

Anleitung für den Einsatz Ihrer Dokumente in LLMs

Inhalt

Einleitung	3
Anforderungen an die Antworten	4
Qualitätssicherung durch Zusatzdokumente	4
Allgemeine Anforderungen an den Content	5
Spezielle Inhalte und Strukturen in Dokumenten	6
Contentqualität, Fragen und Prompts	7
Generelle Schwierigkeiten von LLMs	8
Fazit	9

Einleitung

Die Zukunft des Kundenservices liegt in der intelligenten Nutzung von Large Language Models (LLMs) wie OpenAI und ChatGPT. Doch um das volle Potenzial dieser Technologien zu entfalten, braucht es mehr als nur leistungsstarke Algorithmen – es bedarf der richtigen Dokumentation. Mit KAI, dem virtuellen KI-Assistenten, haben Unternehmen die Möglichkeit, LLMs effizient als „Antwortmaschinen“ einzusetzen.

Damit ein LLM präzise und verlässliche Antworten liefert, müssen die zugrunde liegenden Dokumente perfekt auf die speziellen Anforderungen dieser Technologie abgestimmt sein. LLMs verarbeiten Inhalte anders als menschliche Mitarbeitende, was eine gezielte Aufbereitung der Informationen unerlässlich macht.

Erfahren Sie in diesem Smart Guide, welche Anforderungen Ihre Dokumente erfüllen müssen, um in KAI und anderen LLM-basierten Systemen erfolgreich genutzt zu werden.

Anforderungen an die Antworten

In diesem Dokument wird der Einsatz von LLMs zur Erstellung von Antworten im Customer Service Center beschrieben. Die Anforderungen an die Dokumente ergeben sich aus den Anforderungen an die Antworten.

Korrektheit der Antworten

Die Antworten müssen inhaltlich richtig sein, und es dürfen keine sogenannten „Halluzinationen“ (falsche oder erfundene Informationen) auftreten.

Vollständigkeit

Je nach Thema müssen die Antworten wichtige Informationen wie Sicherheitswarnungen, Verfügbarkeiten, Ober- und Untergrenzen, Selbstbeteiligungen und rechtliche Hinweise enthalten. Diese Anforderungen gelten für verschiedene Fragestellungen und dürfen nicht ausgelassen werden.

Nützlichkeit

Die Antworten sollten in einem gut strukturierten, lesbaren und verständlichen Format gegeben werden. Ein Beispiel hierfür wäre eine Schritt-für-Schritt-Anleitung, bei der die richtige Reihenfolge der Schritte von Bedeutung ist.

Konsistenz

Antworten müssen konsistent und wiederholbar sein, d. h., auf dieselben Fragen sollten immer dieselben Antworten folgen. Abweichungen oder unterschiedliche Varianten sind zu vermeiden. Diese Konsistenz wird durch die richtige Nutzung von Prompts und die Qualität der zugrunde liegenden Dokumente erreicht.

Um diese Qualitätsstandards zu gewährleisten, ist es entscheidend, zuerst die Themen und die potenziellen Fragestellungen klar zu definieren. Ohne eine genaue Vorstellung davon, welche Themen und Fragen behandelt werden sollen, ist es schwierig, die Dokumentenqualität effizient zu beurteilen und zu verbessern.

Qualitätssicherung durch Zusatzdokumente

Es ist nicht immer notwendig, sämtliche Informationen in einem einzigen Dokument zu hinterlegen. Um Begriffe zu erklären oder Zusammenhänge zu hinterlegen kann es sinnvoll sein, zusätzliche Dokumente mit Informations- oder Anweisungscharakter in der

Wissensdatenbank zu hinterlegen. So können beispielsweise zusätzliche Glossare verwendet werden oder auch Prozessbeschreibungen, um die Reihenfolge von Arbeitsschritten deutlich zu machen.

Allgemeine Anforderungen an den Content

Die Funktionen des LLMs können nur so gut sein wie der verfügbare Content. Wenn der Inhalt unvollständig ist oder Widersprüche enthält, können die Antworten nicht die gewünschten Anforderungen erfüllen.

Die Anforderungen an den Content können viele Aspekte betreffen, die auch von der jeweiligen Situation abhängen.

Es gibt zwei Arten von Qualitätsanforderungen:

01 | Generelle Anforderungen

Diese gelten immer, unabhängig von der Situation oder Fragestellung.

02 | Spezifische Anforderungen

Diese können erst geprüft werden, wenn klar ist, welche Fragen gestellt werden und welche Anwendungsfälle (Use Cases) es gibt.

Genereller Inhalt der Dokumente

In internen Servicedokumenten wird oft vorausgesetzt, dass bestimmte Dinge schon bekannt sind. Zum Beispiel, was Agenten in Schulungen lernen oder was für uns Menschen logisch erscheint. Dinge, die „schon immer so waren“, müssen für das LLM (KI-Modell) nicht automatisch verständlich sein, da es dieses Hintergrundwissen nicht hat. Daher ist es sinnvoll, die Inhalte von einer fachfremden Person gegenlesen zu lassen und auf Schlüssigkeit sowie Inhaltslücken überprüfen zu lassen. Auf diese Weise findet man heraus, an welchen Stellen auch das LLM ins Straucheln geraten würde.

Aufbau der Dokumente

Es muss in den Dokumenten klar ersichtlich sein, welches Thema sie behandeln und aus welcher Perspektive. Es ist immer wieder zu beobachten, dass Dokumente in unterschiedlichen Kategorien zugeordnet sind, beispielsweise „Geschäftskunden“ und „Privatkunden“. Dieser Unterschied ist aber im Text nicht entnehmbar. Sofern die Regelungen für verschiedene Kundengruppen widersprüchlich sind, führt dies zu fehlerhaften Beauskunftung, sofern das Metadatum „Kundengruppe“ für das LLM nicht sichtbar ist. Dieser Aspekt gilt auch, wenn die Gliederung der Dokumente unübersichtlich ist, nur teilweise eine Fallunterscheidung vorgenommen wird. Das System kann nicht wissen, in welchen Fällen die Regelung allgemeinen Charakter hat und daher z. B. die Kundengruppe fehlt und in welchen Fällen die Kundengruppe vorausgesetzt oder vergessen wurde.

Verständlichkeit der Begriffe

Es ist sowohl bei der Nutzung eines Inhaltes durch Menschen als auch durch Maschinen zu prüfen, ob die verwendeten Begriffe verständlich sind. Im Falle einer Nutzung durch Menschen ergibt sich diese Bewertung aus dem Ausbildungsstand der Nutzer. Im Falle einer Maschine gilt es zu prüfen, welche internen, ungewöhnlichen oder auch fach- oder firmenspezifischen Begriffe verwendet werden. So kann beispielsweise ChatGPT mit einem Begriff wie einer Produktbezeichnung nichts anfangen, wenn auch jeder im Unternehmen weiß, was das ist. Oder Begriffe werden in vielen Bedeutungen verwendet.

Entsprechend ist zu prüfen, ob der zu verwendende Text derartige Begriffe enthält, ohne dass sie erläutert sind.

„Lost in the middle“ Problem

Je umfangreicher Dokumente und Contexte sind, desto höher ist die Wahrscheinlichkeit, dass eine falsche Beantwortung erfolgt. Bei großen Context Windows steigt die Wahrscheinlichkeit, Inhalte als nicht ausreichend relevant für die Nutzung in einer Bearbeitung zu erkennen. Daher spielt die Lokalität, also wo Inhalte gespeichert sind, eine Rolle. Natürlich kann ein RAG Konzept die Menge der Inhalte reduzieren, ist ein Aspekt jedoch an verschiedenen Stellen in einem oder mehreren Dokumenten verteilt kann das Problem auftreten.

Dokumentenmenge

Häufig argumentieren Kunden, die mit der Antwortqualität Probleme haben damit, dass sie ja noch mehr Dokumente verfügbar hätten und sie nur einen Ausschnitt verwendet haben. Dahinter steckt die Hoffnung, dass durch mehr Dokumente die Qualität der Antworten steigen würde.

Tatsächlich verhalten sich LLMs sehr statistisch – häufige Aussagen werden häufiger verwendet, seltene, widersprüchliche häufiger ignoriert. Ist also die umfangreiche und korrekte Aussage häufiger oder werden durch die häufigeren Meldungen Sachverhalte umfassender dargestellt, dann kann das die Antworten verbessern.

Allerdings zeigt die Erfahrung, dass Aussagen mehrheitlich unscharf sind und wirklich konkrete und detaillierte Aussagen in den Dokumenten eher die Ausnahme darstellen.

Es hat sich gezeigt, dass mehr Dokumente nicht immer zu besseren Antworten führen. Tatsächlich kann eine große Menge an Dokumenten die Antwortqualität sogar verschlechtern. Das liegt daran, dass:

- Aktualisierungen bei einer großen Menge an Dokumenten schwer umzusetzen sind.
- Es oft zu Wiederholungen kommt, was den Überblick erschwert.
- Alte und neue Informationen vermischt werden, was zu unsicheren und ungenauen Antworten führen kann.

Spezielle Inhalte und Strukturen in Dokumenten

Konkretheit von Angaben

Oft referenzieren Texte auf Gegebenheiten, die sich aus der Lebenswelt des Autors oder des erwarteten Lesers ergeben. Beispielsweise Angaben wie „Melde Dich dann bei Ireen“ oder „wie das schon letztes Jahr definiert wurde“, oder auch Angaben, die auf offensichtlich bekannte Regeln verweisen („entsprechend den Regelungen zu zahnärztlichen Leistungen“), die aber das System nicht kennen kann.

Daher müssen die Angaben in den Texten klar und konkret sein. Zum Beispiel sollten statt relativen Zeitangaben („seit letztem Jahr“) genaue Daten genannt werden. So wird sichergestellt, dass sowohl Menschen als auch Maschinen die Informationen richtig verstehen können.

Lokalität

LLMs haben wie oben beschrieben häufig Probleme damit, zusammenhängende Aussagen zu finden, die an verschiedenen Stellen stehen. Das kann passieren, wenn die Informationen in unterschiedlichen Dokumenten oder weit auseinander im selben Dokument stehen. Tests zeigen, dass Inhalte am Anfang oder Ende eines Textes mehr beachtet werden, während Informationen in der Mitte öfter übersehen werden.

Es ist also wichtig, zusammengehörende Inhalte auch gemeinsam darzustellen.

Ein Tipp: Man kann dieses Problem teilweise umgehen, indem man den Text so gestaltet, dass alle wichtigen Inhalte zuerst zusammengefasst werden. Danach kann die Antwort auf Basis dieser Zusammenfassung gegeben werden („Chain of Thought“-Technik).

Contentqualität, Fragen und Prompts

Ein Text muss immer so gut sein, wie die Fragen, die er beantworten soll. Ohne eine klare Übersicht der Themen und Fragen ist es schwer, die Qualität des Inhalts zu bewerten. Es hat sich als sinnvoll gezeigt, die Prompts zur Beantwortung von Fragen und die Bewertung der Dokumentenqualität in einem Zusammenhang zu sehen.

Widersprüche

Oft enthalten Texte Widersprüche, da sie unterschiedliche Redaktionszeiten oder Sichten enthalten. Um Widersprüchlichkeiten zu identifizieren ist es wichtig zu wissen, welche Widersprüchlichkeiten überhaupt relevant sind.

Beispielsweise ist es in einer Projektdokumentation nicht entscheidend, wenn verschiedene Angaben zum Projektende beschrieben sind, wenn an das LLM keine Fragen gestellt werden, die die Kalenderdaten der Projekte betreffen.

Die Relevanz von Widersprüchlichkeiten ergibt sich also aus den erwarteten Fragen. Sollen beispielsweise Zeitangaben erfragt werden, dann müssen diese Widerspruchsfrei hinterlegt sein. Hier haben wir das erste Beispiel von Qualitätsanforderungen, die auf Basis von Use Cases definiert werden können.

Vollständigkeit

Sind Inhalte in einem Content Bestand nicht vollständig, dann wird das LLM entweder Antworten nicht wissen – womit man grundsätzlich leben kann. Oder aber halluzinieren, was für Anwender sehr viel unglücklicher ist.

Dabei gilt auch wieder, wie oben, dass sich die Vollständigkeit aus den zu stellenden Fragen ergibt. Beispiel: Wenn Fragen zu Lebenssituationen eines Bürgers (z. B. „Ich habe geheiratet, was muss ich tun?“) mithilfe von Verwaltungsvorschriften beantwortet werden sollen, müssen diese Vorschriften alle relevanten Informationen enthalten. Sonst fehlt dem LLM die Grundlage, um die Fragen richtig zu beantworten.

Dasselbe gilt, wenn relevante Aussagen (beispielsweise zu Obergrenzen, Prozessschritten, Gültigkeit, Verfügbarkeit etc.) in den Dokumenten teilweise, aber nicht grundsätzlich stehen. Dann können sie auch nicht verlässlich in den Antworten genutzt werden.

Die Liste der geforderten Vollständigkeit ergibt sich also aus den Fragestellungen, die geplant sind.

Generelle Schwierigkeiten von LLMs

LLMs haben eine begrenzte Funktionalität. Es handelt sich um einen stochastischen Prozess, der potenziell, aus Sicht des Anwenders, fehlerbehaftet ist. In diesem Abschnitt beschreiben wir Probleme, die uns im Laufe der Zeit aufgefallen sind.

Wörtliches Verständnis

LLMs nehmen jede Formulierung wörtlich. Wird beispielsweise in einem Satz über eine Produktfunktion die Zukunftsform verwendet kann es sein, dass in der Antwort zu der Frage nach Produktfunktionen diese eine Funktion fehlt, da sie offensichtlich erst in der Zukunft angeboten wird, auch wenn der Schreiber die Verwendung des Futurs so genau gar nicht überdacht hatte.

Beschränkte Gültigkeiten

Ein generelles Problem besteht darin, dass LLMs Texte interpretieren und wiedergeben. Es besteht kein Verständnis des „Sinns“. Daher kann es passieren, dass zeitliche Gültigkeiten nicht korrekt verstanden werden. Beispielsweise kann es sein, dass Regelungen nur in der ersten Jahreshälfte gültig sind, vom LLM jedoch generell als gültig angeführt werden. Ebenso kann es vorkommen, dass als nicht gültig beschriebene Texte dennoch die Statistik prägen und zu nicht korrekten Antworten führen.

Implizite Formulierungen

Es kann passieren, dass Formulierungen, die uns Menschen eingängig sind vom LLM nicht erkannt werden.

Ein Beispiel wäre: „Wir treffen uns am Montag am Hauptbahnhof, Spiegelstraße 1“. Stellt man nun die Frage, wo man sich trifft, ist es kein Problem die korrekte Antwort zu bekommen.

Fragt man jedoch nach der Adresse des Hauptbahnhofs ist das System häufig nicht in der Lage, die Adresse zu melden. Hier benötigt es eine konkrete Formulierung „die Adresse des Hauptbahnhofs ist...“.

Ein weiteres Problem sind Vorsilben, die nur einmal erwähnt und dann vorausgesetzt werden.

Zum Beispiel: „Eine Barauszahlung ist möglich.“ In den folgenden Sätzen steht vielleicht nur noch „Auszahlung“. Dass das LLM versteht, dass weiterhin „Barauszahlung“ gemeint ist, ist nicht selbstverständlich. Es erkennt die Verbindung nicht automatisch.

Fazit

Beim Einsatz von LLMs gibt es viel zu bedenken und dieses Dokument ist bei weitem nicht vollständig, da auch wir kontinuierlich dazu lernen. Der Schlüssel zum erfolgreichen Einsatz von LLMs liegt in der Qualität der Dokumente, auf die das System Zugriff hat. Mit unserer **USU-Qualitätsgarantie** sichern wir zu, dass die Antworten von KAI, unserem KI-Assistenten, stets korrekt, vollständig, nützlich und konsistent sind.

Wenn Sie mehr darüber erfahren möchten, wie Sie Ihre Wissensdatenbank fit für den Einsatz von KI machen, freuen wir uns, von Ihnen zu hören.

Benötigen Sie weitergehende Informationen, eine Live-Demo oder haben Sie Fragen?

Die meisten Fragen lassen sich im direkten Kontakt am besten klären.

Ich freue mich darauf, Ihre Fragen und Wünsche telefonisch zu beantworten.

Jetzt Termin vereinbaren.



Anke Kanwischer
USU Community Developer
Knowledge Management
anke.kanwischer@usu.com