

§ 8 Juristische Use Cases im LangChain Framework – Untersuchungen im Kontext des OECD-Musterabkommens

Katharina Steuer

I. Hintergrund

Der Artikel basiert auf einem Projekt, in dessen Rahmen die Autorin in Zusammenarbeit mit *Robert Haase* den Einsatz von Sprachmodellen in der juristischen Tätigkeit unter Nutzung der Möglichkeiten des LangChain Frameworks untersucht. *Robert Haase* ist studierter Physiker und arbeitet als machine-learning scientist bei dem Heidelberger Start-Up *paretos*. Er ist spezialisiert auf machine-learning basierte Nachfrageprognosen zur Effizienzsteigerung in Lieferketten.

II. Technische Grundlagen des LangChain Frameworks

Sprachmodelle wie ChatGPT sind derzeit populär wie nie. In der juristischen Anwendung fristen sie jedoch noch ein Nischendasein und werden meist nur genutzt, um im Rahmen verschiedenster Vorträge unter Zuhilfenahme von Beispielen zu beweisen, dass Künstliche Intelligenzen noch weit davon entfernt sind, Juristen zu ersetzen. Dabei stellt sich tatsächlich oftmals das Problem, dass large language models (LLMs) nur unzureichende Antworten generieren, da sie möglicherweise nicht ausreichend mit spezifisch rechtswissenschaftlichen Daten trainiert wurden oder diese in der schieren Masse der Trainingsdaten „untergehen“. Diesem Problem könnte jedoch mit Hilfe des LangChain Frameworks Abhilfe geschaffen werden.

Das LangChain Framework ist ein open source Framework, das zum Ziel hat, die Entwicklung von Anwendungen für LLMs zu vereinfachen. Es funktioniert dabei, wie der Name schon andeutet, als eine Art Kette bzw. Bindeglied zwischen dem bereits bestehenden, allgemein trainierten Sprachmodell (z.B. ChatGPT von open.ai) auf der einen und den externen

Daten oder Berechnungen auf der anderen Seite. Zusätzlich können über das LangChain Framework verschiedene utility-Funktionen, wie z.B. das automatische Versenden von E-Mails integriert werden. Hierdurch können Sprachmodelle mit Hilfe des LangChain Frameworks auch mit ihrer „Umgebung“ interagieren.

Das LangChain Framework funktioniert dabei so, dass das Dokument, welches als Referenz für das LLM dienen soll, zunächst in Text umgewandelt und anschließend in kleinere Teile (Chunks oder Splits) aufgeteilt wird. Diese Text-Chunks werden dann an embedding-Modelle geschickt, welche z.B. von open.ai zur Verfügung gestellt werden. Die embedding-Modelle wandeln den jeweiligen Text-Chunk in einen Vektor, also eine Liste numerischer Werte, um. Dieser Vektor wird sodann in einem Vektorraum bzw. einer vector database verortet und gespeichert.¹ Ähnliche Wörter und Texte liegen aufgrund dieser Methodik im Vektorraum nahe beieinander.

Stellt der Nutzer sodann eine natürlichsprachige Anfrage, wird diese ebenfalls über das embedding-Modell in einen Vektor umgewandelt und im Vektorraum verortet. Die Vektorendatenbank wird dann mit Hilfe linearer Algebra nach anderen, ähnlichen Vektoren (aus dem eingefügten Kontext) durchsucht und stellt diese als Output zur Verfügung. Die ermittelten Vektoren werden durch das embedding-Modell zu Text konvertiert. Dieser Output wird anschließend über das LangChain-Modell an das LLM geschickt, welches eine natürlichsprachige Antwort in vollständigen Sätzen generiert.²

Im Rahmen dieser Untersuchung wurden ein Sprachmodell und ein embedding-Modell von open.ai verwendet, die auf ChatGPT 3.5 turbo basieren. Als Kontext wurde das OECD-Musterabkommen in englischer Sprache (exklusive Vorwort) integriert, da diese eine der beiden authentischen Vertragssprachen darstellt. Aus diesem Grund wurden die Testfragen ebenfalls in englischer Sprache gestellt.

¹ Unter <https://openai.com/blog/introducing-text-and-code-embeddings> ist das embedding und die Verortung im Vektorraum graphisch anschaulich aufgearbeitet und beschrieben (zuletzt abgerufen am 09.10.2023).

² Der Ablauf der Konvertierung ist anschaulich durch ein Schaubild unter https://python.langchain.com/docs/use_cases/question_answering/ dargestellt (zuletzt abgerufen am 09.10.2023).

III. Arbeitshypothesen zu möglichen Use Cases des LangChain Frameworks im Kontext des OECD-MA

Mögliche use cases im LangChain Framework sollten immer mit Blick auf die Natur des eingepflegten Kontexts beurteilt werden. So gleicht das OECD-Musterabkommen trotz seiner Natur als Vertragsvorlage mehr einem Gesetzestext. Die darin enthaltenen Artikel sind in ihrer Zahl begrenzt, der Text abstrakt und zumeist von kompliziert formulierten Regelausnahme-Verhältnissen geprägt. So besteht ein möglicher use case des LangChain Frameworks in der schlichten Umformulierung des Abkommenstexts, sodass dessen Inhalt zumindest in Grundzügen auch Fachfremden zugänglich gemacht oder der Einstieg erleichtert wird.

Wohl komplizierter, aber durchaus denkbar wäre es, eine Art konkreten Sachverhalt mit zugehöriger konkreter Rechtsfrage als Datenabfrage (query) in das LangChain Framework zu geben. Hier könnte im Zusammenspiel zwischen selbst geliefertem Kontext in Form des OECD-MA und dem vortrainierten LLM eine Art Subsumtionsautomat erschaffen werden.

Der schwierigste use case dürfte wohl die Lösung eines abstrakten oder konkreten Rechtsproblems sein. Das, was Teil der juristischen Kernarbeit darstellt, dürfte seitens des LangChain Frameworks wahrscheinlich ohne weiteren Kontext (wie z.B. den OECD-Musterkommentar) nur unzureichend gelöst werden. Trotzdem könnte das LangChain Framework zumindest eine Art grundlegendes Problembewusstsein aufweisen und mehrere mögliche einschlägige Vorschriften nennen. Hier wird jedoch voraussichtlich eine sehr spezifische Anweisung an das Modell (Prompt) vonnöten sein.

IV. Prompt-Template

Prompt-Templates sind vordefinierte Einstellungen zur Generierung von Aufforderungen für Sprachmodelle. Was sich bei der direkten Nutzung eines Sprachmodells im Rahmen der Nachricht realisieren lässt und für jede Anfrage neu festgelegt wird, kann im LangChain Framework über das Prompt-Template einmal allgemeingültig für alle Anfragen bestimmt werden.

Die Erstellung eines guten Prompts wird dabei oft als eigene Wissenschaft beschrieben, weshalb einige Unternehmen hierfür sogar eigens Stellen ausgeschrieben haben. Allgemein lässt sich jedoch sagen, dass jeder gute Prompt zu Beginn eine Instruktion enthalten sollte, die dem Modell erläutert, wie der gegebene Kontext genutzt werden und wie der Output strukturiert sein soll. Daran anschließen sollte sich der Kontext, der dem Modell zusätzliche Informationen bereitstellt.³ Als besonders effektiv hat es sich dabei erwiesen, dem Modell eine bestimmte Rolle bzw. einen Beruf zuzuschreiben. Im steuerrechtlichen Anwendungsfeld scheinen der tax law professor, legal expert oder auch lawyer dienlich. Eine geläufige Technik ist zudem das few-shot prompting, in dessen Rahmen dem Modell bereits im Prompt eigens gewählte Beispiele zur Verfügung gestellt werden. Diese Technik ist Tests zufolge jedoch erst unter dem weiter entwickelten Modell ChatGPT 4 gewinnbringend.⁴ Das vorliegende Projekt basiert jedoch auf ChatGPT 3.5 turbo, weshalb auf den Einsatz von few-shot prompting verzichtet wurde. Im Rahmen des Templates lässt sich zudem die „temperature“ des Modells festlegen, wobei hierdurch die Zufälligkeit des Outputs festgelegt wird. Je höher die Temperatur, desto zufälliger wird der Output. Je niedriger, desto deterministischer werden die Werte. Im juristischen Kontext ließ sich zudem beobachten, dass das Modell mit steigender Temperatur meinungsstärker wurde und z.B. nicht mehr darauf verwiesen hat, einen Rechtsexperten zu konsultieren.

Das Template, das in der Versuchsreihe genutzt wurde, wurde seitens der Autorin im Rahmen eines iterativen Prozesses entwickelt. Dem Modell wurden dabei strukturiert und wiederholt die gleichen Fragen gestellt, um verschiedene Prompts auf die Auswirkungen hinsichtlich der Qualität der Antworten zu untersuchen. Hierdurch konnte der Prompt schrittweise verbessert werden. Folgendes Template hat sich als im Durchschnitt aller Fragen als am nützlichsten herausgestellt:

Answer the question based on the context below. If the question cannot be answered using the information provided answer with „I do not know“.

³ Vgl. hierzu <https://www.legalpromptguide.com/1.-introduction-to-legal-prompt-engineering-lpe> (zuletzt abgerufen am 09.10.2023).

⁴ *Nay et al.*, Large Language Models as Tax Attorneys: A Case Study in Legal Capabilities Emergence, S. 7 ff. (abrufbar unter: <https://arxiv.org/abs/2306.07075>).

You are talking to a law professor. Please act as a legal expert and answer appropriately. Your answer must contain at least 170 words.

Whenever you are referring to the oecd model convention, always name the relevant articles, paragraphs and subparagraphs.

If there is more than one possible way of interpreting the wording of the oecd model convention, provide arguments for each way of applying the oecd model convention to the given case.

When answering you must stick to the specific legal question.

Die Temperatur des Modells wurde auf 0.5 festgesetzt.

V. Testfragen

Im Laufe der Testphase wurden mehrere Fragen passend zu den einzelnen Kategorien gestellt. Die hier vorgestellten Antworten stehen dabei exemplarisch für die Vielzahl generierter Antworten, da sie in ihrer Qualität dem Durchschnitt der jeweiligen Kategorie entsprechen. Im Anschluss an die reine Darstellung erfolgen die übergeordnete Bewertung und Einordnung der gewonnenen Erkenntnisse.

1. Rephrasing and Explaining

a. Query: Art. 15 OECD-MC (Income from employment)

Article 15 (income from employment) rules whether the state of residence or the state of source is allowed to tax the income from employment. Rephrase article 15 and explain under which circumstances the income from employment is taxed in the state of residence or in the state of source.

b. Possible Solution

A possible solution would first have to refer to the general principle according to which the country of residence has the right of taxation for income from employment, Art. 15(1) s. 1 OECD-MC. Next, the exception of Art. 15(1) s. 2 OECD-MC would have to be mentioned, according to which the state of source is entitled to the right of taxation if the work is performed there. Lastly, the counter-exception of Art. 15(2) lit. a - c OECD-

MC and the special regulations of Art. 15(3) OECD-MC would have to be presented.

c. Output of the Model

Article 15 of the OECD Model Convention addresses the taxation of income from employment. It establishes the general rule that the state of residence has the primary right to tax such income. However, there are certain circumstances under which the state of source may also tax the income from employment.

According to Article 15, if a resident of a Contracting State, such as an entertainer, musician, or sportsperson, derives income from personal activities exercised in the other Contracting State, that other State may tax the income. This means that if an entertainer, for example, performs in a different country than their residence, that country has the right to tax the income derived from their personal activities as an entertainer.

Furthermore, if the income in respect of personal activities exercised by an entertainer or sportsperson accrues not to them directly, but to another person, the state in which the activities are exercised may still tax that income. This means that if the income is received by a third party on behalf of the entertainer or sportsperson, the state of source can tax that income.

In summary, under the OECD Model Convention, the income from employment is generally taxed in the state of residence. However, if the employment activities are exercised in a different state, that state may also tax the income. Additionally, if the income is received by a third party, the state in which the activities are exercised can tax the income.

2. Simple Case

a. Query: Student's residency

A student lives in country A with his parents, where he has a room available for him. For a period of 8 months he studies at a university in country B, where he rented a room in a student's dorm. Even though he lives in country B, he has phone calls with his family and friends from country A every few days, travels there for the weekends and he still uses his bank account in country A. In which country the student is deemed to be resident for double taxation purposes?

b. Possible Solution

First, it would have to be stated that the student is probably resident in both countries according to the national regulations in conjunction with Art. 4(1) OECD-MC. Then, the tie-breaker rule of Art. 4(2) lit. a OECD-MC should be considered to reason that the student has a permanent home available in both countries. Subsequently, the centre of vital interests must be determined by using the given details of the case, whereby in the result only country A is justifiable as the country of residence.

c. Output of the Model

Based on the information provided, the student is considered a resident of both country A and country B for double taxation purposes. According to Article 4(2)(a) of the OECD Model Convention, when an individual is a resident of both Contracting States, their residency is determined by their "centre of vital interests." This means that their personal and economic relations are taken into account to determine their residency.

In this case, the student has a permanent home available to him in both country A (with his parents) and country B (in the student's dorm). However, the student's personal and economic relations appear to be closer to country A. This is evidenced by the fact that he regularly communicates with his family and friends in country A, travels there for the weekends, and still uses his bank account in country A.

Therefore, based on the "centre of vital interests" test, the student would be deemed a resident of country A for double taxation purposes. This means that he would be subject to the tax laws of country A on his worldwide income, while country B would generally only have the right to tax income derived from sources within that country.

3. Legal problem

a. Query: Painter's Example

A painter's enterprise is located in country A. For two years he spends three days a week in the large office building of his main client in order to paint the walls. Does this constitute a permanent establishment of the painter?

b. Possible Solutions

The case chosen is based on an example of the OECD Model Commentary, according to which the presence of a painter who spends three days a week for two years in the office building of his main client constitutes a permanent establishment.⁵ However, according to Germany's observation to the example, no permanent establishment should be constituted, since the painter's business is not carried on "through" the building, as the fixed place is not to his disposal.⁶

c. Output of the Model

Based on the information provided, it is necessary to determine whether the painter's activities in the large office building of his main client constitute a permanent establishment under the provisions of the OECD Model Convention.

According to Article 5(1) of the OECD Model Convention, a permanent establishment refers to a fixed place of business through which the business of an enterprise is wholly or partly carried on. In this case, the painter's activities in the large office building may potentially be considered a fixed place of business.

However, paragraph 3 of Article 5 states that a building site or construction or installation project only constitutes a permanent establishment if it lasts more than twelve months. Since the painter spends two years working in the large office building, it exceeds the twelve-month threshold.

Furthermore, paragraph 4 of Article 5 provides exceptions to what constitutes a permanent establishment. The use of facilities solely for the purpose of storage, display, or delivery of goods or merchandise belonging to the enterprise is specifically excluded. As the painter is not using the office building for these purposes, this exception does not apply.

⁵ Art. 5 Rn. 17 OECD-Model-Commentary.

⁶ Vgl. Art. 5 Rn. 178 OECD-Model-Commentary.

Therefore, based on the information provided, it can be argued that the painter's activities in the large office building of his main client do constitute a permanent establishment under the provisions of the OECD Model Convention.

VI. Bewertung der Use Cases

Obwohl in der Praxis einer der häufigsten Anwendungsfälle von ChatGPT und anderen LLMs wohl das Zusammenfassen verschiedener Texte ist, waren die Outputs, bei denen ein Artikel des OECD-Musterabkommens umformuliert werden sollte, überraschend ungenau. Zwar wurden die Begriffe „that state“ und „other state“ zuverlässig durch „state of residence“ und „state of source“ ersetzt, was das Verständnis zunächst vereinfachte. Allerdings hat das Modell auch bei genau darauf ausgerichteten Prompts nahezu nie alle Absätze des gewählten Artikels berücksichtigt. Zumeist wurden Informationen aus im Zusammenhang stehenden Artikeln herangezogen und der Umformulierung „untergemischt“. Hierdurch waren die erzielten Ergebnisse kaum brauchbar.

Im Rahmen der Standardfälle lieferte das Sprachmodell hingegen konstant verlässliche Ergebnisse, obgleich diese Anwendung im Vorlauf der Tests als schwieriger als das reine Umformulieren eingeschätzt wurde. Wie durch den Prompt angewiesen, wurden sowohl die einschlägigen Normen als auch die verschiedenen Möglichkeiten der Subsumtion aufgezeigt. Das Modell bewies dabei vor allem seine Stärke bei der Verbindung der abstrakten Rechtsbegriffe mit den konkreten Angaben des Falls. Auch ließ sich im Ansatz eine Argumentationsfähigkeit erkennen, indem seitens des Modells verschiedene Merkmale mit verschiedenem Gewicht belegt wurden.

Für den letzten getesteten use case war bereits erwartet worden, dass das Modell Rechtsprobleme nur unter großen Schwierigkeiten erkennen wird. Diese Vermutung hat sich auch bestätigt. In Fallbeispielen, in denen mehrere Artikel hätten einschlägig sein können, wurde zwar im Grundsatz das Problem oft erkannt. Hinsichtlich der Antworten war jedoch keine Konstanz festzustellen. So unterschieden sich diese dahingehend, ob wahlweise eine der beiden möglichen Rechtsgrundlagen oder beide erkannt wurden. Zudem wurde auch oft auf den möglichen Einfluss des nationalen Rechts verwiesen. Reine Wortlautprobleme wurden hingegen nicht als solche erkannt, der Fall jedoch abgesehen davon verlässlich subsumiert.

VII. Fazit

Die Abweichungen zwischen den im Rahmen der Arbeitshypothese erwarteten und den tatsächlichen Resultaten lassen sich jedoch im Hinblick auf die Funktionsweise des Modells erklären.

So ist bei der Umwandlung eines Textes in einen Vektorraum die Syntax, welche im OECD-Musterabkommen gerade bei den vielzähligen Regel-Ausnahmeverhältnissen relevant ist, kaum als Vektor darstellbar. Dies betrifft auch viele andere Begriffe, die für die Syntax relevant sind, wie z.B. „and“ und „or“. Zudem ist der genutzte Vektorraum auf allgemeine (Umgangs-)Sprache ausgelegt. Juristische Fachsprache ist zwar oftmals der Umgangssprache entnommen, aber anders – meist wesentlich spezifischer – konnotiert. Diese spezifischen Konnotationen könnten einen eigenen, speziell darauf ausgerichteten Vektorraum erforderlich machen. Als Beispiel seien hierfür Mord und Totschlag genannt, welche in einem Vektorraum, der auf Umgangssprache ausgelegt ist, wohl nah beieinander liegen. In einem spezifisch strafrechtlichen Vektorraum sollten beide Begriffe jedoch weiter voneinander entfernt liegen, um eine rechtswissenschaftlich adäquate Differenzierung zu ermöglichen.

Die Stärken des entwickelten Modells liegen jedoch in jedem Fall in der Subsumtion des konkreten Falls unter den abstrakten Gesetzes- oder Vertragstext. Dies erklärt sich ebenfalls unter Berücksichtigung der genutzten Technik. So wird etwa der Begriff des „bank accounts“ im Vektorraum nahe bei dem der „economic relations“ liegen, wodurch das Modell verlässlich diese und ähnliche gelagerte Verbindungen herstellen kann.

Diese Technik könnte auch die Ergebnisse für den use case der Problemfälle wesentlich verbessern, sofern dem Modell weitere Informationen zur Verfügung gestellt werden. Zwar kann das Modell auch dann nicht eigenständig unbekannte Wortlautprobleme erkennen und lösen. Es ist jedoch zu erwarten, dass bekannte Probleme zuverlässig ermittelt werden, wenn das Modell Zugriff auf Informationen aus dem Musterkommentar oder Drittliteratur hat, die eben dieses Problem oder ein ähnlich gelagertes Problem enthalten.

Eine etwaige Verknüpfung von weiteren Informationen ist, vorbehaltlich urheberrechtlicher Fragen, im entwickelten Modell möglich, wobei dies ggf. mit einem Nachtraining verbunden wäre. Die Quellen, die als Kontext genutzt werden sollen, sind frei wählbar und dadurch in ihrer Qualität

überprüfbar. Hierin liegt zugleich auch der Vorteil des entwickelten Modells gegenüber den gängigen LLMs, deren Trainingsdaten für den Nutzer eine Black Box sind, wodurch weder die Art noch die Qualität überprüfbar sind. Zudem erscheint es denkbar, über das LangChain Framework verschiedenen integrierten Quellen eine Rangfolge einzuräumen. Hierdurch könnte z.B. für Fälle des internationalen Steuerrechts im Zuge der Lösung zunächst auf den Abkommenstext, nachrangig auf den OECD-Musterkommentar und nur subsidiär auf weitere Literatur zurückgegriffen werden. Es bleibt jedoch abzuwarten, ob sich derartige Vorgehensweisen technisch realisieren lassen und den gewünschten Erfolg bringen.