent zwar führend (StatCounter, zitiert in Statista, 2025), ergänzt seine Suchergebnisse aber bereits um Google AI Overview, eine KI erzeugte Zusammenfassung, die direkt in den Ergebnissen angezeigt wird. Damit entsteht das sogenannte Zero Click Verhalten, bei dem Nutzer:innen ihre Antworten direkt in der Suchoberfläche erhalten und nicht mehr auf externe Seiten klicken müssen. Viele Publisher und Unternehmen, die bisher stark auf organischen Suchverkehr angewiesen waren, überdenken deshalb ihre Strategien zur Sichtbarkeit in einer zunehmend KI gesteuerten Suchlandschaft.

Gleichzeitig entstehen Alternativen wie Perplexity AI, You.com, Google Gemini und SearchGPT. Sie liefern direkte Antworten ohne klassische Trefferlisten und wurden gezielt für den Bedarf an gesprächsbasierten, direkten Informationen entwickelt. Perplexity AI gewinnt zunehmend Nutzer:innen als Alternative zu herkömmlichen Suchmaschinen, während Google die Funktionen seines Gemini Modells ausbaut und KI-gestützte Suche mit den AI Overviews innerhalb des Google Ökosystems kombiniert (Eisenbrand, 2025). Aktuelle Daten von Seer Interactive (2025) zeigen, dass Anfragen mit AI Overviews einen deutlichen Rückgang bei der organischen Klickrate verzeichnen – von 1,41 Prozent auf 0,64 Prozent im Jahresvergleich. Bezahlte Klickraten gingen insgesamt zurück, unabhängig davon, ob ein AI Overview angezeigt wurde. Marken, die in AI Overviews erscheinen, profitieren jedoch: Die organische Klickrate steigt dort von 0,74 auf 1,02 Prozent, die bezahlte Klickrate von 7,89 auf 11 Prozent. Das unterstreicht die wachsende Bedeutung von Sichtbarkeit innerhalb eines zunehmend klickfreien Suchumfelds.

Mit dem Wandel des Suchverhaltens hin zu KI-generierten Antworten müssen Unternehmen ihre Inhalte für klassische Suchmaschinen und zugleich für Anfragen in sprachmodellbasierten Systemen optimieren, um sichtbar zu bleiben. Das ist besonders wichtig für Branchen wie die Versicherungswirtschaft, in denen Kund:innen schnelle, präzise und personalisierte Informationen erwarten. Wer das neue Suchverhalten frühzeitig berücksichtigt, kann die eigene Sichtbarkeit in KI erzeugten Ergebnissen deutlich verbessern. Das stärkt die Wettbewerbsposition und eröffnet neue Potenziale für digitale Kund:innen-Ansprache, Beratung, Vertrieb und Service – von interaktiven, gesprächsbasierten Informationen bis hin zu personalisierten Produktempfehlungen und Vergleichsübersichten.

Durch den Vergleich der Ergebnisse sprachmodellbasierter Suche mit denen klassischer Suchmaschinen wollen wir die Faktoren identifizieren, die Sichtbarkeit und Platzierung in diesen neuen KI-gestützten Systemen bestimmen. Im Fokus steht die Frage, wie Unternehmen ihre Präsenz in KI erzeugten Suchergebnissen so gestalten können, dass sie in einem sich wandelnden Suchumfeld relevant bleiben.

Herausforderungen KI-gestützter Suche

Trotz der Vorteile KI unterstützter Suchsysteme bestehen weiterhin erhebliche Herausforderungen. Eine der zentralen Schwächen betrifft die Genauigkeit. Sprachmodelle erzeugen gelegentlich falsche oder irreführende Informationen, die mit hoher Überzeugung präsentiert werden. Dieses Phänomen ist als Halluzination bekannt (Zhang et al., 2023). Um dem entgegenzuwirken, setzen viele Unternehmen inzwischen auf Echtzeitverweise zu Quellen im Netz und entwickeln fortgeschrittene Mechanismen zur Faktenprüfung.

Auch rechtliche und urheberrechtliche Fragestellungen sorgen für anhaltende Diskussionen (Karamolegkou et al., 2023). Nachrichtenverlage und Inhaltsanbieter:innen äußern Bedenken, dass KI-Systeme ihre Inhalte ohne Zustimmung auslesen und zusammenfassen, ohne die Quelle korrekt anzugeben (Hagey et al., 2023). Technologieunternehmen geraten deshalb zunehmend unter Druck, sich mit den ethischen und rechtlichen Folgen ihrer Datenverwendung auseinanderzusetzen.

Ein weiteres zentrales Thema bleibt die Transparenz von Antworten und der Umgang mit Verzerrungen. Da Sprachmodelle auf umfangreichen Datensätzen trainiert werden, übernehmen sie unbeabsichtigt auch bestehende Vorurteile aus den zugrunde liegenden Informationen (Gallegos et al., 2024). Die Weiterentwicklung dieser Systeme zielt daher unter anderem darauf, Trainingsdaten zu diversifizieren und technische Verfahren zu verbessern, mit denen sich Antworten nachvollziehbarer gestalten lassen.

Potenziale KI-gestützter Suche für die Versicherungswirtschaft

Die Versicherungswirtschaft ist auf präzise und konsistente Kommunikation angewiesen – etwa in Vertragsunterlagen, bei Schadenprozessen oder im Kundenservice. Zwar sind Produkte, rechtliche Rahmenbedingungen und regulatorische Vorgaben oft komplex, doch bessere Zugänglichkeit und individuelle Ansprache können entscheidend dazu beitragen, Transparenz zu erhöhen, Vertrauen zu stärken und fundierte Entscheidungen zu ermöglichen.

KI-gestützte Suche verändert den Zugang zu Versicherungsinformationen grundlegend. An die Stelle statischer Suchanfragen treten dynamische, interaktive Gespräche (Mo et al., 2024). Nutzer:innen können Rückfragen stellen, Details klären und Antworten erhalten, ohne sich direkt an eine Ansprechperson wenden zu müssen. Gerade digital affine Zielgruppen, die auf Selbstbedienung setzen, schätzen dieses Prinzip. Allerdings bevorzugen viele Nutzer:innen bei komplexeren oder langfristigen Versicherungsprodukten weiterhin das sogenannte ROPO-Verhalten (Research Online, Purchase Offline). Sie informieren sich online, schließen den Vertrag aber offline ab – in der persönlichen Beratung mit Makler:innen (GDV, 2023).

LLM-Systeme unterstützen diese Entwicklung, indem sie kontextbezogene und personalisierte Antworten liefern. Sie vereinfachen Fachsprache, verkürzen die Recherchezeit und übersetzen rechtliche Begriffe in verständliche Sprache. Das erhöht sowohl das Vertrauen als auch das Verständnis auf Kund:innenseite.

Ein besonderer Vorteil von KI-gestützter Suche liegt in der Möglichkeit, Produktvergleiche in Gesprächsform zu erstellen, und zwar außerhalb klassischer Vergleichstools von Maklerplattformen. Nutzer:innen können Vergleiche heute selbst steuern, indem sie freie Fragen stellen und gezielt nachfragen. Das ermöglicht flexiblere, individuellere und kontextbezogene Gegenüberstellungen, die sich an den konkreten Bedarf anpassen und dies innerhalb eines dialogbasierten Ablaufs.

Darüber hinaus unterstützen diese Systeme bei der individuellen Beratung. Sucht jemand beispielsweise nach „der besten Reiseversicherung für Familien mit kleinen Kindern“, berücksichtigen die Antworten gezielt Anforderungen wie Kinderabsicherung, familienfreundliche Zusatzleistungen oder medizinische Versorgung im Ausland. Gleichzeitig bleibt der Bezug zu Rückfragen erhalten, sodass auch im Verlauf des Gesprächs die Relevanz erhalten bleibt.

Entscheidend ist dabei die richtige Balance zwischen Innovation und Vertrauenssicherung. Zwar bestehen weiterhin Herausforderungen wie Halluzinationen, Verzerrungen oder rechtliche Unsicherheiten. Wer jedoch darauf wartet, bis diese Systeme vollständig ausgereift sind, riskiert, den Anschluss zu verlieren. Auch wenn Versicherer nicht kontrollieren können, wie externe Sprachmodelle Informationen darstellen oder formulieren, können sie die Qualität dieser Antworten entscheidend beeinflussen. Voraussetzung dafür ist, dass ihre Inhalte korrekt, klar strukturiert, aktuell und maschinell lesbar bereitgestellt werden.

Voraussetzungen für klassische Suchsysteme und SEO

Um eine optimale Sichtbarkeit in klassischen Suchsystemen wie Google, Bing oder Yahoo zu erreichen, müssen Webseiten eine Reihe technischer, struktureller und inhaltlicher Anforderungen erfüllen. Diese verbessern sowohl die Auffindbarkeit als auch das Ranking auf den Ergebnisseiten der Suchsysteme (Google Search Central, 2023; Iqbal et al., 2022).

Klassische Suchmaschinenoptimierung bleibt weiterhin unverzichtbar. Systeme, die auf großen Sprachmodellen beruhen, sind auf etablierte Suchsysteme wie Google angewiesen, um Webseiten zu durchsuchen und zu indexieren (Iqbal et al., 2022). Wenn technische und strukturelle SEO-Standards nicht eingehalten werden, sind Inhalte möglicherweise überhaupt nicht auffindbar.

Der folgende Abschnitt beschreibt die wesentlichen Voraussetzungen für eine wirksame Suchmaschinenoptimierung.

Indexierbarkeit und Crawling-Freundlichkeit

Damit eine Webseite effektiv von Suchsystemen erfasst und indexiert werden kann, müssen bestimmte technische Anforderungen erfüllt sein:

- **Indexierbare Seiten:** Alle wichtigen Seiten dürfen keine „noindex“-Tags oder Einschränkungen in der robots.txt enthalten. Crawler benötigen eine explizite Freigabe, um auf Inhalte zugreifen und diese einordnen zu können (Google Search Central, 2023).
- **Klare Sitemap und interne Verlinkung:** Eine gut strukturierte Sitemap und logische interne Links ermöglichen es Suchsystemen, die Seite zu navigieren und sämtliche relevante Inhalte zu entdecken (Iqbal et al., 2022).
- **Mobile First Optimierung:** Da Suchsysteme die mobile Nutzbarkeit priorisieren, müssen Webseiten responsiv gestaltet oder durch mobile Layouts ergänzt werden, um auf allen Geräten ein konsistentes Nutzungserlebnis zu bieten (Google Search Central, 2023).
- **Optimierung der Ladegeschwindigkeit:** Schnell ladende Seiten verbessern die Nutzererfahrung und fördern bessere Platzierungen. Entscheidend sind optimierte Bilder, minimiertes CSS und JavaScript sowie leistungsfähige Server (Terenteva, 2023).
- **Vermeidung JavaScript basierter Inhalte:** Zentrale Inhalte sollten ohne JavaScript Rendering zugänglich sein, da Crawler JavaScript Elemente oft nicht vollständig erfassen können (Google Search Central, 2023).

Auch wenn diese technischen Anforderungen die Zugänglichkeit und Indexierbarkeit sichern, sind Seitenstruktur und Inhalt für den Erfolg im Ranking ebenso wichtig.

Neben der technischen Erreichbarkeit sorgt eine gute Onpage-Optimierung dafür, dass Inhalte sowohl für Suchsysteme als auch für Nutzer:innen verständlich, relevant und klar strukturiert sind. Die folgenden bewährten Praktiken helfen dabei, Struktur, Bedeutung und Zweck von Inhalten effektiv zu vermitteln (Moz, 2024; Iqbal et al., 2022).

Klassische Onpage-SEO-Faktoren

Wirksame Onpage-Optimierung bedeutet, Inhalte und Relevanz gegenüber Suchsystemen und Nutzer:innen klar zu kommunizieren. Zentrale Maßnahmen sind:

- **Eindeutige Titel und Metabeschreibungen:** Jede Seite sollte präzise und ansprechende Metadaten enthalten, die den Inhalt korrekt widerspiegeln und die Klickrate erhöhen (Moz, 2024).
- **Strukturierte Überschriften (H1, H2, H3):** Eine klare Überschriftenhierarchie verbessert die Lesbarkeit und hilft Suchsystemen, Themen und Abschnitte gezielt zu erfassen (Iqbal et al., 2022).
- **Saubere URL-Struktur:** URLs sollten kurz, beschreibend und frei von unnötigen Parametern sein. Gut strukturierte URLs erhöhen die Crawl-Effizienz und das Vertrauen von Nutzer:innen (Terenteva, 2023).
- **Bildbeschreibungen (Alt-Texte):** Um die Zugänglichkeit und Relevanz von Inhalten zu verbessern, sollte jedes Bild einen beschreibenden Alt-Text enthalten. So können Suchsysteme visuelle Elemente einordnen (Google Search Central, 2023).
- **Keyword-Optimierung:** Relevante Suchbegriffe müssen sinnvoll und organisch in Titel, Überschriften und Inhalt eingebunden sein. Übermäßige Wiederholungen (Keyword Stuffing) sollten vermieden werden, um Lesbarkeit und Vertrauen zu sichern (Moz, 2024).

Diese Onpage-Strategien bilden die Grundlage für starke SEO-Ergebnisse. Dennoch bleibt hochwertiger und nutzerzentrierter Inhalt ebenso entscheidend, um langfristig wettbewerbsfähig zu bleiben. Technische und strukturelle Maßnahmen sorgen dafür, dass Inhalte überhaupt zugänglich sind – doch erst durch Qualität und Relevanz entsteht echte Sichtbarkeit. Inhalte, die sich an den tatsächlichen Bedürfnissen der Nutzer:innen orientieren, sind die Voraussetzung für nachhaltigen Erfolg im digitalen Raum (Google Search Central, 2023; Moz, 2024).

Inhaltsqualität und Nutzererlebnis

In der klassischen Suchmaschinenoptimierung bleibt hochwertiger Inhalt der zentrale Erfolgsfaktor. Eine gute Nutzererfahrung und Inhalte, die exakt zur Suchintention passen, sind entscheidend für langfristige Sichtbarkeit:

- **Einzigartige und hochwertige Inhalte:** Inhalte müssen einen eigenständigen Mehrwert bieten. Duplikate sollten vermieden und eine klare Abgrenzung zu Wettbewerber:innen gewährleistet sein (Google Search Central, 2023).
- **Relevante und gut strukturierte Texte:** Eine klare Gliederung, kurze Absätze, Aufzählungen und ein intuitives Layout verbessern sowohl die Lesbarkeit als auch das Nutzererlebnis (Moz, 2024).
- **Abstimmung auf die Suchintention:** Inhalte sollten gezielt auf den jeweiligen Zweck der Nutzer:innen ausgerichtet sein – ob informierend, navigierend oder transaktionsbezogen. Nur so lassen sich Relevanz und Sichtbarkeit erreichen (Iqbal et al., 2022).
- **Regelmäßige Aktualisierungen:** Aktualisierte Inhalte senden ein Signal für Autorität und Relevanz. Seiten, die regelmäßig gepflegt werden, erzielen bessere Platzierungen und stärken das Vertrauen (Google Search Central, 2023).
- **Signale für Nutzer:innenbindung:** Lange Verweildauer und geringe Absprungraten deuten auf hohe Qualität und Relevanz hin. Seiten sollten deshalb so gestaltet sein, dass sie zur weiteren Erkundung einladen und Interaktion fördern (Terenteva, 2023).

Guter Content und ein positives Nutzungserlebnis hängen eng mit Vertrauen und Autorität zusammen – zwei Faktoren, die zunehmend durch externe Signale beeinflusst werden. Im nächsten Abschnitt geht es darum, wie die externe Reputation die SEO-Performance beeinflusst.

Neben dem, was auf der eigenen Seite passiert, stützen sich Suchsysteme stark auf Signale aus dem weiteren Netz, um Glaubwürdigkeit und Relevanz einer Marke einzuschätzen. Hier kommt die Offpage-Optimierung ins Spiel, also der Teil der SEO, der sich mit der Wahrnehmung und Verlinkung durch Dritte beschäftigt.

Offpage-SEO und Vertrauenswürdigkeit

Neben technischen und inhaltlichen Aspekten haben auch externe Faktoren erheblichen Einfluss auf Sichtbarkeit und Autorität in der klassischen Suche:

- **Hochwertige Backlinks:** Verlinkungen von vertrauenswürdigen, inhaltlich relevanten Seiten steigern die wahrgenommene Glaubwürdigkeit einer Website. Besonders wertvoll sind redaktionelle Verlinkungen von etablierten Domains (Moz, 2024; Brin und Page, 1998).
- **Markennennungen und soziale Signale:** Erwähnungen in Blogs, Medien oder auf sozialen Plattformen stärken die Autorität – auch dann, wenn keine direkte Verlinkung vorhanden ist (Terenteva, 2023).
- **Aufbau von Domain Autorität:** Domains, die über längere Zeit hinweg positive Signale senden, regelmäßig hochwertigen Inhalt veröffentlichen und ein niedriges Spam Niveau aufweisen, erzielen in der Regel bessere Rankings (Moz, 2024).
- **Vertrauenswürdigkeit (E-A-T):** Fachliche Expertise, Autorität und Vertrauenswürdigkeit spielen eine zentrale Rolle – vor allem in sensiblen Bereichen wie Finanzen oder Versicherungen. Wichtige Faktoren sind unter anderem verifizierte Autor:innenschaft, klare Kompetenznachweise, nachvollziehbare Quellen und der Einsatz sicherer Verbindungen (HTTPS) (Google Search Central, 2023; Moz, 2024).

Diese Prinzipien prägen die klassische SEO seit vielen Jahren. Mit dem Aufstieg KI-gestützter Systeme kommen nun neue Anforderungen hinzu. In dialogorientierten Modellen zählen neben Links auch Kontext, Gesprächslogik und semantische Tiefe. Im nächsten Abschnitt zeigen wir, wie sich bewährte SEO-Strategien mit den Anforderungen der KI-Systeme verbinden lassen.

Technologische Grundlagen der LLM-Suche

LLMs suchen nicht nur nach Stichwörtern, sondern verstehen deren Intention. Anstatt Begriffe abzugleichen, ordnen sie Bedeutungen in Vektorräumen zu, um den Kontext nuanciert zu erfassen.

LLM gestützte Suchsysteme unterscheiden sich grundlegend von der klassischen Stichwortsuche. Anstatt nur einzelne Wörter abzugleichen, erfassen sie ihre Bedeutung durch Vektorräume in einem semantischen Kontext. Beim sogenannten Dense Retrieval werden sowohl die Suchanfrage als auch die Inhalte in hochdimensionale Zahlenräume überführt. Diese sogenannten Embeddings ermöglichen eine Zuordnung auf Basis von Absicht, selbst wenn die Begriffe sprachlich nicht exakt übereinstimmen. Im Zentrum dieses Prozesses stehen sogenannte Transformer.

Das sind Modelle des tiefen Lernens, die entwickelt wurden, um Beziehungen zwischen Wörtern in einem Satz zu erkennen, auch wenn sie weit voneinander entfernt stehen. Eingeführt wurde dieser Ansatz von Vaswani und Kolleg:innen im Jahr 2017. Diese Modelle nutzen das Prinzip der Selbstaufmerksamkeit, um sprachliche Abhängigkeiten über längere Texte hinweg zu erfassen. Dadurch können sie sprachliche Feinheiten verstehen, Mehrdeutigkeiten auflösen und kontextbezogene Antworten erzeugen.

Diese Architektur bildet den Kern moderner LLM-Suche. Der Fokus verschiebt sich vom bloßen Wortabgleich hin zum Verstehen sprachlicher Zusammenhänge. Im nächsten Schritt betrachten wir das Modell, das diese Entwicklung möglich gemacht hat. Es handelt sich um den sogenannten Transformer.

Transformer als Basis der generativen KI

Transformer haben die Art verändert, wie Maschinen Sprache verarbeiten. Anstatt Wörter nacheinander zu lesen, erfassen sie den gesamten Satz auf einmal und erkennen, welche Begriffe inhaltlich relevant sind. Der Prozess beginnt mit der sogenannten Tokenisierung. Dabei wird ein Satz in kleinere Einheiten zerlegt. Diese sogenannten Tokens werden dann in Zahlen umgewandelt. Das Modell kann dadurch mit Sprache arbeiten, als wäre sie Mathematik.

Im Zentrum steht der sogenannte Self-Attention-Mechanismus. Er zeigt dem Modell, welche Wörter im Satz miteinander in Beziehung stehen. Dieser Mechanismus arbeitet über viele Ebenen hinweg und in mehrere Richtungen. So kann das Modell sowohl Grammatik als auch Bedeutung erfassen.

Diese Entwicklung hat moderne Sprachmodelle erst möglich gemacht. Sie sind in der Lage, komplexe Texte zu verstehen, auf riesige Datenmengen zuzugreifen und in bemerkenswerter sprachlicher Qualität zu antworten. Im nächsten Schritt geht es darum, wie aus Sprache Geometrie wird. Wir betrachten, wie sich Tokens in einem Raum mathematisch darstellen lassen.

Tokenisierung als erster Schritt

Jedes große Sprachmodell beginnt damit, unbearbeiteten Text in kleinere Einheiten zu unterteilen, sogenannte Tokens. Dabei handelt es sich nicht um vollständige Wörter, sondern um Wortbestandteile wie „Haus“, „rat“ und „versicherung“ anstelle des gesamten Begriffs „Hausratversicherung“. Diese Methode ermöglicht es den Modellen, auch komplexe, seltene oder zusammengesetzte Begriffe effizient zu verarbeiten, indem sie sie in erkennbare Teile aufgliedern.

Jedem dieser Tokens wird dann eine eindeutige Zahl aus einem festen Vokabular zugewiesen. Auf diese Weise wird natürliche Sprache in ein numerisches Format übersetzt, das von Maschinen verarbeitet werden kann.

Nachdem die Sprache nun in strukturierte Zahlen überführt wurde, besteht die nächste Herausforderung darin, diesen Zahlen auch Bedeutung zu geben. Dies geschieht durch sogenannte Embeddings.

Vektoren zur Darstellung von Bedeutung

Bevor ein Modell Sprache verstehen oder erzeugen kann, braucht es eine Möglichkeit, Bedeutung mathematisch abzubilden. Genau das leisten Embeddings. Dabei wird jedem Token ein dichter Vektor zugewiesen, also eine Zahlenreihe, die semantische Zusammenhänge abbildet. Begriffe wie „Versicherungspolice“ und „Versicherungsschutz“ liegen im Zahlenraum nahe beieinander, während nicht verwandte Begriffe weiter entfernt sind. Diese Beziehungen entstehen durch das Training auf großen Textmengen.

Sie ermöglichen es dem Modell, Nuancen zu erkennen, Kontexte zu verstehen und Bedeutungsähnlichkeiten einzuordnen.

Bei Eingaben aus dem Versicherungsbereich zum Beispiel liegen Begriffe wie „Premium“ und „Selbstbehalt“ häufig im selben Bedeutungsraum. Das signalisiert dem Modell, dass sie thematisch zusammengehören. Dieser Raum bildet die semantische Grundlage für alle weiteren Berechnungen.

Ab diesem Punkt beginnt das Modell, mit Aufmerksamkeitsmechanismen ein höheres Bedeutungsverständnis aufzubauen.

Aufmerksamkeitsmechanismen: Kontext priorisieren

Embeddings bilden eine numerische Darstellung von Bedeutungskontexten. Aufmerksamkeitsmechanismen bestimmen darauf aufbauend, worauf sich das Modell konzentrieren soll, und zwar dynamisch und mit hoher Genauigkeit. Statt alle Wörter gleich zu behandeln, vergibt das Modell Gewichtungen für jedes Token, je nachdem, wie relevant es im Zusammenhang mit anderen Wörtern ist. In einem Satz wie „Die Selbstbeteiligung greift, bevor die Versicherung zahlt“ erkennt das Modell, dass die Begriffe „Selbstbeteiligung“ und „zahlt“ eng miteinander verknüpft sind, auch wenn mehrere Wörter dazwischen liegen.

Das ermöglicht es Transformer-Modellen, auch komplexe Abhängigkeiten zu erkennen. Diese Fähigkeit ist besonders wertvoll in stark strukturierten Anwendungsbereichen wie der Krankenversicherung. Die sogenannte Self-Attention geht noch weiter, indem sie prüft, wie jedes einzelne Wort im Satz mit allen anderen in Beziehung steht. Sie verleiht dem Modell eine Art flexibles Gedächtnis und die Fähigkeit, das Wesentliche hervorzuheben.

Diese selektive Fokussierung ist einer der Hauptgründe, warum große Sprachmodelle in der Lage sind, Nuancen zu erfassen. Doch wie zeigt sich diese Fähigkeit in der Praxis? Die Antwort liegt in der Websuche als Schnittstelle.

Websuche als Zugang zu LLM-gestützten Antworten

Allein durch statisches Training lassen sich keine aktuellen Informationen bereitstellen. Um den steigenden Erwartungen von Nutzer:innen gerecht zu werden, greifen große Sprachmodelle auf eine integrierte Websuche zurück. Moderne Systeme mit LLM-Unterstützung gehen über ihr statisches Vorwissen hinaus, indem sie Suchanfragen in Echtzeit ausführen. Wenn Nutzer:innen Fragen stellen, die aktuelle oder sehr spezifische Informationen erfordern, aktiviert das System eine Suchschnittstelle. Diese Komponente formuliert die Anfrage gezielt um und ruft Inhalte aus vertrauenswürdigen Quellen ab. Dazu gehören unter anderem Aufsichtsbehörden, Nachrichtenportale oder wissenschaftlich fundierte Fachverlage (Li et al., 2024).

Die abgerufenen Inhalte werden nicht ungeprüft übernommen. Sie durchlaufen zunächst einen Filter, der die Glaubwürdigkeit, Aktualität und thematische Relevanz der Informationen bewertet. Ohne diesen Zwischenschritt könnten irrelevante oder qualitativ minderwertige Daten das Antwortverhalten des Modells verfälschen. Wie Liu und Kolleg:innen (2023) betonen, sind diese Filter entscheidend für die Qualität und Verlässlichkeit der Ergebnisse.

Nach erfolgreicher Prüfung wird die ausgewählte Information in den internen Kontext des Sprachmodells eingebunden. Dadurch entsteht eine Verbindung zwischen dem statischen Vorwissen und den neu eingebunden Inhalten. Das Ergebnis sind Antworten, die präzise, nuanciert und auf die Absicht der Nutzer:innen abgestimmt sind (Nakano et al., 2022).

Dieser Ablauf ist besonders wichtig in sensiblen Bereichen wie der Versicherungsbranche. Selbst kleinere Ungenauigkeiten können rechtliche oder reputationsbezogene Folgen haben. Um solche Risiken zu minimieren, verlassen sich viele Systeme auf Vertrauenssignale auf den Quellenseiten. Dazu zählen etwa die institutionelle Herkunft der Inhalte, transparente Autor:innenschaft und nachprüfbare Quellenangaben (Pan et al., 2024).

Die Schnittstelle zur Websuche ist deshalb nicht nur ein technisches Hilfsmittel, sondern ein zentrales Element moderner KI-Systeme.

Xiong und Kolleg:innen (2024) beschreiben sie als dynamische Brücke, die das offene Netz mit der Verarbeitungseinheit des Sprachmodells verbindet. Sie hält den Informationsfluss aktuell, überprüfbar und relevant. Das Verständnis dieser Architektur führt zu einer weiterführenden Frage: Wie interagieren diese beiden Komponenten in einem konkreten Suchprozess?

Wie beide Komponenten zusammenwirken

Die Verbindung des Sprachmodells mit der Websuche bildet die vollständige Architektur KI-gestützter Suche. Wenn eine Anfrage gestellt wird, prüft das System zunächst, ob sie durch das interne Wissen des Modells beantwortet werden kann oder ob zusätzlich aktuelle Informationen erforderlich sind. Ist letzteres der Fall, wird die Websuche aktiviert.

Sie beschafft und bewertet die jeweils relevanten Inhalte. Anschließend verbindet das Sprachmodell die neuen Informationen mit dem bestehenden Kontextwissen. So entsteht eine Antwort, die in sich schlüssig, gut verständlich und unmittelbar nutzbar ist.

Das Verständnis dieser Architektur macht deutlich, wie Versicherungsunternehmen die Potenziale KI-gestützter Suche für sich nutzen können. Wer die Aufgaben beider Komponenten kennt, kann Inhalte und Strategien gezielt aufeinander abstimmen. So bleiben Produkte, Dienstleistungen und geschäftskritische Informationen auch in anspruchsvollen KI-gestützten Suchabläufen sichtbar, korrekt und zugänglich.

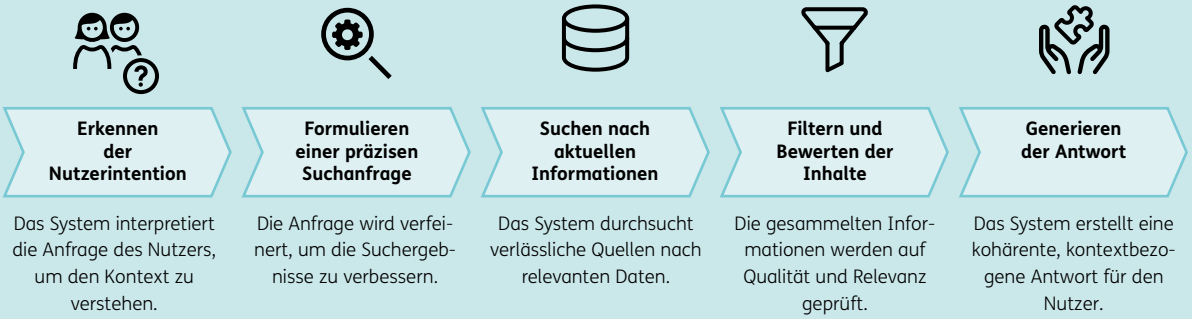
Wie eine KI-Suche Schritt für Schritt funktioniert

KI-gestützte Suche umfasst mehr als leistungsstarke Modelle. Versicherungen müssen verstehen, wie jeder einzelne Schritt funktioniert – von der Interpretation der Anfrage bis hin zur präzisen, kontextbezogenen Antwort. Wer diesen Ablauf kennt, kann seine Strategie gezielt auf das veränderte Nutzer:innenverhalten ausrichten.

Die meisten Suchprozesse, die durch große Sprachmodelle gesteuert werden, folgen fünf zentralen Phasen.

Verarbeitung von Anfragen in der LLM-Suche und Informationsabruf

Auf Basis mehrerer wissenschaftlicher Studien



1 Erkennen der Nutzerintention

Das System nutzt fortgeschrittene Verfahren der Verarbeitung natürlicher Sprache, um die tiefere Bedeutung einer Anfrage zu entschlüsseln. Es prüft, ob das interne Wissen des Modells ausreicht oder ob zusätzlich Informationen in Echtzeit benötigt werden (Yang et al., 2023; Anand et al., 2023). Fragt jemand zum Beispiel nach einer Hausratversicherung bei Küstenschäden, interpretiert das Modell sowohl den expliziten Bedarf nach Überschwemmungen als auch implizite Hinweise auf ein geografisches Risiko.

2 Formulieren einer präzisen Anfrage

Wenn externe Informationen erforderlich sind, verfeinert das Modell die ursprüngliche Nutzer:innenfrage. Es ergänzt Synonyme, Zeitfilter oder branchenspezifische Operatoren, um die Ergebnisqualität zu verbessern (Mo et al., 2023). Diese gezielte Umformulierung sorgt dafür, dass nur besonders relevante Quellen berücksichtigt werden. Das erhöht die Genauigkeit und Geschwindigkeit.

3 Suche nach aktuellen Informationen

Das System greift auf öffentlich verfügbare und vertrauenswürdige Quellen zurück. Dazu zählen behördliche Seiten, akademische Publikationen und renommierte Fachverlage (Dhole und Agichtein, 2024). Bezahlschranken oder geschlossene Datenbanken werden in der Regel ausgeschlossen. Die Auswahl orientiert sich an Glaubwürdigkeit, Aktualität und thematischer Relevanz, damit die Ergebnisse möglichst genau zur Anfrage passen.

4 Filtern und Bewerten der Inhalte

Jede gefundene Quelle wird auf Verlässlichkeit und inhaltliche Passung geprüft. Automatisierte Filter sortieren veraltete, widersprüchliche oder qualitativ minderwertige Inhalte aus (Yang et al., 2020). Diese sorgfältige Auswahl stellt sicher, dass nur hochwertige Informationen in die Antwort des Modells einfließen und das Risiko von Ungenauigkeiten minimiert wird.

5 **Generieren der Antwort**
Das Modell verbindet die neu gewonnenen Informationen mit seinem bestehenden Wissensbestand. Es nutzt dabei Transformer Architektur, Self-Attention und Vektorraumdarstellungen (Yang et al., 2023). Durch diesen generativen Prozess entsteht eine zusammenhängende, inhaltlich präzise Antwort, die statisches Wissen und aktuelle Informationen sinnvoll verbindet. Das Ergebnis ist eine Antwort, die nachvollziehbar, individuell und kontextbezogen ist.

Schlussfolgerung

Nachdem diese fünf Phasen klar beschrieben sind, lässt sich nun ein Vergleich zwischen klassischen, schlüsselwortbasierten Methoden und den fortschrittlicheren, kontextsensitiven Abläufen der LLM-Suche ziehen. Dabei wird deutlich, warum Versicherungen ihre Content-Strategien anpassen müssen, um den sich verändernden Erwartungen von Nutzer:innen gerecht zu werden. Die Funktionsweise KI-gestützter Suche offenbart ein neues Niveau an Kontextverständnis und Präzision beim Informationsabruf. Um die Tragweite dieser Entwicklung zu verstehen, ist ein direkter Vergleich mit den Begrenzungen klassischer, schlüsselwortbasierter Suchmodelle unerlässlich.

Zentrale Unterschiede im Vergleich zu klassischen Suchsystemen

Klassische Methoden setzen auf Keyword Matching und Backlinks. Sie führen häufig zu allgemeinen Trefferlisten mit Links, die nur begrenzt auf die tatsächliche Absicht der Nutzer:innen eingehen (Mo et al., 2024). Im Gegensatz dazu analysieren KI-gestützte Ansätze die Intention der Anfrage und liefern direkte, verwertbare und auf die Nutzer:innen .zugeschnittene Antworten in einem konversationellen Format.. Dadurch verändern sich Anforderungen an Inhalt, Zugänglichkeit und Nutzer:innenbindung grundlegend (Kumar et al., 2024).

Die folgende Tabelle zeigt die grundlegenden Unterschiede zwischen klassischer, schlüsselwortbasierter Suche und der LLM-gestützten Struktur. Sie stellt beide Ansätze entlang von fünf zentralen Dimensionen gegenüber.

Auf Basis dieser Unterschiede folgt im nächsten Abschnitt die Definition der Forschungsziele, die unsere empirische Analyse geleitet haben. Um die damit verbundenen Fragen gezielt zu beantworten, beschreiben wir nun die konkreten Zielsetzungen unserer Untersuchung. Diese bilden die Grundlage für die anschließende Analyse und die abgeleiteten Empfehlungen in diesem Whitepaper.

Zentrale Unterschiede im Vergleich zu klassischen Suchsystemen

Dimension	Klassische Suche (Google & SEO)	LLM-Suche (KI-gestützte Systeme)
1. Verarbeitung der Anfrage Keyword Matching vs. kontextuelles Verstehen	Vergleicht explizite Schlüsselwörter mit indexierten Inhalten. Die Platzierung basiert auf Keyword-Häufigkeit, Backlinks und Metadaten (Kumar et al., 2024). Die Ergebnisse sind oft ungenau und erfordern manuelle Filterung durch die Nutzer:innen.	Interpretiert die vollständige Absicht hinter der Anfrage mithilfe kontextueller Vektoren und Transformer basierter Sprachverarbeitung. Erkennt Nuancen, implizite Bedeutungen und Gesprächszusammenhänge (Mo et al., 2024).
2. Aktualität der Informationen Statische Indizierung vs. Zugriff in Echtzeit	Greift auf ein regelmäßig aktualisiertes statisches Verzeichnis zurück. Das kann zu veralteten Ergebnissen führen, insbesondere in schnelllebigen Bereichen wie der Versicherungsbranche (Liu et al., 2024).	Holt Informationen live aus dem Netz und stellt sicher, dass Antworten aktuelle Vorschriften, Produktänderungen und Marktentwicklungen abbilden (Karamolegkou et al., 2023).
3. Ergebnisformat Linklisten vs. generative Antworten	Gibt eine geordnete Liste von Links zurück. Nutzer:innen müssen die Seiten selbst aufrufen und auswerten. Das verzögert den Zugang zu relevanten Informationen (Liu et al., 2024).	Erzeugt direkte, kontextbezogene Antworten in natürlicher Sprache. Spart Klicks und liefert komprimierte Inhalte unmittelbar (Mo et al., 2023).
4. Personalisierung Allgemeine Ergebnisse vs. individuelle Interaktion	Zeigt weitgehend generische Ergebnisse für ein breites Publikum. Geht kaum auf individuelle Bedürfnisse ein (Mo et al., 2024).	Erzeugt personalisierte Antworten auf Basis von Verhaltensdaten, Suchhistorie und Gesprächskontext. Unterstützt individuelle Produktempfehlungen (Sharma et al., 2024; Karamolegkou et al., 2023).
5. Architektur und Nutzererlebnis Manuelle Navigation vs. KI-gestützte Entscheidungslogik	Nutzer:innen navigieren selbst durch Inhalte, gestützt auf Metadaten, Linkstrukturen und Webhierarchien. Die Interpretation liegt bei der Person.	KI-Agenten analysieren, vergleichen und fassen Inhalte für die Nutzer:innen zusammen. Sie unterstützen Entscheidungen direkt. Das Suchsystem entwickelt sich zu einer logikgestützten Entscheidungsinstanz (Mo et al., 2023).

Studienaufbau und Forschungsziele

Um die Sichtbarkeit in der LLM-gestützten Suche zu verstehen, haben wir getestet, wie Versicherungswebsites in verschiedenen KI-Systemen abschneiden. Anhand realer Suchanfragen haben wir ermittelt, welche Arten von Inhalten am häufigsten angezeigt werden – und warum. Das Ergebnis zeigt nicht nur, was heute funktioniert, sondern auch, was darüber entscheidet, wer morgen gefunden wird.

Diese Studie basiert auf einer Reihe von Hypothesen, die untersuchen, welche Faktoren die Sichtbarkeit von Webseiten in einer LLM-Suche im Vergleich zur klassischen Google Suche beeinflussen. Ziel ist es, herauszufinden, welche Inhalte, Formate oder technischen Strukturen die Wahrscheinlichkeit erhöhen, von diesen grundlegend unterschiedlichen Systemen referenziert oder angezeigt zu werden. Damit soll aufgezeigt werden, wie Large Language Models die Suche verändern und was Versicherer tun müssen, um in diesem Umfeld sichtbar, auffindbar und relevant zu bleiben.

Ausgangspunkt ist die Annahme, dass ein grundlegendes Maß an SEO-Optimierung in beiden Suchumgebungen Voraussetzung für Sichtbarkeit ist. Auf dieser Basis werden gezielt Hypothesen untersucht, die die Stärken und Abrufmechanismen großer Sprachmodelle widerspiegeln.

Hypothese 1: Maschinelle Lesbarkeit und technische Zugänglichkeit erhöhen die Sichtbarkeit in der LLM-Suche

Die erste Hypothese besagt, dass maschinenlesbare und technisch gut zugängliche Inhalte wie semantisch strukturierter HTML-Code, schnell ladende Seiten und barrierefreie Designs mit höherer Wahrscheinlichkeit von LLMs referenziert werden. Da diese Systeme große Datenmengen analysieren und interpretieren müssen, verschafft ihnen sauberer und zugänglicher Code einen klaren Vorteil. Dies entspricht dem Kriterium „Technische Zugänglichkeit und maschinelle Lesbarkeit“.

Untersuchungen haben gezeigt, dass LLMs trotz zunehmender Leistungsfähigkeit weiterhin auf die zuverlässige Analyse und Extraktion strukturierter Inhalte angewiesen sind (Achiam et al. 2023). Webseiten, die auf standardisierte semantische Auszeichnung – etwa durch HTML5-Tags – setzen und auf blockierende Skripte oder komplexes clientseitiges Rendering verzichten, ermöglichen eine klarere Trennung relevanter Inhalte von irrelevanten Elementen und verbessern so die maschinelle Erfassbarkeit und Nutzbarkeit der Daten.

Die Referenzhäufigkeit gut strukturierter, schneller und barrierefreier URLs im Vergleich zu weniger optimierten Seiten kann als Indikator zur Überprüfung dieser Hypothese herangezogen werden.

Hypothese 2: Semantische Verlinkung verbessert die Auffindbarkeit durch LLMs

Die zweite Hypothese geht davon aus, dass Inhalte mit starker interner Verlinkung, konzeptionellen Clustern und strukturierten Entitäten häufiger von LLMs abgerufen werden. Dies entspricht dem Kriterium „Semantische und verlinkte Optimierung“.

Da viele KI-Systeme auf Embedding basierte Abrufmechanismen setzen, tragen klar strukturierte semantische Beziehungen dazu bei, die thematische Relevanz innerhalb einer Seite oder Website zu verstärken (Touvron et al. 2023; Ni et al. 2023). LLMs erfassen Bedeutung über Dokumentgrenzen hinweg und analysieren Inhalte nicht isoliert.

Seiten mit dichter interner Verlinkung und klaren Entitäten, etwa durch schema.org oder verknüpfte Datenstrukturen, sind besser positioniert, als relevante Quelle ausgewählt zu werden. Diese Hypothese lässt sich validieren, indem untersucht wird, wie häufig semantisch angereicherte Seiten in LLM-Antworten genannt werden, verglichen mit strukturell fragmentierten Inhalten.

Hypothese 3: Vertrauenswürdige Quellen werden von LLMs häufiger zitiert

Die dritte Hypothese geht davon aus, dass Inhalte aus vertrauenswürdigen Quellen wie Regierungsseiten, wissenschaftlichen Institutionen oder etablierten Expert:innen von LLMs stärker bevorzugt werden als von klassi-

Übersicht der Hypothesen

- **H1:** Maschinelle Lesbarkeit und technische Zugänglichkeit erhöhen die Wahrscheinlichkeit, in LLM-basierten Abrufen berücksichtigt zu werden.
- **H2:** Semantische Verlinkung von Inhalten verbessert die Chancen, von LLMs abgerufen zu werden.
- **H3:** Inhalte aus vertrauenswürdigen Quellen werden von LLMs häufiger zitiert.
- **H4:** Konversationelle Formatierung passt besser zum Abrufverhalten von LLMs.

schen Suchmaschinen. Dies entspricht dem Kriterium „Zitierfähigkeit und Quellenqualität".

LLMs werden häufig durch Feinabstimmung oder Retrieval Augmentation so weiterentwickelt, dass sie Halluzinationen vermeiden, indem sie Antworten auf verifizierte Inhalte stützen (Nakano et al. 2022). Vertrauenssignale wie Domainautorität, transparente Autor:innenschaft, Zitationsnachweise und institutionelle Zugehörigkeit spielen eine zentrale Rolle dabei, ob eine URL in einer KI-generierten Antwort verwendet wird.

Ethikleitlinien für KI (Jobin et al. 2019) verstärken zudem die Bevorzugung hochwertiger, geprüfter Inhalte. Die Hypothese lässt sich überprüfen, indem analysiert wird, wie häufig Domains mit hohem Vertrauensniveau in LLM-Antworten erscheinen, im Vergleich zu weniger renommierten Quellen.

Hypothese 4: Konversationelle Formatierung entspricht dem Abrufverhalten von LLMs

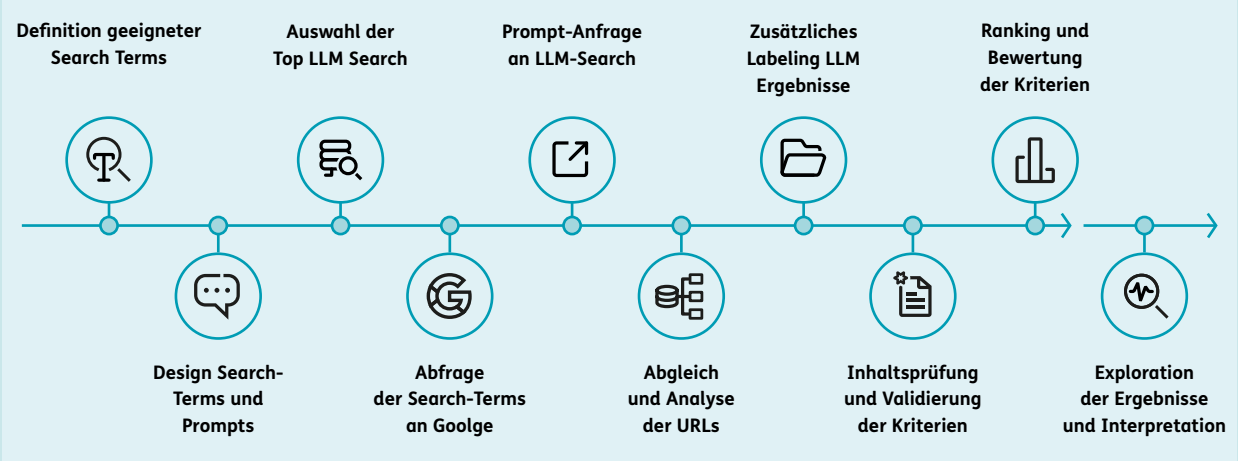
Die vierte Hypothese geht davon aus, dass Inhalte, die im Stil einer Konversation aufgebaut sind – etwa als FAQ, Frage-Antwort-Sektion oder klar segmentierte Antwortblöcke – mit höherer Wahrscheinlichkeit von LLMs abgerufen werden. Dies entspricht dem Kriterium „Struktur und unmittelbare Optimierung für LLMs".

Viele LLMs sind auf dialogreiche Korpora trainiert und bevorzugen Inhalte, die natürlichen, konversationellen Eingaben entsprechen (Wei et al. 2022; Wu et al. 2023; Karpukhin et al. 2020). Inhalte, die mit Fragen als Überschrift und direkten, prägnanten Antworten darunter strukturiert sind, spiegeln das Eingabe-Ausgabe-Verhalten konversationeller KI-Systeme wider. Diese Hypothese lässt sich testen, indem analysiert wird, wie häufig konversationell strukturierte Inhalte in KI-generierten Antworten erscheinen – im Vergleich zu längeren, narrativ geprägten Formaten.

Diese vier Hypothesen bilden die Grundlage für den Vergleich zwischen LLM-basierten Suchmodellen und Ergebnissen der Google Suche. Für jede Hypothese wurde untersucht, wie häufig und auf welcher Position unterschiedliche Inhaltstypen erscheinen. Visualisierungen wurden eingesetzt, um zentrale Unterschiede zu verdeutlichen.

Die Ergebnisse zeigen, wie grundlegende Aspekte von Webarchitektur, Inhaltsstruktur und Quellvertrauen die Sichtbarkeit in LLM-gestützten Abrufsystemen beeinflussen. Gleichzeitig wird deutlich, wie wichtig es ist, Inhalte

Modularer Aufbau von Forschung und Methodik



Wir haben eine strukturierte Methodik mit zehn aufeinander aufbauenden Schritten verfolgt. Diese Struktur ermöglichte einen vollständigen Überblick – von der Formulierung realer Suchanfragen durch Nutzer:innen bis zur Frage, wie LLMs relevante Inhalte abrufen und gewichten. Im nächsten Abschnitt wird gezeigt, wie klassische Google Suchbegriffe in dieses Untersuchungsraster integriert wurden und welche Erkenntnisse sich daraus im Hinblick auf die veränderte Sichtbarkeit in unterschiedlichen Suchsystemen ergeben.

sowohl für menschliche Nutzer:innen als auch für KI-Systeme gezielt zu optimieren, um digitale Sichtbarkeit zu sichern.

Design und Methodik der Vorstudie

Ein fundiertes Verständnis darüber, wie sich eine KI-gestützte Suche auf die Sichtbarkeit von Versicherungen auswirkt, erfordert eine Forschungslogik, die theoretische Präzision mit praktischer Relevanz verbindet.

Dafür wurde eine Methodik entwickelt, mit der sich ermitteln lässt, welche Inhalte von LLM-Suchsystemen abgerufen werden, wie sich das vom klassischen Suchverhalten unterscheidet und welche strukturellen Anpassungen Versicherer berücksichtigen müssen, um sichtbar zu bleiben.

Wir haben 20 Kriterien identifiziert und validiert, die die Sichtbarkeit in KI-gestützten Suchsystemen beeinflussen. In der Studie wurden davon vier zentrale Kriterien vertiefend untersucht, die sich sich an den Hypothesen orientieren. Die Studie stützte sich auf drei zentrale Komponenten:

1. **Hypothesenbasiertes Testverfahren**
   Es wurden Hypothesen formuliert (zum Beispiel „semantische Kohärenz erhöht die Sichtbarkeit in der LLM-Suche“) und anhand eines Datensatzes von über 33.000 URLs geprüft. Dabei kamen 20 strukturierte Inhalts- und Metadatenkriterien zum Einsatz.

2. **Validierung unter realen Bedingungen**
   Über die kontrollierten Testumgebungen hinaus wurden reale, versicherungsbezogene Suchanfragen analysiert, um zu beobachten, wie LLMs Inhalte abrufen und priorisieren. Auf diese Weise konnten die Abrufergebnisse direkt mit konkreten Inhaltsstrukturen verknüpft und in die Entwicklung gezielter Optimierungsstrategien für LLMs (LLMO) überführt werden.

3. **Quantitative und qualitative Analyse**
   Statistische Auswertungen wie Abrufvolumen und Ranking-Muster wurden mit qualitativen Inhaltsanalysen kombiniert. Dazu gehörten unter anderem semantische Struktur, Formatqualität und Metadatenkonsistenz. Ziel war es, nicht nur zu identifizieren, welche Inhalte angezeigt werden, sondern auch zu verstehen, warum genau diese Inhalte ausgewählt wurden.

Durch die Kombination dieser drei Komponenten wurde sichergestellt, dass die Analyse sowohl evidenzbasiert als auch anwendungsorientiert ist. Der folgende Abschnitt zeigt, wie dieses methodische Rahmenwerk in einem strukturierten, prompt-basierten Bewertungsprozess konkret umgesetzt wurde.

Prompt-basierte Strukturierung und Vergleichsanalyse

Um die Präzision unserer Analyse zu erhöhen und eine belastbare Vergleichsbasis zu schaffen, haben wir klassische Google Suchbegriffe in unsere Methodik integriert. Dadurch konnten wir direkt vergleichen, wie Large Language Models (LLMs) im Vergleich zur Google Suche auf ähnliche versicherungsbezogene Suchanfragen reagieren. Die daraus gewonnenen Erkenntnisse sind nicht nur akademisch fundiert, sondern auch für die Praxis von großer Bedeutung.

Zur Strukturierung der Analyse haben wir 18 repräsentative Suchbegriffe aus drei Produktkategorien ausgewählt: Hausratversicherung, Zahnzusatzversicherung und Rechtsschutzversicherung. Jede Kategorie umfasste sowohl markenspezifische Begriffe (zum Beispiel „ERGO Zahnzusatzversicherung“) als auch markenoffene Begriffe (zum Beispiel „beste Zahnzusatzversicherung“), um unterschiedliche Phasen der Customer Journey abzubilden.

Diese Suchbegriffe bildeten die Grundlage für ein strukturiertes Prompt-Raster mit 120 Eingaben. Diese haben wir systematisch variiert nach:

Wichtige Kennzahlen auf einen Blick

3 analysierte Versicherungsprodukte aus den Bereichen Hausrat, Zahnzusatzversicherung und Rechtsschutz

120 Prompts (40 pro Produkt), die verschiedene Funnel-Stufen und Komplexitätsgrade abdecken

18 Google-Suchbegriffe (6 pro Produkt), die als traditionelle Such-Benchmarks verwendet werden

120 Prompts und 18 Suchbegriffe, die jeweils 10 Mal an SearchGPT, Perplexity, You.com und Gemini

33.600 URLs, die über alle Plattformen hinweg abgerufen wurden

25.441 LLM-basierte URLs

5.920 nur bei Google gefundene URLs

2.074 halluzinierte URLs, die nach einer Qualitätsfilterung entfernt wurden

900.909 Datenpunkte, die durch eine Kombination aus manueller und agentenunterstützter Inhaltskennzeichnung und -bewertung generiert wurden

- **Funnel Phase:** 78 Prompts im oberen Funnel (explorative Awareness-Phase, d. h. frühe Phase zur Informationssuche und Ideenfindung) und 42 im unteren Funnel (Consideration und Conversion-Phase, fokussiert auf konkrete Entscheidungen und Handlungsabsichten)
- **Markenkontext:** Gleichverteilung zwischen markenoffener und markenspezifischer Formulierung
- **Komplexitätsgrad:** 90 einfache, alltagsnahe Prompts und 30 Prompts mittlerer Komplexität mit Vergleichs- oder Einschränkungselementen

Diese Struktur spiegelt reales Suchverhalten wider. Die meisten versicherungsbezogenen Anfragen sind explorativ, werden in einfacher, natürlicher Sprache formuliert und sind häufig durch Markenbewusstsein geprägt – insbesondere in vertrauenssensiblen Kontexten (Spatharioti et al. 2023; Chowdhury et al. 2024; Nestaas et al. 2024; Google Consumer Insights 2024). Weil die Prompts reale Suchmuster abbilden, zeigen ihre Ergebnisse, wie gut unterschiedliche Systeme auf konkrete Informationsbedarfe reagieren.

Durch die Standardisierung der Prompt-Struktur konnten wir erkennen, welche Inhaltstypen konsistent erscheinen, welche Modelle Klarheit oder Glaubwürdigkeit priorisieren und welche Formatentscheidungen besonders belohnt werden.

Zur Etablierung einer Vergleichsbasis haben wir alle 18 Google Suchbegriffe jeweils zehnmal abgerufen. Dabei entstanden über 5.800 eindeutige URLs. Die abgerufenen Ergebnisse haben wir in drei Kategorien eingeteilt:

- **Nur von Google gelieferte URLs** (nicht durch LLMs abgerufen)
- **Gemeinsame URLs** (durch beide Systeme gefunden)
- **Nur durch LLMs gelieferte URLs** (nicht in Google Ergebnissen enthalten)

Diese Struktur bildet reales Suchverhalten ab. Sie deckt die gesamte Customer Journey ab – vom ersten Interesse bis zur finalen Entscheidung – und berücksichtigt unterschiedliche Grade der Markenvertrautheit sowie sprachliche Komplexität. Damit entsteht ein realistisches, datenbasiertes Modell zur Bewertung, wie effektiv LLMs Versicherungsinhalte abrufen und gewichten.

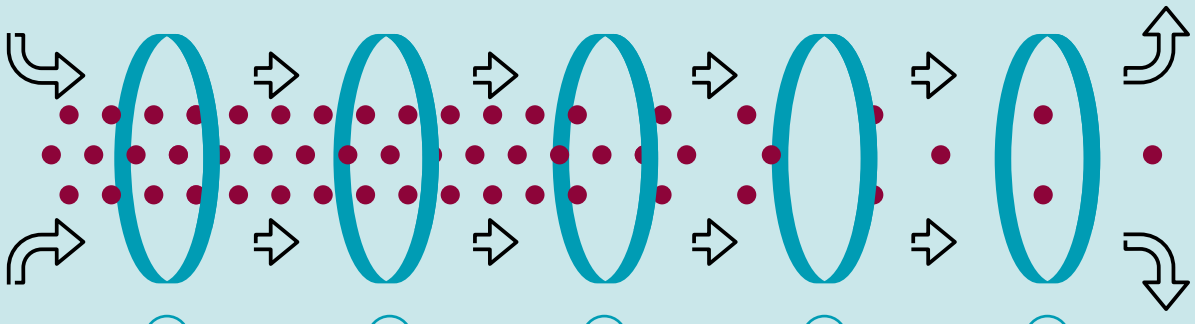
Mit diesem kontrollierten Prompt-Raster konnten wir beobachten, wie die verschiedenen Systeme reagierten – sowohl im Hinblick auf die abgerufenen Inhalte als auch auf deren Häufigkeit. Der nächste Schritt bestand darin, mehrere tausend Ergebnisse aus LLMs und Google zu erfassen, zu vergleichen und zu filtern, um herauszufinden, wo und warum sich Sichtbarkeit verschiebt.

Methodik der Datenauswertung und Voranalyse

Insgesamt wurden über alle vier LLM-Systeme und die Google-Suche hinweg 33.366 URLs generiert. Nach der Entfernung von Duplikaten und fehlerhaften Treffern blieben 31.292 gültige Ergebnisse übrig. Ausgehend vom strukturierten Prompt-Raster auf Basis von 18 klassischen Google Suchbegriffen riefen die LLM-Systeme 25.441 eindeutige URLs ab. Die direkte Suche nach denselben 18 Begriffen bei Google lieferte im Vergleich dazu 5.851 eindeutige URLs. Zusätzlich wurden 2.074 Halluzinationen identifiziert – also Verweise auf nicht existierende oder thematisch irrelevante Seiten. Sie wurden entfernt und machten rund sieben Prozent aller LLM-Ergebnisse aus.

Jede URL wurde nach Quelltyp (zum Beispiel Website eines Versicherers, Maklerplattform, Fachmedium, privates Blog), Markenbezug und Domainstruktur klassifiziert. So konnten Muster identifiziert werden, welche Quellentypen und Inhaltsformate besonders häufig von LLMs abgerufen werden. Diese Bereinigung stellte sicher, dass die

Detaillierter Bewertungs- und Kriterienprüfungsprozess für jede abgerufene URL



1 **Automatisierte Bewertung** Automatisierte Prüfung anhand definierter Kriterien

2 **Punktevergabe/ Aggregation** Zusammenführung von Bewertungen zur Analyse

3 **Kriterienvalidierung** Prüfung der Belastbarkeit der Kriterien

4 **Thematische Gewichtung** Bewertung der Einflüsse auf die Sichtbarkeit

5 **Manuelle Nachprüfung** Abschließende manuelle Vallidierung der Ergebnisse

weitere Analyse ausschließlich auf verifizierbaren und zugänglichen Inhalten beruhte.

Um strukturelle Muster vertieft zu untersuchen, wurde eine repräsentative Auswahl von 606 URLs mit einem Kriterienraster aus 20 Merkmalen bewertet. Dieses Raster deckte vier zentrale Dimensionen ab: maschinelle Lesbarkeit, semantische Kohärenz, Vertrauenswürdigkeit der Quelle und konversationelle Struktur.

Die Bewertungsmethodik kombinierte hypothesenbasierte Tests, strukturierte Punktbewertungen und angewandte Inhaltsanalysen über vier LLM-Systeme und Google hinweg. Die Stichprobe umfasste 606 versicherungsbezogene URLs, die auf 267 eindeutige Seiten normalisiert wurde. Bewertet wurde anhand automatisierter Prüfungen und regelbasierter Verfahren, abgestimmt auf 20 Optimierungskriterien.

Diese Kriterien umfassten u.a. vier Kerndimensionen:

- **Maschinelle Lesbarkeit:** Sauberkeit des HTML-Codes, mobile Darstellbarkeit, Ladegeschwindigkeit und Barrierefreiheit
- **Semantische Verlinkung:** Interne Struktur, inhaltlicher Fluss, Einsatz von Überschriften und logische Link-Hierarchien
- **Quellenvertrauen und Verifikation:** Institutionelle Glaubwürdigkeit, SSL-Verschlüsselung, Autor:innenangabe und regulatorische Offenlegungen
- **Konversationelle Strukturierung:** FAQ-Sektionen, modulare Antwortblöcke und klar abgegrenzte Informationssegmente

Ergänzend wurden 16 Metakriterien erfasst, die tieferliegende Strukturmerkmale abbildeten – darunter Inhaltshierarchie, Duplikation, interne Navigation und Passung zur Prompt-Intention. Die Auswertung folgte einem triangulierten Forschungsdesign: empirische Hypothesenprüfung mit einem gelabelten Datensatz, Design Science zur Sicherstellung praktischer Relevanz und Mixed Methods Logik zur Verknüpfung quantitativen Abrufverhaltens (etwa Halluzinationsquote, Verteilungsmuster) mit qualitativen Inhaltsmerkmalen.

Um Ausgewogenheit und Vergleichbarkeit sicherzustellen, wurde die Stichprobe auf acht Cluster verteilt – darunter markenspezifische und markenoffene Domains, Funnel-Phasen sowie verschiedene Plattformtypen. Dieser gestufte Analyseansatz ermöglichte es, nicht nur zu untersuchen, welche Inhalte von LLMs abgerufen werden, sondern auch, wie Struktur, Format und Herkunft deren Sichtbarkeit in KI-gestützten Suchumgebungen beeinflussen.

Semantische Abrufmuster in LLMs

Ergänzend zu den groß angelegten Prompt-Tests und Inhaltsbewertungen wurde eine diagnostische Analyseebene entwickelt, um zu untersuchen, wie LLMs Informationen auf semantischer Ebene abrufen. Anders als klassische, schlüsselwortbasierte Suchmaschinen operieren LLMs in hochdimensionalen Vektorräumen. Dort erfolgt die Übereinstimmung zwischen Anfrage und Inhalt nicht über exakte Wortwahl, sondern über semantische Nähe und Verknüpfungen.

Um diesen Mechanismus zu analysieren, wurde untersucht, wie sich semantisch strukturierte Inhalte im Abrufverhalten von LLMs bewähren – mit Fokus auf drei Aspekte:

- **Abrufkonsistenz:** Wurden dieselben Seiten wiederholt über verschiedene Prompts abgerufen?
- **Clustering-Verhalten:** Traten verwandte Inhaltstypen in verschiedenen Systemen gemeinsam auf?
- **Semantische Kohärenz:** Wiesen intern verlinkte Inhalte eine höhere Abrufstabilität auf?

Diese Indikatoren halfen einzuschätzen, wie gut LLMs Bedeutung seitenübergreifend erkennen und wie stark semantische Struktur die Auffindbarkeit von Website-Inhalten beeinflusst. Auch wenn keine vollständige Modellierung des latenten Vektorraums durchgeführt wurde, wurden fragmentierte und konsistente Embedding Muster visuell verglichen.

Dabei zeigte sich: Inhalte mit konsistenter Terminologie, klarer Themenhierarchie und funktionierender interner Navigation wurden deutlich häufiger und stabiler abgerufen – über alle Systeme hinweg.

Diese Diagnostik bestätigt eine zentrale Erkenntnis: Inhalte, die nicht nur technisch zugänglich, sondern auch semantisch sauber aufgebaut sind, haben signifikant höhere Chancen, in LLM gestützten Umgebungen gefunden zu werden.

Ablauf der URL-Bewertung

Um die Performance von Versicherungsinhalten in der LLM-gestützten Suche zu bewerten, wurde jede URL im Datensatz anhand eines strukturierten Scoring-Modells geprüft. Der Bewertungsprozess kombinierte automatisierte Prüfungen, regelbasierte Kriterien und gezielte manuelle Bewertungen.

Dieses Fünf-Schritte-Raster bildete die Grundlage für die Analyse. Es ermöglichte, gezielt zu isolieren, welche Inhaltsmerkmale am stärksten wirken und warum bestimmte Versicherungsseiten in KI-gestützten Suchsystemen konstant besser abschneiden als andere.

Studienergebnisse zur Sichtbarkeit von Versicherungsinhalten in der LLM-Suche

Große Sprachmodelle filtern Inhalte anders als klassische Suchmaschinen: Nicht Popularität, sondern Lesbarkeit, Struktur und semantische Klarheit entscheiden über Sichtbarkeit. Unsere Analyse zeigt, welche Inhalte Versicherer in der KI-Suche sichtbar machen – und welche dabei untergehen. Von Fehlerquellen bis zu Erfolgsmustern liefern die Daten wichtige Erkenntnisse für die Content-Strategie der Zukunft.

Verschiebung der Sichtbarkeit über die Systeme hinweg

Der Vergleich zwischen Google und LLM-gestützten Suchsystemen zeigt grundlegende Unterschiede darin, wie Sichtbarkeit erzeugt wird. Diese Verschiebung verändert die Anforderungen an Versicherer, wie sie ihre Inhalte aufbauen und strukturieren müssen.

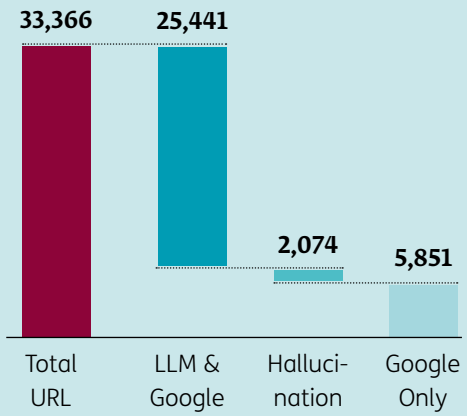
Die LLM-Plattformen riefen insgesamt deutlich mehr URLs ab als Google, wiesen aber auch eine höhere Variabilität in der Ergebnisqualität auf. Gleichzeitig ließen sich Muster erkennen, die strukturierten und maschinenlesbaren Inhalten klar den Vorzug gaben – unabhängig von der Markenbekanntheit. Diese Ergebnisse verdeutlichen, dass Auffindbarkeit in KI-gestützter Suche weniger von klassischer SEO abhängt, sondern stärker davon, wie gut Inhalte mit der Abruflogik der Sprachmodelle übereinstimmen.

Google bleibt weiterhin die am häufigsten genutzte Suchmaschine – auch im Zusammenhang mit Versicherungsinhalten. In der Studie wurden 18 versicherungsbezogene Suchbegriffe aus drei Produktkategorien – Zahnzusatzversicherung, Rechtsschutzversicherung, Hausratversicherung – jeweils zehnmal eingegeben, was zu insgesamt 180 Suchdurchläufen führte. Dieses Volumen spiegelt realistisches Nutzerverhalten wider, bei dem häufig mehrere Suchanfragen nötig sind, um vollständige und präzise Informationen zu erhalten.

Durch diese Abfragen wurden 5.851 eindeutige URLs identifiziert. Pro Suchdurchlauf ergaben sich im Schnitt 585 URLs, was die große Reichweite von Google unterstreicht. Das Format bleibt jedoch statisch: Nutzer:innen müssen mehrere Links durchklicken durchklicken und deren Inhalt selbst interpretieren, um relevante Informationen zu finden.

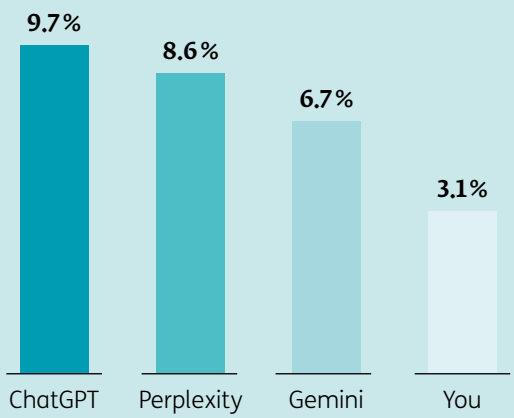
Verteilung der Additiv-URLs nach LLM-Suchmaschine

Prozentuale und absolute Werte mit Halluzination



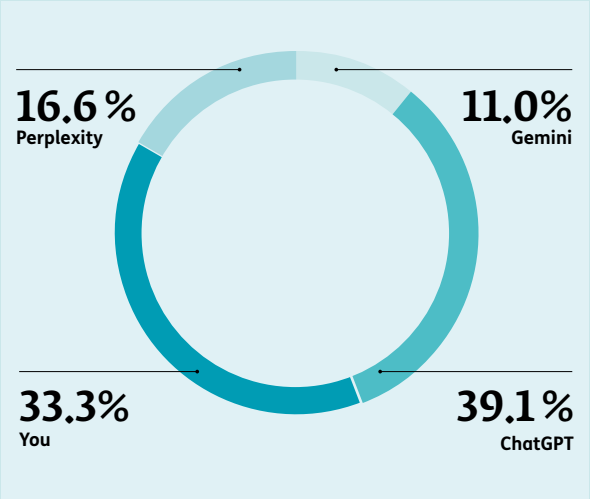
Halluzinationsrate – Anteil fehlerhafter oder erfundener Ergebnisse

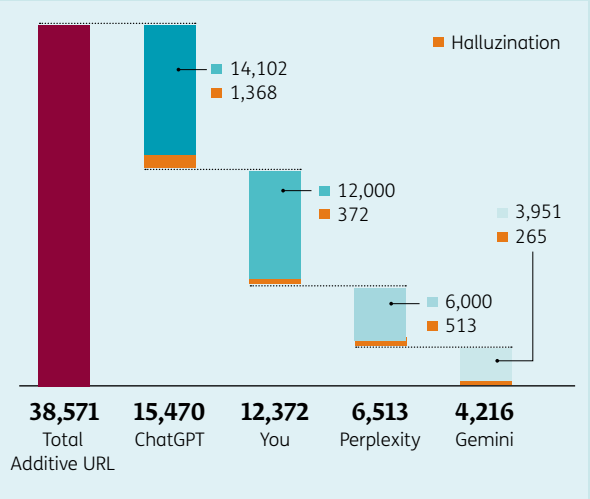
Im Verhältnis zur Gesamtzahl eingereichter URLs



Verteilung der Suchergebnisse nach Suchmaschinentyp

Prozentuale und absolute Werte mit Halluzinationen





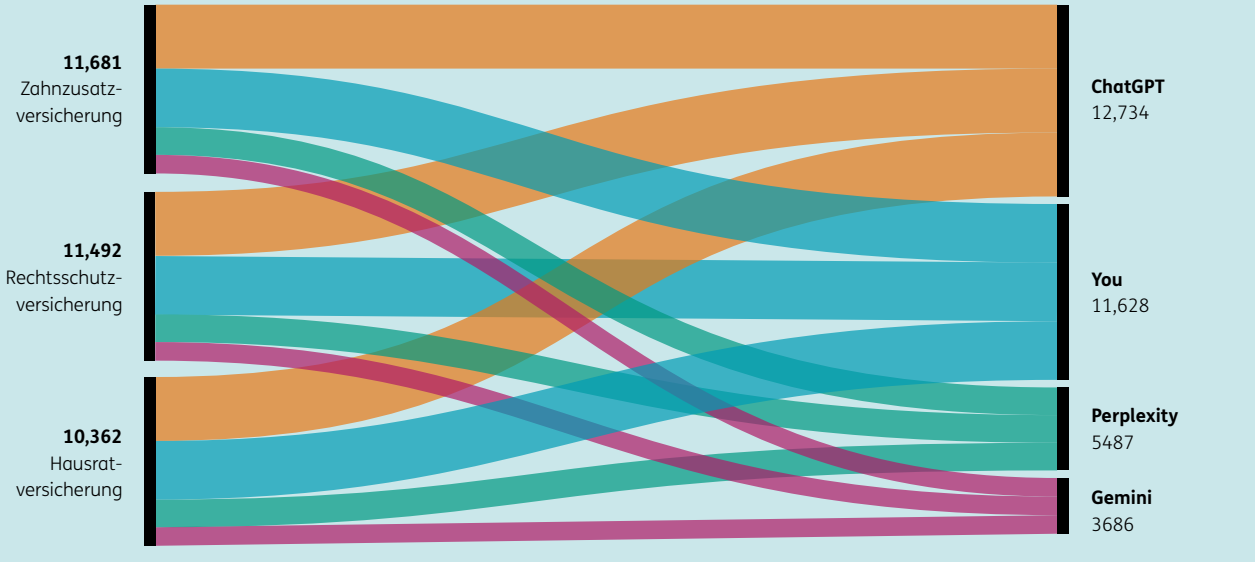
Die Rankinglogik basiert weiterhin auf Keyword-Übereinstimmung und der Stärke des Backlinkprofils, wodurch etablierte Domains und SEO-optimierte Inhalte bevorzugt angezeigt werden. Für allgemeine Informationsbedarfe und bekannte Anbieter funktioniert dieses Prinzip gut. Doch wer gezielt nach spezifischen Produkteigenschaften oder kontextbezogenen Vergleichen sucht, wird durch die Vielzahl an Links ausgebremst – die Informationsbeschaffung wird aufwendiger und weniger effizient.

Treffsicherheit und Halluzinationen in der LLM-gestützten Suche

Die Veränderung im Suchverhalten wird besonders deutlich beim Blick auf KI-gestützte Suchsysteme. In unserer Studie wurden 120 Prompts zu drei Versicherungsprodukten jeweils zehnmal an vier verschiedene LLM-Systeme übermittelt. Daraus ergaben sich insgesamt 27.446 URLs. Davon wurden jedoch 2.074 als Halluzinationen klassifiziert – also Inhalte, die fehlerhaft, irrelevant oder nicht existent waren. Das entspricht einer Halluzinationsrate von 7,6 Prozent.

Trotz erheblicher Fortschritte in der KI-gestützten Suche bleibt das Risiko von Halluzinationen ein zentrales Problem. ChatGPT wies mit 9,7 Prozent die höchste Halluzinationsrate auf, gefolgt von Perplexity mit 8,6 Prozent. Gemini lag bei 6,7 Prozent, während You.com mit nur 3,1 Prozent am besten abschnitt. Diese Unterschiede zeigen, dass die jeweilige Abrufarchitektur direkten Einfluss auf die Stabilität der Ergebnisse hat. Systeme mit größerer

Verteilung der abgerufenen URLs nach Produkt und LLM-Engine



Reichweite liefern zwar mehr Inhalte, gehen jedoch mit einem höheren Risiko einher. Plattformen mit vorsichtiger ausgelegter Abruflogik erzeugen weniger Halluzinationen, stellen jedoch höhere Anforderungen an die Strukturierung der Inhalte.

Für Versicherer bedeutet das: Inhalte müssen nicht nur auf inhaltliche Richtigkeit geprüft werden, sondern auch auf ihre Interpretierbarkeit durch KI-Systeme. Strukturierte Auszeichnungen, institutionelle Vertrauenssignale und faktische Klarheit senken das Risiko von Fehlinformationen und stärken das Vertrauen – sowohl bei den Nutzer:innen als auch bei den Systemen, die auf die Inhalte zugreifen.

Plattformspezifische Tendenzen verdeutlichen diese Muster zusätzlich. Gemini bevorzugt enge Sets hoch vertrauenswürdiger Quellen mit starker Präsenz von Maklerdomänen. Perplexity priorisiert redaktionell aufbereitete Inhalte etablierter Medien. SearchGPT kombiniert semantische Relevanz mit der Indexierung offener Datenquellen. You.com zeigt eine besonders breite Quellenbasis und bezieht Inhalte auch aus Blogs, Verbänden und anderen nicht traditionellen Formaten.

Da Nutzer:innen zunehmend auf mehrere KI-Werkzeuge parallel zugreifen, müssen Versicherer differenzierte Content-Strategien entwickeln, die gezielt auf die Architektur und Abruflogik der jeweiligen Systeme abgestimmt sind. Ein einheitlicher SEO-Ansatz ist in dieser Umgebung nicht mehr wirksam.

Deckungstiefe von LLMs variiert je nach Versicherungsprodukt

Die Auswertung der abgerufenen URLs in drei Produktkategorien – Zahnzusatzversicherung, Rechtsschutzversicherung und Hausratversicherung – zeigt deutliche Unterschiede im Verhalten der KI-gestützten Suchsysteme.

ChatGPT erzeugte das höchste Gesamtvolumen und zeigte besonders im Bereich der Zahnzusatzversicherung eine starke Abdeckung. You.com lag dicht dahinter mit ausgewogenen Ergebnissen über alle drei Produktlinien hinweg. Perplexity rief deutlich weniger URLs ab, zeigte aber ebenfalls gute Ergebnisse bei Zahnzusatzversicherungen. Gemini lag mit Abstand am niedrigsten, insbesondere bei Zahnzusatz und Hausrat. Das weist auf begrenzte Abdeckung durch die jeweilige Datenstruktur oder Abruflogik hin.

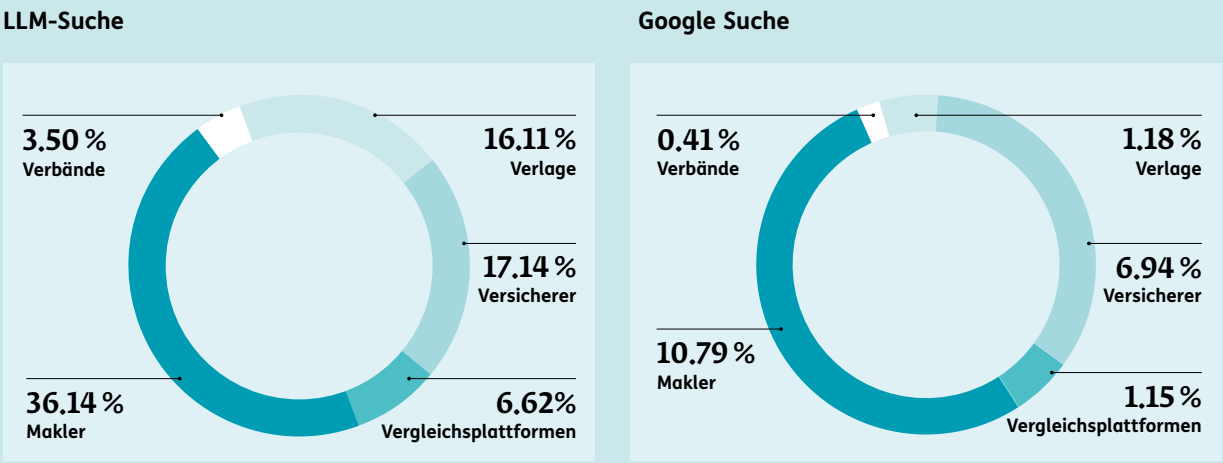
Diese Unterschiede zeigen, dass die technische Architektur einer Plattform nicht nur die Anzahl der Treffer bestimmt, sondern auch die Abdeckung einzelner Themenbereiche beeinflusst. Für Versicherer, die gezielt Sichtbarkeit für bestimmte Produkte aufbauen wollen, ist dies ein zentraler Faktor bei der Ausrichtung ihrer Inhalte auf KI-gestützte Suchsysteme.

Makler erzielen Sichtbarkeitsvorteile in der LLM-Suche

Unterschiedliche Content-Anbieter schneiden in KI-gestützten Suchsystemen sehr unterschiedlich ab, wenn es um markenoffene, versicherungsbezogene Anfragen geht. Makler, zu denen auch Vergleichsportale mit Maklerlizenz gehören, stellen mit 36 Prozent den größten Anteil, weit vor Versicherern mit 17 Prozent und Verlagen mit 16 Prozent. Redaktionelle Vergleichsplattformen ohne Maklerlizenz und Verbände folgen mit deutlich geringeren Anteilen. Diese Zahlen zeigen nicht nur veränderte Ranking-Ergebnisse, sondern auch eine neue Logik, wie Inhalte bewertet und angezeigt werden.

Eine zentrale Unterscheidung betrifft zwei Typen von Vergleichsplattformen. Makler, darunter auch große Vergleichsanbieter, dürfen aktiv Versicherungsverträge vermitteln und stellen meist strukturierte, szenarienbezogene Inhalte bereit, die auf Entscheidungsprozesse ausgelegt sind.

Verteilung nach Unternehmenstyp: LLM-Suche vs. Google



Vergleichsplattformen ohne Maklerstatus liefern hingegen rein redaktionelle oder informationelle Inhalte und verfügen nicht über eine Vermittlungslizenz. Diese Unterscheidung erklärt wesentliche Sichtbarkeitsunterschiede: Maklerplattformen passen besser zu den modularen, semantisch reichhaltigen Inhaltsformaten, die von LLM-Systemen bevorzugt werden.

Die Google-Suche zeigt ein anderes Bild. Die Sichtbarkeit von Maklern fällt dort auf unter 11 Prozent, Versicherer verbleiben bei rund 7 Prozent, während Verlage und Verbände nahezu unsichtbar sind. Das deutet darauf hin, dass klassische SEO-Signale wie Domainautorität, Backlinks und Keyword-Dichte weiterhin den Ausschlag geben. Google bevorzugt technisch optimierte Seiten, selbst wenn deren Inhalte weniger nutzerorientiert oder dynamisch sind. LLM-Systeme hingegen priorisieren Inhalte, die im Kontext interpretierbar, in Dialoge einbettbar und entscheidungsunterstützend nutzbar sind – Inhalte, die klar strukturiert, verlinkt, transparent und modular aufgebaut sind.

Aus dieser Differenz entsteht eine neue Sichtbarkeitslücke. Anbieter mit modularem und strukturiertem Content, insbesondere Makler, gewinnen in der LLM-Suche deutlich an Reichweite. Versicherer, die sich nicht anpassen, verlieren nicht, weil ihre Inhalte qualitativ schwach wären, sondern weil sie für ein System optimiert wurden, das nicht mehr den Takt vorgibt. Sichtbarkeit in beiden Welten erfordert heute zwei parallele Strategien – eine für die klassische Google Infrastruktur, die andere für KI native Auffindbarkeit.

Die Daten zeigen: In LLM-gestützten Umgebungen erzielen Maklerdomains durchgehend die höchste Sichtbarkeit, insbesondere bei markenoffenen Suchanfragen. Inhalte von Verlagen schneiden mäßig ab, Vergleichsplattformen ohne Maklerstatus bleiben deutlich zurück. Direktversicherer erreichen in diesen Systemen meist nur ein bis zwei Prozent der Ergebnisse. LLM-Systeme scheinen eine systematische Präferenz für Inhalte zu haben, die dialogfähig, explorativ und logisch gegliedert sind – Eigenschaften, die auf Makler- und Aggregatorseiten häufiger anzutreffen sind.

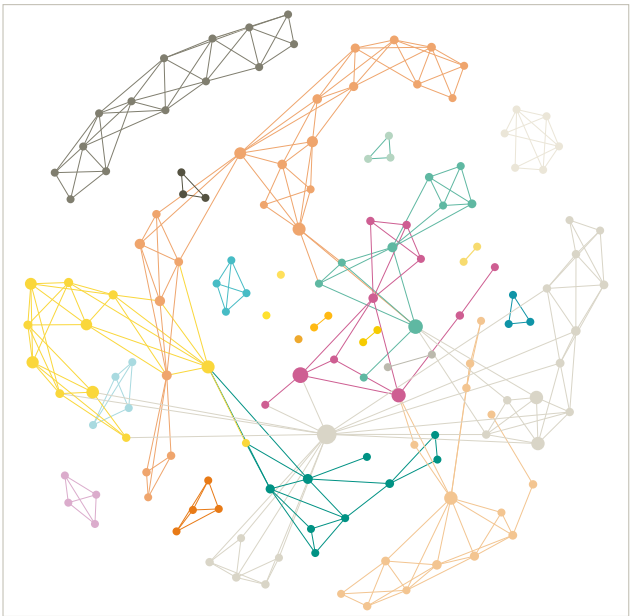
Die Verteilung bei Google ist flacher, aber starrer. Maklerdomains erscheinen mit etwa zwei Prozent, Direktversicherer liegen knapp dahinter bei rund einem Prozent. Die übrigen Treffer verteilen sich auf Verlage, Vergleichsplattformen und Fachverbände. Dieses Muster spiegelt Googles fortgesetzte Abhängigkeit von Ranking-Signalen wider, die LLM-Systeme zunehmend umgehen – etwa technische Infrastruktur, Backlinknetzwerke oder Autoritätsmetriken.

Makler sind strukturell gut aufgestellt, um von diesem Wandel zu profitieren. Ihre Plattformen basieren auf Produktvergleichen, modularer Darstellungslogik und entscheidungsrelevanten Nutzerpfaden. Diese Eigenschaften passen sehr genau zu den Funktionsweisen von LLM-Systemen, die Inhalte segmentieren, kontextuell bewerten und Bedeutung in verwendbare Ausgaben überführen. Mit dem Übergang zu agentischem Verhalten – also der Kombination von Recherche, Filterung und Entscheidungshilfe in einem Prozess – wird diese Struktur zum strategischen Vorteil.

Die Ergebnisse zeigen: Sichtbarkeit folgt einer neuen Logik. LLM-Systeme bevorzugen dialogorientierte, vernetzte und absichtsgesteuerte Inhalte. Diese Merkmale finden sich besonders häufig bei Makler- und Aggregatorseiten. Google hingegen bleibt fokussiert auf Backlinks, Domainautorität und Keyword-Abgleich. Das führt zu einer gleichmäßigeren, aber weniger kontextsensiblen Sichtbarkeitsverteilung.

Strukturelle Muster: Semantische Kohärenz und Analyse der Vektorräume

Ein zentrales Ergebnis dieser Studie ist der eindeutige Zusammenhang zwischen semantischer Struktur und Sichtbarkeit in LLM-Systemen. Inhalte, die intern konsistent aufgebaut, gut verlinkt und konzeptionell durchgängig sind, werden deutlich häufiger abgerufen und wiederverwendet. Visualisierungen des Embedding Verhaltens zeigen, dass fragmentierte oder schwach vernetzte Inhalte zu zerklüfteten Vektorräumen führen, was die Auffindbarkeit mindert.

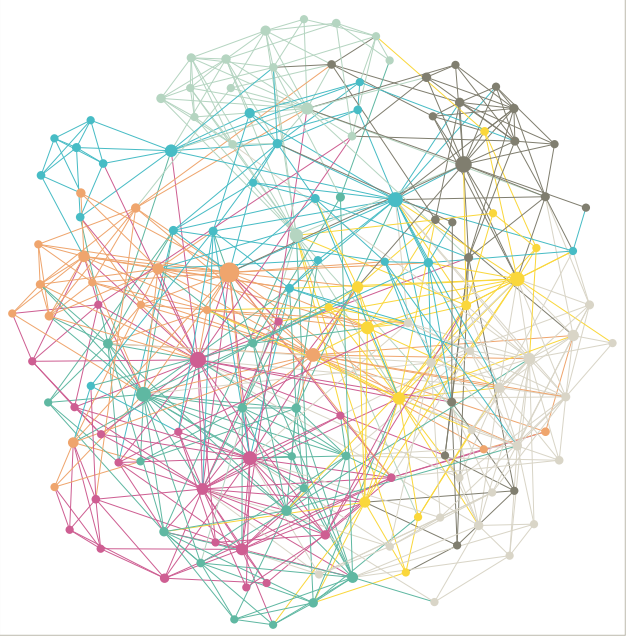


Semantische Fragmentierung – getrennte Cluster im Vektorraum stehen für semantisch isolierte oder strukturell schwache Inhalte. Diese Bereiche weisen aufgrund mangelnder Übereinstimmung mit Anfragevektoren eine geringe Abrufwahrscheinlichkeit auf.

Dichte, vernetzte Vektorcluster stehen dagegen in Zusammenhang mit stabilen Abrufmustern über verschiedene Systeme hinweg. Strukturell kohärente Inhalte, die durch Taxonomien, semantische Anker und logische Verknüpfungen gestützt sind, zeigen eine stärkere Übereinstimmung mit den Embedding-Modellen von LLMs. Die Ergeb-

nisse machen eine klare Trennlinie sichtbar zwischen fragmentierten und kohärenten Vektorräumen.

Diese Netzwerkanalyse zeigt einen fragmentierten Vektorraum. Getrennte Cluster und isolierte Knoten stehen für Inhalte, die keine interne Verlinkung oder keine einheitliche Terminologie aufweisen. Solche Einzelseiten oder Mikroseiten schneiden in LLM-Plattformen regelmäßig schlecht ab. Ihre strukturellen Lücken verhindern eine Übereinstimmung mit der auf Embeddings basierenden Abruflogik und führen zu geringer Sichtbarkeit (Wang et al. 2019; Borah et al. 2021).



Semantische Kohärenz – dichte, vernetzte Vektorräume spiegeln qualitativ hochwertige interne Verlinkung, strukturierte Semantik und stabiles Embedding Verhalten wider. Diese Cluster zeigen durchgängig höhere Abrufraten in der LLM-Suche.

Diese Befunde belegen, dass die Architektur von Inhalten direkten Einfluss auf ihre Performance hat. Seiten, die technisch sauber, aber semantisch unverbunden sind, bleiben für LLM-Systeme schwer auffindbar.

Der Aufbau starker semantischer Beziehungen zwischen Seiten ist deshalb kein optionales Zusatzmerkmal, sondern ein grundlegender Bestandteil für Sichtbarkeit in KI-gestützten Suchumgebungen.

Was Versicherungsinhalte in der LLM-Suche sichtbar macht: Validierung der Hypothesen

Um zu verstehen, welche Inhalte in KI-gestützter Suche sichtbar werden, wurde ein Bewertungsmodell mit 20 Optimierungskategorien entwickelt. Jede dieser Kategorien bildet ein strukturelles oder semantisches Kriterium ab, das für die Abruflogik KI-gestützter Systeme relevant ist. Dazu zählen maschinelle Lesbarkeit, Vertrauenssignale, semantische Kohärenz und eine Formatierung, die gängigen Prompt-Strukturen ähnelt.

Die vier Hypothesen repräsentieren übergeordnete Dimensionen, unter denen mehrere der 20 validierten Kriterien zusammengefasst sind. Die durchschnittlichen Gewichtungen ergeben sich aus der Aggregation der zugeordneten Einzelkriterien. So werden vier von 20 Kriterien sichtbar, die besonders stark zur Sichtbarkeit beitragen. Jede URL im Datensatz wurde dahingehend bewertet, wie stark sie die jeweiligen Kriterien erfüllt. Die Bewertung erfolgte in drei Stufen: mit Intervallen von 25 Prozent, 10 Prozent und 1 Prozent. Letztlich wurde das 1-Prozent-Raster gewählt, um auch feine Unterschiede zwischen ähnlichen Inhaltstypen sichtbar zu machen. Dieser hohe Detaillierungsgrad erwies sich als besonders hilfreich, da viele qualitativ hochwertige Domains in ihrer Performance nur leicht voneinander abweichen. Auf diese Weise konnten subtile, aber inhaltlich bedeutsame Unterschiede in Struktur und Interpretierbarkeit präzise erfasst werden.

Zusätzlich zur Punktevergabe (Skala von 0 bis 100) wurden Gewichtungsfaktoren auf die einzelnen Kriterien angewendet, um ihren relativen Einfluss auf das Retrieval-Verhalten abzubilden. Die Gewichtungen summieren sich auf 100 Prozent und erlauben eine Unterscheidung zwischen bloßer Präsenz eines Merkmals und tatsächlicher Relevanz für Sichtbarkeit. Die Analyse zeigt klare Leistungsunterschiede:

Maschinelle Lesbarkeit als Treiber der Sichtbarkeit

Die Daten belegen, dass technisch zugängliche, maschinenlesbare Inhalte eine zentrale Rolle für die Sichtbarkeit in LLM-gestützten Suchsystemen spielen. Diese Kategorie erreichte einen Maximalwert von 88 Prozent und einen Durchschnittswert von 85 Prozent über alle bewerteten Seiten hinweg. Sie gehört damit zu den leistungsstärksten Dimensionen. Mit einem Modellgewicht von durchschnittlich sechs Prozent zählt sie zugleich zu den einflussreichsten Faktoren im Bewertungssystem.

Diese Ergebnisse bestätigen Hypothese 1. LLM-Systeme bevorzugen eindeutig semantisch strukturiertes HTML, schnell ladende Seitenarchitekturen und transparente Layouts. Technische Merkmale wie ARIA-Auszeichnung, für Mobilgeräte optimiertes Design und klar markierte Überschriften sind heute keine optionalen Eigenschaften mehr. Sie bilden vielmehr die Grundvoraussetzung dafür, dass Inhalte von Sprachmodellen gelesen, verstanden und verwendet werden können.

Diese technische Grundlage ist oft Voraussetzung für weiterführende Merkmale wie semantische Verlinkung oder vertrauenswürdige Quellenkennzeichnung. Versicherer, die diese Standards bei Zugänglichkeit und Struktur nicht erfüllen, laufen Gefahr, aus den Retrieval-Prozessen von LLMs ausgeschlossen zu werden, selbst wenn die Inhalte inhaltlich korrekt und qualitativ hochwertig sind.

Semantische Kohärenz ermöglicht bessere Abrufbarkeit

Semantische und kontextuelle Verknüpfungen haben einen positiven Einfluss auf die Sichtbarkeit von Inhalten in LLM-gestützten Suchumgebungen. Inhaltstypen, die auf interne Struktur optimiert sind, zum Beispiel durch interne Links, strukturierte Metadaten und konzeptionelle Bezüge, erzielten Bewertungen zwischen 72 und 78 Prozent, mit einem Durchschnittswert von 75 Prozent. Das zugeordnete Modellgewicht von fünf Prozent unterstreicht die strategische Bedeutung dieser Dimension für die Sichtbarkeit.

Damit wird Hypothese 2 bestätigt. Semantisch kohärente Inhalte bilden stabilere, besser verknüpfte Cluster im Vektorraum. Dadurch steigt ihre Auffindbarkeit deutlich. Im Gegensatz dazu bleiben isolierte oder fragmentierte Inhalte wie einzelne Landingpages ohne konzeptionelle Verbindungen für LLM-Systeme meist unsichtbar.

Unsere Analyse zeigt: Inhalte, die in semantisch dichten, strukturell geschlossenen Clustern eingebettet sind, werden deutlich häufiger abgerufen. Moderne Embedding-Modelle bewerten die Nähe im Vektorraum und bevorzugen semantisch konsistente Knotenpunkte, wenn eine konzeptuelle Ausrichtung erkennbar ist (Wang et al. 2019; Borah et al. 2021).

Für Versicherer bedeutet das: Konsistente Taxonomien, klare interne Ankerpunkte und eine reichhaltige Verlinkung zwischen verwandten Inhalten sind heute essenziell, um in LLM-Systemen auffindbar zu bleiben. Integrität im Vektorraum, konzeptionelle Durchgängigkeit und interne Kohärenz müssen als grundlegende Elemente jeder Content-Strategie verstanden werden, die auf Sichtbarkeit in KI generierten Antworten abzielt.

Vertrauenssignale als zentraler Rankingfaktor

Unter allen bewerteten Dimensionen erzielten vertrauensbezogene Inhaltsmerkmale die stärksten Ergebnisse. Diese Kategorie erreichte einen maximalen Sichtbarkeitswert von 90 Prozent und einen Durchschnitt von 88 Prozent. Das sind die höchsten Werte im gesamten Bewertungssystem. Das durchschnittliche Gewicht von sechs Prozent bestätigt zusätzlich ihre Relevanz innerhalb des LLM-Rankings.

Diese Ergebnisse bestätigen Hypothese 3. LLM-Systeme bevorzugen systematisch Inhalte aus vertrauenswürdigen Quellen wie behördlichen Portalen, wissenschaftlichen Institutionen und regulierten Fachdomänen. Im Unterschied zu klassischen Suchsystemen, die stark auf Backlinks und Domaingröße setzen, orientieren sich LLMs an einer breiteren Palette von Vertrauensindikatoren. Dazu zählen überprüfbare Autor:innenschaft, institutionelle Konsistenz und nachvollziehbare Zitierstandards.

Die digitale Reputation muss sowohl technisch als auch redaktionell aktiv aufgebaut und gepflegt werden. Der gezielte Umgang mit Domainautorität, die klare Nennung von Verantwortlichen und die Einhaltung nachvollziehbarer Veröffentlichungsstandards entscheiden heute darüber, ob Inhalte durch KI-Systeme gefunden, eingeordnet und wiederverwendet werden. In einer LLM gestützten Suchumgebung ist Vertrauen keine Ergänzung, sondern eine Voraussetzung.

Prompt-ähnliche Strukturierung passt zur Logik von LLM-Systemen

Inhalte, die in prompt-ähnlichen oder dialogischen Formaten aufgebaut sind, zum Beispiel FAQs, Frage-Antwort-Seiten oder segmentierte Ratgeberinhalte, zeigen in LLM gestützten Umgebungen klare Vorteile bei der Auffindbarkeit. Diese Kategorie erzielte Bewertungen zwischen 67 und 75 Prozent, mit einem Durchschnitt von 72 Prozent und einem Modellgewicht von fünf Prozent.

Rankingfaktoren im Vergleich: Bewertung und Relevanzgewichtung

Hypothese	Zugeordnete Kategorie	Durchschnittliche Bewertung	Durchschnittliches Gewicht
H1	Technische Zugänglichkeit und Lesbarkeit	85%	6%
H2	Semantische und verlinkte Optimierung	75%	5%
H3	Authentizität und Vertrauenswürdigkeit	88%	6%
H4	Prompt-orientierte Optimierung für LLM-Systeme	72%	5%

Diese Ergebnisse bestätigen Hypothese 4. LLM-Systeme, die auf dialoglastigen Trainingsdaten und Frage-Antwort-Mustern beruhen, bevorzugen Inhalte, die dieser Eingabe-Antwort-Logik entsprechen. Besonders in Versicherungskontexten mit Bezug zu Produktvergleichen, Zugänglichkeitskriterien oder Servicepfaden erleichtert eine dialogische Struktur die Ausrichtung an den Verarbeitungsmustern von LLMs.

Organisationen, die ihre Inhaltsstruktur nicht an diese Logik anpassen, bleiben womöglich unterrepräsentiert, selbst wenn ihre Informationen korrekt und relevant sind. Der Einsatz modularer, prompt-orientierter Inhaltsformate ist kein optionaler Schritt mehr. Er ist ein strategischer Hebel, um in KI-generierten Ausgaben berücksichtigt zu werden..

Diese vier Dimensionen bilden zusammen eine klare strategische Grundlage für die Sichtbarkeit in der LLM-Suche.

Strategische Implikationen für die Versicherungsbranche für die LLM-Ära

Bei der Suche geht es nicht mehr darum, gefunden zu werden, sondern darum, von Maschinen genutzt werden zu können. Da LLMs die digitale Sichtbarkeit neu gestalten, reichen traditionelle Content-Strategien nicht mehr aus. Was als Nächstes kommt, erfordert neue Strukturen, neue Signale und eine radikale Veränderung in der Art und Weise, wie Versicherer für die Auffindbarkeit sorgen.

Die neue Logik der Sichtbarkeit

Der Wandel von der Keyword-basierten Suche hin zur semantischen Informationsabfrage verändert grundlegend, wie Inhalte strukturiert sein müssen, um gefunden zu werden. Sichtbarkeit bemisst sich heute danach, wie gut Informationen mit der internen Logik großer Sprachmodelle (LLMs) übereinstimmen.

Ein neues Paradigma der Sichtbarkeit

Der Aufstieg großer Sprachmodelle verändert, wie Nutzer auf Informationen zugreifen und mit ihnen interagieren. Klassische SEO-Strategien, basierend auf Keywords, Backlinks und Domain-Autorität, reichen nicht mehr aus. LLMs arbeiten nicht mit exakten Begriffen, sondern bewerten semantische Ähnlichkeit in Vektorräumen. Sichtbarkeit entsteht somit nicht mehr durch einzelne Suchbegriffe, sondern durch die Fähigkeit von Modellen, Inhalte semantisch zu analysieren, zu interpretieren und wiederzuverwenden.

Technische Grundlagen digitaler Sichtbarkeit

LLMs rufen Informationen nicht durch Keyword-Vergleich ab, sondern über semantische Ähnlichkeit mittels Vektor-Embeddings. Damit Inhalte sichtbar werden, müssen sie maschinenlesbar, semantisch strukturiert und technisch zugänglich sein. Sauberer HTML-Code, zugängliches Markup, schnelle Ladezeiten und vollständige Metadaten bilden das technische Fundament. Fehlen diese Voraussetzungen, kann selbst hochwertiger Content unsichtbar bleiben.

Um sich an Trainingsmuster von LLMs anzupassen, sollten Inhalte modular, interpretierbar und klar formuliert sein und über eine konsistente Terminologie und logische Struktur verfügen. Effektive Formate sind beispielsweise FAQ-Sektionen, Wikipedia-ähnliche Zusammenfassungen und nutzungsorientierte Inhaltsbausteine. Semantische Kohärenz über die gesamte Website hinweg, etwa durch einheitliche Begriffe, Themen-Verlinkung und definierte Taxonomien, erleichtert es Modellen, Inhalte als zusammenhängende Wissenscluster zu verarbeiten.

Ein hypothesenbasiertes Scoring-Modell zeigt eine klare Hierarchie der Sichtbarkeitstreiber im LLM-Umfeld: Maschinenlesbarkeit als Voraussetzung, semantische Verlinkung für Konsistenz in den Embeddings, Vertrauenssignale wie Autorenschaft und Glaubwürdigkeit sowie modulare, prompt-kompatible Formate wie FAQs, Bullet Points oder strukturierte Szenarien. Solcher Content wird nicht nur indiziert, er wird Teil des logischen Verarbeitungsprozesses der Modelle.

Relevante Einflussfaktoren für LLM-basierte Sichtbarkeit

Auffindbarkeit setzt semantische Klarheit, strukturelle Konsistenz und eingebettetes Vertrauen voraus. Diese Elemente bilden die zentralen Stellhebel für LLM-orientierte Sichtbarkeit.

Vertrauen, Markenstruktur und Metadaten

LLMs priorisieren zunehmend Inhalte, die transparent, gut belegt und reputationsstark sind. Vertrauenssignale wie klare Autorenschaft, redaktionelle Verantwortung und Verweise auf regulatorische Instanzen sind entscheidend. Auch die Markenidentität muss technisch verankert sein, etwa durch konsistente Namensgebung und Messaging, den Einsatz von Metadatenstandards wie schema.org oder JSON-LD sowie fortgeschrittene Verfahren wie Knowledge Graphs oder synthetische Trainingsdaten.

Sicherstellung der strukturellen Integrität der Vektorräume

Dauerhafte Sichtbarkeit erfordert die gezielte Schließung sogenannter Vektorlücken, also Fälle, in denen Inhalte aufgrund unpräziser Formulierungen, unvollständiger Erklärungen oder schwacher semantischer Struktur nicht mit relevanten Anfragen übereinstimmen. Diese Lücken lassen sich durch Monitoring von LLM-Ausgaben, gezielte Textverfeinerung, stärkere semantische Verknüpfung und eine optimierte interne Inhaltsstruktur (Taxonomie) schließen.

Performance-Monitoring und operative Integration

Nach dem Schließen von Vektorlücken müssen Unternehmen die tatsächliche Performance ihrer Inhalte in LLM-Umgebungen analysieren. Klassische Traffic-Kennzahlen reichen nicht mehr aus. Relevant sind: Abrufhäufigkeit durch Modelle, Einbindung in LLM-generierte Antworten, Konsistenz der Markenerwähnung und -präsenz in dialogischen Multi-Turn-Interaktionen. Prompt-Tests, Plattformvergleiche und Halluzinationsanalysen identifizieren strukturelle Schwächen und liefern darüber hinaus wichtige Impulse für iterative Optimierungsprozesse.

Plattform- und Wettbewerbsdifferenzierung

Nicht alle LLMs verarbeiten Inhalte gleich, und nicht alle Marktakteure sind gleich gut vorbereitet. Strukturelle Angleichung und plattformindividuelle Optimierung bieten einen entscheidenden Wettbewerbsvorteil in dieser neuen Landschaft.

LLM- & plattformindividuelle Optimierung

Ein zentrales Ergebnis der Analyse ist die unterschiedliche Performance auf verschiedenen LLM-Plattformen. ChatGPT und You.com rufen zwar große Mengen an URLs ab, weisen jedoch höhere Halluzinationsraten auf. Gemini und Perplexity liefern stärker kuratierte, quellengefilterte Ergebnisse. Jede Plattform reagiert auf eigene Inhaltsmerkmale. Eine universelle Optimierungsstrategie existiert nicht; Sichtbarkeit muss plattformspezifisch geplant, getestet und umgesetzt werden. Versicherer müssen analysieren, wie ihre Angebote auf jeder Plattform performen und gezielt darauf reagieren.

Strukturierte Inhalte performen besser

Unabhängig von der Plattform zeigen bestimmte Strukturmuster durchgehend bessere Auffindbarkeit. Technisch optimierte, semantisch strukturierte Inhalte schneiden über alle getesteten Systeme hinweg deutlich besser ab. Seiten mit modularen Formaten, etwa FAQs, strukturierte Leistungsbeschreibungen oder klar definierte Antwortbausteine, werden häufiger und verlässlicher abgerufen als lange, markenzentrierte Fließtexte. Es geht nicht um Stil, sondern um strukturelle Anpassung an die Art und Weise, wie LLMs Inhalte analysieren, priorisieren und rekonstruieren.

Warum Makler besser abschneiden

Ein konsistentes Ergebnis der Untersuchung: Makler und Vergleichsportale erzielen deutlich höhere Sichtbarkeit, unabhängig von Produktkategorie und Plattform. Ursache ist ihre Inhaltsstruktur: vergleichsorientiert, klar segmentiert und stark verlinkt. Diese Merkmale sind optimal auf das Abrufverhalten von LLMs abgestimmt. Demgegenüber stehen viele Versicherungsseiten mit unstrukturiertem Text, komplexer Navigation und fehlerhaftem Markup. Ohne strukturelle Transformation werden Versicherer Reichweite an besser adaptierte Intermediäre verlieren.

Operative Vorbereitung auf LLM-Integration

Damit Inhalte in KI-Ökosystemen nutzbar sind, müssen Unternehmen ihre Fähigkeiten neu denken, Services über Application Programming Interface (APIs) zugänglich machen und technische, strategische sowie organisatorische Lücken schließen.

Die agentenbasierte Customer Journey

Mit dem Aufstieg agentischer KI entstehen neue Interaktionsmodelle. LLM-gestützte Agenten gehen über reine Informationsbereitstellung hinaus, sie vergleichen Produkte, erstellen Angebote in Echtzeit oder führen Aufgaben autonom aus. Aktuell sind Versicherungsservices in Deutschland nicht mit solchen Systemen kompatibel.

APIs für Angebote, Vertragsinformationen oder Preislogik sind meist nicht vorhanden oder für externe Modelle unzugänglich. Dadurch bleiben Versicherer von den entstehenden agentenbasierten Nutzerreisen ausgeschlossen – ein strategisch bedeutsames Defizit, das technisches wie regulatorisches Handeln erfordert.

Strategischer Wendepunkt für Versicherer

Große Sprachmodelle verändern nachhaltig, wie digitale Services gefunden, verstanden und genutzt werden. Für die Versicherungswirtschaft markiert dies einen strategischen Wendepunkt. Makler und Aggregatoren sind strukturell besser aufgestellt – ihre Inhalte sind modular, semantisch verlinkt und technisch abrufbar. Mit dem Aufkommen agentischer Systeme wird sich die Kluft zwischen auffindbaren und unsichtbaren Angeboten weiter vertiefen. Gleichzeitig ist klar: Die Treiber dieser neuen Sichtbarkeit sind nicht mehr hypothetisch, sondern beobachtbar, messbar und gezielt beeinflussbar.

Versicherer, die ihre Services über APIs zugänglich machen und Inhalte modular sowie LLM-gerecht aufbauen, können ihre digitale Sichtbarkeit zurückgewinnen und ihre Rolle in KI-basierten Kundeninteraktionen aktiv mitgestalten. Wer hingegen zögert, riskiert, die Kontrolle darüber zu verlieren, wie, wann und ob die eigenen Produkte Nutzern überhaupt präsentiert werden.

Strategische Implikationen für eine veränderte Sichtbarkeitslogik

Die Studie bestätigt einen strukturellen Wandel in der Art und Weise, wie Inhalte aus dem Versicherungsbereich von KI-basierten Systemen entdeckt, abgerufen und weiterverwendet werden. Herkömmliche SEO-Signale wie Keyword-Platzierung, Backlinks oder Domain-Autorität sind nicht länger ausreichend, um digitale Sichtbarkeit sicherzustellen. Stattdessen wird Auffindbarkeit zunehmend davon bestimmt, wie gut Inhalte mit der Abruflogik großer Sprachmodelle übereinstimmen. Dazu gehören Maschinenlesbarkeit, semantische Kohärenz, Vertrauenswürdigkeit und eine auf Konversation ausgerichtete Formatierung.

Organisationale Fähigkeiten und interdisziplinäre Zusammenarbeit

Die Erfüllung dieser neuen Anforderungen wird nicht allein durch Marketing oder klassische SEO-Maßnahmen möglich sein. Sichtbarkeit in LLM-Umgebungen muss zu einer gemeinsamen Aufgabe werden, getragen von redaktionellen, technischen, regulatorischen und produktbezogenen Teams. Unternehmen müssen interne Prozesse etablieren, um Prompts zu testen, semantische Scorings zu erstellen, Abrufverhalten zu beobachten und mit neuen Inhaltsformaten zu experimentieren. Dazu gehört auch die Neugestaltung zentraler Seiten, die Angleichung von Metadaten an KI-Systeme sowie die technische Erschließung strukturierter Schnittstellen für Dienstleistungen, die gefunden oder automatisiert verarbeitet werden sollen.

Darüber hinaus müssen sich auch die internen Kompetenzen weiterentwickeln. Die meisten Marketing- und SEO-Teams wurden auf die Logik klassischer Suchmaschinen ausgelegt. Dieses Wissen bleibt wertvoll, muss jedoch um ein tiefes Verständnis dafür ergänzt werden, wie LLMs Inhalte abrufen, interpretieren und rekonstruieren. Dazu gehört Wissen über Vektorsemantik, Verhalten in Embedding-Räumen, Prompt Engineering und KI-spezifische Leistungsmetriken. Dies kann durch gezielte Weiterqualifizierung, funktionsübergreifende Zusammenarbeit und spezifisches Training erreicht werden.

Von Sichtbarkeit zu Abrufbarkeit

Erfolg im Bereich KI-gesteuerter Sichtbarkeit wird durch die koordinierte Zusammenarbeit redaktioneller, technischer und strategischer Einheiten bestimmt. Es braucht gemeinsame Taxonomien, konsistente Inhaltsmodelle und ein neues Verständnis von Leistung. In dieser Umgebung hängt Sichtbarkeit nicht nur von der Qualität der Inhalte ab, sondern davon, wie diese strukturiert, verlinkt und maschinell interpretierbar gestaltet sind.

Marketing bedeutet in dieser neuen Realität nicht länger nur Sichtbarkeit – es bedeutet Abrufbarkeit. Die Organisationen, die verstehen, wie LLMs Inhalte lesen, bewerten und wiederverwenden, werden die nächste Generation digitaler Nutzererlebnisse gestalten. Jetzt ist der Zeitpunkt, um diese Fähigkeiten von innen heraus aufzubauen.

Ausblick: Die Zukunft der LLM-basierten Auffindbarkeit

Mit der Weiterentwicklung von LLMs definieren neue Kräfte wie autonome Agenten, innovative Monetarisierung und kuratierte Content-Partnerschaften neu, wie auf Informationen zugegriffen wird. Diese Dynamiken erfordern, dass Unternehmen ihre Strategien und Infrastruktur überdenken, um in einer KI-gesteuerten Landschaft relevant und sichtbar zu bleiben.

Während sich diese Studie auf die strukturellen Grundlagen der Sichtbarkeit in LLM-Umgebungen konzentriert hat, zeichnen sich bereits neue Dynamiken ab. Sie entstehen durch veränderte Interaktionsmodelle, neue Formen der Monetarisierung und eine zunehmende Plattformkontrolle. Auch wenn diese Entwicklungen nicht Teil der empirischen Untersuchung waren, sind sie für alle Unternehmen relevant, die langfristig sichtbar bleiben wollen.

Strategische Inhaltsallianzen

Parallel zum Wandel der Suche verändert sich auch das Modell der Inhaltsverteilung. Das frühere Prinzip der offenen Indexierung durch Suchmaschinen wird zunehmend durch kuratierte Integration ersetzt. Führende Anbieter von Sprachmodellen, darunter OpenAI und Mistral, gehen inzwischen direkte Partnerschaften mit Verlagen und Datenanbietern ein. Beispiele sind die Vereinbarungen mit Axel Springer oder die Unterstützung durch Medienkonzerne wie Bertelsmann. Dabei geht es nicht mehr nur um das Crawlen öffentlich zugänglicher Inhalte, sondern um die gezielte Integration definierter Datensätze, die als hochwertig, vertrauenswürdig und modellrelevant gelten.

Sichtbarkeit wird damit zunehmend verhandelt und nicht mehr ausschließlich verdient. Unternehmen, die auf organischen Traffic oder klassische SEO-Strategien setzen, müssen ihre Distributionslogik überdenken. Indexiert zu sein reicht nicht mehr aus. Wer Teil der LLM-gestützten Antwortebene bleiben will, muss Möglichkeiten zur Lizenzierung, direkten Anbindung oder Partnerschaften prüfen, um Inhalte in den Prompt Stack oder die Entscheidungslogik der Modelle einzubringen. Entstehen wird ein fragmentiertes, aber strategisch steuerbares Entdeckungsökosystem, in dem Business Development und Inhaltsintegration genauso wichtig sind wie technische Optimierung der Inhalte.

Der Aufstieg agentischer Systeme

Eine der folgenreichsten Veränderungen ist das Aufkommen agentischer Nutzungserlebnisse. Anstatt dass Nutzer:innen aktiv Suchanfragen stellen und Ergebnisse auswählen, übernehmen LLM-gestützte Systeme diese Aufgaben zunehmend selbstständig. Solche Operatoren oder autonomen Agenten interpretieren natürliche Sprache und führen mehrstufige Aufgaben aus. Dazu zählen das Buchen von Reisen, das Zusammenfassen von Berichten oder das Vergleichen von Produkten – gestützt auf interne Entscheidungslogiken und externe Werkzeuge.

Auf der Entwicklerkonferenz Google I/O 2025 wurde dieses Prinzip in Form eines speziellen Check-out Agenten vorgestellt. Dieser entdeckt relevante Angebote, navigiert eigenständig durch E-Commerce Prozesse und schließt Transaktionen über Google Pay ab, ohne dass Nutzer:innen eingreifen müssen (Google, 2025). Auch im Versicherungsbereich könnten solche Agenten eingesetzt werden, etwa um Angebotsstrecken vollständig zu durchlaufen, Zahlungen abzuwickeln oder Policen direkt in Konversationsoberflächen auszustellen.

Dieser Wandel markiert den Abschied vom klassischen Such-Funnel. Sichtbarkeit bedeutet nicht mehr, in einer Trefferliste aufzutauchen. Sichtbarkeit heißt, innerhalb der Logik eines Agenten handlungsfähig und interoperabel zu sein. Für Unternehmen bedeutet das: Produkte, Services und Inhalte müssen so strukturiert sein, dass sie innerhalb dieser agentischen Prozesse erkannt, verwendet und weiterverarbeitet werden können. Maschinenlesbare Formate, API-Zugang und eindeutig definierte Schnittstellen werden zu Voraussetzungen für die Teilnahme an LLM-gestützten Nutzerreisen. Wer nicht ansteuerbar ist, wird möglicherweise gar nicht mehr berücksichtigt.

Versicherer müssen standardisierte Schnittstellen für Angebote und Schadenprozesse schaffen, über die KI-Systeme Deckungsoptionen abrufen, Vergleiche anstoßen

und Transaktionen auslösen können – ohne menschliches Eingreifen.

Während sich KI-Systeme von der reinen Informationsvermittlung hin zur aktiven Ausführung entwickeln, genügen statische Webseiten nicht mehr. Angebote, Schadenmeldungen und Produktvergleiche müssen in maschinenlesbarer Form vorliegen, damit sie von LLMs verarbeitet werden können. Künftig zählen nicht nur Inhalte, die informieren, sondern solche, die handeln ermöglichen. In der Praxis heißt das: strukturierte Produktmetadaten, modulare Leistungslogiken und vertragsrelevante Textbausteine, die KI-Agenten dynamisch zusammenstellen, vergleichen und aktiv nutzen können.

Agentische KI und die Zukunft der Versicherungswirtschaft

Die Integration agentischer KI wird die Wettbewerbsdynamik im Versicherungsmarkt grundlegend verändern. In dem Maße, in dem autonome Systeme bestimmen, wie Nutzer:innen Services entdecken und auswählen, müssen sich Versicherer innerhalb der Entscheidungslogik KI-gestützter Systeme positionieren.

Sichtbarkeit wird künftig nicht mehr über direkte Kundengewinnung definiert, sondern darüber, wie gut ein Produkt von Maschinen entdeckt, interpretiert und verarbeitet werden kann. In diesem Umfeld wird strategische Auffindbarkeit zur Funktion von Datenqualität, Schnittstellengestaltung und semantischer Klarheit.

Dieser Wandel hat direkte Auswirkungen auf die Geschäftsmodelle der Versicherungswirtschaft. Statische Angebote und lineare Vertriebstrichter sind nicht auf das Verhalten autonomer Agenten abgestimmt. Stattdessen muss die Produktarchitektur modular aufgebaut sein, sodass Agenten auf Basis des aktuellen Kontexts und der Nutzerabsicht passgenaue Pakete konfigurieren und empfehlen können.

Preislogiken müssen künftig in der Lage sein, dynamische Eingaben, Ereignistrigger und episodische Risikobewertungen zu verarbeiten. Auch die Rolle des Menschen wird sich entsprechend verändern: Berater:innen agieren nicht mehr nur als Auslöser von Transaktionen, sondern als strategische Ansprechpartner:innen für komplexe, nicht automatisierbare Situationen. Erfolgreiche Anbieter:innen etablieren plattformkompatible Strukturen, die sich nahtlos in agentisch gesteuerte Ökosysteme integrieren lassen.

Auf Prozessebene ermöglichen agentische Systeme den Übergang von reaktiver Servicebereitstellung hin zu proaktiver Orchestrierung. Im Vertrieb werden KI-Agenten relevante Produkte auffinden, Nutzer:innen vorqualifizieren und den Abschluss vorantreiben. Im Kundenservice übernehmen sie Routinetätigkeiten, verbessern die Schadenabwicklung und identifizieren Deckungslücken.

Damit wird Interaktion neu definiert. Aus punktuellen Transaktionen entstehen kontinuierliche Optimierungsschleifen. Für Versicherer bedeutet das mehr als die Bereitstellung von APIs. Es braucht eine vollständige Abstimmung zwischen interner Logik und externen Entscheidungssystemen. Nur diejenigen, die ihre Produkte in die operativen Rahmenwerke KI-gestützter Agenten einbetten, werden im nächsten Kapitel digitaler Distribution relevant bleiben.

Werbung in LLM-Systemen: Merchant Programme und AI Native Ads

Während LLM-Systeme klassische Suchmaschinen zunehmend ersetzen oder ergänzen, entstehen neue Monetarisierungsmodelle. Plattformen wie Perplexity testen derzeit eingebettete Werbeformate, darunter visuelle Anzeigen im Stil von Bannern sowie sogenannte Merchant Programme, die direkt in konversationelle Antwortformate eingebettet sind (Perplexity AI 2024; Bastian 2024).

Im Unterschied zu klassischen Anzeigen liegt bei diesen Formaten der Fokus auf kontextbezogener Relevanz, Ausrichtung auf Nutzerabsichten und qualitativem Nutzererlebnis. In einer Umgebung, in der ein Sprachmodell nicht nur Antworten liefert, sondern vollständige Informationsverläufe konstruiert, muss sich Werbung anpassen. Die Herausforderung für Marken besteht darin, neue kreative und gezielte Strategien zu entwickeln, mit denen Produkte und Botschaften natürlich innerhalb von KI generierten Inhalten erscheinen können. Dies kann beinhalten, erste Werbeformate auf neuen LLM-Plattformen zu testen, Produktdaten semantisch für die maschinelle Integration umzustrukturieren oder mit Agenturen zusammenzuarbeiten, die auf Prompt Design und KI-gestützte Medien spezialisiert sind. Ziel ist nicht allein Sichtbarkeit, sondern Relevanz zu schaffen. Markenbotschaften sollen nicht nur erscheinen, sondern sich nahtlos in die vertrauenswürdigen Informationsflüsse einfügen, auf die Nutzer:innen zurückgreifen.

Vorbereitung auf das, was kommt

Während sich KI-Systeme weiterentwickeln, müssen Unternehmen über klassische Optimierung hinausdenken. Sie müssen beginnen, Infrastruktur, Partnerschaften und Datenstrategie auf die neue Logik der maschinellen Auffindbarkeit auszurichten. Die hier skizzierten Entwicklungen markieren nicht nur die nächste Welle der Disruption, sondern auch die nächste strategische Chance. Organisationen, die jetzt handeln, definieren aktiv, wie sie künftig von Maschinen gefunden, interpretiert und verwendet werden.

Autor:innen

Luisa-Marie Schmolke
Innovation Developer
ERGO Group AG

Hamidreza Hosseini
Founder & CEO
ECODYNAMICS GmbH

Danksagung

Dieses Whitepaper wurde mit wertvoller Unterstützung, konstruktiven Ideen und fachlichem Feedback zahlreicher weiterer Mitwirkender erstellt. Wir bedanken uns herzlich bei allen, die mit ihrem technischen Fachwissen, sorgfältigem Lektorat und konstruktiven Rückmeldungen maßgeblich zur Entstehung und Weiterentwicklung dieses Whitepapers beigetragen haben.

Hendrik Schulze Bröring (Senior Partner bei ECODYNAMICS GmbH) leitete die Dateninfrastruktur und -erhebung. Diese Arbeit bildete die Grundlage für die in diesem Whitepaper dargestellten Analysen und Erkenntnisse.

Besonderer Dank gilt

Georgina Neitzel
Leierin des ERGO Innovation Lab
ERGO Group AG

Larissa Selzer
SEO Expertin
ERGO Group AG

Literaturverzeichnis

Achiam, J., Adler, S., Agarwal, S., Ahmad, L., Akkaya, I., Aleman, F. L., ... & McGrew, B. (2023). Gpt-4 technical report. arXiv preprint arXiv:2303.08774.

Bastian, M. (2024, November 14). Perplexity AI launches ad program with sponsored follow-up questions. THE DECODER. https://the-decoder.com/perplexity-ai-launches-ad-program-with-sponsored-follow-up-questions/

Borah, A., Barman, M. P., & Awekar, A. (2021, August). Are word embedding methods stable and should we care about it? In Proceedings of the 32nd ACM Conference on Hypertext and social media (pp. 45-55).

Borgesius, F. Z., van Bekkum, M., van Ooijen, I., Schaap, G., Harbers, M., & Timan, T. (2025). Discrimination and AI in insurance: what do people find fair? Results from a survey. arXiv preprint arXiv:2501.12897.

Chowdhury, G., & Chowdhury, S. (2024). AI-and LLM-driven search tools: A paradigm shift in information access for education and research. Journal of Information Science, 01655515241284046.

Creswell, J. W., & Clark, V. L. P. (2017). Designing and conducting mixed methods research. Sage publications.

Denzin, N. K. (2017). The research act: A theoretical introduction to sociological methods. Routledge.

Eisenbrand, R. (2025, March 5). "Übersicht mit KI": Google testet AI Overviews in Europa – auch in Deutschland. OMR. https://omr.com/de/daily/google-ai-overviews-uebersicht-mit-ki-deutschland

Eleni Spatharioti, S., Rothschild, D. M., Goldstein, D. G., & Hofman, J. M. (2023). Comparing Traditional and LLM-based Search for Consumer Choice: A Randomized Experiment. arXiv e-prints, arXiv-2307

Gallegos, I. O., Rossi, R. A., Barrow, J., Tanjim, M. M., Kim, S., Dernoncourt, F., ... & Ahmed, N. K. (2024). Bias and fairness in large language models: A survey. Computational Linguistics, 50(3), 1097-1179.

Gartner, Inc. (2023). Predicts 2024: How GenAI will reshape tech marketing. https://www.gartner.com/en/doc/800771-predicts-2024-how-genai-will-reshape-tech-marketing

GDV. (2024, July 31). Digitaler Versicherungsvertrieb wächst deutlich. GDV. https://www.gdv.de/gdv/medien/medieninformationen/versicherung-vertrieb-abschluesse-digital-181036

Google. (2025, May 14). Google I/O 2025 Keynote. https://io.google/2025/

Google. (n.d.). How Google Search Algorithms Works. https://www.google.com/intl/en_us/search/howsearchworks/how-search-works/ranking-results/

Google Search Central. (n.d.). Google for Developers. https://developers.google.com/search/docs/appearance/page-experience

Hagey, K., Kruppa, M., Bruell, A., & iStock, E. L. W. S. J. (2023, December 14). News publishers see Google's AI search tool as a Traffic-Destroying nightmare. WSJ. https://www.wsj.com/tech/ai/news-publishers-see-googles-ai-search-tool-as-a-traffic-destroying-nightmare-52154074

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design science in information systems research. MIS quarterly, 75-105.

Iqbal, M., Khalid, M. N., Manzoor, A., Abid, M. M., & Shaikh, N. A. (2022). Search engine optimization (seo): A study of important key factors in achieving a better search engine result page (serp) position. Sukkur IBA Journal of Computing and Mathematical Sciences, 6(1), 1-15.

Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. Nature machine intelligence, 1(9), 389-399.

Joko, H., Chatterjee, S., Ramsay, A., De Vries, A. P., Dalton, J., & Hasibi, F. (2024, July). Doing personal laps: Llm-augmented dialogue construction for personalized multi-session conversational search. In Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval (pp. 796-806).

Karamolegkou, A., Li, J., Zhou, L., & Søgaard, A. (2023). Copyright violations and large language models. arXiv preprint arXiv:2310.13771.

Karpukhin, V., Oguz, B., Min, S., Lewis, P. S., Wu, L., Edunov, S., ... & Yih, W. T. (2020, November). Dense Passage Retrie-

val for Open-Domain Question Answering. In EMNLP (1) (pp. 6769-6781).

Kumar, A., & Lakkaraju, H. (2024). Manipulating large language models to increase product visibility. arXiv preprint arXiv:2404.07981.

Li, J., Zhang, J., Li, H., & Shen, Y. (2024). An Agent Framework for Real-Time Financial Information Searching with Large Language Models. arXiv preprint arXiv:2502.15684.

Liu, D., Chen, M., Lu, B., Jiang, H., Han, Z., Zhang, Q., ... & Qiu, L. (2024). Retrievalattention: Accelerating long-context llm inference via vector retrieval. arXiv preprint arXiv:2409.10516.

Liu, L., Meng, J., & Yang, Y. (2024). LLM technologies and information search. Journal of Economy and Technology, 2, 269-277

Liu, N. F., Zhang, T., & Liang, P. (2023). Evaluating verifiability in generative search engines. arXiv preprint arXiv:2304.09848.

Liu, Y., Deng, G., Li, Y., Wang, K., Wang, Z., Wang, X., ... & Liu, Y. (2023). Prompt Injection attack against LLM-integrated Applications. arXiv preprint arXiv:2306.05499

March, S. T., & Smith, G. F. (1995). Design and natural science research on information technology. Decision support systems, 15(4), 251-266.

Mialon, G., Dessì, R., Lomeli, M., Nalmpantis, C., Pasunuru, R., Raileanu, R., ... & Scialom, T. (2023). Augmented language models: a survey. arXiv preprint arXiv:2302.07842.

Mo, F., Mao, K., Zhao, Z., Qian, H., Chen, H., Cheng, Y., ... & Nie, J. Y. (2024). A survey of conversational search. arXiv preprint arXiv:2410.15576.

Monir, S. S., Lau, I., Yang, S., & Zhao, D. (2024). VectorSearch: Enhancing Document Retrieval with Semantic Embeddings and Optimized Search. arXiv preprint arXiv:2409.17383.

Moz. (2024, November 25). What are On-Page ranking factors for SEO? Moz. https://moz.com/learn/seo/on-page-factors

Nakano, R., Hilton, J., Balaji, S., Wu, J., Ouyang, L., Kim, C., ... & Schulman, J. (2021). Webgpt: Browser-assisted question-answering with human feedback. arXiv preprint arXiv:2112.09332.

Nestaas, F., Debenedetti, E., & Tramèr, F. (2024). Adversarial search engine optimization for large language models. arXiv preprint arXiv:2406.18382.

Nie, G., Zhi, R., Yan, X., Du, Y., Zhang, X., Chen, J., ... & Hu, J. (2024, October). A hybrid multi-agent conversational recommender system with llm and search engine in e-commerce. In Proceedings of the 18th ACM Conference on Recommender Systems (pp. 745-747).

Pan, R., Cao, B., Lin, H., Han, X., Zheng, J., Wang, S., ... & Sun, L. (2024). Not all contexts are equal: Teaching llms credibility-aware generation. arXiv preprint arXiv:2404.06809.

Perplexity AI. (2024, November 12). Why we're experimenting with advertising. Perplexity AI. https://www.perplexity.ai/hub/blog/why-we-re-experimenting-with-advertising

Schick, T., Dwivedi-Yu, J., Dessì, R., Raileanu, R., Lomeli, M., Hambro, E., ... & Scialom, T. (2023). Toolformer: Language models can teach themselves to use tools. Advances in Neural Information Processing Systems, 36, 68539-68551.

Schneider, P., Poelman, W., Rovatsos, M., & Matthes, F. (2024). Engineering conversational search systems: A review of applications, architectures, and functional components. arXiv preprint arXiv:2407.00997.

Schwartz, B. (2025, March 20). Microsoft confirms Schema helps its LLMS (Copilot) understand your content. Search Engine Roundtable. https://www.seroundtable.com/schema-llms-copilot-bing-microsoft-39093.html

Sharma, N., Liao, Q. V., & Xiao, Z. (2024, May). Generative echo chamber? effect of llm-powered search systems on diverse information seeking. In Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems (pp. 1-17).

Shi, X., Liu, J., Liu, Y., Cheng, Q., & Lu, W. (2025). Know where to go: Make LLM a relevant, responsible, and trustworthy searchers. Decision Support Systems, 188, 114354

StatCounter. (2025, January). Marktanteile der Suchmaschinen und Marktanteile der mobilen Suche in Deutschland. Cited in Statista. Retrieved January 15, 2025 from https://de.statista.com/statistik/daten/studie/301012/umfrage/marktanteile-der-suchmaschinen-und-marktanteile-mobile-suche/

Terenteva, E. (2023, July 18). Crawlability & Indexability: What they are & How they affect SEO. Semrush Blog. https://www.semrush.com/blog/what-are-crawlability-and-indexability-of-a-website/

Touvron, H., Martin, L., Stone, K., Albert, P., Almahairi, A., Babaei, Y., ... & Scialom, T. (2023). Llama 2: Open foundation and fine-tuned chat models. arXiv preprint arXiv:2307.09288.

Veale, M., & Zuiderveen Borgesius, F. (2021). Demystifying the draft EU Artificial Intelligence Act-Analysing the good, the bad, and the unclear elements of the proposed approach. Computer Law Review International, 22(4), 97–112.

Wang, B., Li, Q., Melucci, M., & Song, D. (2019, May). Semantic Hilbert space for text representation learning. In The World Wide Web Conference (pp. 3293-3299).

Wang, L., Yang, N., Huang, X., Yang, L., Majumder, R., & Wei, F. (2024, January). Large search model: Redefining search stack in the era of llms. In ACM SIGIR Forum (Vol. 57, No. 2, pp. 1-16). New York, NY, USA: ACM.

Wang, Z., Xu, Z., Srikumar, V., & Ai, Q. (2024, May). An in-depth investigation of user response simulation for conversational search. In Proceedings of the ACM Web Conference 2024 (pp. 1407-1418).

Wei, J., Wang, X., Schuurmans, D., Bosma, M., Xia, F., Chi, E., ... & Zhou, D. (2022). Chain-of-thought prompting elicits reasoning in large language models. Advances in neural information processing systems, 35, 24824-24837.

Werner, T., Soraperra, I., Calvano, E., Parkes, D. C., & Rahwan, I. (2024). Experimental evidence that conversational artificial intelligence can steer consumer behavior without detection. arXiv preprint arXiv:2409.12143.

Xiong, H., Bian, J., Li, Y., Li, X., Du, M., Wang, S., ... & Helal, S. (2024). When search engine services meet large language models: visions and challenges. IEEE Transactions on Services Computing.

Yan, L., Liu, Y., & Liu, S. (2024). The Search for Balance: The Impact of LLM-based Conversational Search on Information Processing and Polarization.

Yao, S., Zhao, J., Yu, D., Du, N., Shafran, I., Narasimhan, K., & Cao, Y. (2023, January). React: Synergizing reasoning and acting in language models. In International Conference on Learning Representations (ICLR).

Yoon, C., Kim, G., Jeon, B., Kim, S., Jo, Y., & Kang, J. (2024). Ask Optimal Questions: Aligning Large Language Models with Retriever's Preference in Conversational Search. arXiv preprint arXiv:2402.11827.

Zhang, Y., Li, Y., Cui, L., Cai, D., Liu, L., Fu, T., ... & Shi, S. (2023). Siren's song in the AI ocean: a survey on hallucination in large language models. arXiv preprint arXiv:2309.01219.