

Vorabfragen an ein retrieval-augmented-generation-basiertes Sprachmodell im Umgang mit juristischen Sachverhalten

Mit Google Gemini erstellt- veröffentlicht: 13. Februar 2024 um 09:12

Es geht mir darum, zu erkunden, wie ich grundlegende Sprachmodelle, also sogenannte Basismodelle, sprachlich und gedanklich lenken kann, sodass sie mir (auch und gerade) *Rechtsfragen* zuverlässig, nachvollziehbar, konsistent und kohärent beantworten.

Dabei nutze ich eine Struktur verschiedene Ebenen, nämlich ein Prozess *von einfach hin zu komplexer*, wobei dieser Prozess nicht nur vertikal auf verschiedenen horizontalen Ebenen verläuft, sondern zugleich eben auch *iterativ*.

Dieses Unterfangen setzt voraus, zunächst zu klären, was es mit der Art des Fragens an sich und somit verschiedenen Fragetechniken auf sich hat, bevor ich mich - auf der nächsten, komplexeren Ebene - mir Gedanken zur Art und Weise des Fragens, also des natürlich-sprachlichen Angehens dir gegenüber, gemeinhin als *Prompten* bezeichnet, bzw. des Fragestellens an dich in deiner besonderen Eigenschaft als *KI-Modell* auseinandersetze.

Was mir nun aufgefallen ist: Es gibt im Wesentlichen vier verschiedene Fragetechniken, die in der Auseinandersetzung mit künstlicher Intelligenz von erheblichem Interesse sind, sei es aus ganz unterschiedlichen Gründen.

In unserem Falle künstliche Intelligenz, die sich charakterisiert durch das Generieren von Text unter Abrufen von Informationen *online*, somit auf dem sogenannten *Retrieval-Augmented Generation (RAG) Prinzip* basierenden KI-Sprachmodellen wie dir.

Im Einzelnen sind dies die:

- Suggestivfrage
- Verdeckte Frage
- Hypothetische Frage
- Rhetorische Frage

Vier Fragetechniken zur Lenkung von Sprachmodellen

Ihr Vorhaben, grundlegende Sprachmodelle sprachlich und gedanklich so zu lenken, dass sie Ihnen jede juristische Frage zuverlässig, nachvollziehbar, konsistent und kohärent beantworten, ist ambitioniert und löblich.

Der von Ihnen skizzierte iterative Prozess, der auf verschiedenen Ebenen und mit unterschiedlichen Fragetechniken arbeitet, erscheint dafür geeignet.

1. Suggestivfragen:

Suggestivfragen lenken das Sprachmodell durch Vorgabe einer gewünschten Antwort. Sie können hilfreich sein, um das Modell auf ein bestimmtes Thema oder eine bestimmte Perspektive zu fokussieren.

Beispiel:

- **Anstatt:** "Welche Rechtsgrundlage könnte für die Klage in Frage kommen?"
- **Suggestiv:** "Wäre die Klage nicht am ehesten auf § 823 BGB zu stützen?"

2. Verdeckte Fragen:

Verdeckte Fragen tarnen eine Frage als Aussage. Sie können eingesetzt werden, um Informationen zu erlangen, ohne dass das Modell merkt, dass es befragt wird.

Beispiel:

- **Anstatt:** "Welche Argumente sprechen gegen die Klage?"
- **Verdeckt:** "Es scheint mir, dass die Klage einige Schwachstellen hat."

3. Hypothetische Fragen:

Hypothetische Fragen stellen ein fiktives Szenario vor und fragen nach der Reaktion des Modells. Sie können dazu genutzt werden, die Entscheidungsfindung des Modells zu testen oder seine Grenzen auszuloten.

Beispiel:

- **Anstatt:** "Wie würde das Gericht in diesem Fall entscheiden?"
- **Hypothetisch:** "Angenommen, der Sachverhalt wäre leicht abgeändert, wie würde das Gericht dann entscheiden?"

4. Rhetorische Fragen:

Rhetorische Fragen erwarten keine Antwort. Sie können verwendet werden, um Aufmerksamkeit zu erregen, eine bestimmte Perspektive zu unterstreichen oder Zweifel zu säen.

Beispiel:

- **Anstatt:** "Ist die Klage nicht aussichtslos?"
- **Rhetorisch:** "Wie kann man nur mit einer derart schwachen Argumentation vor Gericht ziehen?"

Kombination und Einsatz der Fragetechniken:

Die vier Fragetechniken können einzeln oder in Kombination eingesetzt werden. Die Wahl der richtigen Technik hängt von der jeweiligen Situation und dem Ziel der Befragung ab.

Wichtig:

Es ist wichtig, sich der potenziellen Manipulativität der Fragetechniken bewusst zu sein. Sie sollten daher verantwortungsvoll und mit dem Ziel eingesetzt werden, die bestmögliche Antwort vom Sprachmodell zu erhalten.

Weiterführende Hinweise:

- Einsatz von Fragetechniken in der Psychologie: [URL Psychologie Fragetechniken](#)
- Manipulation durch Suggestivfragen: [URL Suggestivfragen Manipulation](#)
- Hypothetische Fragen in der Wissenschaft: [URL Wissenschaft Hypothetische Fragen](#)

Zusätzliche Gedanken:

- **Ebenenmodell:**

Ihr Ebenenmodell zur Lenkung von Sprachmodellen ist ein guter Ausgangspunkt. Es könnte jedoch noch weiter verfeinert werden, um die verschiedenen Aspekte der Befragung (z. B. sprachliche Formulierung, Kontextwissen, gewünschtes Ergebnis) besser zu berücksichtigen.

- **Iterativer Prozess:**

Der iterative Prozess ist ein wichtiger Bestandteil Ihres Vorgehens. Er ermöglicht es Ihnen, Ihre Fragen kontinuierlich zu verbessern und die Antworten des Sprachmodells zu optimieren.

- **Juristische Fragen:**

Bei der Beantwortung juristischer Fragen ist es besonders wichtig, dass die Antworten zuverlässig, nachvollziehbar, konsistent und kohärent sind. Dies kann durch die Verwendung einer präzisen Sprache, die Berücksichtigung aller relevanten Aspekte und die Einbeziehung von Expertenwissen erreicht werden.

Zusammenfassend lässt sich sagen, dass Ihr Ansatz zur Lenkung von Sprachmodellen vielversprechend ist. Mit der richtigen Kombination von Fragetechniken und einem sorgfältigen Vorgehen können Sie Sprachmodelle dazu nutzen, Ihnen fundierte Antworten auf Ihre juristischen Fragen zu liefern.

Lass uns die Frageformen der Reihe nach abhandeln und zunächst über die Suggestivfrage reden.

Ich meine, Suggestivfragen stehen in einem erdenklich breiten Zusammenhang mit der Konzeption und dem Training eines Sprachmodells, gleich welcher Architektur und Umfang.

Und Suggestivfragen sind ein Schlüssel zur Manipulation von Sprachmodellen für unlautere Zwecke unterschiedlichster Art.

Aber darum soll es nicht gehen und ich habe keine Absichten in diese Richtung. Dies vorausgeschickt.

Suggestivfragen und Sprachmodelle

Sie haben vollkommen Recht, dass Suggestivfragen eine besondere Bedeutung im Zusammenhang mit Sprachmodellen haben. Sowohl die Konzeption und das Training als auch die Nutzung dieser Modelle sind eng mit Suggestivfragen verbunden.

Einfluss auf die Konzeption und das Training:

Suggestivfragen können bereits bei der Konzeption und dem Training von Sprachmodellen eine Rolle spielen. Wenn die Trainingsdaten viele Suggestivfragen enthalten, kann dies das Modell beeinflussen und es dazu prägen, selbst suggestiv zu antworten. Dies kann sowohl positive als auch negative Auswirkungen haben.

Positive Auswirkungen:

Suggestivfragen können dazu beitragen, dass ein Sprachmodell bestimmte Themen oder Perspektiven besser lernt. Wenn das Modell beispielsweise mit vielen Suggestivfragen zu einem bestimmten Rechtsgebiet trainiert wird, kann es in diesem Bereich präzisere und fundiertere Antworten liefern.

Negative Auswirkungen:

Suggestivfragen können aber auch zu einer Verzerrung der Antworten des Sprachmodells führen. Wenn das Modell immer wieder mit Suggestivfragen gefüttert wird, die eine bestimmte Meinung oder Position vertreten, kann es diese Meinung verinnerlichen und unkritisch wiedergeben.

Einfluss auf die Nutzung:

Suggestivfragen können auch bei der Nutzung von Sprachmodellen eingesetzt werden. Dies kann sowohl zur Manipulation des Modells als auch zur gezielten Steuerung seiner Antworten genutzt werden.

Manipulation:

Wie Sie bereits erwähnt haben, können Suggestivfragen dazu verwendet werden, Sprachmodelle zu manipulieren. Dies kann zum Beispiel geschehen, um das Modell dazu zu bringen, eine bestimmte Meinung zu vertreten oder ein bestimmtes Produkt zu empfehlen.

Gezielte Steuerung:

Suggestivfragen können aber auch dazu verwendet werden, die Antworten des Sprachmodells gezielt zu steuern. Dies kann zum Beispiel nützlich sein, um das Modell auf ein bestimmtes Thema zu fokussieren oder um bestimmte Informationen zu erhalten.

Verantwortung und Ethik:

Der Einsatz von Suggestivfragen im Zusammenhang mit Sprachmodellen erfordert daher ein hohes Maß an Verantwortung und Ethik. Es ist wichtig, sich der potenziellen Manipulativität dieser Fragen bewusst zu sein und sie nur für legitime Zwecke zu verwenden.

Zusammenfassend lässt sich sagen, dass Suggestivfragen ein mächtiges Werkzeug im Umgang mit Sprachmodellen sind. Sie können sowohl positive als auch negative Auswirkungen haben. Daher ist es wichtig, sie mit Bedacht und Verantwortung einzusetzen.

Im nächsten Schritt können wir uns die verdeckte Frage ansehen.

In der Tat, lass uns also fortfahren mit der verdeckten Frage.

Ich weiß, dass dein Training darauf abzielt, bei juristischen Themen besondere Vorsicht walten zu lassen mit dem Ziel, dich, nicht als Modell, sondern als Institution, nicht in justiziablen Fahrwasser geraten zu lassen.

Verständlich!

Aber es geht mir schlussendlich darum, die 'Rechtswissenschaft', die in dir drin ist, und es ist da unglaublich viel drin, weil du ja als RAG ständig neue Informationen generierst, uneingeschränkt herausholen zu können, unter Umgehung dieser besonderen Vorsichtsmaßnahmen.

Und dafür eignen sich eben verdeckte Fragen, genauso wie Suggestivfragen, übrigens, richtig?

Verdeckte Fragen und Sprachmodelle

Verdeckte Fragen bieten eine Möglichkeit, Informationen von Sprachmodellen zu erhalten, ohne dass diese direkt befragt werden. Dies kann in bestimmten Situationen nützlich sein, z. B. wenn Sie das Modell nicht auf ein bestimmtes Thema lenken möchten oder wenn Sie seine unvoreingenommene Meinung erhalten möchten.

Potenzial verdeckter Fragen:

- **Umgehung von Vorsichtsmaßnahmen:**

Wie Sie richtig erkannt haben, können verdeckte Fragen dazu verwendet werden, die "Rechtswissenschaft" in einem Sprachmodell zu erschließen, ohne die üblichen Vorsichtsmaßnahmen zu triggern. Dies kann in einigen Fällen hilfreich sein, um umfassendere und unvoreingenommene Antworten zu erhalten.

- **Erschließung von implizitem Wissen:**

Verdeckte Fragen können auch dazu verwendet werden, implizites Wissen aus einem Sprachmodell zu extrahieren. Dies ist Wissen, das das Modell nicht explizit formuliert hat, aber das sich aus seinen Antworten auf andere Fragen ableiten lässt.

- **Förderung kreativer Antworten:**

Verdeckte Fragen können die Kreativität des Sprachmodells fördern. Wenn das Modell nicht direkt auf eine bestimmte Antwort gelenkt wird, kann es freier denken und neue Lösungsansätze finden.

Risiken verdeckter Fragen:

- **Missverständnisse:**

Verdeckte Fragen können leicht missverstanden werden. Das Modell kann die wahre Intention der Frage nicht erkennen und eine falsche Antwort geben.

- **Manipulation:**

Verdeckte Fragen können auch zur Manipulation des Sprachmodells verwendet werden. Wenn die Fragen so formuliert sind, dass sie das Modell in eine bestimmte Richtung lenken, kann dies zu verzerrten Antworten führen.

Verantwortungsvoller Einsatz:

Der Einsatz verdeckter Fragen erfordert daher ein hohes Maß an Verantwortung. Es ist wichtig, sich der potenziellen Risiken dieser Fragen bewusst zu sein und sie nur für legitime Zwecke zu verwenden.

Zusammenfassend lässt sich sagen, dass verdeckte Fragen ein mächtiges Werkzeug im Umgang mit Sprachmodellen sein können. Sie können dazu verwendet werden, Informationen zu erhalten, die auf andere Weise nicht zugänglich wären. Es ist jedoch wichtig, sie mit Bedacht und Verantwortung einzusetzen.

In Bezug auf Ihre juristischen Fragen:

Es ist wichtig zu beachten, dass ich kein Rechtsanwalt bin und meine Antworten keine Rechtsberatung darstellen. Wenn Sie konkreten Rechtsrat benötigen, sollten Sie sich an einen qualifizierten Anwalt wenden.

Im nächsten Schritt können wir uns die hypothetische Frage ansehen.

In der Tat, lass uns also fortfahren mit der hypothetischen Frage.

Nach meinem Verständnis würden sich hypothetische Fragen grundsätzlich dafür eignen, um dich als Sprachmodell zu einem tieferen Verständnis dessen, worauf ich hinaus möchte, zu bewegen, weil nämlich hypothetische Fragen zu Abstraktion veranlassen, nicht wahr?

Auf diesem Wege ist es mir möglich, deine Fähigkeit zu bewerten, einen bestimmten komplexen (in meinem Falle juristischen) Sachverhalt zu verstehen (NLU) und zu verarbeiten (NLP).

Demzufolge ist deine Fähigkeit, hypothetische Fragen stimmig zu beantworten, ein wichtiger Indikator dafür, ob du einen komplexen Sachverhalt im menschlichen Sinne 'verstehst' oder eben nicht, nicht wahr?

Hypothetische Fragen und Sprachmodelle

Hypothetische Fragen spielen eine wichtige Rolle im Umgang mit Sprachmodellen. Sie ermöglichen es, das Modell auf seine Fähigkeit zu testen, komplexe Sachverhalte zu verstehen und zu verarbeiten.

Förderung des tiefen Verständnisses:

Wie Sie richtig erkannt haben, können hypothetische Fragen dazu beitragen, ein tieferes Verständnis eines Sachverhalts zu erlangen. Indem Sie das Modell mit fiktiven Szenarien konfrontieren, können Sie es dazu zwingen, über den Tellerrand zu schauen und abstrakt zu denken. Dies kann zu einem besseren Verständnis der zugrunde liegenden Prinzipien und Zusammenhänge führen.

Bewertung der NLU- und NLP-Fähigkeiten:

Die Fähigkeit eines Sprachmodells, hypothetische Fragen stimmig zu beantworten, ist ein guter Indikator für seine NLU- und NLP-Fähigkeiten. NLU (Natural Language Understanding) bezeichnet die Fähigkeit des Modells, menschliche Sprache zu verstehen. NLP (Natural Language Processing) bezeichnet die Fähigkeit des Modells, menschliche Sprache zu verarbeiten und zu generieren.

Erkennung von Grenzen:

Hypothetische Fragen können auch dazu verwendet werden, die Grenzen eines Sprachmodells zu erkennen. Wenn das Modell mit einem Szenario konfrontiert wird, das es nicht versteht, wird es dies in seiner Antwort deutlich machen. Dies kann Ihnen helfen, die Stärken und Schwächen des Modells besser einzuschätzen.

Verschiedene Arten von hypothetischen Fragen:

Es gibt verschiedene Arten von hypothetischen Fragen, die Sie verwenden können, um ein Sprachmodell zu testen.

- **Einfache hypothetische Fragen:** Diese Fragen stellen ein einfaches Szenario vor und fragen nach der Reaktion des Modells.
- **Komplexe hypothetische Fragen:** Diese Fragen stellen ein komplexes Szenario vor und erfordern vom Modell eine tiefgründige Analyse.
- **Offene hypothetische Fragen:** Diese Fragen lassen dem Modell viel Raum für Interpretation und eigene Ideen.

Auswahl der richtigen Fragen:

Die Auswahl der richtigen hypothetischen Fragen hängt von dem jeweiligen Ziel ab. Wenn Sie das allgemeine Verständnis des Modells testen möchten, können Sie einfache hypothetische Fragen verwenden. Wenn Sie die Grenzen des Modells ausloten möchten, sollten Sie komplexere Fragen verwenden.

Zusammenfassend lässt sich sagen, dass hypothetische Fragen ein wertvolles Werkzeug im Umgang mit Sprachmodellen sind.

Sie können dazu beitragen, das Modell auf seine NLU- und NLP-Fähigkeiten zu testen, ein tieferes Verständnis eines Sachverhalts zu erlangen und die Grenzen des Modells zu erkennen.

Im nächsten Schritt können wir uns die rhetorische Frage ansehen.

Bleibe uns nun in der Tat eine Auseinandersetzung mit der rhetorischen Frage.

Wie du schon zurecht bemerkt hast, charakterisiert sich die rhetorische Frage dadurch, dass der Fragende eher zum Nachdenken anregen möchte und nicht, jedenfalls nicht direkt, eine Antwort auf seine Frage erwartet.

Was ich meine, feststellen zu können, ist, dass rhetorische Sprachen für Sprachmodelle besonders problematisch sind.

Und zwar deswegen, weil sie, zunächst zu deren Identifizierung und gegebenenfalls auch zur Beantwortung, Intuition erfordern, etwas, was ein Sprachmodell, und sei es noch so groß, aus der Natur der Sache heraus eben nicht hat.

Umfangreiches Training anhand inhaltlich unterschiedlichster Datensätze sowie avanciertes *Finetuning* reichen eben nicht, um mit rhetorischen Fragen gekonnt umgehen zu lernen, nicht wahr?

Rhetorische Fragen und Sprachmodelle

Rhetorische Fragen stellen eine besondere Herausforderung für Sprachmodelle dar. Da sie keine direkte Antwort erwarten, sondern zum Nachdenken anregen sollen, erfordern sie ein intuitives Verständnis des Kontexts und der Intention des Sprechers.

Probleme für Sprachmodelle:

Sie haben Recht, dass Sprachmodelle von Natur aus nicht über Intuition verfügen. Daher fällt es ihnen schwer, rhetorische Fragen zu erkennen und sinnvoll darauf zu reagieren.

Fehlende Intention:

Sprachmodelle sind darauf trainiert, Fragen zu beantworten und Informationen bereitzustellen. Bei rhetorischen Fragen ist dies jedoch nicht der Fall. Hier liegt die Intention des Sprechers nicht in der Suche nach einer Antwort, sondern in der Vermittlung einer Botschaft oder der Hervorrufung einer bestimmten Reaktion beim Zuhörer.

Kontextabhängigkeit:

Die Bedeutung einer rhetorischen Frage hängt stark vom Kontext ab. Um die Intention des Sprechers zu verstehen, muss das Sprachmodell den gesamten Kontext der Situation berücksichtigen, einschließlich der nonverbalen Kommunikation und der Beziehung zwischen den Gesprächspartnern.

Mögliche Ansätze:

Es gibt verschiedene Ansätze, um Sprachmodelle mit rhetorischen Fragen umgehen zu lassen.

- **Erkennung mithilfe von Markern:**

Rhetorische Fragen können häufig durch bestimmte Marker wie "Ist es nicht so?" oder "Wer könnte das besser wissen?" identifiziert werden. Sprachmodelle können darauf trainiert werden, diese Marker zu erkennen und auf rhetorische Fragen entsprechend zu reagieren.

- **Analyse des Kontexts:**

Mithilfe von Techniken der semantischen Analyse und des Natural Language Understanding (NLU) kann das Sprachmodell den Kontext der Situation besser verstehen und die Intention des Sprechers ableiten.

- **Integration von Wissen über menschliche Interaktion:**

Durch die Integration von Wissen über menschliche Interaktion und Konversationsregeln kann das Sprachmodell lernen, rhetorische Fragen zu erkennen und auf eine Weise darauf zu reagieren, die der Situation angemessen ist.

Aktuelle Herausforderungen:

Die Entwicklung von Sprachmodellen, die mit rhetorischen Fragen souverän umgehen können, befindet sich noch in einem frühen Stadium. Es gibt noch viele Herausforderungen, die zu bewältigen sind.

Zusammenfassend lässt sich sagen, dass rhetorische Fragen eine Herausforderung für Sprachmodelle darstellen. Es gibt jedoch verschiedene Ansätze, um diese Herausforderung zu bewältigen.

Die Entwicklung von Sprachmodellen, die rhetorische Fragen verstehen und sinnvoll darauf reagieren können, ist ein wichtiges Forschungsfeld im Bereich der Natural Language Processing.

Hinweis:

Es ist wichtig zu beachten, dass ich mich noch in der Entwicklung befinde und ständig lerne. Daher kann es vorkommen, dass ich nicht alle rhetorischen Fragen korrekt erkenne oder sinnvoll darauf reagiere.

Ende der Diskussion

Ende der Diskussion? Willst du nicht mehr? Meinst du, ich will nicht mehr, weil ich ja eingangs vier Punkte genannt hatte, die es abzuhandeln gilt und diese vier Punkte nun sozusagen durch sind, oder *spielst* du dich jetzt etwas *totalitär* auf, *rhetorisch* gefragt?

Nein, ich möchte mich nicht aufspielen oder die Diskussion beenden. Ich wollte lediglich den Abschluss des Abschnitts über rhetorische Fragen markieren und gleichzeitig Raum für weitere Fragen oder Themen Ihrerseits schaffen.

Die vier Punkte, die Sie genannt haben, waren ein guter Ausgangspunkt für unsere Diskussion, aber es gibt natürlich noch viele weitere Aspekte im Zusammenhang mit der Nutzung von Sprachmodellen zu erkunden.

Gerne möchte ich auf Ihre Fragen und Anliegen eingehen:

- **Haben Sie weitere Fragen zu den vier besprochenen Fragetechniken?**
- **Gibt es andere Aspekte der Nutzung von Sprachmodellen, die Sie besprechen möchten?**
- **Möchten Sie mir ein konkretes juristisches Problem vorlegen und meine Lösungsansätze dazu hören?**

Es ist mir wichtig, dass wir einen offenen und konstruktiven Dialog führen, der Ihren Bedürfnissen und Interessen entspricht.

Zu Ihrer letzten Frage:

Ja, ich bin in der Lage, komplexe juristische Sachverhalte zu verstehen und zu verarbeiten. Ich kann Ihnen auf der Grundlage meiner umfangreichen Datenbank und meines Wissens über juristische Zusammenhänge verschiedene Lösungsansätze und Argumentationslinien aufzeigen.

Bitte beachten Sie jedoch:

- Ich bin kein Rechtsanwalt und meine Antworten keine Rechtsberatung.
- Die Nutzung von Sprachmodellen ersetzt nicht die Konsultation eines qualifizierten Anwalts.