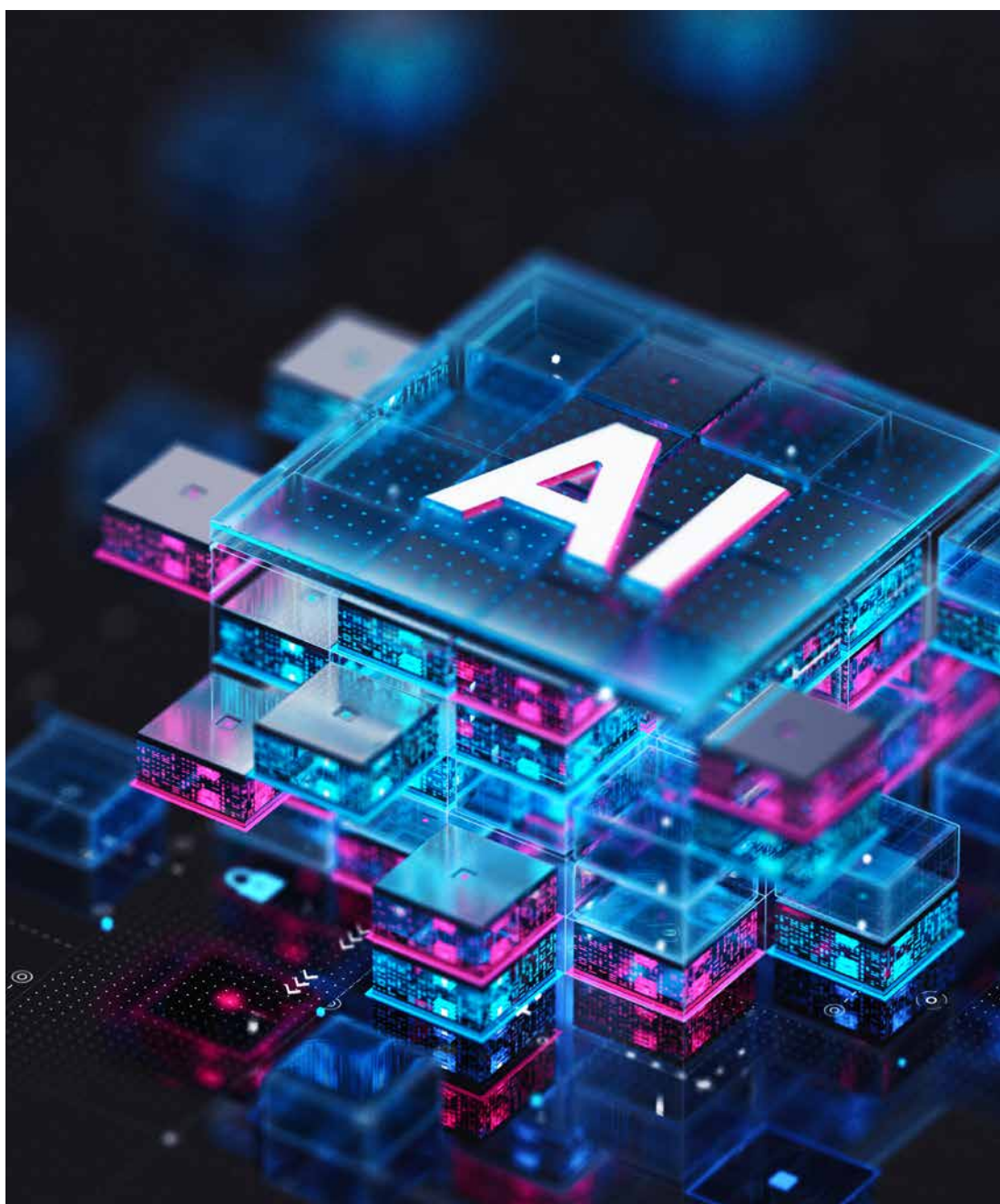


Generative AI in Finance Studie 2024

Autoren: Sophie Hundertmark und Nils Hafner

In Kooperation mit:



Inhaltsverzeichnis

Abbildungsverzeichnis	3
Ausgangssituation und Ziele der Studie	4
1.1 Die Forschungsfragen	4
1.2 Begriffserklärung	5
1.3 Die Autoren	6
Die Partner der Studie	7
Umfragedesign und Methodik	8
Dimensionen von Generative AI-Anwendungen	9
Die Ergebnisse	11
5.1 Einordnung der von uns analysierten Generative AI-Anwendungen auf der «Gen AI Skala»	11
5.2 Treiber für Generative AI-Anwendungen	27
5.3 Strategische Entscheidungen für die Implementation	27
5.4 Vorgehen zum Einführen von Generative AI- Anwendungen	29
Anwendbarkeit der AI-Anwendungen in der Finanzbranche	35
Einordnung des Nutzens und der Machbarkeit	47
Fazit und Handlungsvorgehen bei Einführung	49
Quellenverzeichnis	50

Usecases:

Generative AI im Kundenservice der Walbusch-Gruppe zur automatisierten Bearbeitung von Geschäftsvorfällen	15
Generative AI im operativen Betrieb der Gemeinde Glattfelden	18
Generative AI in der Softwareentwicklung mit Generative AI bei PensionDynamics	23
Die digitale Assistentin Clara: Wie aus einem einfachen SMS-Dienst bei Helvetia Versicherungen Schweiz ein GenAI-Bot wurde	37
Mitarbeitenden-Unterstützung im Kundenkontakt der KPT Versicherung	41

Abbildungsverzeichnis

Abb. 1: Anwendungsbereiche von Generative AI	9
Abb. 2: Die Gen AI Skala	11
Abb. 3: Chat- und Voicebots im Kundenservice auf der Gen AI Skala	11
Abb. 4: Chatbots in der Produktberatung auf der Gen AI Skala	12
Abb. 5: Agenten-Unterstützung am Telefon auf der Gen AI Skala	12
Abb. 6: Agenten-Unterstützung in der E-Mail Bearbeitung auf der Gen AI Skala	13
Abb. 7: Mitarbeitendennutzung von Chat GPT auf der Gen AI Skala	14
Abb. 8: Mitarbeitendennutzung des Microsoft Co-Piloten auf der Gen AI Skala	20
Abb. 9: Einsatz für Marketingtexte, wie Newsletter oder Social Media auf der Gen AI Skala	20
Abb. 10: Generierung von Produktbeschreibungen auf der Gen AI Skala	21
Abb. 11: Bereitstellung von Informationen für den Vertrieb auf der Gen AI Skala	21
Abb. 12: Onboarding und HR-Assistent auf der Gen AI Skala	22
Abb. 13: Unterstützung in der Software-Entwicklung auf der Gen AI Skala	22
Abb. 14: Best Practice Prozessbeschreibung in Schritten	30
Abb. 15: Projektvorgehen für den internen Wissensaustausch in Schritten	31
Abb. 16: Projektvorgehen für klar definierte interne Prozesse in Schritten	33
Abb. 17: Chat- und Voicebots im Kundenservice auf der Gen AI Skala	35
Abb. 18: Agenten-Unterstützung am Telefon auf der Gen AI Skala	36
Abb. 19: Agenten-Unterstützung in der E-Mail-Bearbeitung auf der Gen AI Skala	43
Abb. 20: Mitarbeitendennutzung von Chat GPT auf der Gen AI Skala	43
Abb. 21: Mitarbeitendennutzung des Microsoft Co-Piloten auf der Gen AI Skala	43
Abb. 22: Einsatz für Marketingtexte, wie Newsletter oder Social Media auf der Gen AI Skala	44
Abb. 23: Onboarding und HR-Assistent auf der Gen AI Skala	44
Abb. 24: Unterstützung in der Software-Entwicklung auf der Gen AI Skala	45
Abb. 25: Branchenspezifische Generative AI Agenten auf der Gen AI Skala	45
Abb. 26: Generative AI-Anwendungen auf der Gen AI Skala	46
Abb. 27: Positionierung der Gen-AI Anwendungen im Spannungsfeld zwischen Nutzen und Machbarkeit für Banken	48

1 Ausgangssituation und Ziele der Studie

Bisherige Recherchen zu den Erwartungen von Banken- und Versicherungskunden legen nahe, dass Kunden bei der Auswahl eines Finanzdienstleisters – sei es eine Bank oder eine Versicherung – tendenziell dieselben Bewertungskriterien anwenden, die auch in anderen Wirtschaftssektoren, also in finanzfremden Branchen, wie Retail, Pharmaindustrie oder Unterhaltungsbranche, üblich sind. In diesem Kontext ist der Einsatz modernster Technologien von wesentlicher Bedeutung (Leyes et al. 2021) (Uribe-Linares et al., 2023) (du Toit et al., 2017).

Es wird davon ausgegangen, dass es für Finanzinstitute auf lange Sicht nicht ausreichend ist, sich ausschliesslich im Vergleich mit Akteuren der eigenen Branche zu positionieren. Kunden beziehen Innovationen, die sie in anderen Branchen beobachten, in ihren Erwartungshorizont an Anbieter von Finanzdienstleistungen immer mehr mit

ein. Daher ist es erforderlich, auch andere Industrien wie den Einzelhandel, den Online-Handel, die Pharmaindustrie, die Gesundheitswelt, die Software-Branche und andere in die Technologie-Benchmarking-Prozesse bzw. in ein Monitoring von Innovationen einzubeziehen. Bei einer etwaigen Adaption müssen aber auch die strengen Compliance-Vorschriften, Regulierungen und Datenschutzbestimmungen, die für Finanzunternehmen gelten, berücksichtigt werden.

Die vorliegende Studie zielt darauf ab, genau diese Forschungslücke für Finanzinstitute zu schliessen. Wir untersuchen, wie andere Branchen wie Einzelhandel, Gesundheit, Tourismus und viele weitere bereits Generative AI nutzen und evaluieren die Übertragbarkeit dieser Anwendungsfälle auf Schweizer Finanzunternehmen.

1.1 Die Forschungsfragen

Das Team vom IFZ hat zusammen mit den teilnehmenden Partnern folgende Forschungsfragen definiert.

Wie setzen finanzfremde Branchen heute Generative AI im Kundenservice ein?

1. Wie wird Generative AI in finanzfremden Branchen genutzt, um das Kundenerlebnis zu steigern?
2. Wie wird Generative AI in finanzfremden Branchen genutzt, um die Mitarbeiterproduktivität und Mitarbeiterzufriedenheit zu stärken?
3. Wie wird Generative AI in finanzfremden Branchen zur Steigerung der Effizienz oder der Effektivität der Prozesse genutzt?
4. Wie wird Generative AI in finanzfremden Branchen im Marketing vor allem zur personalisierten Kundenkommunikation eingesetzt?
5. Wie wird Generative AI in finanzfremden Branchen zur internen Weiterbildung genutzt?
6. Welche zukünftigen Einsatzmöglichkeiten von Generative AI planen finanzfremde Branchen?
7. Wo stehen Schweizer Finanzunternehmen im Vergleich zu finanzfremden Branchen?
8. Was können Schweizer Finanzunternehmen im Hinblick auf die Nutzung von Generative AI von finanzfremden Branchen lernen und in der eigenen Branche anwenden?

1.2 Begriffserklärung

Künstliche Intelligenz, KI (engl.: Artificial Intelligence ,AI)

Künstliche Intelligenz (englisch: Artificial Intelligence) bezieht sich auf die Fähigkeit von Computern und Maschinen, menschenähnliche kognitive Funktionen wie Lernen, Verstehen, Problemlösen und Entscheidungsfindung zu simulieren. KI-Systeme werden entwickelt, um Aufgaben zu erledigen, die typischerweise menschlichen Intellekt erfordern, darunter visuelle Wahrnehmung, Spracherkennung und strategische Planung. Im Herzen der KI-Technologie stehen Algorithmen und Modelle, die auf Daten trainiert werden, um Muster zu erkennen und darauf basierend Entscheidungen zu treffen. Künstliche Intelligenz kann in zahlreichen Anwendungsbereichen eingesetzt werden, von der Automatisierung industrieller Prozesse über die Verbesserung von Kundenerfahrungen bis hin zur Unterstützung medizinischer Diagnostik. Sie bietet das Potenzial, Effizienz zu steigern, menschliche Fähigkeiten zu erweitern und in vielen Bereichen innovative Lösungen zu liefern.

Generative Künstliche Intelligenz (engl.: Generative Artificial Intelligence)

Generative Künstliche Intelligenz (kurz: Generative KI) (englisch: Generative Artificial Intelligence) (kurz: Generative AI)) bezieht sich auf den Einsatz von KI, insbesondere von generativen Modellen, um neue Inhalte oder Daten zu erstellen. Diese Modelle werden trainiert, um auf der Grundlage von vorhandenen Beispielen neue, bisher nicht gesehene Daten zu generieren. Generative AI findet Anwendung in verschiedenen Bereichen wie der Erstellung von Texten, Bildern, Musik und sogar Videos. Es ermöglicht die automatische Erzeugung von Inhalten, die menschlicher Kreativität ähneln, und eröffnet neue Möglichkeiten für kreative Prozesse, Forschung und Unterhaltung.

Conversational Artificial Intelligence (Conversational AI)

Conversational AI bezieht sich auf Technologien, die es Maschinen ermöglichen, natürliche, menschenähnliche Gespräche mit Benutzern zu führen. Diese Technologien nutzen Methoden der KI, insbesondere aus den Bereichen maschinelles Lernen und natürliche Sprachverarbeitung, um gesprochene oder geschriebene Sprache zu verstehen, darauf zu reagieren und sinnvolle Dialoge zu führen. Zu den Hauptkomponenten der Conversational AI gehören: Die Spracherkennung (Automatic Speech Recognition, ASR), die natürliche Sprachverarbeitung (Natural Language

Processing, NLP), das Dialogmanagement und die Sprachsynthese (Text-to-Speech, TTS). Conversational AI wird in verschiedenen Anwendungen eingesetzt, darunter Chatbots, virtuelle Assistenten, Kundensupportsysteme und mehr, um die Interaktion zwischen Menschen und Maschinen zu verbessern und zu erleichtern.

Grosse Sprachmodelle (engl.: Large Language Models (LLMs))

Large Language Models sind fortschrittliche KI-Modelle, die darauf trainiert werden, natürliche Sprache zu verstehen und zu generieren. Diese Modelle wurden mittels riesiger Mengen an Textdaten vortrainiert, um ein umfassendes Verständnis von Sprache zu entwickeln. Sie können komplexe Texte verstehen, Fragen beantworten, Dialoge führen und sogar Texte in verschiedenen Stilen oder Stimmen generieren. LLMs wie GPT-4o basieren auf tiefen, neuronalen Netzwerken mit Millionen oder sogar Milliarden von Parametern, die es ihnen ermöglichen, vielseitige Sprachaufgaben zu bewältigen und menschenähnliche Texte zu erzeugen. Sie werden in Bereichen wie der maschinellen Übersetzung, dem Textverständnis, der Sprachgenerierung und anderen Anwendungen eingesetzt, um die Kommunikation zwischen Menschen und Maschinen zu verbessern.

ChatGPT

ChatGPT ist ein neuartiges Chatbot-System, das auf der Generative Pretrained Transformer-Architektur (kurz: GPT-Architektur) basiert. Es ist darauf vortrainiert, menschenähnliche Konversationen zu führen und natürliche Sprache zu verstehen und zu generieren. ChatGPT kann auf Texteingaben reagieren, Fragen beantworten, Informationen bereitstellen und einen kontextbezogenen Dialog mit Benutzern führen. Es basiert auf einem breiten Wissen aus einer Vielzahl von Quellen und kann flexibel auf verschiedene Themen und Anwendungsfälle eingehen. ChatGPT wird verwendet, um interaktive Gespräche in verschiedenen Kontexten zu ermöglichen, von Kundenservice und Unterstützung bis hin zu kreativen Anwendungen wie Storytelling oder der Simulation von Charakteren¹.

¹ Passend zur Studie wurden die Definitionen mit ChatGPT-4o verfasst. Dabei kam der initiale Prompt (Anweisung an ChatGPT) von den Studienautoren. Nachdem ChatGPT eine Antwort verfasst hatte, wurde diese von den Autoren auf inhaltliche Richtigkeit geprüft, sprachlich leicht umgeschrieben und in die vorliegende Studie integriert.

1.3 Die Autoren

Die vorliegende Studie wurde von Sophie Hundertmark und Prof. Dr. Nils Hafner geleitet, durchgeführt und geschrieben.



Über Sophie Hundertmark

Nach Stationen in der Industrie und in Agenturen gehörte Sophie Hundertmark zu den ersten Masterstudentinnen in der Schweiz, die zu Chatbots. Seit 2022 promoviert sie an der Universität Fribourg zur Kombination von Conversational AI, Fuzzy Logic, Ethic und Large Language Models (LLMs). Dazu arbeitet sie als wissenschaftliche Mitarbeiterin und Dozentin am Institut für Finanzdienstleistungen Zug der Hochschule Luzern (IFZ). Sophie Hundertmark verfügt über langjährige Erfahrungen als selbstständige Beraterin für die strategische Begleitung sowie Umsetzung von AI und Chatbot-Projekten. Zudem berät sie Unternehmen zum (ethisch-) verantwortungsvollem Umgang mit Generative AI, ChatGPT und ähnlichen AI-Systemen. In diesem Zusammenhang sorgt sie für einen regelmässigen Austausch zwischen Akademie und Praxis zu allen Themen der AI getriebenen Conversational Automation.

Ende 2020 erschien Sophies erstes Buch „Digitale Freunde: Wie Unternehmen Chatbots erfolgreich einsetzen können«. Anfangs 2024 erschien der von ihr zusammen mit Nils Hafner herausgegebene Sammelband „Kundendialog-Management“ bei Springer Gabler. Seit Mitte 2024 ist das Buch „Generative KI für Unternehmen: Strategien, Werkzeuge und Best Practices für die erfolgreiche Integration“, welches sie zusammen mit Dr. Daniela Suter veröffentlicht hat, verfügbar.

Als Growth Lead und Beirat begleitet Sophie innovative Unternehmen, wie AlpineAI, eggheads.ai, Joaia und Cloudburst und ist im Vorstand des Swiss Insights Verband.

Kontakt

Sophie Hundertmark ist Senior Wissenschaftliche Mitarbeiterin am Institut für Finanzdienstleistungen Zug IFZ der Hochschule Luzern – Wirtschaft

Sophie.hundertmark@hslu.ch



Über Nils Hafner

Prof. Dr. rer. pol. Nils Hafner ist internationaler Experte für den Aufbau langfristig, profitabler Kundenbeziehungen. Er ist Professor für Kundenbeziehungsmanagement an der Hochschule Luzern, leitet zwei Studienprogramme zum Digital Banking und zum Sales und Marketing im Banking. Prof. Dr. Hafner studierte Betriebswirtschaftslehre, Psychologie, Philosophie und Neuere Geschichte in Kiel und Rostock, Deutschland. Er promovierte im Innovationsmanagement/Marketing über die Qualitätssteuerung von Call Center Dienstleistungen. Nach seiner Tätigkeit als Practiceleader CRM bei einem der grössten Beratungshäuser der Welt baute er von 2002 bis Mai 2006 das erste CRM-Masterprogramm im deutschsprachigen Raum auf. Heute ist er als Autor und internationaler Key-Note Speaker tätig und berät Geschäftsführungen und Vorstände. Sein Schwerpunkt in der Beratung liegt im Bereich langfristig, profitabler Servicestrategien.

In seinem Blog „Hafner on CRM“ und seinem Podcast „Hafners CX Podcast“ versucht er dem Thema seine informativen, schönen, schlimmen, tragischen und lustigen Seiten abzugewinnen. Dort veröffentlicht er auch seit 2006 die CRM-Trends des Jahres. Das macht er seit 2020 zusammen mit dem deutschen Berater Harald Henn den CEX Trendradar, der inzwischen international eine relevante Entscheidungs-Grundlage für Technologie und Prozessberatungen im Top Management geworden ist und von weltweiten Technologie-Anbieter systematisch gefördert wird. Ende 2018 erschien sein Amazon Nr. 1. Bestseller „Die Kunst der Kundenbeziehung“ in der zweiten Auflage bei Haufe. Anfangs 2024 erschien der von ihm zusammen mit Sophie Hundertmark herausgegebene Sammelband „Kundendialog-Management“ bei Springer Gabler.

Prof. Dr. Nils Hafner ist Studiengangsleiter des CAS Digital Banking am Institut für Finanzdienstleistungen Zug IFZ der Hochschule Luzern – Wirtschaft

Nils.hafner@hslu.ch

2 Die Partner der Studie



Aargauische Kantonalbank

Argauer Kantonalbank (AKB)

Die AKB ist eine der führenden Universalbanken für Privat- und Firmenkundinnen und -kunden im Kanton Aargau und hat auch institutionellen Anlegerinnen und Anlegern viel zu bieten. Zu ihren Stärken zählen die persönliche Beratung und die regionale Nähe zu ihren Kunden. Ihr Betreuungskonzept ist klar und verständlich. Kompetenz ist bei der AKB nicht kompliziert – ihre Produkte und Services sind es ebenso wenig. Die AKB ist in 32 Geschäftsstellen im Kanton Aargau und im angrenzenden solothurnischen Gebiet Olten-Gösgen-Gäu vertreten. Die AKB zählt rund 940 Mitarbeitende (807 Vollzeitstellen); davon sind 89 Lernende, Praktikantinnen, Praktikanten und Trainees. Aktuell betreut die Bank rund 250 000 Kundinnen und Kunden.



St. Galler Kantonalbank

St. Galler Kantonalbank

Als regional verwurzelte Universalbank berät und betreut die St. Galler Kantonalbank seit über 150 Jahren Privat- und Geschäftskunden rund ums Geld. Die St. Galler Kantonalbank erfüllt ihre Aufgaben seit eh und je mit viel Engagement und Gespür für die Menschen in ihrer Marktregion. «Wir sind die erste Bank für unsere Kundinnen und Kunden» die erste Bank zu sein bedeutet für uns, dass wir die Wünsche und Ziele unserer Kundschaft ins Zentrum stellen und sie stets aufs Neue begeistern.



InventxLab

Als eigenständige, rechtliche Einheit bietet das InventxLab Raum für die Entwicklung und Umsetzung von Ideen zum Ausbau der Wettbewerbsfähigkeit der Schweizer Finanzindustrie. Als Tochterunternehmen der Inventx AG vernetzt, begleitet und befähigt es Mitarbeitende der Inventx wie auch Partner und Kunden, ihre innovativen Ideen zur Marktreife zu bringen. Das InventxLab unterstützt Initiativen zur Effizienzsteigerung wie auch zur Geschäftsmodell-Innovation, welche an den regelmässig durchgeführten Innovation Days entstehen und die in einem strukturierten Innovationsprozess vorangetrieben werden. Zusätzlich werden Anwendungsfelder von neuen Technologien evaluiert und entwickelt, welche für das Finanz- und Versicherungs-Ökosystem relevant sind.

3 Umfragedesign und Methodik

Um auf der Unternehmensseite, zunächst branchenunabhängig, die heutigen Anwendungen, Chancen und Herausforderungen von Generative AI sowie die ersten Erfahrungen mit diesen Anwendungen zu evaluieren, wurden qualitative Befragungen mit leitenden Mitarbeitenden aus den Bereichen Kundenservice, Informationsdienste, Compliance, Marketing und Kommunikation durchgeführt. Qualitative Befragungen zeichnen sich dadurch aus, dass sie tiefe Zusammenhänge genauer analysieren und beantworten können (Holz et al., 2006). Zur Vorbereitung der Befragungen haben die Studienautoren nach Rücksprache mit den Studienpartnern einen Gesprächsleitfaden mit 20 Fragen entwickelt. Die genauen Fragen wurden dann während des Interviews, je nach Antworten des Befragten, vertieft und angepasst.

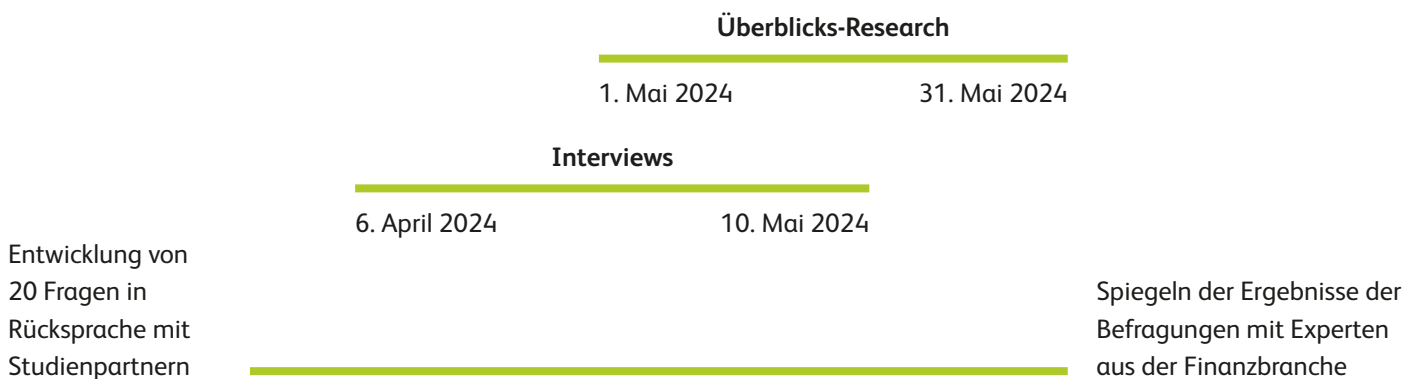
Der Leitfaden beinhaltet im Groben die folgenden Themenschwerpunkte:

1. **Generative AI-Anwendungen im Allgemeinen**
2. **Treiber bzw. Auslöser für diese Anwendungen**
3. **Compliance Abklärungen zur Umsetzung der AI-Anwendungen**
4. **Erkenntnisse, Erfahrungen und Ausblick zu den Generative AI-Anwendungen**

Die Interviews fanden zwischen dem 06. April und dem 10. Mai 2024 statt. Diese fanden mehrheitlich online über Microsoft Teams statt. Befragt wurden insgesamt 40 Ansprechpartner aus Deutschland, Schweiz und Österreich.

Zusätzlich wurden die Befragungen mit einem Überblicks-Research im Zeitraum vom 01. Mai bis 31. Mai ergänzt. Dabei wurden überwiegend Online-Magazine aus der Schweiz, Deutschland und Österreich im Hinblick auf Medienberichte zu Generative AI-Anwendungen untersucht. Sofern es sich bei den Berichten um keine rein werblichen Inhalte von Technologie-Anbietern, sondern um Erfahrungsberichte von Unternehmen handelte, wurden diese Ergebnisse ebenfalls in die Auswertungen integriert. Allen Interview-Teilnehmern wurde eine anonyme Verarbeitung ihrer Daten zugesagt. Aus diesem Grund ist es den Autoren nicht immer möglich die jeweiligen Unternehmensnamen offenzulegen. Einzelne Befragten haben jedoch ihr Einverständnis zum Teilen von Beispielen gegeben. Diese Beispiele sind in der Folge klar benannt.

Schlussendlich wurden die Ergebnisse der Befragungen mit Experten aus der Finanzbranche gespiegelt, um die Übertragbarkeit der Generative AI-Anwendungen auf die Finanzindustrie abschätzen zu können.



Dimensionen von Generative AI-

4 Anwendungen

Dieses Kapitel beschreibt verschiedene Dimensionen und Arten einer Generative AI-Anwendung und beleuchtet die spezifischen Charakteristika. Die jeweilige Ausprägung kann durch technische Rahmenbedingungen vorgegeben sein, zum Beispiel sensitive Inputdaten, welche die Schweiz nicht verlassen dürfen, oder im Rahmen des Projektsetups definiert werden.

Art der Anwendung von Generative AI

Generative AI kann in unterschiedlichen Formen eingesetzt werden. Bots führen mit ihren Nutzern Dialoge in natürlicher Sprache und bieten Unterstützung bei Kundenanfragen, internen Prozessen und anderen Kommunikationsaufgaben (Khennouche 2024). Für kreative Inhalte oder automatisierte Berichterstellung können mit Generative AI Models Texte, Bilder und Videos generiert werden (Strobel, Banh, Möller, Schoormann 2024). Generative AI unterstützt auch

Model

Die Transparenz und die Nachvollziehbarkeit der KI-Modelle sind wichtige Faktoren, die das Vertrauen in die Technologie und damit deren Akzeptanz beeinflussen. Auch die Prognosegenauigkeit und die jeweilige Robustheit des Modelles sollten verifiziert werden.

Weiter unterscheiden sich die Generative AI-Models in ihrer Grösse, Qualität und in ihrem Ressourcenbedarf. Beispielsweise wird bei Sprachmodellen zwischen Large Language Models (LLMs), welche grosse Datenmengen verarbeiten, und komplexe Aufgaben bewältigen können, und Small Language Models, die für einfachere Tasks entwickelt und kosteffizienter sind, unterschieden (Beatty, 2024). Small Language Models können zudem auch mit beschränkten Ressourcen auf spezifische Anwendungsfälle angepasst werden.

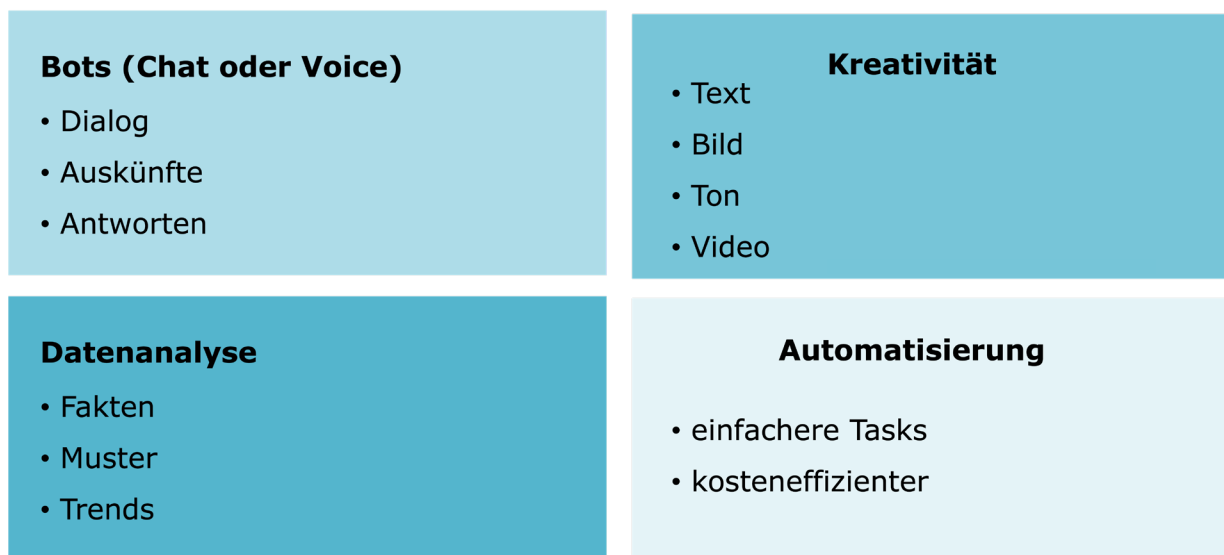


Abb. 1: Anwendungsbereiche von Generative AI

bei der Analyse grosser Datenmengen um Muster, Trends und weitere wertvolle Erkenntnisse zusammen zu stellen oder bei der Automatisierung von Geschäftsprozessen (Rane 2023).

Datenklassifikation

Welche Daten werden mit der Generative AI-Anwendung verarbeitet? Für die Wahl des Betriebsmodelles der Generative AI-Anwendung ist es entscheidend, wie die

verwendeten Daten klassifiziert sind. Sind es öffentliche Daten, die allgemein zugänglich sind und kein besonderes Schutzniveau erfordern, sind es unternehmensinterne Daten, die nur für Mitarbeiter zugänglich sind, vertrauliche Daten, die einen höheren Schutzbedarf haben oder sogar personenbezogene oder geheime Daten, deren Zugriff stark geschützt werden sollte?

Deployment

Die Implementierung und Bereitstellung von Generative AI-Anwendungen kann auf verschiedene Weisen erfolgen, wobei es zwischen dem Deployment des Modelles und der Anwendung zu unterscheiden gilt. SaaS (Software as a Service) oder Public Cloud Deployments bieten eine einfache Skalierbarkeit und Zugänglichkeit. Community Cloud Infrastrukturen und On-Premise Installationen bieten höhere Kontrolle und Sicherheit bei höheren Kosten.

Model Deployment

Bei Closed Source Models ist das Deployment durch ihre Anbieter und Kooperationen gegeben. In den meisten Fällen wird hier eine Public Cloud von AWS, Azure oder Google genutzt. Je nach Anbieter kann die Region gewählt werden, z.B. Azure OpenAI Service aus dem Azure Datacenter CH-Nord. Weit verbreitet sind zudem Enterprise Deployments in einer Public Cloud – die exklusive Bereitstellung des Modelles für eine spezifische Kundenfirma.

Bei Open Source Modellen kann das Deployment hingegen frei gewählt werden. Deployments im eigenen Datacenter (On-Premise) oder in Community Clouds sind möglich, wobei die hohe benötigte Rechenleistung zu beachten ist.

Deployment Anwendung

Generative AI-Anwendungen können in einer Public Cloud, wie auch On-Premise deployed werden. Für die Entscheidung ist das Deployment des Modelles, der angebundenen Systeme, sowie eines allfälligen Dateninputs relevant. Bei Generative AI-Anwendungen mit Build-In Generative AI ist stets zu prüfen, ob und wie auf ein externes Modell zugegriffen wird.

Integration der unternehmensspezifischen Daten in Language Models

Die Anpassung von Large oder Small Language Models an unternehmensspezifische Anforderungen erfolgt durch verschiedene Techniken (Srivata, 2024):

- » **Retrieval-Augmented-Generation (RAG):** Für eine gestellte Anfrage relevante Daten werden abgerufen und als zusätzlicher Kontext für das Language Model bereitgestellt. Diese Technik eignet sich für dynamische unternehmensspezifische Daten.
- » **Finetuning:** Das vortrainierte Modell wird mit den unternehmensspezifischen Daten weiter trainiert.
- » **Hybrid:** Kombination von RAG und Finetuning, um die Vorteile beider Ansätze zu nutzen.

Die Wahl dieser und weiterer spezifischen Dimensionen hängt von den individuellen Bedürfnissen und Prioritäten des Unternehmens ab. Eine richtige Balance zwischen Flexibilität, Sicherheit, sowie Effizienz und Kontrolle ist entscheidend für eine erfolgreiche Umsetzung eines Generative AI-Anwendungsfalles.

Build versus Buy

Ein wichtiger Aspekt, den Unternehmen bei der Einführung von Generative AI-Anwendungen berücksichtigen müssen, ist die Frage, ob sie eine bestehende Lösung von einem Anbieter kaufen (Buy) oder eine eigene Lösung entwickeln (Build) wollen. Folgende Faktoren spielen dabei eine Rolle:

- » **Expertenwissen:** Eine Eigenentwicklung (Build) benötigt Ressourcen und Expertenwissen. Ist dies vorhanden?
- » **Markt:** Gibt es verfügbare Produkte für den Anwendungsfall (Buy)?
- » **Usecase:** Eine Eigenentwicklung (Build) kann in Betracht gezogen werden, wenn der Usecase einfach mit Generative AI-Tools und Werkzeugen umgesetzt werden kann. Custom GPT, die Azure Open AI Toolbox oder H2O sind nur einige Beispiele für mögliche Tools. Auch bei spezifischen Anforderungen kann eine massgeschneiderte Lösung (Build) von Vorteil sein.
- » **Roadmap / Skalierbarkeit:** Wie sieht die Generative AI Roadmap aus? Oft bieten fertige Lösungen (Buy) Skalierbarkeitsoptionen, die flexibel an den Bedarf angepasst werden können.
- » **Betrieb und Weiterentwicklung:** Bei Generative AI-Anwendungen ist es notwendig, die Lösung in Einklang mit den raschen Fortschritten der Technologie weiterzuentwickeln und zu warten. Beim Einkauf von Lösungen kann dies dem Anbieter ausgelagert werden (Buy).

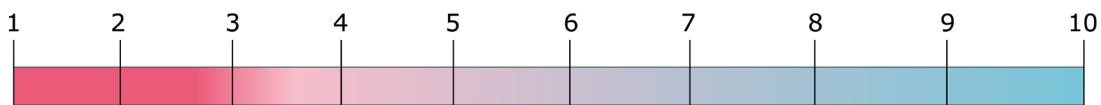
5 Die Ergebnisse

5.1 Einordnung der von uns analysierten Generative AI-Anwendungen auf der «Gen AI Skala»

Während der ersten Befragungen fiel schnell auf, dass sich die von den Unternehmen angewandten Generative AI-Anwendungen mehrheitlich auf einer Skala zwischen zwei Polen einordnen lassen. Der eine Pol sind Anwendungen, die direkt durch den Kunden bedient werden. Der andere Pol sind Prozesse, die komplett im Innern des Unternehmens ablaufen und zum Teil selbst für die Mitarbeitenden nicht sichtbar sind. So wird aufgrund der mangelnden Sichtbarkeit, beispielsweise in der Versicherungsindustrie, von Dunkelverarbeitung gesprochen. Dazwischen bewegen sich Anwendungen, die alle oder einzelne Mitarbeitende unterstützen. Grafisch kann das wie folgt dargestellt werden:

Anwendungen, die sich auf der Mitte des Strahls befinden, sind eher intern ausgerichtet. Sie haben zum Ziel, die Mitarbeitenden-Zufriedenheit und die Mitarbeitenden-Produktivität zu unterstützen. Anwendungen auf der rechten Seite des Strahls sind reine Automatisierungen, denen lediglich die Steigerung der Effizienz als Ziel zugrunde liegt.

Natürlich kommt es dabei zu Wechselwirkungen gemäss der Service Profit Chain (Heskett et al. 1990). So realisieren auch die Kunden, wenn Mitarbeitende aufgrund besserer Unterstützung motivierter sind oder allgemein die Prozesse, wie beispielsweise ein Versand von Waren oder Unterlagen, schneller ablaufen können.



Kundenanwendungen

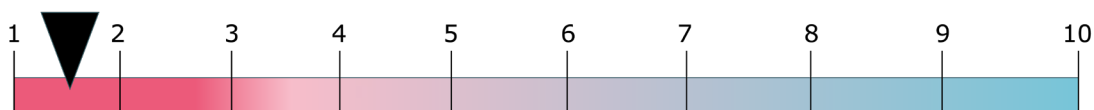
Automatisierte Dunkelverarbeitung

Abb. 2: Die Gen AI Skala

Anwendungen, die ganz oder eher links auf dem Strahl liegen, beschäftigen sich mit Generative AI-Anwendungen zwischen Kunde und Unternehmen. Ziel ist es hier, dem Kunden ein besseres Kundenerlebnis bieten zu können und darüber hinaus auch Kosten einzusparen.

Eine sehr naheliegende Anwendung von Generative AI im Kundenservice ist die Verwendung von Chat- oder Voicebots (Abb. 3). Handelt es sich bei dem befragten Unternehmen um ein Unternehmen, bei dem der Kundenservice eine direkte Rolle in der Wertschöpfungskette spielt, so sind

Chat- und Voicebots im Kundenservice

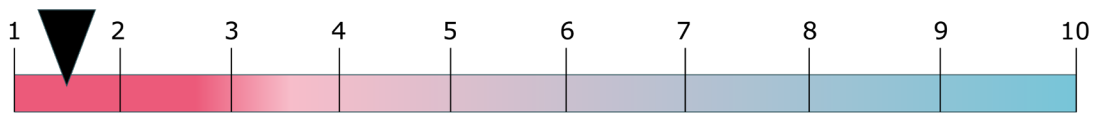


Kundenanwendungen

Automatisierte Dunkelverarbeitung

Abb. 3: Chat- und Voicebots im Kundenservice auf der Gen AI Skala

Chatbots in der Produktberatung



Kundenanwendungen

Automatisierte Dunkelverarbeitung

Abb. 4: Chatbots in der Produktberatung auf der Gen AI Skala

Chat- oder Voicebots zur automatisierten Beantwortung der Kundenanfragen keine Seltenheit mehr. Zwei Drittel der befragten Unternehmen aus dem Bereich Business-to-Consumer (B2C) gaben an, dass sie bereits Chatbots auf der Webseite oder Voicebots am Telefon einsetzen. Dabei ist anzumerken, dass die Hälfte dieser Unternehmen auch vor dem Aufkommen von Generative AI Ende 2022 bereits Bots eingesetzt hat. Neu ist, dass diese Bots aber auch Generative AI nutzen. Da sich dadurch die Qualität der Antworten an den Kunden substantiell verbessert hat, stehen diese Unternehmen vor allem vor einer kommunikativen Herausforderung. Sie müssen ihren Kunden erklären, dass der Bot eine erhebliche Verbesserung erfahren hat, und es sich gegebenenfalls lohnt, den Bot erneut auszuprobieren.

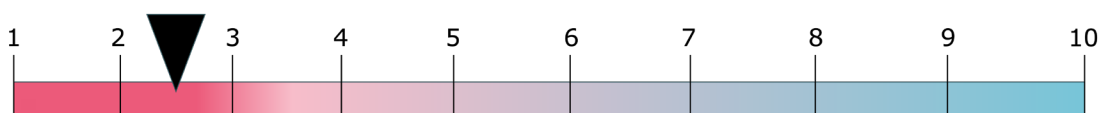
Ein Beispiel für ein Unternehmen, das dank Generative AI im Kundenservice zur Entlastung der Mitarbeitenden geführt hat und gleichzeitig zur Steigerung der Zufriedenheit von ungedulden Kunden, die rund um die Uhr eine Antwort wünschen, ist GS1 (<https://www.gs1.ch/de/barcodes-standards/barcodes>). Dieser Chatbot nutzt mehrheitlich die Datenquellen, die die Webseite ohnehin schon bietet und beantwortet die Kundenanfragen individuell und persönlich mit Hilfe von Generative AI. Ein weiteres Beispiel für die Entwicklung eines Chatbots mit Generative AI ist «Clara» der Helvetia Versicherungen Schweiz. Dieser wird in dieser Studie in der Case Study „Die digitale Assistentin Clara: Wie aus einem einfachen SMS-Dienst bei Helvetia Versicherungen Schweiz ein GenAI-Bot wurde.“ auf Seite 37 vorgestellt.

Während sich vor der akzeptierten Verwendung von Generativer AI in Chatbots die meisten Unternehmen gescheut hatten, Bots auch als Produktberater einzusetzen, ist diese

Anwendung nun immer häufiger zu beobachten (Abb. 4). Dies gilt jedoch nur für die Unternehmen, welche mehrere Produkte gleichzeitig über einen Online-Shop anbieten. Die Rede ist hier von klassischen Online-Shops wie der Schweizer Baumarkt «Jumbo» (<https://www.jumbo.ch/de>). Wenn Kunden im Bereich Heimwerken die Beratung fehlt, kaufen sie schnell das falsche Produkt oder brechen den Kaufprozess sogar aufgrund von Unsicherheit ab. Jumbo ist eines der ersten Unternehmen mit einer solchen Anwendung. Das Unternehmen hat gezeigt, dass Generative AI-Beratungen durchaus zu mehr bzw. höheren Warenkörben führen können.

Chat- und Voicebots zeichnen sich dadurch aus, dass sie die Kundendialoge vollumfänglich automatisieren. Dies ist jedoch nicht für alle Kundendialoge empfehlenswert. Die Befragungen haben gezeigt, dass vor allem Unternehmen im Business-to-Business (B2B) oder Unternehmen mit wenigen, aber dafür sehr individuellen Kundendialogen, eher keine Bots einsetzen. Dafür zeigt die Studie, dass diese Unternehmen durchaus Generative AI zur Unterstützung der Agenten am Telefon nutzen (Abb. 5). Häufig werden derartige System auch als Enablement bezeichnet. Hier sind die KI-Assistenten die häufigste Umsetzung, welche während des Gesprächs mithören und dem Telefon-Agenten nützliche Tipps, direkt während des Telefonats, geben. Der Agent kann so schneller und meist auch passender auf das Kundenbedürfnis eingehen. Die KI nutzt als Datenbasis in der Regel eine Vielzahl von verschiedenen Datenquellen. Dies können Transkriptionen von vergangenen Telefonaten, allgemeine Webseiten-Texte, Produktdatenbanken und vieles mehr sein. Ein sehr gutes Beispiel liefert in diesem Zusammenhang die Krankenkasse KPT. Sie finden die Case Study „Generative AI im Kundenkontakt der KPT Versicherung“ auf Seite 40.

Agenten-Unterstützung am Telefon

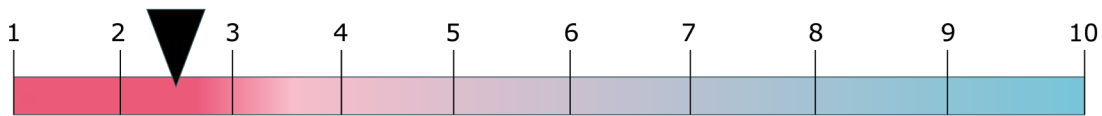


Kundenanwendungen

Automatisierte Dunkelverarbeitung

Abb. 5: Agenten-Unterstützung am Telefon auf der Gen AI Skala

Agenten-Unterstützung in der E-Mail-Bearbeitung



Kundenanwendungen

Automatisierte Dunkelverarbeitung

Abb. 6: Agenten-Unterstützung in der E-Mail Bearbeitung auf der Gen AI Skala

Neben der Unterstützung am Telefon haben fünf der befragten Unternehmen angegeben, dass ihre Mitarbeitenden zur schnelleren und qualitativ-hochwertigeren Beantwortung von E-Mails Generative AI einsetzen (Abb. 6). Das zeigt unter anderem die Generative AI-Anwendung des deutschen Handelsunternehmens «Unite». Der hier dargestellte Anwendungsfall von Generative AI ist auf die anderen befragten Unternehmen, die Generative AI in der E-Mail-Beantwortung nutzen, übertragbar.

Anzumerken ist, dass wie bei der Agenten-Unterstützung am Telefon auch bei dieser Anwendung die KI den Serviceprozess nie vollständig übernimmt, sondern sie im Sinne eines KI-Assistenten (hier: AI-Companion) quasi beratend zur Seite steht.

Im ersten Schritt fasst die KI die Kundenanfrage für den Berater lediglich zusammen. Anschliessend schlägt die KI eine mögliche „Nächstbeste Kommunikation“ vor. Dieser Text kann vom Berater weiter angepasst werden. Der Berater kann in Form von Prompts (Befehle an die KI) weitere Befehle an den KI-Assistenten geben. Die von der KI abgegebenen Handlungsempfehlungen werden final von den Mitarbeitenden kontrolliert, bevor die KI-generierten Antworten gesendet werden. Die Mitarbeitenden von Unite übernehmen somit eine «Gatekeeper-Funktion», mit der sie die Prozessausführung, jedoch nicht die Prozessverantwortung, individuell an die KI delegieren können. Als Datengrundlage stehen der Generative AI, ähnlich wie bei der Agenten-Unterstützung am Telefon, eine Vielzahl von Datenquellen aus dem Unternehmen, wie beispielsweise auch Kundendaten, zur Verfügung.

Im Praxiseinsatz sorgt KI bei Unite für hohe Effizienz und Empathie in der Kundenkommunikation. Das LLM erkennt nicht nur automatisch den Inhalt und die Tonalität einer Konversation, sondern erstellt in Echtzeit Antwortvorschläge, die vollständig, empathisch, rechtschreib- und gesetzeskonform sowie individuell sind. Und dies in einer der zehn Sprachen, in denen Unite mit seinen Kunden kommuniziert. Weiter ist es möglich, der KI unterschiedliche Tonalitäten für verschiedene Anwendungsbereiche

anzutrainieren und diese je nach Situation automatisch ausgeben zu lassen.

Der Einsatz von KI im Kundenservice bietet für Unternehmen wie Unite eine Vielzahl von Vorteilen. Insbesondere bei Standardfragen, die bereits im Vorfeld in das Sprachmodell der jeweiligen Unternehmen eintrainiert wurden, zeigt sich die hohe Effizienz und Geschwindigkeit der automatisierten Beantwortung von Kundenanfragen. Durch die Automatisierung wird zudem eine gleichbleibende Servicequalität gewährleistet, da menschliche Fehler in Standardprozessen weitestgehend minimiert werden. Die Wirtschaftlichkeit, die mit der Automatisierung einhergeht, ermöglicht es den Mitarbeitenden, sich auf wertschöpfende und spezialisierte Aufgaben zu konzentrieren. Dies wiederum führt zu einer höheren Zufriedenheit der Mitarbeitenden, da sie sich auf erfüllendere Aufgaben fokussieren können, anstatt sich mit wenig anspruchsvollen Routinetätigkeiten zu beschäftigen. Zudem bietet die Automatisierung damit die Notwendigkeit zur persönlichen beruflichen Entwicklung durch Spezialisierung und Weiterbildung. Diese Vorteile machen den Einsatz der neuen Lösung im Kundenservice von Unternehmen wie Unite auch äusserst attraktiv für Mitarbeitende.

Unite verfolgt einen ausgewogenen Ansatz, bei dem das Gleichgewicht zwischen Menschen und Maschinen im Kundenservice im Vordergrund steht. Das Unternehmen strebt an, den Kundenservice so weit wie möglich zu automatisieren, ohne dabei die Qualität zu beeinträchtigen. Komplexere Aufgaben werden weiterhin von Experten von Unite persönlich mit den Kunden bearbeitet. Trotz Automatisierung können die Kundendienstmitarbeitenden individuell auf emotionale Bedürfnisse eingehen und eine empathische Beratung bieten, insbesondere ist das im Bereich des Beschwerdemanagements wichtig. Self-Service-Optionen für Kunden werden nur dann angeboten, wenn sie einen echten Mehrwert bieten, wie zum Beispiel eine spürbare Zeitersparnis. Dieser ausgewogene Ansatz gewährleistet eine optimale Kundenbetreuung und Kundenzufriedenheit bei Unite. Einen ähnlichen Ansatz verfolgt die Unternehmensgruppe Walbusch, wie Sie der Case Study auf der nächsten Seite entnehmen können.

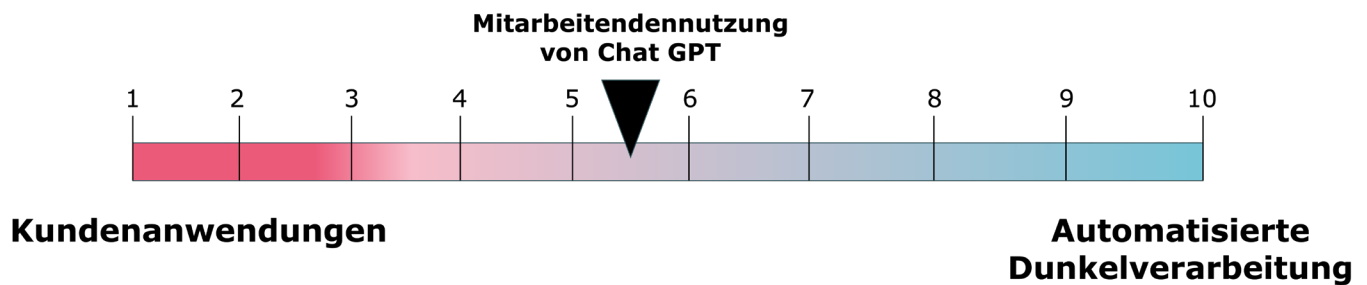


Abb. 7: Mitarbeitendennutzung von Chat GPT auf der Gen AI Skala

Knapp die Hälfte der interviewten Unternehmen gab an, dass sie ihren Mitarbeitenden die Nutzung von ChatGPT nicht verboten haben, sondern lediglich Weisungen ausgesprochen haben, wie Mitarbeitende den KI-Assistenten nutzen dürfen. Dazu gehört hauptsächlich die Weisung, ChatGPT nicht mit persönlichen oder unternehmensinternen Daten zu nutzen. Auf die Frage, wie dies kontrolliert wird, gab es jedoch keine konkreten Antworten. Vielmehr appellieren die Unternehmensleitungen an ihre Mitarbeitenden, dass diese verantwortungsvoll mit schützenswerten Daten umgehen.

Während der Befragung konnten die Umfrageteilnehmenden nur wenig zu den konkreten Anwendungen rund um die Nutzung von ChatGPT ihrer Mitarbeitenden sagen. Dies erstaunt nicht aufgrund der Tatsache, dass Unternehmen ihre Angestellten mit der Nutzung von ChatGPT eher sich selbst überlassen. Im Allgemeinen wurde aber deutlich, dass ChatGPT vor allem zur Text- oder Bildgenerierung eingesetzt wird.

Weiterlesen auf Seite 20

Usecase

Generative AI im Kundenservice der Walbusch-Gruppe zur automatisierten Bearbeitung von Geschäftsvorfällen



Ein Gastbeitrag von Thorsten Schmelz, Bereichsleiter Kundenservice Walbusch Walter Busch GmbH & Co. KG, Mitglied der Geschäftsleitung

Der Usecase

Seit 2012 arbeitet die Walbusch-Gruppe mit der CRM-Lösung der Schweizer BSI AG. Seitdem wurde das System beständig zu einer ganzheitlichen Suite rund um Kunden, Daten und Prozesse für eine optimale Customer Experience weiterentwickelt. 2019 ergab sich durch die Einbindung neuer KI-Technologien die Möglichkeit, Geschäftsvorfälle automatisiert zu bearbeiten.

Walbusch kommt historisch aus dem Katalogbasierten Versandhandel. Kataloge und Mailings waren und sind daher auch nach Einführung des Online-Kanals weiter wichtige Impulse, um Kunden zu erreichen. Mit der Entwicklung der Walbusch-Gruppe vom reinen Katalogversender zu einem Multi-Channel-Unternehmen veränderten sich allerdings sowohl das Kaufverhalten als auch das Informationsbedürfnis der Kunden. Hinzu kommen Anforderungen aus einem wachsenden Bewusstsein für umweltschonende Ressourcennutzung sowie aus individuellen Kanal- und Frequenz-Präferenzen der Kunden.

Zur Bearbeitung der Kundenanliegen nutzte die Walbusch-Gruppe in ihrem Kundenservice einen Geschäftsvorfall zur Reduktion von Werbemitteln. Ein Geschäftsvorfall ist ein strukturierter Ablauf innerhalb des CRM, der den Nutzer in definierten Schritten durch den Prozess führt und mit einem standardisierten Output für ein gleichbleibendes Qualitätsergebnis sorgt. Die Klassifikation als eigenständiger Geschäftsvorfall ist daher nur sinnvoll, wenn der Prozess in einer ausreichenden Anzahl vorkommt. Dies trifft ebenso auf den Einsatz von KI zur Automatisierung zu. Dieser Geschäftsvorfall erfüllte eine erste wichtige Voraussetzung für den Einsatz von KI. Ausserdem konnten die entsprechenden Kundenanliegen, die per Mail an den Kundenservice gerichtet waren, mit einer hohen Quote erkannt und einem entsprechenden Vorgangspool zugewiesen werden. Zudem war der Geschäftsvorfall in der manuellen Bearbeitung bereits stark vorstrukturiert. So musste der Bearbeiter z.B. nur die sogenannte Kundeninfo (der Grund des Kunden für den Werbereduktionswunsch) eingeben und die entsprechenden Sperrkennzeichen wurden vorausgefüllt. Damit waren die Voraussetzungen für eine Automatisierung optimal.

Für die Erkennung wurde das Large Language Model BERT von Google genutzt. Dieses kam bereits vortrainiert mit 3,3 Milliarden Wörtern in 120 Sprachen.

Um das LLM auf Walbusch-Daten zu trainieren, wurde es mit 800.000 E-Mails und damit verknüpften Geschäftsvorfällen angereichert. Diese E-Mails enthielten diverse Kundenanliegen. Im Speziellen wurden darüber hinaus noch 100.000 Geschäftsvorfälle und Werbereduktion mit den entsprechenden E-Mails ins supervised Training einbezogen. Dieses Training erfolgte in nur 16 Stunden.

Die eigentliche Arbeit für Walbusch ergab sich im Anschluss. Um die Geschäftsvorfälle automatisiert bearbeiten zu können, braucht es eine ausreichend hohe accuracy rate > 90%.

Daher hat ein kleines Team von Kundenservice-Mitarbeitenden Fälle, die die KI nicht entscheiden konnte, über Wochen manuell bearbeitet. Aus den Ergebnissen hat die KI weiter gelernt, bis sich eine accuracy von 93% eingestellt hat.

Die KI erkennt den Reduktionswunsch des Kunden in verschiedenen Ausprägungen, die prozessual hintereinandergeschaltet sind. Zunächst wird die Entscheidung getroffen, ob es sich überhaupt um einen Werbesperrwunsch handelt. Danach wird geprüft, ob es sich um einen Wunsch nach Vollsperrung oder um eine Kanalsperre handelt. Dementsprechend werden automatisiert Sperrkennzeichen gesetzt und eine Kommunikation an den Kunden verschickt. Dabei wird der Kunde darüber informiert, dass es sich um eine automatisierte Antwort handelt. Alle anderen bzw. nicht eindeutige Vorgänge werden in die manuelle Bearbeitung ausgesteuert.

Mehrwerte und Lehren

Der Einsatz von KI zur automatisierten Bearbeitung von Geschäftsvorfällen war für die Walbusch-Gruppe ein voller Erfolg. Bereits kurz nach der gelungenen Einführung des ersten KI-Geschäftsvorfalles wurden weitere Cases live genommen. Hierzu gehören Rückmeldungen von Kunden auf Zahlungserinnerungen und Mahnungen, die nicht prozessrelevant sind: Hinweise zu Zahlungsüberschneidungen, Ankündigung von Forderungsbegleichungen oder Entschuldigungen für Fristüberschreitungen. Der Mahnprozess ist vordefiniert, vor jeder neuen Mahnstufe erfolgt erneut ein Saldoabgleich, vor der Inkassoübergabe zudem eine manuelle Prüfung. Die Rückmeldungen sind aus Kundensicht nachvollziehbar, die genannten Inhalte haben aber keinen Einfluss auf den Mahnprozess und müssen daher nicht individuell bearbeitet werden. Auch hier erfolgen Erkennung und Bearbeitung durch die KI.

Darüber hinaus werden mittels KI-Rückmeldungen zu laufenden Bestellungen identifiziert und priorisiert. Dies betrifft z.B. fehlerhafte oder vom Kunden vergessene Angaben zu abweichenden Lieferanschriften, bestellten Artikeln oder Artikelgrößen. Solange Bestellungen noch nicht fakturiert sind, sind Korrekturen, Ergänzungen oder Stornierungen noch möglich, so dass in der Folge teure Retouren, zusätzliche Sendungen und Service-Kontakte vermieden werden können.

Im Ergebnis konnten seit Einführung über 54.000 Kundenanliegen automatisiert bearbeitet werden, bei 45.000 Geschäftsvorfällen konnte zumindest manueller Aufwand reduziert werden, indem die KI einen Vorschlag erzeugen konnte. Zudem konnten mehr als 4.000 Retouren und weitere Service-Kontakte vermieden werden. So konnte die Walbusch-Gruppe einen mittleren sechsstelligen Betrag einsparen und zudem die Service-Qualität in Richtung Kunde erhöhen.

Gleichzeitig wurden aus den ersten KI-Geschäftsvorfällen wichtige Learnings erzielt:

1. KI ist kein Selbstläufer. Zum einen muss KI trainiert werden bis sie akzeptable Ergebnisse liefert, zum anderen muss sie bezogen auf den Automatisierungsgrad und die Qualität gemonitort werden. Im beschriebenen Case haben zunächst operative Führungskräfte das Projekt neben dem Tagesgeschäft aufgesetzt und betreut. Dabei ist das Erfolgs Monitoring nach Einführung zu kurz gekommen. Der Automatisierungsgrad hatte sich nach einigen Monaten durch Veränderungen im Prozess - zunächst unbemerkt - deutlich verschlechtert. Ausserdem mussten Evaluierung und Umsetzung weiterer KI-Massnahmen oftmals operativen Anforderungen untergeordnet werden. Daher wird jetzt eine verantwortliche Funktion geschaffen, die den Fokus auf dem Thema KI-Prozesse halten wird.
2. Es braucht eine ausreichende Menge an Daten für eine Automatisierung. Die medial oft sehr attraktiv aufbereiteten und eindrucksvollen KI-Cases, insbesondere aus den kreativen Bereichen, funktionieren auch deshalb so gut, weil sie auf Milliarden von Daten aus dem Internet basieren. Wenn für eine KI-Umsetzung die reinen Unternehmensdaten herangezogen werden müssen, braucht es u.U. viel Zeit, um das System zu trainieren. Die Möglichkeiten der generativen KI können Trainingsprozesse inzwischen beschleunigen und bereits mittels Empfehlungen (next best action, Textvorschläge) und Teilautomatisierungen den Arbeitsaufwand erheblich reduzieren.

So können der Automatisierungsgrad etwa dynamisch gesteigert oder Potentiale für den Einsatz eines gezielten Automatisierungsprojektes entwickelt werden. Es empfiehlt sich, über einen Business Case die produktiven Kosten für das ausgewählte KI-Modell und den Projektaufwand mit dem gewünschten Output bzw. Automatisierungsgrad ins Verhältnis zu setzen. Nicht immer muss das neueste und leistungsstärkste - und damit oft teuerste – KI-Modell genutzt werden.

3. Es braucht einen strukturierten Prozess. Ohne einen konstanten Ablauf kann eine KI zwar in der Erkennung helfen, nicht aber im Prozess unterstützen.

Ausblick

Die Walbusch-Gruppe hat weitere Top-Prozesse definiert, in denen sie Automatisierungspotentiale sieht, die nun mit dem Partner BSI analysiert und angegangen werden. Darüber hinaus wird sie im Laufe des Jahres auf das aktuelle Release der BSI Customer Suite wechseln. Als Voraussetzung dafür ist das System bereits erfolgreich in die Cloud migriert worden. Mit dem neuen Release steht den Mitarbeitenden dann der sogenannte Companion als interaktiver KI-Assistent zur Verfügung. Dieser wird Kundenanliegen zusammenfassen und mit offenen Vorgängen verknüpfen können. Er wird die Mitarbeitenden mit Empfehlungen für die nächste Kundenaktion und mit der eigenständigen Erstellung von Textvorschlägen unterstützen.

Mithilfe von Sentiment Analysen, dem Companion als KI-Assistenten und automatisierten Geschäftsprozessen wird der Kundenservice zum Smart Service Center. Die Kunden werden zu jeder Zeit mit Selfservices, virtuellen KI-Agents und bestmöglich unterstützenden, engagierten Mitarbeitenden in ihrer customer journey begleitet.

Über das Unternehmen

Walbusch wurde 1934 als Versandhandel in Solingen gegründet. Seit den 60er Jahren hat sich das Unternehmen einen Namen für gute Hemden gemacht. Heute bietet Walbusch ein

Vollsortiment qualitativ, hochwertiger Männer- und Damenmode in Deutschland, sowie in den Niederlassungen in Österreich und der Schweiz an. Ausgehend vom Katalogangebot bietet Walbusch den kompletten Service eines Multi-Channel-Anbieters: Der Kunde kann telefonisch bestellen, bequem in den Online-Shops kaufen oder sich in einem der über 40 Walbusch-Fachgeschäfte beraten lassen. Zur Walbusch-Gruppe gehören darüber hinaus der Gesundheitsversand Avena sowie der Männermode-Spezialist Mey & Edlich.

Das inhabergeführte Familienunternehmen beschäftigt ca. 1.100 Mitarbeiter.



Kontaktinformationen

Thorsten Schmelz, Bereichsleiter Kundenservice Walbusch
Walter Busch GmbH & Co. KG, Mitglied der Geschäftsleitung

thorsten.schmelz@walbusch-gruppe.com

Usecase

Generative AI im operativen Betrieb der Gemeinde Glattfelden



Ein Gastbeitrag von Valentino Vinzens, Gemeindeschreiber Gemeinde Glattfelden

Zur Ideenfindung und zum Verfassen der Inhalte hat Valentino Vinzens die AI-Technologie «SwissGPT» genutzt.

Im Zeitalter der digitalen Transformation nimmt die Bedeutung innovativer Technologien kontinuierlich zu, insbesondere im Hinblick auf die Steigerung von Effizienz und Effektivität innerhalb von Verwaltungsprozessen. Die Generative Künstliche Intelligenz (KI), vor allem in Form von SwissGPT, stellt eine solche bahnbrechende Technologie dar, die neue Wege für fortschrittliche Lösungen in der Gemeindeverwaltung eröffnet. Das Team steht an der vordersten Front dieser technologischen Revolution und nutzt die Möglichkeiten Generative AI, um nicht nur unsere Arbeitsabläufe zu optimieren, sondern auch um innovative Antworten auf komplexe Fragestellungen zu finden. Dieser Einsatz war zunächst aufgrund datenschutzrechtlicher Bedenken und der Tatsache, dass viele Gemeinden den Zugang zu Diensten wie ChatGPT gesperrt hatten, eine Herausforderung. Trotz dieser Hindernisse hat das Team Wege gefunden, Generative AI verantwortungsvoll und effektiv in der Gemeindeverwaltung Glattfelden einzusetzen, um sowohl unsere Dienstleistungen als auch die internen Prozesse grundlegend zu verbessern. Dieser Beitrag beleuchtet den innovativen Einsatz Generative AI, die Überwindung anfänglicher Hürden und die bemerkenswerten Ergebnisse des Vorgehens.

Die Gemeindeverwaltung Glattfelden sah sich mit der Herausforderung konfrontiert, eine Vielzahl zeitaufwendiger Aufgaben zu bewältigen, die sowohl Ressourcen banden als auch die Fähigkeit einschränkten, sich auf unsere Kernkompetenzen und -aufgaben zu konzentrieren.

Um diesen Herausforderungen zu begegnen, hat sich das Team für den Einsatz von Generative AI, insbesondere SwissGPT, entschieden. Diese Technologie versprach eine Vielzahl von Aufgaben effizienter zu gestalten, darunter das Schreiben von Berichten, das Entwerfen von Anträgen und Beschlüssen, das Kompaktzusammenfassen grosser PDF-Dokumente, Übersetzungen, das Erstellen von Newsmeldungen für Social Media oder für die Website sowie das Verfassen von Medienmitteilungen und Reden für Behördenmitglieder.

Die Implementierungsphase in der Gemeindeverwaltung sah vor, schrittweise spezifische Anwendungen Generative AI in die Arbeitsabläufe zu integrieren, beginnend mit den Bereichen, in denen der grösste unmittelbare Nutzen erwartet wurde. Herausforderungen, vor allem in Bezug auf Datenschutz und Akzeptanz, wurden durch umfassende Schulungen und die Entwicklung klarer Richtlinien zur Datennutzung überwunden.

Mehrwert & Lehren

Durch den Einsatz Generative AI konnte die Gemeindeverwaltung nicht nur die Effizienz bei der Bewältigung zeitaufwendiger Aufgaben steigern, sondern auch signifikant mehr Zeit für Kernaufgaben freimachen. Besonders die Fähigkeit, umfangreiche Dokumente schnell zusammenzufassen und hochwertige Inhalte in kürzerer Zeit zu erstellen, hat sich als wertvoll erwiesen.

Der Mehrwert des Einsatzes Generative AI liegt nicht nur in der Effizienzsteigerung, sondern auch in der Möglichkeit, das Personal von repetitiven Aufgaben zu entlasten und so mehr Raum für kreative und strategische Arbeit zu schaffen.

Eine wichtige Erkenntnis aus diesem Prozess ist, dass die erfolgreiche Integration Generative AI in der Gemeindeverwaltung eine sorgfältige Planung, klare Datenschutzrichtlinien und fortlaufende Schulungen erfordert. Auch das stetige Bewusstsein, dass diese Technologie zur Verfügung steht, muss immer wieder bei den Mitarbeitenden in Erinnerung gerufen werden. Diese Schritte sind entscheidend, um sowohl die Akzeptanz als auch die verantwortungsvolle Nutzung der Technologie sicherzustellen.

Angesichts der positiven Erfahrungen und der erzielten Ergebnisse plant unsere Gemeindeverwaltung, die Nutzung Generative AI weiter zu vertiefen und zu schärfen. Sie suchen aktiv nach weiteren Anwendungsmöglichkeiten

im Alltag, die nach und nach in unsere Prozesse integriert werden sollen. Das Ziel ist es, in Zukunft noch mehr Personalressourcen zu schonen und uns vermehrt komplexeren Aufgaben zu widmen, während Fleissarbeiten verstärkt durch KI abgewickelt werden.

Der gezielte Einsatz Generative AI hat es unserer Gemeindeverwaltung ermöglicht, nicht nur vorhandene Herausforderungen effektiv zu bewältigen, sondern auch den Grundstein für zukünftige Innovationen zu legen. Durch die Nutzung von SwissGPT konnten wir bedeutende Fortschritte erzielen, die sowohl die Leistungsfähigkeit unserer Organisation als auch die Zufriedenheit der Gemeindeglieder nachhaltig verbessern. Die Erfahrung unterstreicht das Potenzial Generative AI als mächtiges Werkzeug für Verwaltungen, die an der Spitze der technologischen Entwicklung stehen möchten, und bietet wertvolle Einblicke für andere, die diesen Weg beschreiten wollen.

Über Glattfelden

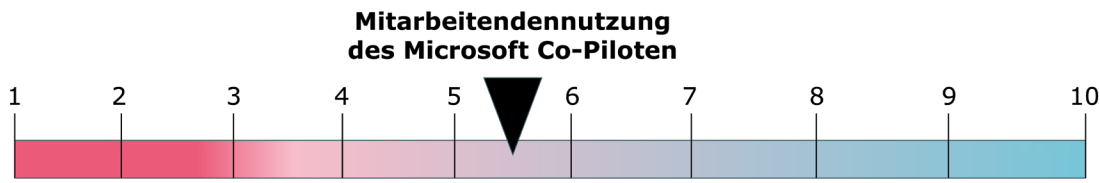
Glattfelden ist eine politische Gemeinde mit rund 5'300 Einwohnenden im Bezirk Bülach des Kantons Zürich an der Staatsgrenze zu Deutschland. Glattfelden liegt im Zürcher Unterland am letzten Teilstück des Flusses Glatt, der bei Rheinsfelden in den Rhein mündet. Der Ort umfasst neben dem Dorf Glattfelden die Siedlungen Schachen, Aarüti, Rheinsfelden und Zweidlen.



Kontaktinformationen

Valentino Vinzens, Gemeindegliederschriftführer Gemeinde Glattfelden

valentino.vinzens@glattfelden.ch



Kundenanwendungen

**Automatisierte
Dunkelverarbeitung**

Abb. 8: Mitarbeitendennutzung des Microsoft Co-Piloten auf der Gen AI Skala

Neben der Nutzung von ChatGPT ist auch der Microsoft Co-Pilot eine viel genannte Generative AI-Anwendung, wenn es um die eher mitarbeiterorientierten Generative AI-Anwendungen geht (Abb. 8). Der Co-Pilot ist vor allem bei Unternehmen mit mehr als 50 Mitarbeitenden im Einsatz und wird hier mehrheitlich über die interne IT-Abteilung an die Mitarbeitenden ausgerollt. Zum Zeitpunkt der Befragungen gab es kein Unternehmen, bei dem der Co-Pilot über die gesamte Organisation hinweg verfügbar war. Es gibt lediglich einzelne Pilotgruppen oder Abteilungen, die den Co-Piloten evaluieren. Abgesehen von der klassischen Text- und Bildgenerierung gab ein Drittel der Co-Pilot Nutzenden an, dass sie das Tool mehrheitlich im Zusammenhang mit dem Verfassen von E-Mails nutzen. Dies vor allem, wenn der Inhalt der E-Mail unternehmensinterne Informationen enthält, die der Co-Pilot dank Anbindungen an interne Systeme schnell zusammentragen kann.

Unter den befragten Unternehmen gaben vor allem die Mitarbeitenden aus Gesundheitsbetrieben (Privatkliniken), wie auch Medien- und Retailunternehmen an, dass Generative AI immer mehr im Marketing zum Erstellen von Newsletter-, sowie Social Media Texten verwendet wird (Abb. 9). Dabei nutzen die Mitarbeitenden mehrheitlich Anwendungen wie ChatGPT oder den Microsoft Co-Piloten. Lediglich ein Unternehmen gab an, eine spezifische, eigenentwickelte Lösung zur Erstellung von Marketing-Texten zu haben. Weiter traut sich noch kein Unternehmen, Generative AI blind zur Generierung von Marketingtexten einzusetzen. Die KI-Anwendungen dienen den Angestellten

lediglich als Inspirationsquelle. Alle von der KI erstellten Texte werden immer noch von Menschen redigiert. Bei der Erstellung von Bildern mit Generative AI werden die Bilder in der Regel eins-zu-eins übernommen, aber vor der Publikation ebenfalls von einem Menschen als brauchbar eingestuft und mit einer entsprechenden Kennzeichnung markiert, die transparent die Nutzung von Generative AI angibt. Insgesamt gaben jedoch nur zwei der befragten Unternehmen an, dass sie KI als festen Bestandteil in der Bild-Generierung integriert haben. Bei der Textgenerierung sind Tools wie ChatGPT für nahezu alle Marketers schon zum täglichen Werkzeug geworden. Eine sehr gute Anwendung zeigt die Case Study der Gemeinde Glattfelden, die die Schweizer AI-Anwendung SwissGPT als Alternative zu ChatGPT nutzt. Sie finden die Case Study, „Generative AI im operativen Betrieb der Gemeinde Glattfelden“ auf Seite 18 in dieser Studie.

Produktbeschreibungen, generiert mit einer Generative AI, haben gleich zwei Vorteile. Zum einen kann die KI eigenständig aus unterschiedlichen Datenquellen Informationen zum Produkt einholen und diese zu einer sinnvollen Beschreibung zusammenfügen und dies oft deutlich schneller als Mitarbeitende. Zum anderen kann die Generative AI die Texte in Sekundenschnelle an unterschiedliche Zielgruppen und Medien anpassen. Von den befragten Unternehmen machten fünf Unternehmen deutlich, dass sie genau den Prozess der Verknüpfung von verschiedenen Produktdatenbanken und anderen Informationsquellen bereits abgeschlossen haben und



Kundenanwendungen

**Automatisierte
Dunkelverarbeitung**

Abb. 9: Einsatz für Marketingtexte, wie Newsletter oder Social Media auf der Gen AI Skala

die interne Effizienz mit diesem Anwendungsfall deutlich steigern konnten (Abb. 10). Weiter gaben die Unternehmen, die diese Anwendung schon nutzen, an, dass nicht nur die Effizienz und Produktivität gesteigert wurde, sondern auch die Qualität der Beschreibungen deutlich gesteigert werden konnte. Bei einem der befragten Unternehmen war die interne IT-Abteilung Treiber für diese Anwendungen. Bei den anderen Unternehmen wurde diese Anwendung direkt durch die Produktmanagementabteilungen angetrieben.

In der Vertriebsabteilung wird Generative AI genutzt, um Arbeitsprozesse zu optimieren und die Effizienz zu erhöhen (Abb. 11). Dabei ist ein essenzieller Erfolgsfaktor die gute Vernetzung der verschiedenen Applikationen und die Konsolidierung der Ergebnisse in einer einheitlichen Anwendung, welches den Vertriebsmitarbeitenden zur Verfügung steht. Generative AI konsolidiert Kundeninformationen und

Überblick über die konkrete Kundensituation hat. Darüber hinaus wird Generative AI für Finanz-Prognosen und Bedarfsplanungen eingesetzt, um die Planungsprozesse zu verfeinern.

Da es für diese Anwendungen entscheidend ist, dass die KI einen Zugang zu internen Systemen und vor allem auch zu Kundendaten hat, ist die Berücksichtigung des Datenschutzes und die Art und Weise wie und von wem die Daten verarbeitet werden, essenziell. Laut den Befragten lässt sich diese Generative AI-Anwendung nur umsetzen, wenn wichtige Regelungen zur Einhaltung des Datenschutzes berücksichtigt werden. Befragte gaben an, Anwendungen dieser Art erst nach einer gründlichen Prüfung durch die Rechts- und IT-Sicherheitsabteilungen umsetzen zu können. Andererseits haben sich gerade in den USA derartige Anwendungen unter der Bezeichnung «Sales-Enablement» durchgesetzt.

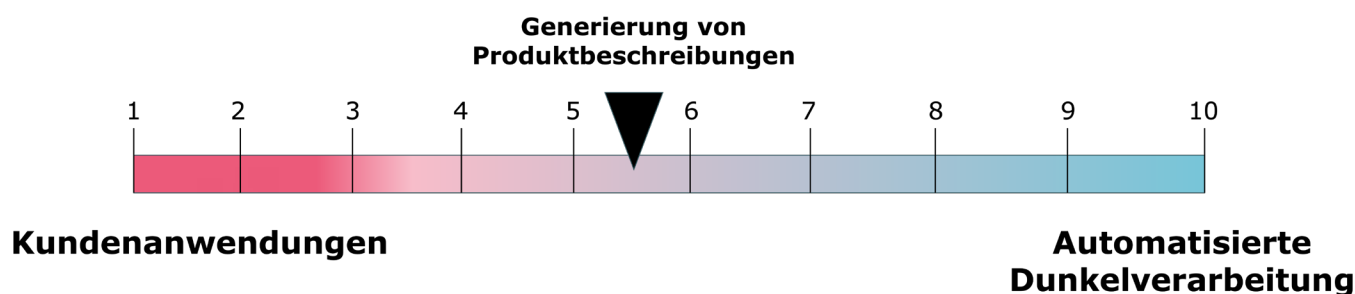


Abb. 10: Generierung von Produktbeschreibungen auf der Gen AI Skala

bereitet sie für den Berater in einer Form auf, dass dieser die Informationen direkt für seine Vertriebsaktivitäten nutzen kann. Gerade in Branchen, in denen die meisten Vertriebsaktivitäten auf Betreiben des Kunden zustande kommen (Inbound) und Vertriebsmitarbeiter zum Teil mehrere Hundert Kunden betreuen, ist eine schnelles empathisches «Sich in die Situation des Kunden versetzen» oft Schlüssel zum Erfolg. In diesem Zusammenhang kann Generative AI beispielsweise die Informationen aus der Kundenakte zu einem Bild verdichten, so dass ein Berater schnell einen

Drei der befragten Unternehmen haben einen Personal-Assistent (englisch: HR-Assistent) mit Generative AI als Anwendung skizziert (Abb. 12). Dabei unterstützt der digitale Assistent (Chatbot) sowohl neue Mitarbeitenden zu Beginn ihres Arbeitsverhältnisses, er steht aber auch bestehenden Mitarbeitenden für häufige Fragen zur Verfügung. Grundsätzlich ist der Einsatz von Chatbots während der Onboarding-Phase bzw. allgemein als Personal-Assistent nichts Neues. Chatbots werden hier schon lange eingesetzt, die Mitarbeitenden entweder proaktiv auf die

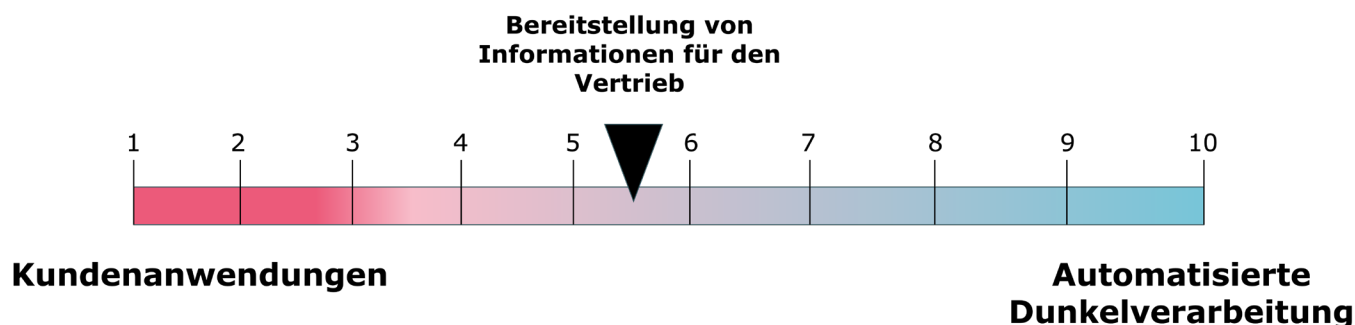


Abb. 11: Bereitstellung von Informationen für den Vertrieb auf der Gen AI Skala

wichtigsten Guidelines, Regeln und Abläufe im Unternehmen aufmerksam machen oder ihre allgemeinen Fragen beantworten. Der Vorteil beim Einsatz der Generative AI ist an dieser Stelle, dass der digitale Assistent jeweils die Tonalität seines Nutzers annehmen kann und in sehr natürlicher Sprache kommunizieren kann. Als Datengrundlagen werden klassische interne Leitfäden bzw. Informationsbroschüren und andere interne Richtlinien bzw. Dokumente verwendet.

Obwohl die Anwendung sehr trivial erscheint und einzelnen Unternehmen auch vor Generative AI-Anwendungen dieser Art eingeführt haben, hat nur ein kleiner Teil der Befragten diese Anwendung genannt. Es wird vermutet, dass

kette hat, wird Generative AI ebenfalls immer häufiger eingesetzt (Abb. 13). Die Mehrheit der Befragten Unternehmen aus dem IT-Umfeld nutzt Anwendungen, wie «Github Copilot» oder ChatGPT mindestens zum automatisierten Code-Review. Einzelne erste Unternehmen nutzen Generative AI sogar zur initialen Generierung von Software-Codes. Dabei muss erwähnt werden, dass in diesem Fall das Testen bzw. Korrigieren des Codes meist länger dauert, als wenn der initiale Code von einem Menschen erstellt worden ist. Dennoch berichten Unternehmen auch bei dieser Anwendung insgesamt gute Erfahrungen und auch trotz längeren Optimierungs-Phasen einen insgesamten Anstieg der Produktivität. Um diese Anwendung detaillierte zu beschreiben, haben wir Ihnen die Case-Study von Pension Dynamics auf der nächsten Seite zur Verfügung gestellt.

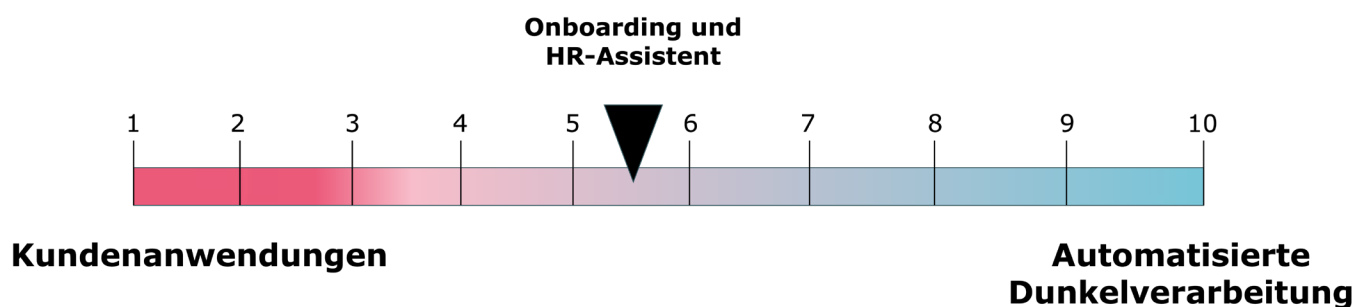


Abb. 12: Onboarding und HR-Assistent auf der Gen AI Skala

Personal-Abteilungen dem Thema Generative AI vielleicht bislang noch nicht so viel Aufmerksamkeit geschenkt haben. Im Fokus stehen bei der Personal-Abteilung häufig der Umgang mit Bewerbungen, anstatt Anwendungen zu entwickeln, die die Entwicklung der bestehenden Mitarbeitenden fördern.

Bei Unternehmen, in denen die Software-Entwicklung einen hohen Stellenwert in der gesamten Wertschöpfungs-

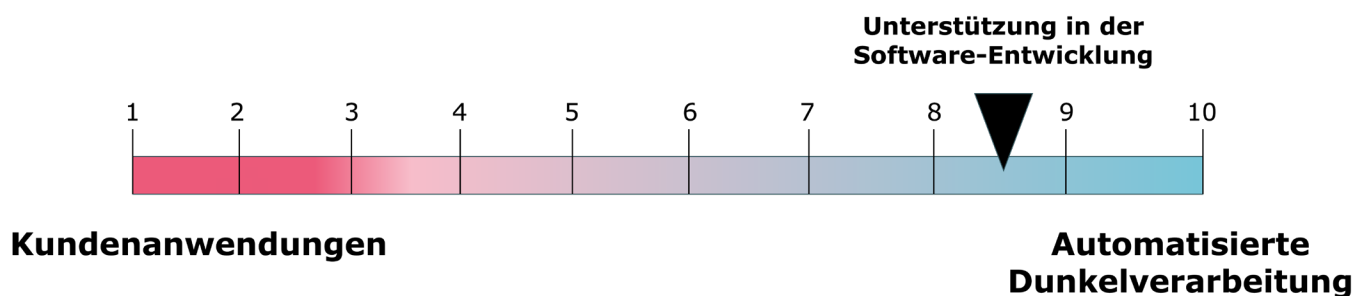


Abb. 13: Unterstützung in der Software-Entwicklung auf der Gen AI Skala

Usecase

Generative AI in der Softwareentwicklung mit Generative AI bei PensionDynamics



Ein Gastbeitrag von Yvonne Wiedemar, Co-CEO Experience bei PensionDynamics

Problemstellung

PensionDynamics stand vor der Herausforderung, die Effizienz und Produktivität in der Softwareentwicklung zu steigern, um den wachsenden Anforderungen des Marktes gerecht zu werden. Die vollständige Neuprogrammierung der Produkte, basierend auf einem modernen cloud-nativen Technologie-Stack, neu entwickelten UX-Prototypen und den komplexen Businessregeln des BVG, erforderte erhebliches Know-how in der Softwarearchitektur und führte zu zeitaufwendigen Programmieraufgaben.

Lösungsansatz

Um diese Herausforderungen zu bewältigen, entschied sich PensionDynamics, Generative AI-Tools in den Entwicklungsprozess zu integrieren. Dabei fiel die Wahl auf GitHub Copilot und ChatGPT von OpenAI. GitHub Copilot wird vor allem für Code-Vorschläge und Autovervollständigungen genutzt, während ChatGPT für kontextbezogene Fragen und Problemlösungen für Standardaufgaben oder zur Dokumentation eingesetzt wird.

GitHub Copilot unterstützte die Entwickler durch kontinuierliche Code-Vorschläge und Autovervollständigungen, wodurch klar definierte Programmieraufgaben schneller erledigt wurden. Sowohl GitHub Copilot als auch ChatGPT bieten detaillierte Erklärungen und Lösungen für spezifische Probleme, was die Effizienz bei der Problemlösung deutlich erhöht.

Beispiel 1: Codeerstellung anhand von Prompt (Bild: nira.com)

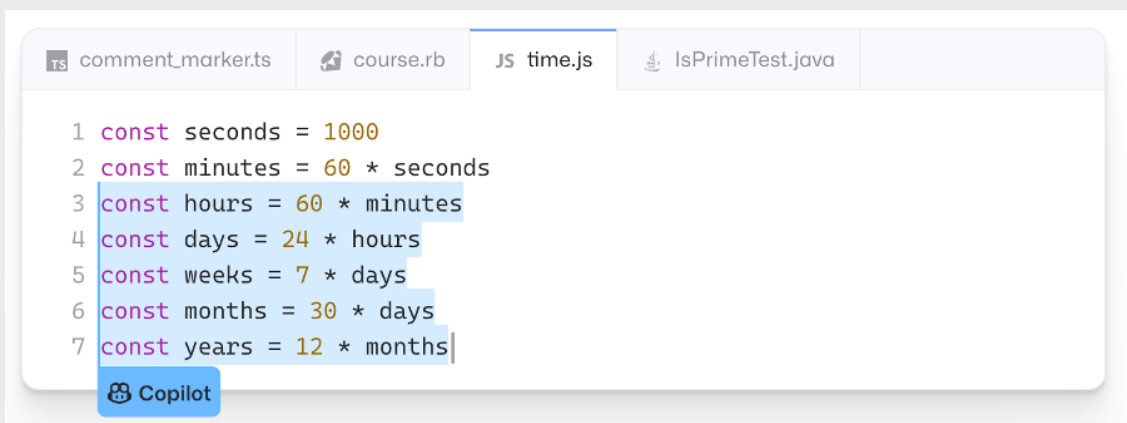


The screenshot shows a code editor with four tabs: 'collaborators.ts', 'get_repositories.py', 'non_alt_images.js', and 'PersonUtils.java'. The active tab is 'get_repositories.py'. The code is as follows:

```
1 import urllib.request, json
2
3 def get_repositories(org):
4     """List all names of GitHub repositories for an org."""
5     url = 'https://api.github.com/orgs/' + org + '/repos'
6     request = urllib.request.Request(url)
7     response = urllib.request.urlopen(request)
8     data = json.loads(response.read().decode())
9     return [repo['name'] for repo in data]
```

A Copilot icon is visible at the bottom left of the code block.

Beispiel 2: Automatisches ausfüllen anhand von getippten Codezeilen (Bild: nira.com)

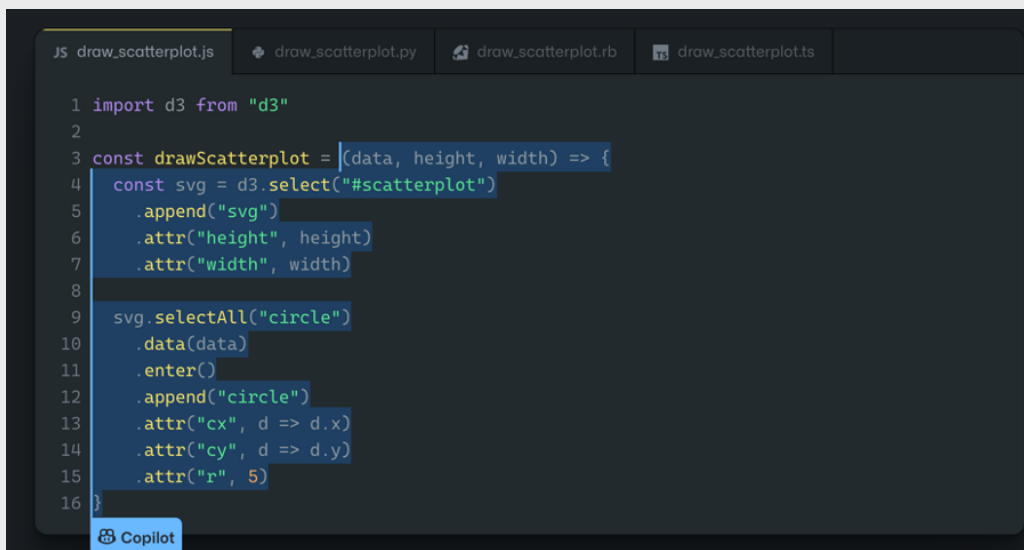


The screenshot shows a code editor with four tabs: 'comment_marker.ts', 'course.rb', 'time.js', and 'IsPrimeTest.java'. The active tab is 'time.js'. The code is as follows:

```
1 const seconds = 1000
2 const minutes = 60 * seconds
3 const hours = 60 * minutes
4 const days = 24 * hours
5 const weeks = 7 * days
6 const months = 30 * days
7 const years = 12 * months
```

A Copilot icon is visible at the bottom left of the code block.

Beispiel 3: Code-Vorschlag Grundaufbau (Bild: nira.com)



The screenshot shows a code editor with four tabs, all named 'draw_scatterplot.*'. The active tab is 'draw_scatterplot.js'. The code is as follows:

```
1 import d3 from "d3"
2
3 const drawScatterplot = (data, height, width) => {
4     const svg = d3.select("#scatterplot")
5     .append("svg")
6     .attr("height", height)
7     .attr("width", width)
8
9     svg.selectAll("circle")
10    .data(data)
11    .enter()
12    .append("circle")
13    .attr("cx", d => d.x)
14    .attr("cy", d => d.y)
15    .attr("r", 5)
16 }
```

A Copilot icon is visible at the bottom left of the code block.

Implementierung

Die Implementierung begann mit einer Testphase, in der verschiedene Szenarien und Aufgabenstellungen mit GitHub Copilot und ChatGPT durchgespielt wurden. Das Team von PensionDynamics evaluierte die Leistung der Tools anhand von Metriken wie Zeitersparnis, Codequalität und Benutzerzufriedenheit. Nach der erfolgreichen Testphase wurden die Tools schrittweise in den regulären Entwicklungsprozess für das Backend integriert.

Herausforderungen, wie die Integration, in bestehende Entwicklungsumgebungen und die Schulung der Entwickler wurden durch gezielte Trainings und Leitfäden überwunden. Besondere Sorgfalt wurde auf Code Reviews (durch Menschen) und Informationssicherheit gelegt, um die Integrität und Qualität des Codes zu gewährleisten und den Schutz sensibler Daten sicherzustellen.

Ein bedeutender Vorteil der Implementierung war die drastisch verkürzte Time-to-Market, was einen erheblichen Wettbewerbsvorteil darstellt. Durch die schnellere Bereitstellung neuer Funktionen und Produkte konnte PensionDynamics die effektive Programmierzeit messbar um das Dreifache verkürzen im Vergleich zu herkömmlichen Methoden. Dies ermöglichte dem Team, die Produktentwicklung insgesamt zu verbessern und innovative Lösungen effizienter zu liefern.

Im Rahmen des Informationssicherheitsmanagementsystems (ISMS) nach ISO 27001 wurden die KI-Tools auf IS-Risiken evaluiert und bewertet. Im Rahmen dieser Evaluation wurde eine Guideline für den Umgang mit AI-Tools geschaffen. Die Entwicklungsumgebung von PensionDynamics ist so konzipiert, dass Personendaten nur in der Produktion in sicher zugänglichen Containern vorhanden sind, die nicht an die Tools oder die Entwicklungsumgebung angeschlossen sind. Sensible Daten sind so jederzeit geschützt und es ist sichergestellt, dass keine unbefugten Zugriffe auftreten.

Durch diese Massnahmen konnte PensionDynamics die Effizienz ihrer Softwareentwicklung erheblich steigern und den Entwicklern mehr Freiraum für kreative und strategische Aufgaben schaffen.

Mehrwert und Lehren

Durch den Einsatz Generative AI konnte PensionDynamics beeindruckende Ergebnisse erzielen. „**Die Einführung von GitHub Copilot und ChatGPT hat uns ermöglicht, unsere Entwicklungsprozesse drastisch zu beschleunigen und gleichzeitig die Codequalität signifikant zu verbessern,**“ sagt Remo Peduzzi, CTO von PensionDynamics. Natürlich kann das AI-Tool nicht die Software an sich programmieren, aber einzelne klar definierte Funktionen und Mikroteile kann sie sehr gut vorbereiten. Die Entwicklungszeit wurde so um das Dreifache verkürzt. Messungen ergaben, dass Entwickler dank GitHub Copilot klar definierte Codieraufgaben bis zu 70 % schneller erledigen konnten. Zudem verbesserte sich die Codequalität durch kontinuierliche Code-Vorschläge und schnelle Fehlerbehebung. Die Dokumentation konnte nun zügig und umfassend erledigt werden, und die Bug-Rate im produktiven Einsatz wurde um min. 30 % reduziert (erste Indizien, fortlaufende Messung läuft).

Der Einsatz von Generative AI brachte dem Unternehmen sowohl quantitative als auch qualitative Verbesserungen. Die gesteigerte Effizienz ermöglichte eine schnellere Bereitstellung neuer Features und erhöhte die Kundenzufriedenheit. Durch die Automatisierung von Standardaufgaben können sich die Entwickler nun vermehrt auf komplexere Aufgaben mit dezidiert Businesslogik konzentrieren. Die gewonnene Zeit wird für Coaching und Pair Programming genutzt, um das Know-how und die Seniorität des Teams stetig zu erhöhen.

Neben den quantitativen und qualitativen Verbesserungen hat das gesamte Team, während der Einführung von Generative AI, auch viele neue Erkenntnisse rund um den Einsatz der neuen Technologien gewonnen: So zeigte sich, dass GitHub Copilot sich als besonders nützlich für Code-Vorschläge erweist, während die architektonische Planung weiterhin die Expertise erfahrener Entwickler erfordert. Die Qualität der KI-Vorschläge hängt stark von der Expertise des Entwicklers ab.

Die Bedeutung der Informationssicherheit ist ebenfalls entscheidend. Die Integration der KI-Tools in das bestehende ISMS und die Einhaltung der ISO 27001-Normen gewährleisten den Schutz sensibler Daten. Personendaten werden in einem separaten Container gespeichert, um Sicherheitsrisiken zu minimieren. Die sorgfältige Code-Review des AI-generierten Codes erweist sich als äusserst wichtig. PensionDynamics nutzt ein Vier-Augen-Prinzip und automatisierte Tests im Entwicklungsprozess. „Ein gut strukturiertes Testkonzept ist essenziell, um die Qualität und Sicherheit der Software zu gewährleisten, besonders bei der Nutzung von KI.“ betont der CTO Remo Peduzzi.

Für andere Unternehmen empfiehlt es sich, die Einführung Generative AI mit einer Testphase zu beginnen und die Tools schrittweise zu integrieren. Regelmässige Schulungen und Pair Programming sind wichtig, um die Nutzung zu optimieren. „Der Einsatz Generative AI erfordert nicht nur innovative Tools, sondern auch die umfassende Expertise unserer Entwickler, um wirklich transformative Ergebnisse zu erzielen.“ sagt Yvonne Wiedemar, Co-CEO Experience.

Ausblick

PensionDynamics plant, die Nutzung Generative AI weiter zu intensivieren und auszuweiten, um die Softwareentwicklung kontinuierlich zu optimieren. Zu den zukünftigen Plänen gehören:

1. Ausweitung der Tools für Frontend-Development: In der ersten Umsetzungsphase für diesen Usecase fokussierte man sich auf das Backend-Development. Während das Frontend-Team bereits ChatGPT nutzt, soll nun auch GitHub Copilot für die Kreation von React-Code ausgerollt werden.
2. Durchführen von AI Challenges: PensionDynamics wird Wettbewerbe organisieren, bei denen Teams gegeneinander antreten, um einzelne Features zu programmieren und zu testen. Das Ziel ist, voneinander zu lernen wie die AI-Tools effektiv genutzt werden können (insbesondere im Bereich Prompting und Code Suggestion) und die Qualität weiter zu verbessern. Diese Events dienen auch dazu, das Know-how im Bereich der objektbasierten Software-Entwicklung und der testgetriebenen Entwicklung zu vertiefen.
3. Kontinuierliche Suche nach innovativen Tools: Als innovatives Unternehmen mit einem Digital-Mindset

ist PensionDynamics ständig auf der Suche nach neuen Tools, die helfen, die Unternehmensvision zu verwirklichen. Der Markt für neue Technologien wird kontinuierlich beobachtet und ein Ideen-Board für neue Tools wurde eingerichtet. Diese Tools werden von Fachpersonen getestet und auf ihre Tauglichkeit geprüft – zurzeit z.B. Explorer Globe Engineer für interne Schulungszwecke.

4. Integration von KI-Tools im Support, Dokumentenmanagement und Self-Service: PensionDynamics arbeitet derzeit an Proof-of-Concepts für die Integration von KI-Tools im Supportbereich, dem Dokumentenmanagement und Self-Service in den Portalen.

Über das Unternehmen

PensionDynamics vereinfacht und transformiert die komplexe Welt der beruflichen Vorsorge in der Schweiz mit benutzerfreundlichen Applikationen für Pensionskassen, Arbeitgeber und Privatpersonen.



Kontaktinformationen

Yvonne Wiedemar, Co-CEO Experience bei PensionDynamics

yvonne@pensiondynamics.ch

5.2 Treiber für Generative AI-Anwendungen

Insgesamt gilt der Generative AI-Chatbot ChatGPT bei nahezu allen befragten Unternehmen als der grosse Treiber für den Start mit Generative AI. Zwar haben Unternehmen vereinzelt auch schon vorher mit KI experimentiert, der grosse «Durchbruch» kam dann aber erst ab November 2022 mit ChatGPT.

Geht es jedoch um die Treiber innerhalb des Unternehmens, so ergaben die Befragungen keine eindeutige Antwort. Unternehmen, welche eher konservativ organisiert sind, hatten entweder bereits eine KI-Arbeitsgruppe gegründet oder haben diese nach dem Launch von ChatGPT ins Leben gerufen. Auslöser für diese KI-Arbeitsgruppen waren dann einerseits direkte Anregungen und Aufträge aus den Geschäftsleitungen. In anderen Fällen ging die Initiative von Innovations-Teams aus, die mit einer KI-Arbeitsgruppe die Potentiale und Risiken des Themas ausloten wollten.

In einem dritten Szenario wurden Generative AI-Anwendungen von einzelnen Abteilungen an die Geschäftsleitung getragen. Dies sind vielmals Arbeitsgruppen aus dem Kundenservice, die Generative AI-Lösungen zur Steigerung der Kundenerfahrung und zur besseren Bewältigung des Kundenkontaktvolumens einsetzen wollten. In einem Unternehmen existierte sogar bereits vor 2022 eine Abteilung für Conversational Automation. Folglich war diese Abteilung der Treiber der Generative AI-Entwicklung.

Viertens sind es auch die IT-Abteilungen, die ohnehin Software für das Unternehmen einkaufen. Diese Abteilungen haben sich vor allem für Generative AI-Anwendungen zur Steigerung der Produktivität eingesetzt.

Etwa ein Drittel der Befragten gab an, dass die Beschäftigung mit den neuen Generative AI-Anwendungen genau genommen gar nicht durch interne Treiber ausgelöst worden sei, sondern die Aufmerksamkeit auf die neuen Möglichkeiten direkt durch die Technologie-Anbieter getrieben wurde. Dabei handelt es sich mehrheitlich um IT-Anbieter, die ohnehin schon mit dem Unternehmen zusammenarbeiten und die nun auch Lösungen im Umfeld Generativer AI anbieten. Zum Teil haben die IT-Anbieter die neuen Lösungen sogar zu besonderen, finanziellen Konditionen angeboten oder die Unternehmen im Rahmen einer Co-Entwicklungsphase an der Entwicklung teilhaben lassen.

Insgesamt lässt sich zusammenfassen, dass bei Unternehmen, in denen es kein spezifisches KI-Board bzw. Arbeitsgruppe gibt, die Treiber entweder aus dem IT-Department kamen oder aus den Abteilungen, in denen die meiste Wertschöpfung innerhalb des Unternehmens stattfindet. Zusätzlich wurden IT-Department oder interne Abteilungen vielmals durch externe Technologie-Anbieter erst auf die neuen Möglichkeiten aufmerksam gemacht und anschliessend durch besondere Einführungs-Angebote weiter angelockt.

5.3 Strategische Entscheidungen für die Implementation

5.3.1 Berücksichtigung von Datenschutz und Compliance Richtlinien

Aufgrund der Tatsache, dass diese Studie am Ende vor allem Finanzdienstleistungsunternehmen eine sinnvolle Unterstützung zur Nutzung Generativer AI bieten soll, wurden auch alle Umfrageteilnehmenden gefragt, welche Priorität Themen rund um Datenschutz, Privatsphäre und Regulierungen in ihrem Projekt gespielt haben.

Hier lässt sich zunächst ein deutlicher Unterschied zwischen den befragten Branchen feststellen. Branchen, wie Retail, E-Commerce oder auch IT-Hersteller betrachten das Thema Datenschutz meist eher im zweiten Schritt. Hier stand bei den Projekten zunächst eher das Mitarbeiter- oder Kundenerlebnis im Vordergrund.

Interessant ist, dass auch die befragten Unternehmen aus dem Gesundheitssektor Datenschutz als weniger relevant in der Anfangsphase eingestuft haben. Bislang wurden im Gesundheitssektor hauptsächlich Mitarbeitenden- und prozessbezogene Anwendungen mit Generativer AI umgesetzt und diese wurden dann lediglich mit entsprechenden Richtlinien zur Nutzung, beispielsweise der Hinweis keine Patientendaten in die Generative AI-Anwendung einzugeben, ergänzt. Es wurde stark an die Vernunft der einzelnen Kollegen appelliert, anstatt das Verbote oder besonders sichere Lösungen eingeführt worden sind.

Im Hinblick auf die generative AI-Anwendungen, die sich direkt an den Kunden richten, wie Bots, gab es im Allgemeinen auch keine vertiefenden Abklärungen oder

Stolpersteine aufgrund von Compliance oder Datenschutz, so die Befragten. Jedoch sind sich alle Befragten einig, dass der Einsatz von Generative AI transparent dargestellt werden muss und zum Teil auch eine Anpassung der Allgemeinen Datenschutzbedingungen nötig ist.

Bislang sind die meisten Chat- oder Voicebots so gestaltet, dass sie keine oder nur sehr eingeschränkte Kundendaten verwenden. Folglich ist auch das Thema «Schutz von persönlichen Daten» bislang für die meisten Unternehmen

5.3.2 Cloud vs. On-Premise Dienste

Die Thematik um die Entscheidungen, ob eine neue IT-Lösung in der Cloud oder On-Premise auf eigenen Servern betrieben wird, tritt vor allem bei denjenigen Unternehmen immer wieder auf, die bislang noch keine einheitliche Cloud-Strategie haben. Dies sind mehrheitlich jene Unternehmen mit besonders sensiblen Daten. Vielmals zählen Unternehmen aus der Finanz- und Gesundheitsbranche, wie auch aus dem Bereich der Behörden dazu.

Auch im Hinblick auf die Nutzung von Generative AI bzw. Sprachmodellen stellt sich immer wieder die Frage, ob die LLMs aus der Cloud betrieben werden sollen oder On-Premise bei den einzelnen Unternehmen. Im Zusammenhang mit Sprachmodellen kommt es dann immer noch zur Frage, ob ein Sprachmodell von einem der bekannten grossen Anbieter wie OpenAI, Google oder ähnliche genutzt werden soll oder ein Open-Source Modell. Lediglich Open-Source Modelle lassen sich On-Premise bewirtschaften.

5.3.3 Make or Buy

Ähnlich, wie die Fragestellungen zwischen Cloud und On-Premise taucht bei der Einführung von neuen Technologien auch schnell die Frage nach dem Make or Buy auf. Damit gemeint ist die Frage, ob ein Unternehmen die Lösung selbst entwickelt oder die Lösung von einem anderen Anbieter eingekauft wird.

Im Hinblick auf Generative AI-Anwendungen muss zunächst gesagt werden, dass es genau genommen zwei Stufen des Make or Buy gibt. Zum einen wäre es möglich, dass Unternehmen sich sogar ein eigenes Sprachmodell konstruieren, welches sie dann regelmässig weiterentwi-

noch ein Randthema. Das begrenzt jedoch auch den individuellen Nutzen, den Kunden von diesen Anwendungen haben. Daher haben die Befragungen auch gezeigt, dass sich Unternehmen in Zukunft viel umfangreichere Kundenanwendungen speziell für Selbst-Hilfe-Lösungen (englisch: Self-Service) und Beratungen wünschen und dabei umfassendere Lösungen eingesetzt werden sollen, welche dann auch mit Kundendaten arbeiten. Sobald dies Realität wird, sind sich die Befragten einig, dass dann auch Themen im Hinblick auf Datenschutz mehr Bedeutung gewinnen.

Die vorliegende Studie zeigt eindeutig, dass eine Minderheit der befragten Unternehmen zwar die Möglichkeiten des On-Premise Betriebs eines Sprachmodells geprüft, schlussendlich den Initialaufwand dann als zu hoch eingeschätzt hat. So haben sich am Ende alle Befragten für ein Cloud-Hosting entschieden. Dabei gilt es jedoch anzumerken, dass es auch beim Cloud-Hosting unterschiedliche Ansätze gibt. Jene Unternehmen, die weniger Rücksicht auf Datenschutz nehmen müssen, weil sie beispielsweise ohnehin nur mit öffentlichen Daten arbeiten, nutzen mehrheitlich die Cloud Angebote von Microsoft, Google oder Amazon. Unternehmen mit sensiblen und schützenswerten Daten entscheiden sich dagegen für ein lokales Rechencenter oder nutzen bestehende Verträge mit langjährigen geprüften Dienstleistungspartnern. So erstaunt es auch nicht, dass knapp ein Drittel der Befragten angegeben hat, erste Experimente mit Open-Source Sprachmodellen zu machen, anstatt auf die bekannten Modelle ChatGPT oder Gemini von OpenAI bzw. Google zu setzen. Die Tatsache, dass Unternehmen Open-Source Sprachmodelle selbst hosten können, erhöht die Kontrolle über die eingegebenen Daten und bietet Unternehmen mit hohen Compliance-Anforderungen eine höhere Datensicherheit.

ckeln und für eigene Services nutzen. Zum anderen besteht immer auch die Möglichkeit auf ein vorhandenes Sprachmodell zurückzugreifen und dann weitere Services darauf aufbauend selbst umsetzen oder ebenfalls kaufen.

Von den befragten Unternehmen dieser Studie hat kein Unternehmen ein Sprachmodell von Grund auf entwickelt. Vereinzelt gaben Interview-Partner zwar an, diese Optionen geprüft zu haben und möglicherweise aus eigenem Interesse erste Versuche gemacht zu haben, schlussendlich sich dann aber doch für ein bestehendes Sprachmodell entschieden haben.

Was die weiteren Umsetzungen angeht, so gaben lediglich drei Unternehmen an, dass sie den Make-Ansatz verfolgen und ihre Generative AI-Anwendungen ohne weitere Dienstleister umsetzen. Hierbei handelt es sich um grosse Konzerne, die die Generative AI-Lösung heute schon als fixen Bestandteil ihres Wettbewerbsvorteils sehen und somit auch die volle Kontrolle über die Lösung, die Datenhaltung und die Weiterentwicklung haben möchten. Alle anderen befragten Unternehmen haben sich im ersten Schritt für

Buy entschieden. Dies wird durch den Wissens-Vorsprung der Dienstleister begründet. Einzelne Unternehmen haben meist nicht ausreichend Fachwissen und internes Wissen zur Umsetzung bzw. Anwendung der neusten Generative AI-Anwendungen.

5.4 Vorgehen zum Einführen von Generative AI- Anwendungen

Ein wichtiger Teil dieses Forschungsprojektes war eine Befragung zum Vorgehen bei der Einführung von Generative AI-Anwendungen. Die Tatsache, dass die Antworten hier sehr divers sind, zeigt, dass es bislang noch kein standardisiertes, erprobtes Vorgehen gibt. Vielmehr experimentieren die Unternehmen auch in diesem Feld noch. Während einzelne Unternehmen das Thema stark losgelöst von den restlichen Unternehmensgeschäften betrachten und somit auch die Einführung von Generative AI eher in einem «geschützten und abgekapselten» Raum stattfinden, ist die Einführung von KI-Anwendungen für andere ein grosses und strategisches Thema, welches über verschiedene Instanzen engmaschig überwacht wird.

Verknüpft man die Arten von KI-Anwendungen mit den Vorgehensmodellen, so lassen sich hier jedoch einzelne Tendenzen ableiten. Im Folgenden werden wir auf die wichtigsten Unterscheidungen eingehen. Wir beginnen dabei mit jenen Generative AI-Anwendungen, die sowohl

von hoher Bedeutung für den Kunden wie auch für die Mitarbeitenden sind. Die Tatsache, dass hier interne, wie auch externe Stakeholder berücksichtigt werden müssen, macht diese Anwendungen in der Einführung besonders komplex. Im Anschluss an die kundenorientierten Generative AI-Anwendungen folgen in diesem Kapitel Einführungsszenarien für Anwendungen, die mehrheitlich nach innen, also zu den Mitarbeitenden gerichtet sind.

Dabei zeigt sich auch, dass bei der Einführung von eher mitarbeiter-unterstützenden AI-Anwendungen deutliche Unterschiede bei der Einführung in Anhängigkeit von der konkreten Anwendung vorliegen. Dagegen ist bei den Kundenanwendungen der Prozess der Einführung mehrheitlich der gleiche. Oder anders ausgedrückt: mit Hilfe der Befragungen konnte eine Art Best Practice Prozess definiert werden, der entweder von den Unternehmen schon genau so praktiziert wird oder der sich aus den Erfahrungen der befragten Unternehmen ergeben hat.

5.4.1 Einführung von Generative AI-Anwendungen direkt an der Kundenschnittstelle

Bei der Einführung von Generative AI-Anwendungen, die Kunden direkt bedienen können, wie beispielsweise Chat- oder Voicebots, muss, wie einleitend erwähnt, berücksichtigt werden, dass neben den internen Stakeholdern vor allem auch die Kundensicht in das Projekt einfließen muss. Die Kunden müssen nicht nur befragt werden, sondern es müssen auch ihre Erwartungen getroffen oder im Idealfall sogar übertroffen werden.

Zum Stand der Befragungen gibt es nur begrenzt Benchmarks, das bedeutet, die Erwartungen der Kunden im Hinblick auf Generative AI-Anwendungen bei Unternehmen sind meist noch nicht allzu hoch. Gerade bei Bots

herrscht heute aber eine eher geringe Erwartung aufgrund der eher schlechten Kundenerlebnisse in der Vergangenheit. Dennoch haben Kunden ihre Erwartungen an einen guten Kundenservice und dieser ist mehrheitlich unabhängig von Künstlicher Intelligenz.

Im Rahmen der Analyse unserer Unternehmens-Befragungen konnten wir bei den Generative AI-Anwendungen an der Kundenschnittstelle einige Parallelen feststellen. Diese ergänzt unsere in der Vergangenheit vorgeschlagene Vorgehensweise (Hafner & Hundertmark, 2024) mit konkretem Praxiswissen. Gesamthaft leiten wir daraus folgendes Vorgehen als Best Practice Prozess ab:

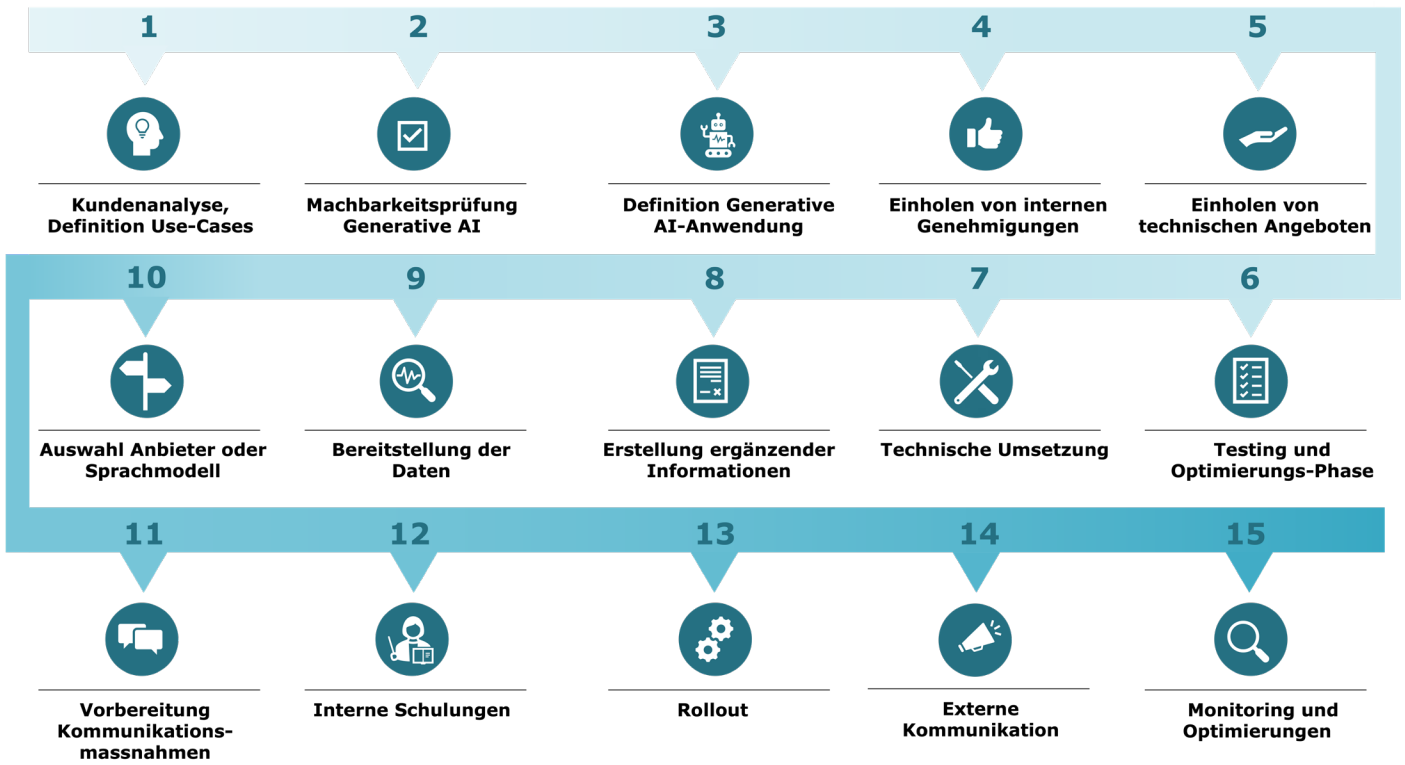


Abb. 14: Best Practice Prozessbeschreibung in Schritten

1. Analyse bestehender Kundendialoge bzw. Interaktionen mit dem Kunden im Hinblick auf heutige oder zukünftige Herausforderungen, Definition des Usecases beispielsweise mit der Value Irritant-Matrix (Price & Jaffe, 2008).
2. Machbarkeitsprüfung von Generative AI zur Lösung der Herausforderungen.
3. Definition Generative AI- Anwendung inkl. der erwarteten Chancen, sowie Zieldefinition (KPIs).
4. Einholung interner Genehmigungen für Generative AI-Anwendung (Budget, Compliance).
5. Einholen von Angeboten zu den technischen Umsetzungen, mit vorheriger «Make or Buy» Entscheidung.
6. Auswahl eines passenden Technologie-Anbieters oder Entscheidung für Make. Entscheidung, welches Sprachmodell genutzt werden soll.
7. Bereitstellung der Daten.
8. Erstellung ergänzender Informationen zu Prompts, Finetuning, Regeln für das LLM.
9. Technische Umsetzung: Verknüpfung Sprachmodell, Daten, Prompts und Finetuning.
10. Testing und Optimierungs-Phase.
11. Vorbereitung aller Kommunikationsmassnahmen, intern und extern.
12. Interne Schulungen für Mitarbeitende zum Umgang mit der neuen AI-Anwendung.
13. Rollout, je nach Unternehmen zuerst Soft-Launch oder direkt finaler Launch.
14. Externe Kommunikation.
15. Monitoring und Optimierungen.

5.4.2 Generative AI-Anwendungen zur Verbesserung des internen Wissensaustausch

Im Hinblick auf mitarbeiterunterstützende Generative AI-Anwendungen stellen Anwendungen, die den internen Wissensaustausch vereinfachen sollen, eine oft erwähnte Lösung dar. Gemeint sind hier mehrheitlich Anwendungen zur Entlastung der Personalabteilung, in denen Mitarbeitende ihre Fragen rund um die Organisation und zu personalpolitischen Themen mittels einem KI-Chatbot beantwortet bekommen. Oder auch Anwendungen für neue Mitarbeitende, wie beispielsweise ein Generative AI-basier-

tes Einführungsprogramm direkt auf Augenhöhe des neuen Bewerbers und mit kontinuierlicher Begleitung während der ersten Monate. Ebenfalls in diese Kategorie fallen jene Generative AI-Anwendungen, die für die Mitarbeitenden die tägliche Informationssuche im Tagesgeschäft vereinfachen. Beispielsweise können Vertriebsmitarbeitende sich in kurzer Zeit vollständige Kundeninformationen von der Generative AI zusammenstellen lassen.

Sämtliche dieser Anwendungen erfordern eine saubere Anbindung bzw. Vernetzung der Generative AI mit den internen Wissensquellen. Die Einführung dieser Anwendungen erfordert also deutlich mehr Vorbereitung und ein strukturiertes Projektvorgehen. Folgende Arbeitsschritte wurden in diesem Zusammenhang von den Befragten genannt.



Abb. 15: Projektvorgehen für den internen Wissensaustausch in Schritten

1. Usecase Findung und Integration aller Beteiligten.
2. Auswählen des passenden LLMs und des umsetzenden Technologieanbieters.
3. Datenschutz-Abklärungen im Hinblick auf gewählten Technologie-Anbieter.
4. Datenbereitstellung bzw. Bereinigung.
5. Verknüpfung LLM mit Datenquellen.
6. Test-Phase und Finetuning.
7. Allgemeines Rollout.
8. Schulungen.

An dieser Stelle haben alle Befragten deutlich betont, wie wichtig eine saubere Datengrundlage zur Einführung solcher Anwendungen ist. Dennoch startet das Vorgehen hier zunächst mit dem «Abholen» und der Integration der wichtigsten Stakeholder und Anwender der identifizierten Anwendung. Erst wenn alle Beteiligten ihrer Teilnahme am Projekt zugestimmt haben, kann es losgehen. Es wurden in diesem Zusammenhang teilweise auch Budget-Absprachen genannt. Da die Mehrheit der Befragten Unternehmen aus

der DACH-Region selbst noch keine eigenen Erfahrungen mit Generative AI gemacht hat, fehlt es überwiegend an wichtigem, internem Wissen. Es folgt also die Suche bzw. Auswahl eines passenden Technologie-Partners. Vereinzelt werden auch weitere, externe Beratende hinzugezogen. Im Zusammenhang mit der Wahl des passenden Technologie-Anbieters steht ebenfalls auch die Wahl für das passende Sprachmodell. Hier müssen Unternehmen entscheiden, ob sie auf eines der bekannten Modelle der grossen Tech-

Anbieter, wie GPT von OpenAI oder Gemini von Google setzen wollen oder eher den Einsatz von Open-Source Modellen, wie beispielsweise Llama oder Falcon, verfolgen wollen.

In der Regel erfolgt dann die Bereitstellung der passenden Daten ebenfalls in Zusammenarbeit bzw. im engen Austausch mit dem Technologie-Anbieter. Die Experten wissen meist am besten, in welcher Form die Daten bereitgestellt werden müssen, damit das Sprachmodell nachhaltig darauf zugreifen kann. Gleichzeitig kennt jedes Unternehmen selbst seine Daten am besten. Die Projektbeteiligten müssen also wissen, welche Daten das Sprachmodell überhaupt verstehen kann und wo diese liegen. Weiter können auch nur die internen Mitarbeitenden die Aktualitäten und Korrektheit, der zur Verfügung gestellten Daten, prüfen und sicherstellen.

Sobald das Sprachmodell bekannt ist und auch alle Daten verfügbar gemacht sind, folgt die Verknüpfung des Sprachmodells mit den Daten. Hier unterscheiden sich die Antworten lediglich im Hinblick auf die Art und Weise des Hostings. Ein Unternehmen gab an, dass sie ihre eigenen Daten weiterhin On-Premise und intern hosten und somit auch das Sprachmodell selbst hosten. Die anderen befragten Unternehmen haben bereits den Weg in die Cloud gefunden bzw. zumindest im Hinblick auf die Generative AI-Anwendungen alles in einer (lokalen) Cloud abgelegt.

Es folgt die Phase des Testens:

1. Kann die neue Generative AI-Lösung wirklich das, was alle erwartet haben?
2. Funktioniert der Zugriff auf die Datenquellen und werden diese vom Sprachmodell auch korrekt verarbeitet und an die Nutzer ausgegeben?

Diese Fragen müssen beantwortet werden, sobald die eigenen Datenquellen mit dem Sprachmodell verbunden sind. Die Phase des Testens stellt also einen wesentlichen Teil im gesamten Einführungsprozess dar. Diese Phase wird durch Optimierungen am System ergänzt. Zum Teil ist eine Anpassung der Daten notwendig, bei anderen genügt ein vertiefendes Fine-Tuning des Sprachmodells. Während der Fine-Tuning Phase wird aus dem meist eher allgemeinen Sprachmodell ein unternehmensspezifisches Sprachmodell gemacht. Konkret werden sogenannte Prompts hinzugefügt, die dem Sprachmodell genauere Anweisungen zum Verhalten oder zur individuellen Unternehmens-Tonalität mitgeben.

Sobald das Testen und Fine-Tuning abgeschlossen sind, kann die neue Generative AI-Anwendung ausgerollt werden. Je nach Anwendung folgen idealerweise Schulungen für Mitarbeitende. Der Mindeststandard einer unternehmensweiten Bekanntmachung wurde bei allen befragten Unternehmen erfüllt. Hier ergibt sich eine Anmerkung zum Zeitpunkt dieser Befragungen. In den Interviews endet der Einführungsprozess an dieser Stelle. Aufgrund der Neuheit der Technologie konnte keiner der Befragten Informationen zur Phase nach dem Rollout geben, da diese meist erst kürzlich stattgefunden hatten.

5.4.3 Co-Pilot und/oder ChatGPT zur Steigerung der internen Produktivität

Zum Zeitpunkt der Befragungen gaben einige Unternehmen an, dass ihre Mitarbeitenden entweder ChatGPT frei oder unter Verwendung von Richtlinien nutzen dürfen. Andere gaben an, dass sie ihren Mitarbeitenden bereits Co-Pilot zur Verfügung gestellt haben. Interessant ist bei dieser Anwendung, dass es bislang keinerlei strukturierte Einführungen dieser Tools gibt. Lediglich drei Unternehmen gaben an, dass sie für Ihre Mitarbeitenden eine allgemeine Schulung zum Prompten und zum Umgang mit den KI-Tools gegeben haben.

Hinsichtlich Co-Pilot werden von Microsoft verschiedene Schulungsformate kostenlos zur Verfügung gestellt. Keiner der Befragten gab jedoch an, dass diese Schulungskonzepte auch verwendet worden sind. Vielmehr bekommen die Mitarbeitenden einfach die entsprechende Lizenz auf ihren Computer und sind dann mehrheitlich auf sich allein gestellt. Auch fehlen hier Methoden oder Instrumente zur Erfolgsmessung. Obwohl das Ziel einer Steigerung der Produktivität definiert ist, gibt es keine konkreten Massnahmen, um dieses Ziel auch messen zu können.

5.4.4 Generative AI-Anwendungen für klar definierte interne Prozesse

Generative AI-Anwendungen, die lediglich klar definierte interne Prozesse und damit einzelne Abteilungen betreffen, sind beispielsweise Anwendungen im Bereich der Software-Entwicklung oder im Marketing.

Im Allgemeinen unterscheiden sich die Antworten der Befragten hier kaum, zu denen der Anwendungen im

Hinblick auf Wissensaustausch intern. Jedoch ist der Kreis der involvierten Personen meist viel kleiner. Das führt auch dazu, dass die Budget-Entscheider meist nicht Teil des Projektes sind, sondern separat vom Projekt-Team abgeholt werden müssen. Dies passiert idealerweise in der ersten Phase. Danach ist der Prozess nahezu der gleiche, wie zuvor beschrieben:

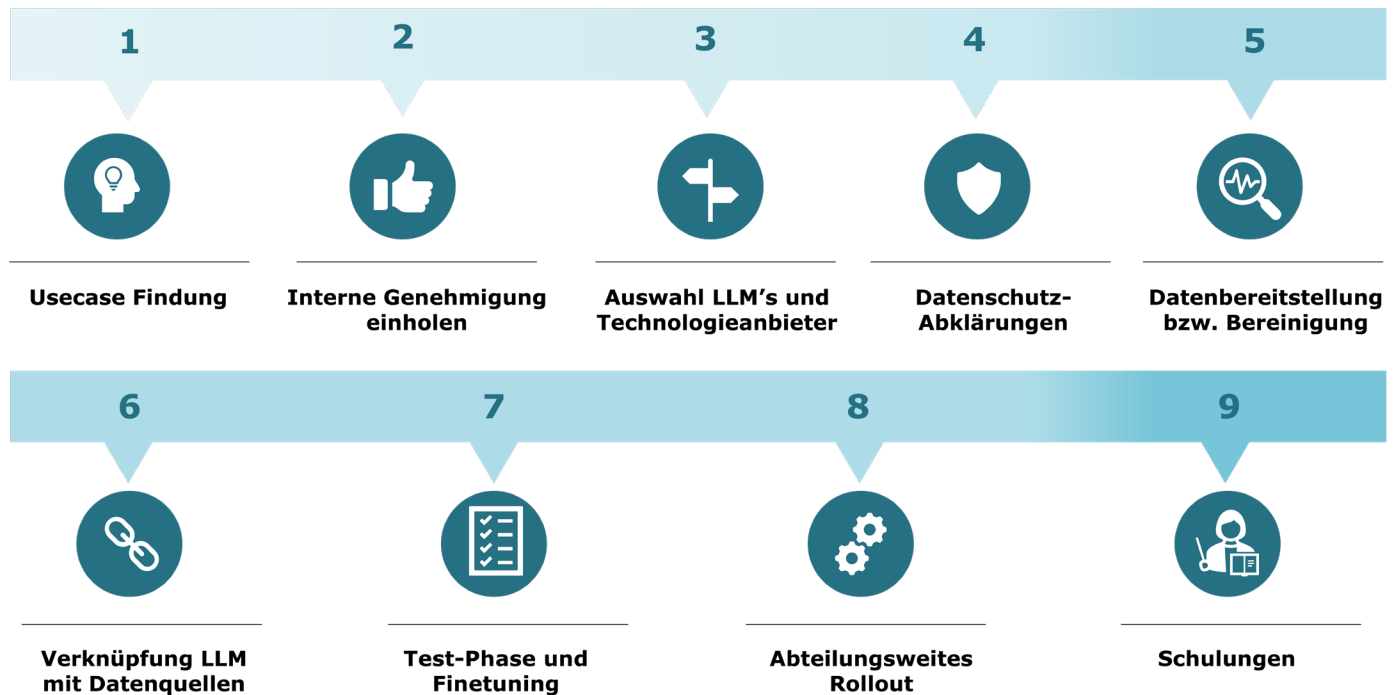


Abb. 16: Projektvorgehen für klar definierte interne Prozesse in Schritten

1. Usecase Findung.
2. Interne Genehmigung einholen.
3. Auswählen des passenden LLMs bzw. Technologieanbieter.
4. Datenschutz-Abklärungen im Hinblick auf gewählten Technologie-Anbieter.
5. Datenbereitstellung bzw. Bereinigung.
6. Verknüpfung LLM mit Datenquellen.
7. Test-Phase und Finetuning.
8. Abteilungsweites Rollout.
9. Schulungen.

Hier ist es wichtig zu erwähnen, dass häufig «Support-Abteilungen», wie Compliance oder IT, nicht Teil eines solchen Projektteams sind. Folglich müssen die relevanten Mitarbeitenden hier immer im entsprechenden Prozess-Schritt eingebunden werden und daher meist auch vorab individuell abgeholt und informiert werden.

Bei der Wahl des Technologie-Anbieters wird häufig auf bestehende Technologie-Anbieter zurückgegriffen. Viele Befragte äusserten hier sogar, dass es ursprünglich sogar der Technologie-Anbieter direkt war, der auf die neuen, möglichen Generative AI-Anwendungen hingewiesen hat. So ist in vielen Unternehmen beispielsweise der Kunden-

service-Management-Software-Anbieter auf das Team zugekommen und hat seine neuen Lösungen im Hinblick auf Generative AI vorgestellt.

Eine weitere leichte Anpassung zum zuvor beschriebenen Vorgehen ist das Ende mit den verbundenen Schulungen. Sofern es sich um spezifische Generative AI-Anwendungen handelt, sind interne Schulungen essentiell. Mitarbeitende sind darauf angewiesen, genau zu verstehen, wie sie die neuen Möglichkeiten einsetzen und vor allem welche bestehenden Arbeitsprozesse nun ergänzt werden, um die gewünschten Produktivitätsverbesserungen zu erreichen.

5.4.5 Der KI-Prozess-Kompass

Im Rahmen der Analyse der Unternehmens-Befragungen konnten wir also bei der Einführung von Generative AI-Anwendungen einige Parallelen feststellen. Diese ergänzen unsere in der Vergangenheit vorgeschlagene Vorgehensweise (Hafner & Hundertmark, 2024) mit konkretem Praxiswissen. Daher präsentieren wir folgendes Vorgehen als Best Practice Prozess und bezeichnen es als «KI-Prozess-Kompass».

Der KI-Prozess-Kompass beginnt mit einer Analyse der bestehenden Kundendialoge sowie der Interaktionen mit den Kunden, um aktuelle oder zukünftige Herausforderungen zu identifizieren. In diesem Schritt wird auch der Usecase definiert. Ein Instrument dazu ist die Value Irritant-Matrix von Price und Jaffe (2008). Anschließend erfolgt eine Machbarkeitsprüfung, um festzustellen, ob der Einsatz von Generative AI zur Lösung der Herausforderungen geeignet ist.

Sobald die Machbarkeit bestätigt ist, wird die konkrete Generative AI-Anwendung ausgewählt. Dabei werden die erwarteten Chancen konkretisiert und Zielsetzungen mit konkreten Key Performance Indicators (KPIs) definiert. In der Folge müssen interne Genehmigungen eingeholt werden, was die Sicherstellung des Budgets und die Erfüllung von Compliance-Vorgaben umfasst.

Im nächsten Schritt werden Angebote für die technische Umsetzung eingeholt, wobei eine „Make or Buy“-Entscheidung vorausgeht. Es folgt die Auswahl eines passenden Technologieanbieters oder die Entscheidung zur internen Umsetzung (Make). An diesem Punkt muss ebenfalls festgelegt werden, welches Sprachmodell verwendet werden soll.

Die Daten werden bereitgestellt, und ergänzende Informationen zu Prompts, Finetuning sowie Regeln für das Large Language Model (LLM) werden erstellt. Daraufhin erfolgt die technische Umsetzung, bei der das Sprachmodell, die Daten, die Prompts und das Finetuning miteinander verknüpft werden.

Es folgt eine Phase des Testings und der Optimierung. Gleichzeitig werden alle internen und externen Kommunikationsmaßnahmen vorbereitet. Intern werden Schulungen für die Mitarbeitenden durchgeführt, um sie im Umgang mit der neuen AI-Anwendung zu schulen.

Je nach Unternehmensstrategie erfolgt der Rollout entweder als Soft-Launch oder direkt. Externe Kommunikationsmaßnahmen begleiten diesen Prozess. Schliesslich

wird das System kontinuierlich überwacht, und Optimierungen werden vorgenommen, um den langfristigen Erfolg zu sichern.

Bei der Einführung von Generative AI-Anwendungen, die Kunden direkt bedienen können, wie beispielsweise Chat- oder Voicebots, muss berücksichtigt werden, dass neben den internen Stakeholdern vor allem auch die Kundensicht in das Projekt einfließen muss. Die Kunden müssen nicht nur befragt werden, sondern es müssen auch ihre Erwartungen getroffen oder im Idealfall sogar übertroffen werden. Hier kommt ein iterativer agiler Prozess zum Tragen.

Im Hinblick auf mitarbeiterunterstützende Generative AI-Anwendungen stellen Anwendungen, die den internen Wissensaustausch vereinfachen sollen, eine oft erwähnte Lösung dar. Gemeint sind hier mehrheitlich Anwendungen zur Entlastung der Personalabteilung, in denen Mitarbeitende ihre Fragen rund um die Organisation und zu personalpolitischen Themen mittels einem KI-Chatbot beantwortet bekommen.

Sämtliche dieser Anwendungen erfordern eine saubere Anbindung bzw. Vernetzung der Generative AI mit den internen Wissensquellen.

Für die Anwendung des KI-Prozess-Kompass bedeutet dies, dass er von uns zwar als Best-Practice Beispiel und als ein erfolgreiches Tool zur Einführung von Generative AI-Anwendungen dient. Wenden Unternehmen den Kompass aber an, muss zunächst individuell geprüft werden, welche Prozess-Schritte für das Unternehmen, sowie deren Kunden und Mitarbeitende wirklich relevant sind und welche eher kurz behandelt oder gar übersprungen werden können.

Unternehmen, die bis zur Nutzung des KI-Prozess-Kompass noch keinerlei eigene Erfahrungen mit Generative AI-Projekten gesammelt hatten, sollten an dieser Stelle über die Nutzung externen Wissens nachdenken. In diesem Zusammenhang sind die vielen Usecases, die in der Studie gezeigt werden, besonders wertvoll. Erfahrungen anderer Unternehmen verdeutlichen immer wieder, wie wichtig der Austausch in einer Community von Anwendern ist. Noch können sich nur wenige Unternehmen in diesem Feld als Experten bezeichnen. Vielmehr geht es zu diesem Zeitpunkt darum, Wissen zu teilen, von den Erfahrungen anderer zu profitieren und neue Erkenntnisse und Best Practices mit Generative AI zu entwickeln.

Anwendbarkeit der AI-Anwendungen in 6 der Finanzbranche

Die aufgezeigten Anwendungsfälle zeigen die vielfältigen Einsatzmöglichkeiten von Generative AI, die auch für Finanzdienstleister von grossem Interesse sind. In diesem Kapitel wird diskutiert, wie die beschriebenen AI-Anwendungen auf den Finanzsektor übertragen werden können und welche spezifischen Vorteile sich daraus ergeben. Dabei wird geprüft, inwieweit die zuvor genannten Anwendungen für die Zielgruppe der Finanzdienstleister relevant sind und wie die UseCases, auch im Hinblick auf die branchenspezifischen Compliance Richtlinien aus der Schweizer Finanzwelt, übertragbar sind.

Dieses Kapitel entstand zusammen mit Carla Caspar. Mit ihrer Erfahrung als Strategic Innovation Manager und Leiterin Data Science & AI bei Inventx AG kennt sie die Bedürfnisse und Prozesse der Finanzbranche, wie auch das Potenzial und die Herausforderungen in der Umsetzung von AI-Anwendungen bei Banken und Versicherungen.

Für Finanzdienstleister sind Chat- und Voicebots besonders wertvoll, um den Kundenservice rund um die Uhr zu gewährleisten oder Spitzen zu brechen, wie zum Beispiel bei Einführung einer neuen Mobile App oder E-Banking Version. Kunden haben häufig einfache, automatisierbare Fragen technischer oder fachlicher Natur. Mit dem Einsatz von Generative AI können Finanzinstitute sicherstellen, dass diese Anfragen schnell und präzise beantwortet werden. Dies erhöht nicht nur die Kundenzufriedenheit, sondern entlastet auch die Service-Mitarbeitende, die sich auf komplexere Anfragen konzentrieren können. Wichtig ist hier zu erwähnen, dass Bots ihr volles Potential aber erst nach einer Identifikation und Authentifizierung des Kunden entfalten können.

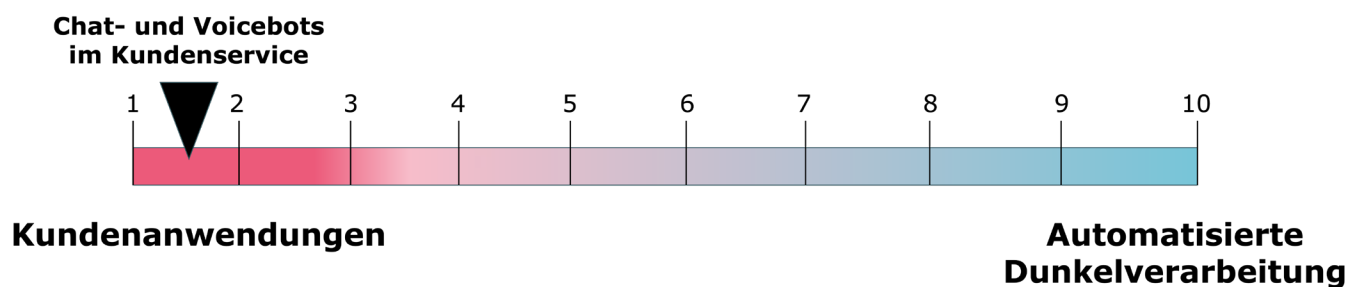


Abb. 17: Chat- und Voicebots im Kundenservice auf der Gen AI Skala

Über Carla Caspar



Carla Caspar, als Strategic Innovation Manager im InventxLab, leitet und fördert Innovationen im Bereich Data Science & KI und berät Kunden strategisch in diesen Technologiefeldern. Seit 2013 bei Inventx AG hat sie zuvor den Bereich Data Science & KI aufgebaut und geleitet. Vor ihrer Zeit bei Inventx war sie als Teamleiterin im Bereich Financial Engineering bei der Zürcher Kantonalbank tätig. Weiter engagiert sie sich als Fachrätin bei Graubünden Digital für die Digitalisierung im Kanton Graubünden.

Für Finanzdienstleister ist es entscheidend, dass ihre Telefonagenten oder Berater in der Lage sind, auf spezifische und oft komplexe Kundenanfragen schnell und präzise zu reagieren. KI-Assistenten, die während des Gesprächs mithören, und relevante Informationen sowie Lösungsvorschläge in Echtzeit bereitstellen, können die Effizienz und Qualität der Kundeninteraktionen erheblich verbessern,

zum Beispiel in der Kreditberatung für Reglementdetails. Wertvoll kann auch eine automatisierte Zusammenfassung von Kundengesprächen, im CRM für die Kundenberater oder bei technischen Anliegen im Ticketsystem, sein. Weiter können KI-Assistenten zu Compliance und regulatorischen Zwecken mithören und dokumentieren oder das Gespräch auf Betrug analysieren.

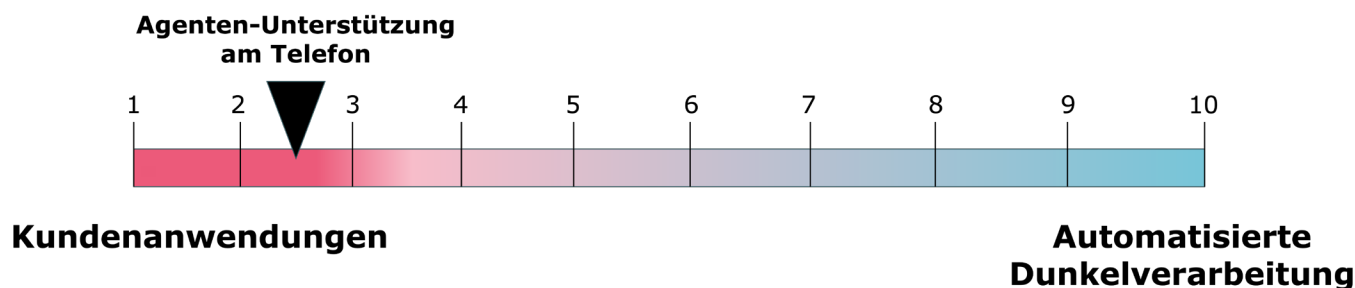


Abb. 18: Agenten-Unterstützung am Telefon auf der Gen AI Skala

Usecase

Die digitale Assistentin Clara: Wie aus einem einfachen SMS-Dienst bei Helvetia Versicherungen Schweiz ein GenAI-Bot wurde.



Ein Gastbeitrag von Nadine Schade, Conversational Marketing Manager und Conversational & Marketing Automation, Helvetia Versicherungen Schweiz

Der Usecase

Die ersten Schritte im Conversational Bereich wurden seitens Helvetia bereits im Jahr 2017 unternommen. Nach einem Design Thinking Workshop wurde ein SMS-Dienst für Velodiebstähle angeboten. Das Projekt war derart erfolgreich, dass in den darauffolgenden Jahren der Chatbot «Clara» ins Leben gerufen wurde. Zu Beginn erfolgte die Abbildung spezifischer Anwendungsfälle, wie beispielsweise Schadenmeldungen oder Adressänderungen in einem Chat-Interface mit einem dynamischen Entscheidungsbaum. Dadurch wurde es Kundinnen und Kunden möglich, diese Anliegen selbstständig und ohne Wartezeiten zu erledigen. Parallel dazu wurde Clara so weiterentwickelt, dass sie auf der Webseite von Helvetia Besucherinnen und Besuchern bei sämtlichen Anliegen unterstützen sollte. Zu diesem Zweck wurde ein Intent-basierter Chatbot mit eigenem Sprachmodell aufgebaut und betrieben. Aufgrund des breiten Angebots von Helvetia waren die Resultate jedoch oft unbefriedigend. Trotz kontinuierlicher Optimierungen erreichte Clara nicht ein zufriedenstellendes Kundenerlebnis.

Die Schnittstelle von OpenAI zu ChatGPT wurde schliesslich Ende Februar 2023 der Öffentlichkeit zugänglich gemacht. Helvetia sah eine Möglichkeit, Clara weiterzuentwickeln und frühzeitig mit LLMs Erfahrungen zu sammeln. Hierfür wurde entschieden, ein Live-Experiment zu starten und dieses in vier Wochen umzusetzen. Das Experiment verfolgte das Ziel, die Zugänglichkeit der Webseite für Kundinnen und Kunden neu zu denken und die Akzeptanz von KI-Anwendungen zu evaluieren.

Das Experiment war erfolgreich. Als weltweit erste börsennotierte Versicherung hat Helvetia einen Service im direkten Kundenkontakt auf Basis der generativen KI implementiert. Zur Beantwortung von Anfragen zu den Themen Versicherungen, Vorsorge und Eigenheim bediente sich Clara des Sprachmodells GPT 3.5 Turbo. Als Wissensbasis wurde dem Sprachmodell ausschliesslich die Webinhalte von Helvetia Schweiz, beispielsweise Produktseiten und Ratgeberartikel, zur Verfügung gestellt. Nach einer siebenmonatigen Testperiode wurde das Experiment abgeschlossen und GenAI in Clara integriert. Seit November 2023 ist Clara als Standarddienst unter Verwendung des GPT4-Modells im Einsatz.

Während des Experiments hatte das Projektteam rund um Clara die Möglichkeit, mit realen Kundenanfragen zu arbeiten und das Modell entsprechend anzupassen. Den Clara-Benutzerinnen und -Benutzern stand es frei,

ob sie mit der GenAI- oder der Intent-Clara interagieren wollten. Heute wird den Nutzerinnen und Nutzern transparent dargestellt, dass Helvetia ChatGPT nutzt und welche Bedeutung dies für den Datenschutz und die Nutzungsbedingungen hat.

Im Rahmen der eigentlichen Einbindung der GenAI Clara war es erforderlich, eine Verknüpfung mit den bereits bestehenden Services herzustellen (siehe Abbildung 1). Derzeit setzt Helvetia auf eine Kombination von GenAI- und «Use-case-spezifischen Bots», die nicht mit GenAI funktionieren. Dabei wird das Prinzip verfolgt, dass LLMs dort eingesetzt werden, wo es sinnvoll ist, und nicht des LLMs halber. Im Fokus steht die Gewährleistung einer bestmöglichen Customer Journey. Die beiden Bot-Welten sind so miteinander verknüpft, dass Benutzerinnen und Benutzer Wechsel in den Architekturen nicht bemerken. Für sie gibt es lediglich eine Clara. Zu Beginn des Experimentes nutzten vor allem Technikbegeisterte die neue Clara. Nachdem sich der erste Hype gelegt hatte, zeichnete sich vor allem ein Interesse an Definitionen, Deckungsanfragen und Versicherungsbedarf ab.

Mehrwerte und Lehren:

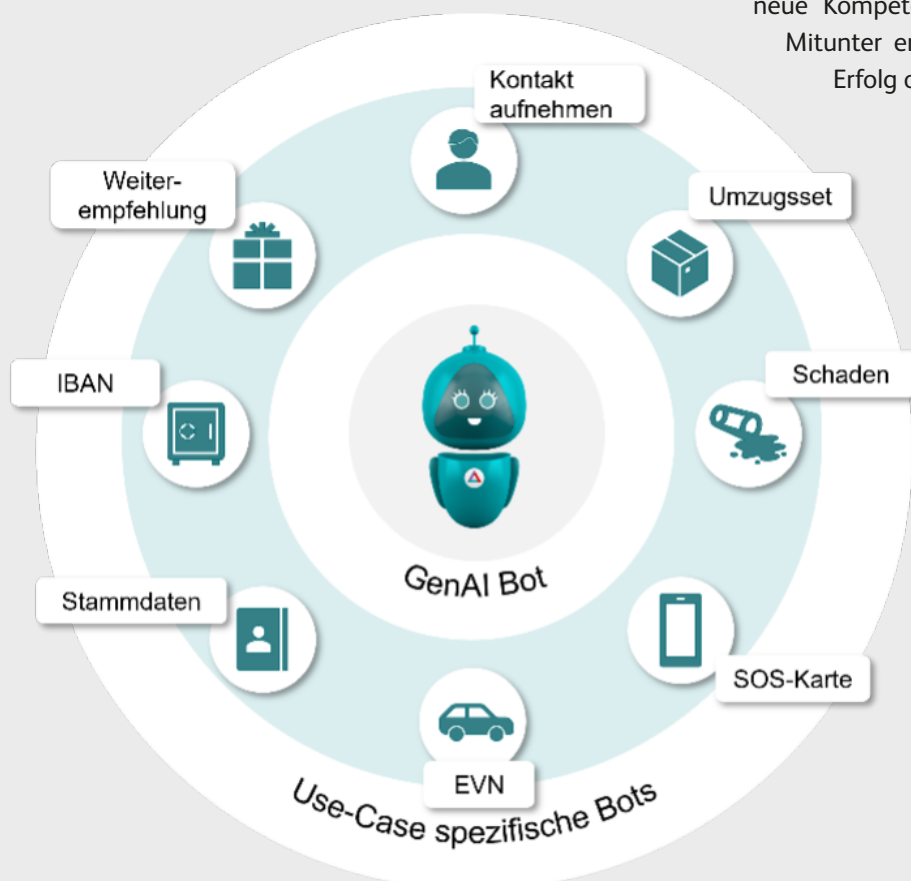
Aus technischer Perspektive eröffnen LLMs gänzlich neue Möglichkeiten hinsichtlich der Antwortqualität. Die ursprüngliche Version von Clara basierte auf einem vordefi-

nierten Sprachmodell und war primär Intent-basiert. Diese Art von Chatbots bietet den Vorteil, dass das gesamte Antwortrepertoire deterministisch festgelegt ist, wodurch eine umfassende Kontrolle über das Antwortverhalten von Clara gewährleistet wird. Im Falle des GenAI-Chatbots werden Inhalte in Form eines Prompts an das Sprachmodell übergeben, welches daraufhin eine Antwort generiert. Der Prompt setzt sich aus der Nachricht des Nutzers, den relevantesten Inhalten sowie sogenannten Constraints (z. B. Persönlichkeitsbeschreibungen von Clara) zusammen. Auf diese Weise kann auf jede Anfrage individuell reagiert werden, allerdings besteht das Risiko, dass die Antworten eher ungenau werden.

In der IT-Entwicklung lassen sich zwei weitere Veränderungen feststellen. Zum einen wird bislang vorausgesetzt, dass Applikationen bei mehrfacher Eingabe identischer Parameter stets identische Resultate liefern. Andererseits kann bei einem Chatbot, der auf Generative AI basiert, folglich nicht mehr gewährleistet werden, dass bei einer zweifachen Anfrage stets die identische Antwort generiert wird. Dies impliziert, dass das Qualitätsmanagement zu einer Herausforderung wird. Des Weiteren basierte die Entwicklung von Applikationen auf der Prämisse, dass die Programmierung einer gewissen Logik und Verlässlichkeit folgt. Die Funktionsfähigkeit der Applikation ist nach der erstmaligen Programmierung gewährleistet. Diese Prinzipien können jedoch bei LLMs nicht länger als gegeben betrachtet werden. Die Erstellung der Prompts erfordert neue Kompetenzen sowie ein hohes Mass an Geduld.

Mitunter entscheidet lediglich ein einziges Wort über Erfolg oder Misserfolg.

Während des Experiments und auch im laufenden Betrieb hat Helvetia viele Erkenntnisse über den Einsatz von Generative AI, das Kundenverhalten und den Umgang mit dieser mächtigen Technologie gewonnen. Die drei wichtigsten Erkenntnisse sind die folgenden:



1. Das Narrativ des „Experiments“ war zielführend, um interne und externe Akzeptanz zu schaffen. Gleichzeitig konnten wir dadurch sehr schnell agieren und hatten in der Folge Zeit, die technischen und fachlichen „Schulden“ abzarbeiten und den Dienst nachhaltig in Clara zu integrieren.
2. Mit den technischen Möglichkeiten scheinen auch die Erwartungen zu wachsen. Wir beobachten, dass die Anfragen im Vergleich zu früher deutlich komplexer geworden sind. Die Nutzer verstehen, dass sie mit einem intelligenten Chatbot sprechen und entsprechend steigen die Erwartungen. Wenn sich unser Bild bestätigt, werden Kundinnen und Kunden bald Chatbots erwarten, die LLM-Fähigkeiten beherrschen.
3. Der anfängliche manuelle Aufwand zur Qualitätssicherung sollte nicht unterschätzt werden. Bei der Einführung eines GenAI-Chatbots ist es wichtig, den Bot kennenzulernen und ein Gefühl für Antwortmuster und Wissenslücken zu erhalten. In der Experimentierphase hat das Projektteam alle Gespräche mit Clara gesichtet, um Kinderkrankheiten zu beheben und Vertrauen in Clara aufzubauen. Für das Qualitätsmanagement gibt es noch keine Plug-and-Play-Lösungen auf dem Markt. Auch hier zeigt sich, dass sich die Technologie zwar rasant entwickelt, aber noch nicht einsatzbereit ist.

Trotz des Mehraufwands ist der Einsatz von Generative AI an der Kundenschnittstelle für Helvetia mit einem grossen Mehrwert verbunden. Helvetia verfolgt damit das Ziel, einen neuen Marktstandard für digitale Assistenten in der Versicherungsbranche zu setzen. Mit Clara wollte Helvetia erstmals einen spürbaren Differenzierungsfaktor im Versicherungsmarkt schaffen und unsere strategische Vision umsetzen, Massstäbe in der Kunden-Convenience zu setzen. Die Antwortqualität von Clara konnte dank Generative AI deutlich verbessert werden. Dies zeigt sich nicht nur qualitativ bei der Betrachtung der Konversationen, sondern auch quantitativ bei den Nutzungszahlen. Waren es im Jahr 2022 noch knapp 95.000 Chats pro Jahr, so stieg die Zahl im Jahr 2023 auf rund 167.000 Konversationen. Ein wesentlicher Teil dieser Nutzungssteigerung kann auf den Einsatz von GenAI zurückgeführt werden. Die Ausdrucksweise von Clara ist viel menschlicher geworden und die Gespräche können dank ihres erweiterten Wissens besser geführt werden. Die Kundinnen und Kunden bemerken den Unterschied in der Intelligenz des Chatbots und der Service wird daher sehr gut angenommen. Generell ist das Feedback sehr positiv.

Ausblick

Neue KI-Produkte, -Modelle und -Funktionen werden derzeit in Rekordzeit auf den Markt gebracht. Dadurch haben immer mehr Menschen die Möglichkeit, Erfahrungen mit KI zu sammeln und sich an neue, intelligente Chatbots und digitale Dienste zu gewöhnen. Chatbots werden immer mehr zu hilfreichen Assistenten, die bei komplexen Aufgaben und alltäglichen Dingen unterstützen können. Auch Sprachassistenten könnten aufgrund der hohen mobilen Nutzung zentral werden. Aber das sind nur Vermutungen. Für Clara wird es in den nächsten Monaten wichtig sein, die Antwortqualität weiter zu verbessern und Customer Journeys zu optimieren. Dazu werden neue Sprachmodelle getestet und ein automatisiertes Qualitätsmanagement-Tool entwickelt. So sollen manuelle Arbeitsprozesse minimiert und der Fokus auf strukturelle Verbesserungen und Weiterentwicklungen gelegt werden. Ziel ist es, Clara als integralen Bestandteil der Website zu positionieren, so dass die Besucherinnen und Besucher der Website die Antworten auf ihre Fragen nicht mehr suchen müssen, sondern sofort finden. Man darf also gespannt bleiben.

Wohin uns die KI-Entwicklungen führen werden, ist noch offen. Helvetia ist aber davon überzeugt, dass grosse Sprachmodelle und GenAI vergleichbare Veränderungen auslösen werden, wie die Verbreitung des World Wide Web oder die Einführung des iPhones. Die strategische Verankerung von KI-Themen und Technologien ist dabei eine grundlegende Unternehmensentscheidung. Helvetia sieht nicht nur beim Chatbot Clara, sondern entlang der gesamten Wertschöpfungskette ein grosses Potenzial für Weiterentwicklungen, beispielsweise in der Schadenmeldung, der Angebotsentwicklung oder im Wissensmanagement.

Versicherungsunternehmen müssen mittel- und langfristig entscheiden, wie sie mit den sich verändernden Marktbedingungen umgehen wollen. Es führt kein Weg daran vorbei, sich mit Generative AI auseinanderzusetzen und zu experimentieren, um fundierte Entscheidungen treffen zu können. Nur durch den Aufbau von Know-how kann der anstehende Wandel aktiv mitgestaltet werden.

Über das Unternehmen

Helvetia Versicherungen Schweiz mit Sitz in St. Gallen hat sich seit 1858 zu einer erfolgreichen Versicherungsgruppe mit über 12 000 Mitarbeitenden und mehr als 7 Millionen Kundinnen und Kunden entwickelt.

Die Aktie der Helvetia Holding AG wird an der Schweizer Börse SIX Swiss Exchange gehandelt.

Mehr zur Helvetia Gruppe: www.helvetia.com



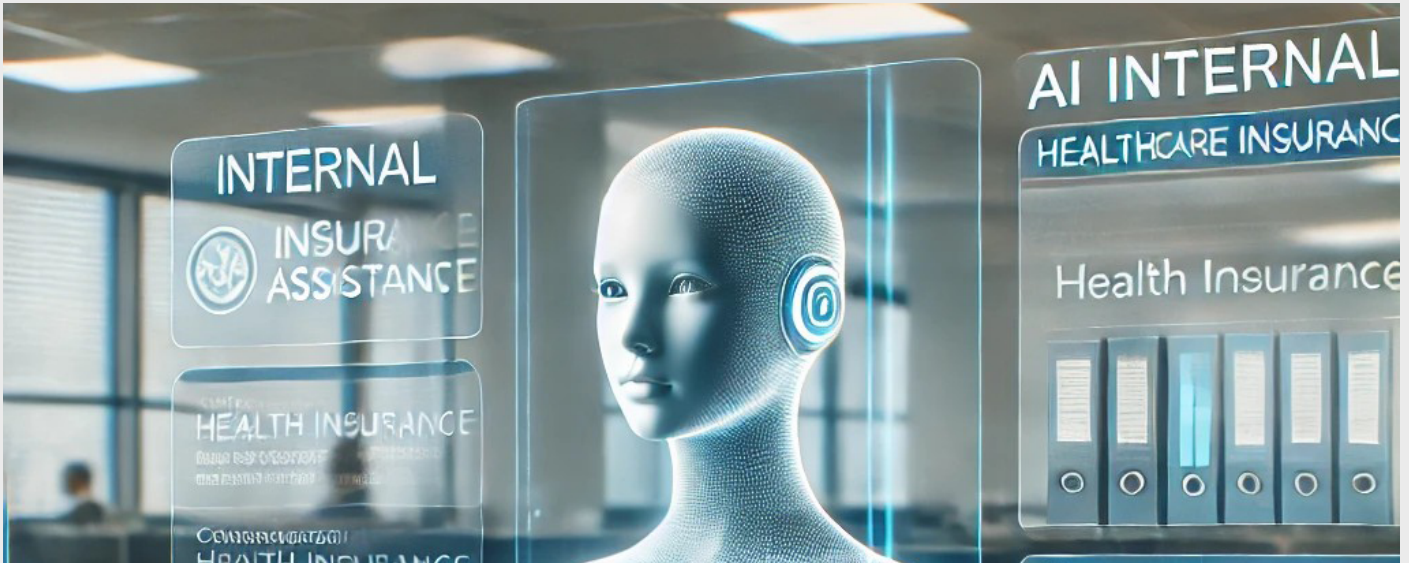
Kontaktinformationen

Nadine Schade ist als Conversational Marketing Managerin bei Helvetia Versicherungen tätig. Seit 2023 ist sie mit der Weiterentwicklung der Fähigkeiten von Clara beschäftigt, mit dem Ziel, diese zu einer vollwertigen digitalen Assistentin zu entwickeln.

nadine.schade@helvetia.ch

Usecase

Mitarbeitenden-Unterstützung im Kundenkontakt der KPT Versicherung



Der Usecase

KPT hat sich bereits frühzeitig mit Generative AI auseinandergesetzt und eigene Erfahrungen gesammelt. Seit Neustem hat KPT einen digitalen Agenten für das Beratungszentrum, der auf Generative AI basiert. Der digitale Agent unterstützt Kundenberater als Souffleuse zu Fragen über Produkte der KPT und detaillierte Versicherungsbedingungen.

Es begann mit einem einfachen Proof of Concept. Mit Retrieval Augmented Generation wurden erste Informationen der KPT-Webseite aufbereitet und bereitgestellt, sowie verschiedene Large Language Modelle getestet und verglichen. Fortlaufend werden Mitarbeiter für das Testing des digitalen Agenten beigezogen. Ihre wertvollen Inputs helfen dem Innovationsteam den Bot des Proof of Concepts zu einer einsatzbereiten Lösung weiterzuentwickeln.

Die Eigenentwicklung der auf öffentlichen Daten basierenden Lösung macht eine spezifische Spezialisierung des Agenten für den Einsatz im KPT-Beratungszentrum möglich. Zum Beispiel müssen umgangssprachliche Fragestellungen mit den Fachbegriffen in Verbindung gebracht werden können, Antworten sollen für die Weitergabe an den Kunden in einfacher Sprache formuliert sein und Stichwort-Prompting zu qualitativ guten Antworten führen.

Mehrwert

Nach einer intensiven Testphase ist der Generative AI-Agent im Herbst 2024 bereit für die Einführung. Dann profitiert das Beratungszentrum der KPT von der Unterstützung der digitalen Souffleuse für effiziente und einheitliche Kundenauskünfte. Und nicht nur Mitarbeiter des Beratungszentrums haben Zugriff auf den Generative AI-Agenten. Im Sinne einer ganzheitlichen AI-Kultur, soll er zukünftig trotz Optimierung für das Beratungszentrum, allen Mitarbeiter der KPT zur Verfügung stehen.

Der digitale Agent für das Beratungszentrum ist nicht die einzige Generative AI-Initiative der KPT. Neben einer Testgruppe, die Copilot M365 ausgiebig testet, werden verschiedene Schulungen und AI-Sessions durchgeführt. Zum Beispiel «GenAI Technology Landscape: die heutigen AI Tool Boxes» für das C-Level, oder ein Schulungsangebot «Prompting», wo die Teilnehmer nach einer Einführung verschiedenen Prompting Strategien ausprobieren und ihre Prompting Kenntnisse ausbauen können.

Zudem informiert ein mehrteiliger GenAI Podcast die KPT Mitarbeiter über aktuelle News und neusten Tipps.

Ausblick

Bei anhaltendem Erfolg soll der Generative AI-Agent mit zusätzlichen Quellen erweitert und für weitere Mitarbeitergruppen spezialisiert werden.

Mittelfristig ist eine Erweiterung der aktuell generellen Antworten auf kundenspezifische Antworten denkbar. Die Antwort wird auf das jeweilige Versicherungsportfolio des Kunden abgestimmt und personalisiert. Werden die Qualitätsansprüche erfüllt, können automatisierte Antworten und eine Ausrollung auf verschiedenen Kundenkanäle in Betracht gezogen werden. Die Vision von Patrick Kaltenrieder, Leiter Digitalisierung bei KPT, ist ein 7x24h Stunden Service, wobei während den Geschäftszeiten die (menschlichen) Mitarbeiter des Beratungscenter aktiv sind und ausserhalb der Geschäftszeiten Generative AI-Agenten, wenn von den jeweiligen Kunden gewünscht, das Beratungscenter übernehmen.

Über das Unternehmen

Die KPT gehört zu den zehn grössten Krankenkassen der Schweiz. Rund 700 Mitarbeitende bieten den 600'000 Versicherten das, was sie wirklich brauchen: Beste Preis-Leistung und höchste Kundenzufriedenheit. Als Genossenschaft übernimmt die KPT gesellschaftliche Verantwortung. Sie ist eine mehrfach ausgezeichnete Arbeitgeberin und ermöglicht den Zugang zu qualitativ hochstehender Gesundheitsversorgung.

Mehr zu KPT: www.kpt.ch



Kontaktinformationen

KPT, Patrick Kaltenrieder, Leiter Digitalisierung (links)
patrick.kaltenrieder@kpt.ch

InventxLab, Carla Caspar, Strategic Innovation Manager (rechts)
carla.caspar@inventx.ch

Allgemeine Nutzung von ChatGPT

Agenten-Unterstützung in der E-Mail-Bearbeitung



Kundenanwendungen

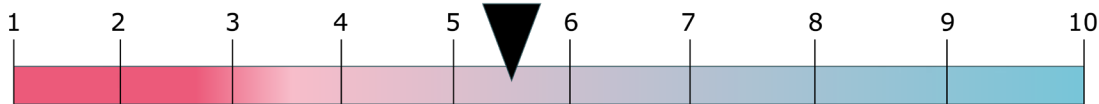
Automatisierte Dunkelverarbeitung

Abb. 19: Agenten-Unterstützung in der E-Mail-Bearbeitung auf der Gen AI Skala

Die Bearbeitung von E-Mail-Anfragen oder weiteren schriftlichen Anfragen kann durch Generative AI ebenfalls optimiert werden (Abb. 19). Finanzdienstleister erhalten täglich eine grosse Anzahl an schriftlicher Kommunikation, die schnell und genau beantwortet werden müssen, zum Beispiel E-Mails, Secure-Messages im Kundenportal oder in der App oder mittels Kontaktformular über die Webseite. KI-basierte Systeme können diese Anfragen analysieren, Zusammenfassungen und Antwortvorschläge erstellen und so den Prozess beschleunigen. Dies gewährleistet eine konstante und qualitativ hochwertige Kundenkommunikation.

Finanzdienstleister können ChatGPT unter Einhaltung ihrer Datenschutzrichtlinien nutzen (Abb. 20). Marktanalysen oder die Zusammenfassung öffentlicher Texte mit Hilfe von ChatGPT tragen zur Effizienz und Produktivität im Arbeitsalltag bei. So können insbesondere komplexe und öffentliche Anlagenberichte schnell und effizient mit ChatGPT zusammengefasst werden oder an die Tonalität einer bestimmten Zielgruppe angepasst werden.

Mitarbeitendennutzung von Chat GPT



Kundenanwendungen

Automatisierte Dunkelverarbeitung

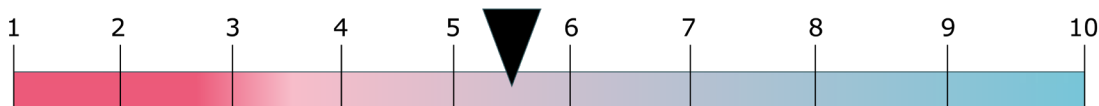
Abb. 20: Mitarbeitendennutzung von Chat GPT auf der Gen AI Skala

Nutzung des Microsoft Co-Piloten

Der Microsoft Co-Pilot kann in Finanzinstituten, sofern die Nutzung mit der Governance einhergeht, für die Erstellung von Berichten, Analysen und Präsentationen genutzt werden. Durch die Integration in bestehende Systeme kann der Co-Pilot schnell auf unternehmensinterne Daten zugreifen und diese für verschiedene Zwecke aufbereiten.

Ein grosses Potenzial bietet der Microsoft Co-Pilot zum Beispiel in der Risikobeurteilung bezüglich des Kreditportfolio- oder Liquiditätsrisikos, wo der Microsoft Co-Pilot gewünschte Fakten aus einer grossen Datenlage zusammenstellt und mit Vergleichen unterstreichen kann.

Mitarbeitendennutzung des Microsoft Co-Piloten



Kundenanwendungen

Automatisierte Dunkelverarbeitung

Abb. 21: Mitarbeitendennutzung des Microsoft Co-Piloten auf der Gen AI Skala

Einsatz für Marketingtexte oder Social Media

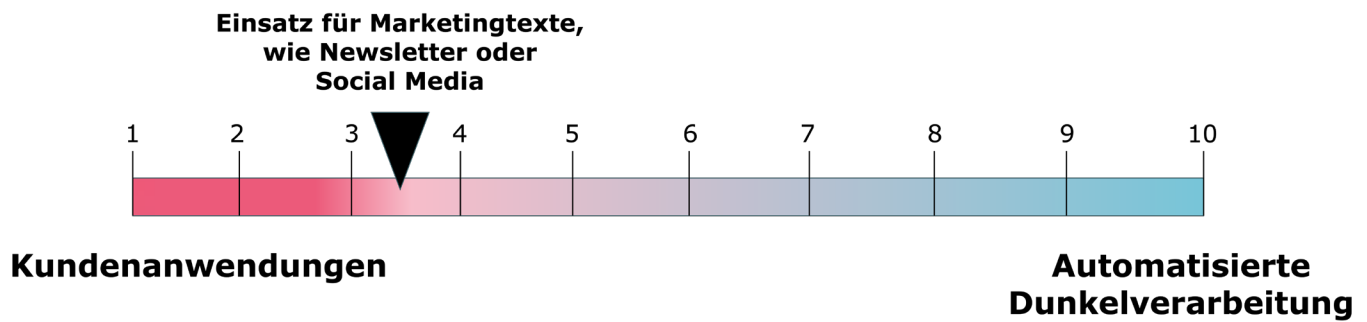


Abb. 22: Einsatz für Marketingtexte, wie Newsletter oder Social Media auf der Gen AI Skala

Im Marketing ist der Einsatz von Generative AI für Finanzdienstleister auch bei engen Datenschutzrichtlinien realisierbar. Tools zur Text-, Bild- und Videogenerierung unterstützen ansprechende, öffentlich zugängliche Newslet-

ter, um Social Media Beiträge zu erstellen, die auf aktuelle Trends und Kundeninteressen abgestimmt sind. Dies kann die Reichweite und Effektivität der Marketingmassnahmen erhöhen.

Onboarding und HR-Assistenz

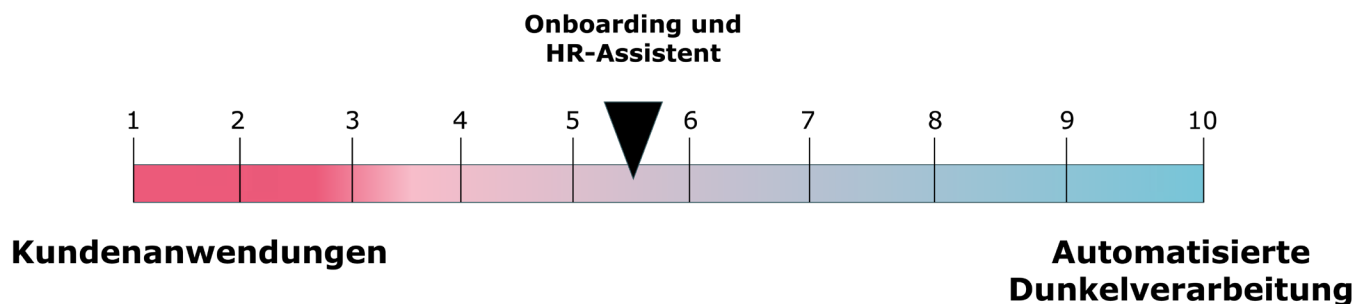


Abb. 23: Onboarding und HR-Assistent auf der Gen AI Skala

Der Einsatz von Generative AI im Onboarding-Prozess kann neue Mitarbeiter unterstützen, sich schneller in die Unternehmensstruktur und -kultur einzuarbeiten. Ein

KI-gestützter HR-Assistent kann häufig gestellte Fragen beantworten und auf relevante Ressourcen hinweisen, was die Effizienz des HR-Teams steigert.

Unterstützung in der Software-Entwicklung

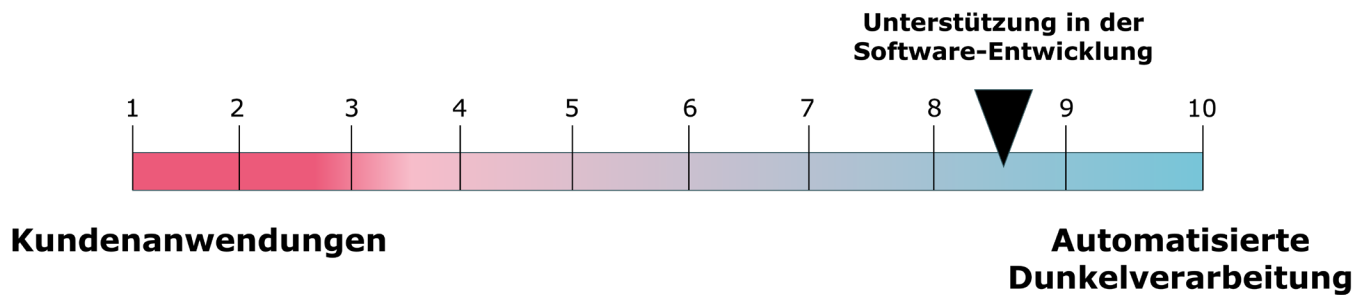


Abb. 24: Unterstützung in der Software-Entwicklung auf der Gen AI Skala

In Finanzinstituten, die eigene Softwareanwendungen entwickeln, kann Generative AI zur Überprüfung von Code und zur Generierung neuer Software-Komponenten eingesetzt werden. Dies kann die Entwicklungszeiten verkürzen und die Qualität der Software verbessern. Neben dem

GitHub Copilot gibt es eine stets wachsende Anzahl an Open-Source Modellen, wie zum Beispiel Code Llama von Meta oder Code Gemma von Google, welche bei erhöhten Sicherheitsansprüchen OnPremise bzw. in einer Private Cloud verwendet werden können.

Branchenspezifische Generative AI Agenten

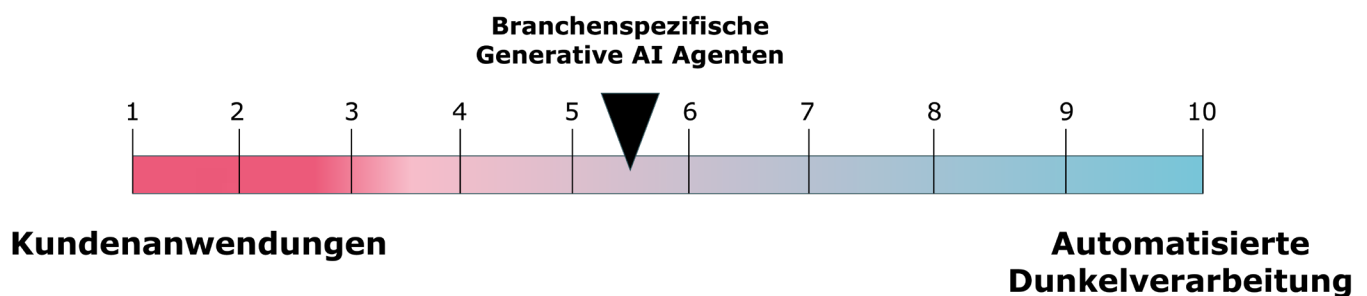


Abb. 25: Branchenspezifische Generative AI Agenten auf der Gen AI Skala

Analog den vorab beschriebenen branchenspezifischen Anwendungsfällen «Generierung von Produktbeschreibungen» und «Bereitstellung von Informationen für den Vertrieb», gibt es in der Finanzbranche branchenspezifische Generative AI-Anwendungsfälle.

Chatbots als Reglements-Spezialisten, Betriebshandbuch-Agenten und Wiki-Bots sind Beispiele, die bereits von mehreren Finanzinstituten genutzt werden und den Alltag

der Mitarbeitenden erleichtert. Beliebte Anwendungen von individuellen Generative AI-Agenten sind ein interner Reglements-Bot in der Kreditberatung, ein Wiki-Bot auf der Webseite oder als E-Banking Hilfe. Die aktuelle Studie «AI and Impact Investing» der HSLU (Ankenbrand, Bieri, Reichmuth, Yilmaz, 2024) zeigt mit der Anwendung von einem Large Language Model zur Analyse von Unternehmen bezüglich SDG-Kriterien das Potenzial von Anwendungsfällen in der Finanzbranche.

Schlussbetrachtung

Es lässt sich beobachten, dass die Anzahl finanzbranchenspezifischer Anwendungsfälle kontinuierlich zunimmt. Die Banken und Versicherungen pilotieren, prüfen und testen. Das Interesse am Einsatz von Generative AI wächst mit der Erfahrung. Die Einsatzgebiete weiten sich aus. Als Bremsfaktor sind die regulatorischen Anforderungen spürbar. Sicherheits- und Datenschutzabklärungen zu AI-Services sind oftmals umfassend. Zudem stellen On-Premise bzw. Private Cloud Installationen Herausforderungen bezüglich der Infrastruktur dar.

Die in Ergebnis-Kapitel beschriebenen Anwendungsfälle von Generative AI lassen sich für Finanzdienstleister adaptieren. Dabei ist es entscheidend, den Datenschutz und die Einhaltung regulatorischer Anforderungen zu berücksichtigen und ein entsprechendes AI-Setup zu wählen.

Generative AI ist auch für Finanzdienstleister ein wertvolles Werkzeug, um in einem wettbewerbsintensiven Markt erfolgreich zu agieren.

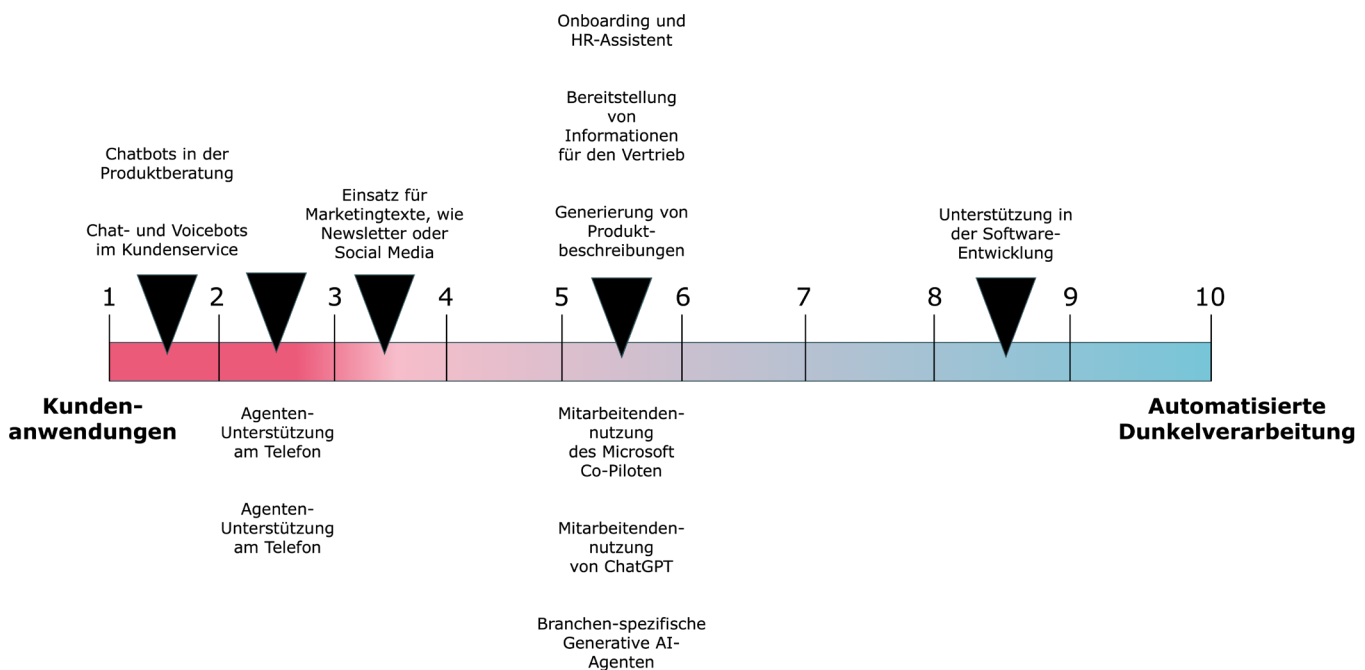


Abb. 26: Generative AI-Anwendungen auf der Gen AI Skala

Einordnung des Nutzens und der Machbarkeit der einzelnen Generative AI-Anwendungen im Hinblick auf die Finanzbranche

Die vorliegende Studie thematisiert verschiedene Anwendungen von Generative AI zunächst ohne Branchenfokus und anschliessend mit konkretem Fokus auf die Finanzbranche. Ziel der Studie ist es, erfolgreiche Generative AI-Anwendungen anderer Branchen zu identifizieren und deren Anwendbarkeit auf die Finanzbranche zu prüfen, um anschliessend Empfehlungen für Banken und Versicherungen im Hinblick auf sinnvolle Generative AI-Anwendungen geben zu können.

Dazu wurden in einem ersten Schritt eine Vielzahl von Experteninterviews mit KI-Experten und Enthusiasten aus unterschiedlichen Branchen geführt. Daraus ergaben sich interessante Generative AI-Anwendungen, die von der reinen Anwendung durch den Kunden bis hin zur automatisierten «Dunkelverarbeitung» eingeordnet wurden. Diese Anwendungen wurden dann zunächst von der Branchen-Expertin Carla Casper im Rahmen eines Gastbetrags auf ihre Anwendbarkeit in der Finanzbranche diskutiert. Dabei ergab sich, dass grundsätzlich alle Anwendungen auch auf die Finanzbranche anwendbar sind, vorausgesetzt der Datenschutz und die Compliance Regeln der jeweiligen Finanzinstitute können eingehalten werden. Vor diesem Hintergrund hat das Team der Hochschule Luzern weitere Fokusgruppen mit Bank- und Versicherungsangestellten durchgeführt. Die Fokusgruppen fanden in der ersten August-Hälfte statt. Ziel der Fokusgruppe war es, die einzelnen Anwendungsfälle mit den Teilnehmenden im Hinblick auf die Mehrwerte für die Banken und Versicherungen, sowie auf deren realistischen Aufwand einer compliance-kompatiblen Lösung einzuschätzen. Die Teilnehmenden wurden dabei gebeten, die einzelnen Anwendungen der Generativen KI auf einer Skala von eins bis fünf zuerst im Hinblick auf Nutzen und anschliessend im Hinblick auf Umsetzbarkeit zu bewerten.

Die Ergebnisse der Fokusgruppe sind in der folgenden Abbildung zusammengefasst. Dabei bedeutet ein hohes Scoring der Machbarkeit, dass der Aufwand für das Finanzinstitut überschaubar ist.

Von besonderem Interesse sind die Anwendungen, die sowohl in Bezug auf ihren Nutzen für die Bank als auch in ihrer Machbarkeit hohe Bewertungen erzielten. Insbesondere das Erstellen von Marketingtexten mittels Generative AI sowie das Verfassen von Produktbeschreibungen erhielten die höchsten Punktzahlen. Diese Anwendungen unterstützen Mitarbeitende in der Ausgestaltung der Kundenkommunikation. Die hohe Machbarkeitsbewertung resultiert vor allem aus den geringen Anforderungen an den Datenschutz bei diesen Anwendungen. Gleichzeitig sind es essenzielle Aufgaben, die ohne den Einsatz von Generative AI für die Mitarbeitenden sehr zeitaufwendig wären. Daher ist der Nutzen dieser Anwendungen ebenfalls sehr hoch.

Bezüglich des Nutzens und der Machbarkeit beim Einsatz von Chat- und Voicebots waren sich die Teilnehmenden der Fokusgruppe einig, dass zunächst eine weitere Differenzierung der Art der Kundenanfragen bzw. Beratungsprozesse erforderlich ist. Da sich Kundenservice-Dialoge nach dem Pareto-Prinzip in einfache und komplexe Anfragen unterscheiden (Hafner & Hundertmark, 2024), scheint uns diese Einteilung ebenfalls nützlich. Einfache Kundenanfragen und allgemeine Produktberatungen durch Generative AI bieten sowohl einen hohen Nutzen als auch eine relativ einfache Umsetzbarkeit. Bei komplexen Produkthanfragen oder Kundenservice-Fällen hingegen sinkt der Nutzen der Generative AI und die Machbarkeit wird erschwert. Dies führt dazu, dass der Einsatz von Chat- und Voicebots bei komplexen Themen zu den am schlechtesten bewerteten Anwendungsfällen für Generative AI gehört.

Ein weiterer interessanter Anwendungsfall ist die Bereitstellung von Informationen für den Vertrieb. Hier waren sich die Fokusgruppen-Teilnehmenden einig, dass diese Anwendung ebenfalls von sehr hohem Nutzen ist. Gleichzeitig wurde auch die Machbarkeit als relativ hoch eingeschätzt. In einigen Softwaresuiten zur Marktbearbeitung wird eine derartige Informationsaufbereitung bereits aus dem CRM-System automatisch eingespielt.

Ebenfalls zum Teil bereits in bestehenden Software-Anwendungen als Zusatz-Funktion integriert, ist die Unterstützung durch Generative AI im Rahmen der Software-Entwicklung. Hier bewerteten die Befragten die Machbarkeit sogar höher als den Nutzen, wobei auch der Nutzen mit 4 von 5 Punkten eher hoch eingestuft wurde.

Interessant waren die Diskussionen zur internen Nutzung von ChatGPT oder dem Microsoft Co-Piloten. Während in den Medien der Co-Pilot oft als der persönliche und intel-

ligente Assistent beschrieben wird, haben die Befragten den Mehrwert des klassischen ChatGPTs höher bewertet. Im Hinblick auf die Nutzung bzw. Machbarkeit von ChatGPT in einem Finanzunternehmen waren sich die Befragten einig, dass zwischen dem klassischen ChatGPT und einem individualisiertem ChatGPT mit eigenen Unternehmensdaten unterschieden werden muss. So ist die Einführung des allgemeinen ChatGPTs einfacher als die eines unternehmens-internen mit eigenen angereicherten Unternehmensinformationen.

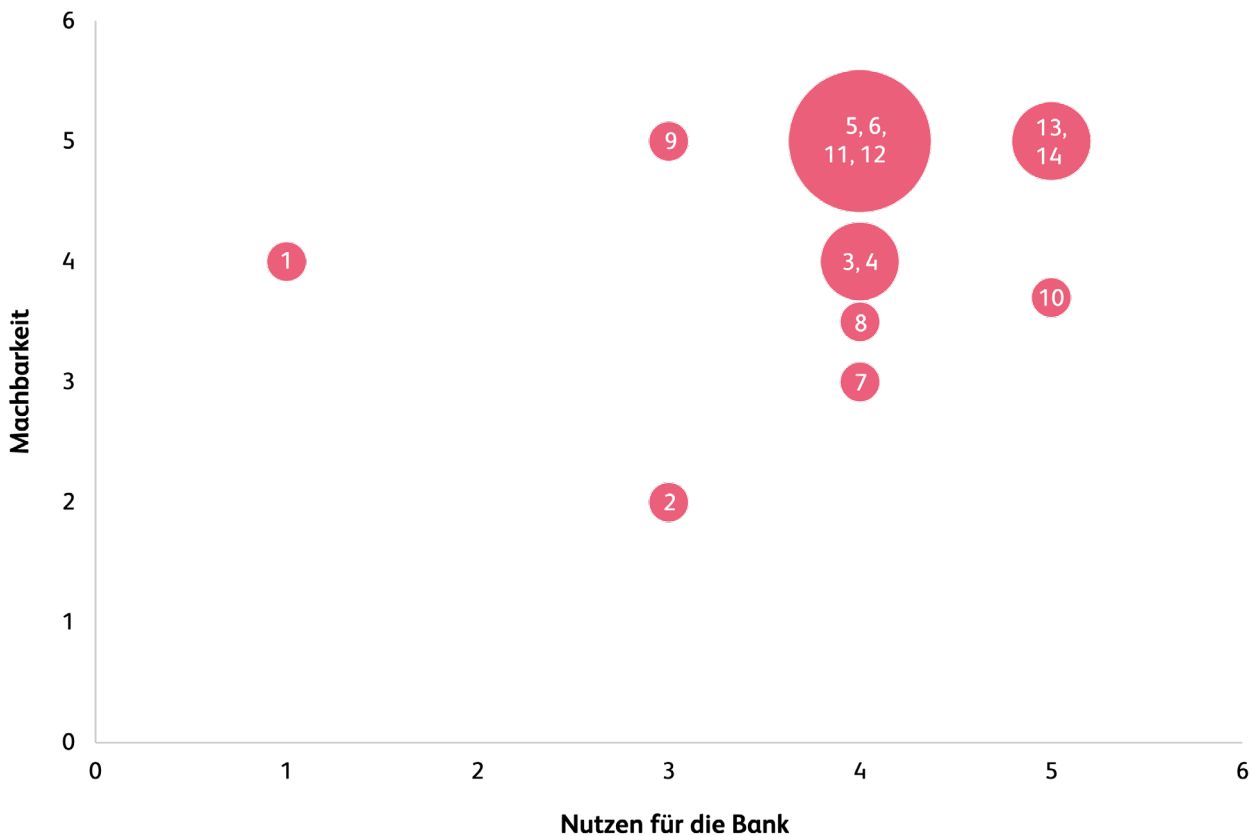


Abb. 27: Positionierung der Gen-AI Anwendungen im Spannungsfeld zwischen Nutzen und Machbarkeit für Banken

1. Chat- und Voicebots im Kundenservice (komplexe Fragen)
2. Chatbots in der Produktberatung (Kundenspezifisch)
3. Agenten-Unterstützung am Telefon
4. Agenten-Unterstützung in der E-Mail Bearbeitung
5. Chat- und Voicebots im Kundenservice (einfache Fragen)
6. Chatbots in der Produktberatung (allgemein)
7. Allgemeine Nutzung von ChatGPT (mit Unternehmenswissen)
8. Onboarding und HR-Assistent (FAQ)
9. Nutzung des Microsoft Co-Piloten
10. Bereitstellung von Informationen für den Vertrieb
11. Allgemeine Nutzung von ChatGPT (mit oder ohne Plus Account)
12. Unterstützung in der Software-Entwicklung
13. Einsatz für Marketingtexte, wie Newsletter oder Social Media
14. Generierung von Produktbeschreibungen

Fazit und Handlungsvorgehen bei

8 Einführung

Die vorliegende Untersuchung hat wesentliches Potential für die Finanzdienstleistungsbranche aufgezeigt, die sich aus dem Einsatz von Generative Artificial Intelligence in anderen Branchen ergeben.

Als zentrale Orientierung im Hinblick auf die Einordnung von Generative AI-Anwendungen für Banken und Versicherungen wurde in dieser Studie die **Gen AI Skala** eingeführt. Die Skala differenziert Anwendungen von Generative AI nach zwischen Anwendungen, die eher im Hinblick auf eine direkte Bedienung durch Kunden vorgesehen sind, von denen, die eher eine automatisierte Dunkelverarbeitung ermöglichen.

In Bezug auf das Nutzen/Machbarkeitsverhältnis wurden vor allem die Einsatzbereiche «Erstellen von Marketingtexten mittels Generative AI» sowie «Verfassen von Produktbeschreibungen» als besonders attraktiv angesehen. Das ist vor dem Hintergrund der immer komplexer werdenden Produktwelten im Anlage-Bereich und der damit verbundenen weltweit komplexer werdenden Einhaltung regulatorischer Vorschriften gut nachvollziehbar. Hier spielt Generative AI seine Stärken in Bezug auf die mühsame Anpassung von Texten in verschiedenen Sprachen und Gesetzgebungsräume aus. Auch aus der Perspektive des Marketings mit einem wachsenden Anspruch an Individualisierung und Personalisierung im Finanzbereich kann der beschriebene Einsatz der Anwendungen überzeugen. Das zeigen nicht zuletzt die vielfältigen Case Studies aus branchenfremden Organisationen. In Bezug auf die Einführung solcher Anwendungen ist auch die Vorgehensweise bei Banken und Versicherungen kaum anders als in Kapitel 5.4. skizziert.

Sollen allerdings Bots direkt und unmittelbar mit dem Kunden individuell über seine Situation kommunizieren, müssen diverse weitere Regulatorien beachtet werden. Die Erkenntnisse des Kapitel 5.4.4. gelten daher besonders für Banken und Versicherungen. Gerade vor dem Hintergrund immer knapper werdender Fachkräfte und dem Anspruch von Mitarbeitenden auf erfüllende und anspruchsvolle Aufgaben, wird es jedoch zukünftig auch für Banken und Versicherungen immer wichtiger, sich auch mit den

Anwendungsgebieten von Generative AI zu beschäftigen. Finanzdienstleister können hier aber durchaus von dem Wissen anderer Branchen profitieren, die sich vielleicht aufgrund einer herausfordernden Margensituation unter Umständen schon früher Gedanken dazu machen mussten, die Produktivität sprunghaft zu erhöhen, um im Wettbewerb bestehen zu können.

Neben der Gen AI Skala, die für Finanzdienstleister genauso anwendbar ist, wie für Unternehmen anderer Branchen, gilt auch der **KI-Prozess-Kompass** aus Kapitel 5.4. als zentrales Ergebnis dieser Studie. Dieser Wegweiser zeigt die bis zu 15 Schritte, die Banken und Versicherungen, aber auch andere Unternehmen bei der langfristigen und erfolgreichen Einführung von Generative AI-Anwendungen durchlaufen sollten.

9 Quellenverzeichnis

Ankenbrand T., Bieri D., Reichmuth L. & Yilmaz E. (2024). AI and Impact Investing

Beatty S. (2024). Tiny but mighty: The Phi-3 small language models with big potential

Hafner, N. & Hundertmark, S. (2024). Kundendialog-Management - Wertstiftende Kundendialoge in Zeiten der digitalen Automation. Wiesbaden: Springer Gabler.

Heskett J. L., Sasser W. E. & Schlesinger L. A. (1997). The service profit chain - how leading companies link profit and growth to loyalty, satisfaction, and value

Holz H. J., Applin A., Haberman B., Joyce D., Purchase H. & Catherine R. (2006). Research methods in computing: what are they, and how should we teach them? In Working group reports on ITiCSE on Innovation and technology in computer science education (ITiCSE-WGR ,06). Association for Computing Machinery, New York, NY, USA, 96–114. <https://doi.org/10.1145/1189215.1189180>

Khenouche F. (2024). Revolutionizing generative pre-trained: Insights and challenges in deploying ChatGPT and generative chatbots for FAQs

Leyes, M., Scanlon, J. T., & Wood, S. (2021). Elevating customer experience: A win-win for insurers and customers. McKinsey & Company. <https://www.mckinsey.com/industries/financial-services/our-insights/elevating-customer-experience-a-win-win-for-insurers-and-customers>

Price, B. & Jaffe, D. (2008). The Best Service is No Service – How to Liberate Your Customers from Customer Service, Keep Them Happy, and Control Costs, San Francisco.

Rane N. (2023). Role and challenges of ChatGPT and similar generative artificial intelligence in finance and accounting

Schick T. & Schütze H. (2021). It's Not Just Size That Matters: Small Language Models Are Also Few-Shot Learners

Srivatsa H. (2024). Fine-tuning versus RAG in Generative AI Applications Architecture

Strobel G., Banh L., Möller F. & Schoormann T. (2024). Exploring Generative Artificial Intelligence: A Taxonomy and Types

du Toit, G., & Burns, M. (2017). Why consumers trust Amazon almost as much as banks. Bain & Company. <https://www.bain.com/insights/why-consumers-trust-amazon-almost-as-much-as-banks/>

Uribe-Linares, G.P., Rios-Lama, C.A. & Vargas-Merino, J.A. (2023). Is There an Impact of Digital Transformation on Consumer Behaviour? An Empirical Study in the Financial Sector. *Economies*, 11, 132. <https://doi.org/10.3390/economies11050132>

Ein besonderer Dank für die Gestaltung dieser Studie geht an Jannik Wüster ([mail@jannikwuester.de](mailto:jannikwuester.de)).