kauz

Linguistic Technologies

Die KI-Assistenten von Kauz für Webseiten und Internen Self-Service

Ein Whitepaper über ChatGPT und kontrollierbare KI

September 2023

Einführung

In diesem Whitepaper beschreiben wir, warum und wie Generative KI (Künstliche Intelligenz), zum Beispiel in Form von ChatGPT den Einsatz von externen Chatbots und internem Wissensmanagement/Self-Service für Mitarbeitende revolutioniert und wie dies alles mit der Kontrollierbaren KI von Kauz zusammenhängt.

Fangen wir, bevor es technischer wird, mit Herausforderungen an, die uns allen geläufig sind und mit denen wir alle täglich zu tun haben:

- Immer mehr Informationen müssen immer schneller von uns sortiert, eingeordnet und verarbeitet werden (Information Overload).
- Die Komplexität im Geschäftsleben nimmt weiter zu, der Innovationszyklus kontinuierlich ab.
- Unternehmen sind gezwungen, kontinuierlich ihre Effizienz zu steigern und einen zunehmenden Personalmangel zu kompensieren.
- Wir als Konsumenten halten uns bei unseren Einkäufen nicht mehr an die Öffnungszeiten, sondern suchen nach Produkten, bewerten und kaufen Leistungen und Produkte an jedem Wochentag und zu jeder Uhrzeit.
- Auch als Mitarbeitende eines Unternehmens ist unsere Arbeitszeit nicht mehr ausschließlich von 08:00 bis 17:00 Uhr, sondern wir arbeiten auch am Wochenende, abends, von zu Hause aus oder aus dem Zug heraus.

Die Informationsflut und die Individualität in der Aufgabenerledigung, auch zu Zeiten, in denen ich keine Kollegin oder Kollegen fragen kann, fordert auch eine Eigenständigkeit in der erfolgreichen Erledigung. Hierbei unterstützt ein Assistent, basierend auf neuester KI-Technologie.

ChatGPT und Generative KI

Diese Herausforderungen sind nicht neu. Neu sind die technologischen Antworten darauf. Ein Meilenstein kam Ende des Jahres 2022 an die Öffentlichkeit: ChatGPT.

ChatGPT (oder allgemein: Sprachmodelle, Large Language Models LLMs)) wird unser Arbeitsleben verändern. Das liegt vor allem an 2 Aspekten:

ChatGPT ist ein **G**enerative **P**re-Trained **T**ransformer-**S**prachmodell. GPT-basierte Anwendungen wie ChatGPT können verschiedene Arten von Text erstellen, abwandeln, ergänzen, übersetzen, zusammenfassen, klassifizieren und Informationen extrahieren.

Dazu können sie Programmcode erstellen und mit Bildern in der Eingabe und Ausgabe arbeiten, kurz sie generieren neuen Content. Man spricht von Generativer KI.

ChatGPT ermöglicht eine simple Interaktion über ein Dialogsystem wie bei einem ChatBot, mit einer einfachen Frage- & Antwortlogik. Man benötigt keine KI-Kenntnisse mehr, um KI zu nutzen.

Dies führt zu einer „Demokratisierung der KI“ und zum Durchbruch in sehr vielen Anwendungsgebieten.

Das folgende Schaubild zeigt einige der Fähigkeiten

ChatGPT hat zum Beispiel folgende Fähigkeiten:

- Texte zu generieren, z. B. E-Mails, Konzepte, Rechtstexte, Gedichte, Programmcode.
- Texte in andere Formen/Formate zu übertragen, z. B. Texte in andere Sprache übersetzen, einen Code (z. B. C+) in einen anderen Code (z. B. Python) umwandeln.
- Texte zusammenzufassen, Inhalte zu extrahieren.
- Excel-Tabellen zu generieren und Analysen von Tabellen in Form von Zusammenfassungen zu generieren.

Diese Fähigkeiten können sehr gut für den Aufbau und Betrieb eines externen KI-Assistenten auf einer Webseite oder für KI-Assistenten für Mitarbeitende genutzt werden.

Bevor wir zu dem „Wie“ hierfür kommen, wollen wir noch einen anderen Aspekt einführen: den Aspekt der Kontrollierbarkeit.

Kontrollierbare KI

ChatGPT hat auch, zumindest im Unternehmens-Kontext, einige Nachteile, die eliminiert oder zumindest eingegrenzt werden müssen. Hierzu nutzen wir die NLU-Engine von Kauz (NLU: Natural Language Understanding). Die Nutzung der NLU-Engine von Kauz erzielt die gewünschte Kontrollierbarkeit.
Diesen kombinierten Ansatz bezeichnen wir als hybriden Ansatz.

Kontrollierbarkeit der Informationsquellen

ChatGPT wurde zwar mit 300 Milliarden Wörtern aus dem Internet und anderen digitalen Quellen trainiert, aber Ihre speziellen Unternehmensinformationen werden schwerlich hierin enthalten sein. Wenn Sie als Marketer ChatGPT beauftragen möchten: „*Bitte erstelle mir einen Text für unseren Marketingflyer für unser neues Produkt xyz. Der Stil soll informativ-sachlich sein und der Text soll maximal 200 Wörter umfassen*“, so wird ChatGPT dies nicht zu Ihrer Zufriedenheit ausführen können, da es sicherlich keine Informationen über Ihr neues Produkt hat. Wir müssen also den sehr fähigen Algorithmus ChatGPT mit Ihren Unternehmensdaten „zusammenbringen“, d.h. die Antwort soll aufgrund Ihrer speziellen Unternehmensinformationen erfolgen.

Hierzu lesen wir die gewünschten Unternehmensinformationen in die Kauz-NLU-Engine ein. Dies kann durch Einlesen („Crawlen“) Ihrer Webseite erfolgen, dies kann aber auch beliebige Dokumente aus Ihren unternehmensinternen Quellen beinhalten (Dokumente, pdfs, Excel-Tabellen, Formeln, Bilder,...). Die eingelesenen einzelnen Fragmente werden durch ein sogenanntes “Embedding“ über den Kauz Editor mit ChatGPT „verbunden“. Technisch funktioniert es so:

“Damit ein großes Sprachmodell mit Texten umgehen kann, müssen Texte zunächst in eine maschinenlesbare Form umgewandelt werden. Jedem Wort (oder auch einzelnen Wortbestandteilen), Satz- und Sonderzeichen wird dann ein numerischer Wert zugeordnet, mit dem der Computer operieren kann. Zusätzlich sollen ähnliche Worte auch ähnliche numerische Werte aufweisen. Dazu werden Wörter mit Hilfe von Vektoren in einer Art mehrdimensionalem Raum abgebildet. Ähnliche Wörter mit ähnlichen Bedeutungen werden möglichst nahe beieinander im Vektorraum positioniert. Die Ähnlichkeit bestimmt sich durch das gemeinsame Auftreten mit anderen, begleitenden Worten. Diese sogenannte Worteinbettung (Word Embedding) in einen multidimensionalen Raum ist die Grundlage für die Sprachverarbeitung durch große Sprachmodelle. “

Quelle: „What Is ChatGPT Doing … and Why Does It Work?“, Stephen Wolfram Writings, 14. Februar 2023. https://writings.stephenwolfram. com/2023/02/what-is-chatgpt-doing-and-why-does-itwork/

Dies mag technisch erscheinen, hat aber 2 wesentliche Auswirkungen:

- Bei einer Anfrage an einen externen oder internen KI-Assistenten wird ChatGPT die Antwort nur aufgrund dieser ihm zur Verfügung gestellten Unternehmensinformationen generieren.
- Die Unternehmensinformationen sind zwar mit ChatGPT „verbunden“, aber nicht direkt in das Modell integriert, in den Algorithmus „reintrainiert“ D. h. bei neuen Informationen fügen wir diese mit neuen Embeddings hinzu, bei Wegfall von Informationen (z. B. das Produktportfolio ändert sich) „kappen wir die Verbindung“. Der Algorithmus muss nicht neu antrainiert werden. Einige neue Testfragen, um das Verhalten zu überprüfen, reichen aus. ChatGPT wird nicht mehr auf die alten Unternehmensinformationen zugreifen.

Hybrider Ansatz der Kauz GmbH:

Kontrollierbarkeit des Dialogflusses

ChatGPT ist zwar ein gutes Dialogsystem, die Dialoge sind aber in einer Frage-Antwort-Systematik. In manchen Anwendungsszenarien ist es wichtig, geführte Dialoge und Aktionen/Prozesse zu integrieren. Dies kann ChatGPT nicht. Hierzu werden im Kauz-Editor sogenannte Dialogpläne (no code) erstellt, mit denen Dialoge gesteuert werden, zum Beispiel ein Ticket-Aufnahmeprozess oder eine Schadenmeldung. Hierüber werden auch Backend-Systeme integriert, zum Beispiel CRM oder ITSM-Systeme.

Kontrollierbarkeit der Antworten

ChatGPT generiert Text/Antworten anhand der ihm zur Verfügung gestellten Informationsquellen. Diese werden in vielen Fällen richtig sein. In manchen Fällen erscheinen sie glaubhaft, sind aber nicht 100% korrekt. Wir empfehlen, von ChatGPT in sensitiven Themenbereichen generierte Antworten (Preise, Vertragsbedingungen, ...) von Menschen kontrollieren und eventuell überarbeiten zu lassen. Hierzu bietet der Kauz-Editor ein entsprechendes Redaktionssystem, in dem Sie die von ChatGPT generierten Antworten(vorschläge) einsehen und verändern können. Diese redaktionell überarbeiteten Antworten werden in der Wissensbasis des KI-Assistenten gespeichert. Bei einer neuen Frage zu diesem Themengebiet wird immer zuerst die KI-Assistent-interne Wissensbasis durchsucht und die dort durch Sie gespeicherte Antwort ausgegeben. So behalten Sie die Kontrolle über das Antwortverhalten.

Kontrollierbarkeit der Modelle

ChatGPT ist sicherlich zurzeit das bekannteste Sprachmodell, es ist aber nicht das einzige. Es gibt bereits viele weitere Sprachmodelle und es kommen neue hinzu. An dieser Stelle seinen nur beispielhaft folgende weitere genannt: BERT (Google), BART (Facebook), LaMDA (Google), LLaMa (MetaAI), MPT (MosaicML), Falcon, Luminous (Aleph Alfa).

Sprachmodelle selbst sind ein Thema für ein eigenständiges Whitepaper. An dieser Stelle sollen folgende Informationen für unseren Kontext genügen:

Jedes Sprachmodell hat seine Vor- und Nachteile. Das eine halluziniert wenig, das andere versteht sehr gut Deutsch, das eine ist auch on-premise lauffähig, andere nur in der Cloud usw. Auch können sich manche Sprachmodelle für manche Einsatzszenarien besonders gut eignen. Insgesamt ist ChatGPT, das auf keinen speziellen Zweck antrainiert wurde, aktuell für den allgemeinen Business-Einsatz (noch?) deutlich am besten geeignet.

Als anwendendes Unternehmen von Sprachmodellen bzw. von Anwendungen, die auf Sprachmodellen basieren, ist es wichtig, zukünftig unabhängig von speziellen Modellen zu sein. Zu rasant ist die derzeitige Entwicklung in diesem Thema, als dass man zukünftige technologische Entwicklungen auch nur auf absehbarer Zeit übersehen könnte. Der KI-Assistent von Kauz basiert standardmäßig auf ChatGPT, ist aber auch mit sehr geringen Anpassungen auf anderen Sprachmodellen lauffähig

Kontrollierbarkeit Datenschutz

Das Thema Datenschutz bei der Nutzung von Sprachmodellen und im speziellen von ChatGPT ist komplex und variantenreich. Es hängt auch davon ab, welchen Zugang man zu ChatGPT nutzt. Für Unternehmen, die die OpenAI-API nutzen wollen, gibt es durchaus Möglichkeiten, die GPT-Technologie datenschutzkonform in die eigenen Produkte und Dienste zu integrieren.

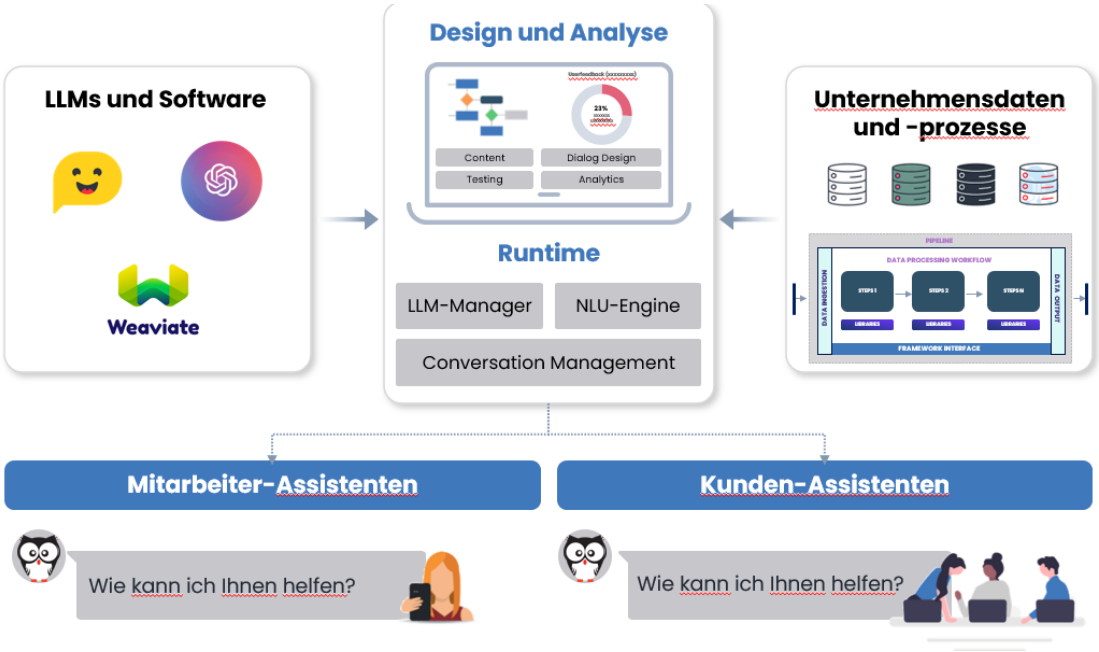
Der KI-Assistent von Kauz nutzt allerdings nicht die OpenAI-API, sondern die Azure Open AI Services von Microsoft. Hierdurch ist gewährleistet, dass die Daten nicht in die USA gesendet werden, sondern auf europäischen Servern verbleiben und nicht für Trainingszwecke der Modelle genutzt werden. Dieses Vorgehen ist DSGVO-konform. Trotzdem raten wir davon ab, persönliche und vertrauliche bzw. sicherheitsrelevante Daten (z. B. Geodaten) in diesem Kontext zu nutzen

Zusammenfassung

Die folgende Architekturfolie fasst das Gesagte zusammen und stellt es nochmals grafisch gut dar:

Der hybride Ansatz des KI-Assistenten von Kauz verbindet Sprachmodelle mit Unternehmensdaten und – Prozessen

Die NLU-Engine von Kauz stellt die Kontrollierbare KI sicher und führt das Design sowie die Analyse für die kontinuierliche Optimierung durch



Kontaktieren Sie uns jetzt für eine Demo oder um Möglichkeiten für den Einsatz von KI in Ihrem Unternehmen zu besprechen.

Sven Wilms, Vertriebsleiter bei Kauz

Kauz GmbH
Erasmusstraße 15
40223 Düsseldorf
Tel: +49 (0) 211 3004 9622
info@kauz.net
www.kauz.net