

MASCHINENLERNEN

KI-Sprachmodelle im Kanzleieinsatz?

von StB Jürgen Derlath, Münster

| In der heutigen digitalen Ära dringen KI-Sprachmodelle in immer mehr Bereiche vor und versprechen innovative Lösungen selbst für komplexe Herausforderungen. Von der Automatisierung von Routineaufgaben bis zur Analyse umfangreicher steuerlicher Daten – die Einsatzmöglichkeiten dieser leistungsstarken Sprachmodelle in der Steuerberatung wären vielfältig. Erfahren Sie in diesem Beitrag, welche Einsatzmöglichkeiten sich KI-Sprachmodelle schon heute in Kanzleien bieten und wo die Stärken – aber auch die Schwächen – liegen. |

„Chatbot-Systeme“ – Regelbasiert oder generisch?

Regelbasierte Chatbots

Chatbot ist nicht gleich Chatbot. Regelbasierte Chatbots verwenden vordefinierte Regeln, um mit Benutzern zu interagieren. Sie sind in der Lage, einfache Fragen schnell und genau zu beantworten, da sie auf bekannte Fragen und Antworten programmiert sind. Diese Chatbots eignen sich gut für Standardfragen wie Öffnungszeiten oder Produktdetails. Regelbasierte Bots können selbst programmiert oder mithilfe von Chatbot-Baukästen erstellt werden. Einige konkrete Beispiele für den Einsatz von regelbasierten Chatbots sind die Beantwortung häufig gestellter Kundenfragen, Unterstützung zur Vereinbarung von Terminen oder zur Lead-Generierung (Adressgewinnung).

Ein Beispiel aus dem Rechtsbereich sind Vertragsgeneratoren. Der Nutzer wird vom Chatbot zwar „gefragt“, welche Klauseln er verwenden möchte, jedoch arbeitet das Programm im Hintergrund einen vorab festgelegten Entscheidungsbaum ab. Regelbasierte Chatbots sind nur so schlau, wie sie programmiert wurden. Sollen sie schlauer werden, muss neu- oder umprogrammiert werden. Ähnliches galt bislang für Expertensysteme im Rechts- und Steuerbereich. Sie basierten ebenfalls auf vordefinierten Regeln und sollten komplexe rechtliche oder steuerliche Probleme lösen können. Diese Systeme wurden entwickelt, um auf Basis von programmierten Regeln und Wissen Entscheidungen zu treffen oder Empfehlungen auszusprechen. Die Funktionsweise dieser Expertensysteme beruhte darauf, dass sie das vorhandene Wissen in Form von Regeln und Logik abbildeten. Durch die Eingabe von Informationen konnten sie Schlussfolgerungen ziehen und Lösungen vorschlagen.

Generische Chatbots

Eine andere Gruppe von Chatbots basiert auf dem Ansatz der KI-Sprachmodelle (Large Language Models, LLMs). Diese Modelle funktionieren nach dem Prinzip des maschinellen Lernens, insbesondere des Deep Learnings, und können aus Erfahrungen lernen, um menschenähnliche Gespräche zu führen. Im Gegensatz zu regelbasierten Chatbots, die vordefinierte Regeln verwenden, sind LLMs flexibler und können auf vielfältige Benutzeranfragen reagieren. Die Funktionsweise von LLMs beruht auf umfangreichen Trainingsdaten

Regelbasierte
Chatbots sind
deterministisch

Generische
Chatbots sind (noch)
„stochastische
Papageien“

und Modellen, die es ihnen ermöglichen, kontextbezogene Antworten zu generieren. Diese Modelle wie GPT-4 nutzen Technologien wie Natural Language Processing (NLP) und können aus einer Wissensdatenbank kreative Antworten ableiten. Sie sind in der Lage, den Chatverlauf zu berücksichtigen und Folgefragen angemessen zu beantworten. Im Vergleich zu regelbasierten Chatbots sind LLMs in der Lage, komplexere Anfragen zu bearbeiten und menschenähnlichere Konversationen zu führen.

■ Large Language Models (LLMs)

ChatGPT ist wohl das bekannteste große Sprachmodell, das auf künstlicher Intelligenz basiert. Entwickelt wurde ChatGPT vom US-amerikanischen Unternehmen OpenAI. ChatGPT steht seit dem 30.11.22 der Öffentlichkeit kostenlos zur Verfügung. Eine einfache Registrierung auf der Website reicht aus. Daneben gibt es eine leistungsfähigere Version, die für 20 EUR im Monat genutzt werden kann. Schnell zogen andere Anbieter nach und brachten eine ganze Reihe von weiteren Sprachmodellen an den Markt. Einige bemerkenswerte LLMs neben OpenAIs GPT-Serie von Modellen (z. B. GPT-3.5 und GPT-4, die in ChatGPT und Microsoft Copilot verwendet werden) sind Googles PaLM und Gemini (vormals BARD), Metas LLaMA-Familie von Open-Source-Modellen, Anthropics Claude-Modelle und die Open-Source-Modelle von Mistral AI. Daneben gibt es Modelle zur Bild-, Video- und Musikerstellung, die auf einer ähnlichen Architektur aufbauen.

Halluzinationen bei generischen Chatbots

Eine spannende Frage ist, ob diese Art von KI darauf beschränkt ist, nur statistische Korrelationen aus den Trainingsdaten zu reproduzieren oder ob sie den Daten, die sie konsumiert, Bedeutung entnimmt und in den Daten, die sie generiert, Bedeutung weitergibt. Vor diesem Hintergrund wurde der Begriff „stochastischer Papagei“ von der Linguistin Emily Bender geprägt, um auf die Funktionsweise von LLMs hinzuweisen. Diese Bezeichnung bezieht sich darauf, dass LLMs zwar in der Lage sind, menschenähnliche Texte zu generieren, jedoch den eigentlichen Sinn der Sprache nicht verstehen. Sie neigen dazu, Inhalte aus ihren Trainingsdaten zu wiederholen und zufällige Elemente hinzuzufügen, ohne den tieferen Sinn zu erfassen. Das Adjektiv „stochastisch“ bedeutet hier „zufällig“ oder „mit Bezug auf Chance oder Wahrscheinlichkeit“, was darauf hinweist, dass die Ausgaben dieser Modelle auf Wahrscheinlichkeiten basieren und nicht unbedingt einen tieferen Sinn widerspiegeln. Obwohl LLMs wie GPT-4 beeindruckende Texte generieren können, bleibt ihre Fähigkeit zur tiefgreifenden Bedeutungsinterpretation begrenzt. Sie reproduzieren hauptsächlich das Gelernte aus ihren Trainingsdaten und fügen zufällige Elemente hinzu, was zuweilen zu unvorhersehbaren oder repetitiven Antworten führen kann.

Halluzinationen entstehen also, wenn die Modelle aufgrund ihrer Funktionsweise nicht sinnvolle bzw. faktisch nicht fundierte Ausgaben generieren. Diese Erfahrung mussten bereits Anwälte in den USA machen, die ihre Schriftsätze von einem Sprachmodell erstellen ließen, dass „praktischerweise“ Präzedenzentscheidung im Sinne der Berater mitsamt den Fundstellen erfand. Neben der Blamage waren auch Strafen fällig.

Eine Erklärung für
Halluzinationen

Begriffsbestimmung:
Halluzinationen

Ansätze gegen Halluzinationen

Es gibt verschiedene Möglichkeiten für den Anwender, die Wahrscheinlichkeit des Auftretens von Halluzinationen zumindest zu verringern. In der Bezahl-Variante von ChatGPT ist es z. B. möglich, dem Sprachmodell Quellen (PDFs, URLs) mitzugeben, damit es für seine Antworten nur Informationen aus diesen Quellen berücksichtigt. So kann – bis zu einem gewissen Grad – das Halluzinieren unterdrückt werden. Kürze Quellen können auch über den Prompt (z. B. aus Pressemitteilung, Berichterstattung) mitgegeben werden. Daneben existieren noch weitere Maßnahmen, wie die Anpassung des Temperaturparameters (das Maß für die „Kreativität“ der Texterstellung) oder bestimmte Prompting-Techniken.

Eigene Wissensbasis mitgeben (PDF, URL, Text im Prompt)

■ Beispiel für die Mitgabe einer Wissensbasis im Prompt

■ Prompt:

Du bist Steuerberater und willst gegen einen Bescheid des Finanzamts Einspruch einlegen, der gegen deinen Mandanten ergangen ist.

Es geht um folgenden Sachverhalt: Die Finanzverwaltung verlangt im Rahmen einer Außenprüfung vom Mandanten die Vorlage des elektronischen Gesamtjournals. Du beziehst Dich in deinem Einspruch auf eine Entscheidung des FG Hamburg, die dem BFH zur Revision vorliegt: Nach einem Urteil des FG Hamburg (23.3.23, 2 K 172/19; Rev. BFH XI R 15/23) hat die Finanzverwaltung keinen Anspruch auf Vorlage eines elektronischen Gesamtjournals, welches nach den Vorgaben der Finanzverwaltung Informationen zu jeder einzelnen empfangenen bzw. versandten E-Mail des Steuerpflichtigen enthalten soll. Die Aufforderung zur Vorlage eines Gesamtjournals, in dem auch nicht nach § 147 Abs. 1 AO aufbewahrungspflichtige E-Mails aufgelistet bzw. nach den Vorgaben der Finanzverwaltung dargestellt werden sollen, überschreitet danach die Befugnisse der Finanzverwaltung aus § 147 Abs. 6 AO und ist damit rechtswidrig.

Formuliere nach diesen Angaben einen Einspruch und beantrage das Ruhen des Verfahrens.

■ Ausgabe (ChatGPT 4, inhaltlich nicht bearbeitet):

Sehr geehrte Damen und Herren,
hiermit lege ich im Namen meines Mandanten Einspruch gegen den Bescheid des Finanzamts vom [Datum des Bescheids] ein, mit dem die Vorlage des elektronischen Gesamtjournals im Rahmen einer Außenprüfung angefordert wird.

Der Einspruch begründet sich darauf, dass nach der Entscheidung des Finanzgerichts Hamburg vom 23. März 2023 (Aktenzeichen 2 K 172/19, Revision beim Bundesfinanzhof unter dem Aktenzeichen XI R 15/23) die Aufforderung zur Vorlage eines elektronischen Gesamtjournals, das nach den Vorgaben der Finanzverwaltung Informationen zu jeder einzelnen empfangenen bzw. versandten E-Mail des Steuerpflichtigen enthalten soll, die Befugnisse der Finanzverwaltung aus § 147 Abs. 6 der Abgabenordnung (AO) überschreitet. Das Finanzgericht Hamburg hat in seiner Entscheidung festgestellt, dass die Aufforderung zur Vorlage eines Gesamtjournals, in dem auch nicht nach § 147 Abs. 1 AO aufbewahrungspflichtige E-Mails aufgelistet bzw. nach den Vorgaben der Finanzverwaltung dargestellt werden sollen, rechtswidrig ist.

Auf Grundlage dieser Entscheidung, die derzeit zur Revision beim Bundesfinanzhof anhängig ist, argumentieren wir, dass die Aufforderung zur Vorlage des elektronischen Gesamtjournals in der vom Finanzamt geforderten Form keinen Bestand haben kann und somit unserem Einspruch stattgegeben werden sollte.

Darüber hinaus beantrage ich hiermit das Ruhen des Verfahrens bis zur endgültigen Entscheidung des Bundesfinanzhofs in der Rechtssache XI R 15/23, um eine einheitliche Rechtsanwendung zu gewährleisten und meinem Mandanten unnötige Belastungen zu ersparen.

Wir bitten um schriftliche Bestätigung des Eingangs dieses Einspruchs und des Antrags auf Ruhen des Verfahrens sowie um Mitteilung über die weiteren Schritte. Für Rückfragen stehe ich Ihnen selbstverständlich zur Verfügung.

Mit freundlichen Grüßen,

[Ihr Name]

Steuerberater

[Adresse]

[Telefonnummer]

[E-Mail-Adresse]

[Kanzlei-Name]

Sprachmodelle für steuerfachliche Aufgaben

Natürlich wäre es wünschenswert, wenn man ein Sprachmodell wie einen steuerlichen Assistenten einsetzen und ihm komplette Aufgaben übertragen könnte. Anwendungsbereiche in der Kanzlei gäbe es genug: Beantwortung einfacher Mandantenanfragen, fachliche Unterstützung bei Anschreiben an Mandanten, Analyse von Mandantendaten etc. Allerdings sind die Erfahrungen mit den allgemeinen Sprachmodellen eher ernüchternd. Sie liefern keine stabil guten Ergebnisse. Wer sie nutzt, muss auf jeden Fall genau gegenlesen. Einerseits werden diese Modelle nicht speziell an deutschen steuerlichen Inhalten geschult, andererseits wird die Datenbasis nur in größeren Abständen aktualisiert. Und auch wenn sich das Sprachmodell die Daten für die Antwort aus dem Internet zieht (z. B. ChatGPT in Verbindung mit der Bing-Suchmaschine), dann ist keineswegs sichergestellt, dass nur zuverlässige Quellen ausgewertet werden (Derlath/Ozimek, KP 24, 28).

Ein weiteres Problem ist die mögliche Verletzung von Berufsgeheimnissen, wenn echte Mandantendaten eingegeben werden. Denn natürlich dient alles, was eingegeben wird, als Trainingsdaten und lässt sich mit den richtigen Prompts später sogar teilweise wieder rekonstruieren.

Aber: Die großen Beratungsgesellschaften geben zurzeit erhebliche Summen aus, um ihre eigenen Geschäftsmodelle mit KI weiterzuentwickeln. Die Wirtschaftsprüfungsgesellschaften verfolgen mit ihren Investitionen in künstliche Intelligenz Ziele wie Effizienzsteigerung, Qualitätsverbesserung der Prüfungen, Nutzung von automatisierten Prozessen zur Datenanalyse und Interpretation sowie Anpassung an die digitale Transformation, um den sich wandelnden Anforderungen gerecht zu werden.

Nach wie vor:
Hoher Kontroll-
aufwand

Verstoß gegen das
Verschwiegenheits-
gebot

Die WP-Gesellschaf-
ten investieren
erhebliche Summen

Sprachmodelle für nicht-steuerfachliche Aufgaben

Besser sieht es im nicht-steuerfachlichen Bereich aus. Hier können KI-Sprachmodelle bereits in vielen Situationen eingesetzt werden, um Bestehendes zu überarbeiten oder um kreative Prozesse zu unterstützen. Aber auch hier gilt: Kritisch gegenlesen!

Texte überarbeiten, abändern

Für diese Anwendung sind mehrere Szenarien denkbar:

- Ein Anschreiben, eine E-Mail sprachlich verbessern (treffendere Formulierungen, zu lange Sätze kürzen, Wortwiederholungen vermeiden etc.) Hierfür bietet sich z. B. der spezialisierte KI-Schreibassistent von DeepL aus Köln an (<https://www.deepl.com/write/write-mobile>).
- Längere Texte gliedern und mit (Zwischen-)Überschriften versehen.
- Manchmal kann es erforderlich sein, Texte abzuwandeln, weil man sie z. B. gegenüber einer Laienzielgruppe verwenden möchte. In diesem Fall ist es möglich die sprachliche Textschwierigkeit durch Zusätze wie „Benutze im Text umgangssprachliche Wörter.“ oder „Erkläre es einem fünfjährigen Kind.“ zu senken.

Neue Texte generieren

- Einen Text für die Kanzleiwebseite als Kurznachricht erstellen.
- Einen Social-Media-Post mit Smileys oder Hashtags aus einem bestehenden Text ableiten.
- Einen Text aus einer Fremdsprache oder in eine Fremdsprache übersetzen (z. B. mit www.deepl.com).
- Vorlagen für Standardbriefe, Checklisten etc. erzeugen.

■ Beispiel: Text für Kanzleiwebseite/Mandantenanschreiben

■ Prompt:

Du bist Steuerberater und hast eine eigene Unterseite auf deiner Kanzlei-Homepage, auf der du deine Mandanten informierst. Für diese Informationsseite benötigst Du einen Text zur Umstellung auf die E-Rechnung, die für 2025 geplant ist. Gehe in dem Text auf die Gründe ein, warum die E-Rechnung in Deutschland eingeführt werden soll, wer künftig E-Rechnungen ausstellen muss und wie der Zeitplan für die Einführung aussieht. Formuliere den Text in leicht verständlicher Sprache. Nutzen für Deine Antwort aktuelle Internet-Seiten.

■ Antwort (ChatGPT 4, inhaltlich nicht bearbeitet):

Liebe Mandantinnen und Mandanten,

wir stehen vor einer wichtigen Veränderung in der Rechnungsstellung, die ab dem Jahr 2025 unser aller Arbeitsalltag beeinflussen wird: die Umstellung auf die E-Rechnung. Diese Neuerung ist ein wesentlicher Schritt in Richtung Digitalisierung und Effizienzsteigerung im Geschäftsverkehr und bringt für alle Unternehmen Veränderungen mit sich.

Sprachlich
verbessern

Zusammenfassen,
Gliedern

Warum wird die E-Rechnung eingeführt?

Die Einführung der E-Rechnung in Deutschland folgt dem Ziel, den Umsatzsteuerbetrug zu bekämpfen und die sogenannte Mehrwertsteuerlücke, die aktuell mehrere Milliarden Euro beträgt, signifikant zu reduzieren. Durch die Nutzung von E-Rechnungen, also elektronisch übermittelten Rechnungen in strukturierten Datenformaten wie XRechnung oder ZUGFeRD, soll die Verarbeitung vereinfacht und automatisiert werden. Dies führt nicht nur zu schnelleren Prozessen und geringeren Kosten, sondern auch zu mehr Sicherheit und Effizienz in der Eingangsrechnungsverarbeitung [3 Quellen].

Wer muss E-Rechnungen ausstellen?

Ab 2025 betrifft die Pflicht zur Ausstellung von E-Rechnungen alle inländischen B2B-Umsätze zwischen in Deutschland ansässigen Unternehmen. Das bedeutet, dass ab diesem Zeitpunkt alle Unternehmen elektronische Rechnungen empfangen können müssen. Ein Jahr später, also ab 2026, müssen sie zudem in der Lage sein, selbst E-Rechnungen zu erstellen und zu versenden [2 Quellen].

Zeitplan und Übergangsfristen

- **Ab 1. Januar 2025:** Empfang von E-Rechnungen muss möglich sein.
- **Ab 1. Januar 2026:** Unternehmen müssen selbst E-Rechnungen ausstellen und versenden können.

Für kleine und mittlere Unternehmen (KMU) könnten vorübergehende Erleichterungen eingeräumt werden, um die Umstellung zu erleichtern. Diese Erleichterungen sind jedoch noch nicht endgültig spezifiziert [1 Quelle].

Was ist eine E-Rechnung genau?

Eine E-Rechnung ist nicht einfach ein PDF-Dokument, das per E-Mail versandt wird, sondern ein strukturierter Datensatz, der elektronisch übermittelt wird und eine automatische Verarbeitung ermöglicht. In Deutschland sind hierfür Formate wie XRechnung und ZUGFeRD entwickelt worden, die einen standardisierten Aufbau besitzen und sowohl für den Sender als auch für den Empfänger erhebliche Vorteile bieten [1 Quelle].

Empfehlungen für die Umstellung

Es ist wichtig, dass Sie sich frühzeitig mit der Thematik auseinandersetzen und Ihre internen Prozesse entsprechend anpassen. Überlegen Sie, ob Sie die Umstellung mit eigenen Mitteln bewerkstelligen können oder ob der Einsatz von externen Dienstleistern sinnvoller ist. Die kontinuierliche Anpassung und Verbesserung Ihrer Prozesse wird entscheidend sein, um auf dem Laufenden zu bleiben und die Einhaltung der gesetzlichen Anforderungen sicherzustellen [1 Quelle].

Wir empfehlen Ihnen, die kommenden Monate zu nutzen, um sich auf diese Veränderung vorzubereiten. Sollten Sie Unterstützung benötigen oder weitere Fragen haben, stehen wir Ihnen gerne zur Verfügung.

Mit besten Grüßen
Ihr Steuerberatungsteam

Quellen am Ende
der Absätze prüfen!

KI als Sparringspartner für Texte, Grafiken, Bilder

Gerade wer nicht regelmäßig schreibt, für den kann ein leeres Blatt ein echtes Hemmnis darstellen. Hier können Sprachmodelle weiterhelfen, indem sie einen ersten Wurf erzeugen, der dann zwar noch umgearbeitet werden muss, aber wenigstens gibt es dafür jetzt eine Vorlage. Zu denken wäre hier etwa an den Gliederungsentwurf für einen Vortrag (z. B. im Rahmen der internen Weiterbildung oder für einen Mandantenabend) und Entwürfe für die einzelnen Folieninhalte auf der Grundlage des Gliederungsentwurfs oder an eine Stellenbeschreibung als Grundlage für eine Stellenanzeige.

Sinngemäß gilt das auch für Bilder und Grafiken. Auch hier können die Anwendenden in natürlicher Sprache beschreiben, was auf dem Bild oder in der Grafik zu sehen sein soll. Allerdings ist das Prompten von Bildern schon ziemlich anspruchsvoll, wenn man selbst bereits eine relativ genaue Vorstellung davon hat, was man will.

Bekämpfung des „Leeres-Blatt-Syndroms“

■ Beispiel: Ein einfacher Prompt für eine Tortengrafik

■ Prompt:

Fertige ein Tortendiagramm an. Verwende weit auseinander liegende Grautöne, damit sich die Segmente der Torte gut für das Auge unterscheiden. Der Titel des Tortendiagramms lautet: Kanzleiumsätze mit A-, B- und C-Mandaten. Ergänze das Diagramm um eine Legende. Schreibe die prozentualen Anteile in das jeweilige Tortensegment. Hier sind die kommaseparierten Daten für das Diagramm:

Mandat, EUR

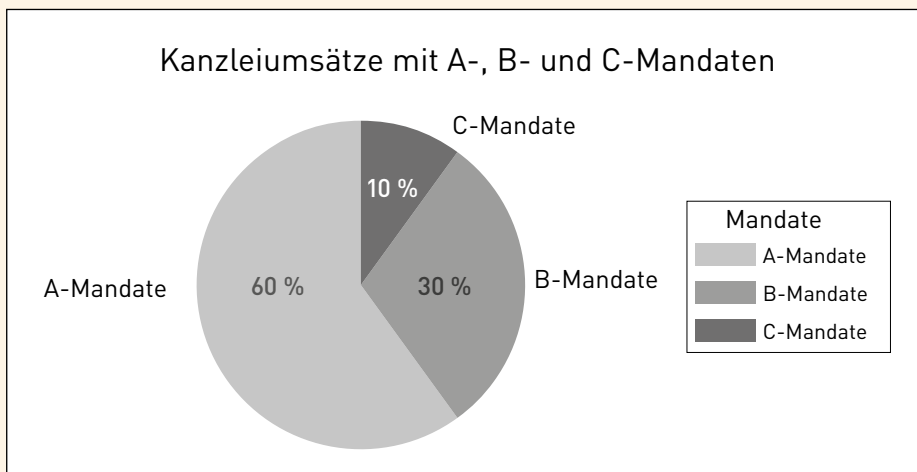
A, 600.000

B, 300.000

C, 100.000

■ Antwort (ChatGPT 4, inhaltlich nicht bearbeitet):

(Der Prompt wird von ChatGPT 4 in Python-Code übersetzt und ausgeführt.)



↘ WEITERFÜHRENDE HINWEISE

- Geschäftsmodell Steuerberatung – Wie wird sich KI auf die Steuerberatung auswirken? Versuch einer ersten Einschätzung (Hausmann, KP 24, 48)
- Chatten statt googeln? Websuche mit den Sprachmodellen nach steuerlichen Inhalten – ein kleiner Praxistest (Derlath/Ozimek, KP 24, 28)
- Digitalisierung/Maschinenlernen – Wie ChatGPT-4 schon jetzt den Alltag in der Steuerberatung unterstützen kann (Ozimek, KP 23, 113)