

Andreas Schultz/Melanie Pospisil/Kathleen Jennrich/Matthias Grabmair

KI-gestützte Gesetzgebung

Mit LegisLLM von der legislativen Aufgabe zum Gesetzesentwurf

Kann Künstliche Intelligenz (KI) Gesetze schreiben? Der Beitrag gibt einen Überblick über Ziele und Gang der Machbarkeitsstudie „Digitaltaugliches Steuerrecht“ und stellt erste Erkenntnisse vor. Ausgangspunkt ist eine Interviewreihe mit Legistinnen und Legisten auf Bundes- und Landesebene, deren Ergebnisse Einblicke in Herausforderungen einer digitaltauglichen Gesetzgebung geben. Aufbauend darauf wird mit „LegisLLM“ ein Anwendungsprototyp vorgestellt, der den Prozess der Formulierung von Änderungsgesetzen begleitet. Anhand eines Beispiels aus dem Steuerrecht wird die Funktionsweise der Anwendung erläutert. Der Beitrag schließt mit einer kritischen Einordnung bestehender (technischer) Herausforderungen und blickt auf Potenziale zur Weiterentwicklung.

I. Einleitung

Das deutsche Rechtssystem ist hochkomplex und vielschichtig verflochten. Stand 2024 sind allein auf Bundesebene rund 1.800 formelle Gesetze mit 52.000 Einzelnormen sowie rund 2.900 Rechtsverordnungen mit 44.000 Einzelnormen in Kraft.¹ Diese sind häufig durch unbestimmte Rechtsbegriffe und (teilweise implizite) Querverweise gekennzeichnet, was eine Herausforderung für diejenigen darstellt, die Gesetze formulieren oder anwenden müssen. Gesellschaftliche Veränderungen und aufkommende gesetzliche Lücken sind Auslöser für immer neue Gesetzesinitiativen. Impulse liefern dabei u. a. Koalitionsverträge, Eckpunktepapiere der Regierung, Stellungnahmen der Verbände und Evaluationsberichte.

Um gesetzgeberische Ziele präzise und systematisch umzusetzen, braucht es ein strukturiertes Gesetzesdesign mit klaren Begriffen, konsistenter Rechtssprache und durchdachter Normtechnik.² Zu diesem Zweck kommen zunehmend neuartige Arbeitsmethoden, wie etwa das „Design Thinking“ zum Einsatz.³ Auch Visualisierungen können helfen, die zu regelnden Prozess- und Verfahrensschritte sichtbar zu machen.⁴ Ein zukunftsweisender Ansatz zur Weiterentwicklung der Gesetzgebung ist das Konzept „Rules as Code“.⁵ Dieses zielt darauf ab, gesetzliche Regelungen nicht nur in natürlicher Sprache, sondern gleichzeitig in einer maschinenlesbaren, regelbasierten Form zu repräsentieren. Dies eröffnet insbesondere erhebliche Potenziale für die Effizienz und Transparenz im Verwaltungsbereich, etwa iRd automatisierten Anspruchsprüfung.⁶ Damit gesetzliche Regeln maschinenlesbar erfasst werden können, müssen sie bereits in der Entwurfsphase ausreichend strukturiert repräsentiert werden.⁷

Zum Formulieren eines technisch anschlussfähigen Gesetzestextes bedarf es geeigneter Methoden und Werkzeuge. Erste innovative Ansätze weisen hier bereits den Weg. Neue Bundesgesetze werden inzwischen regelmäßig einem Digitalcheck unterzogen⁸ und das Zentrum für Legistik unterstützt den Wandel hin zu einer besseren Rechtssetzung.⁹ Aber auch Software-Pilotprojekte werden auf den Weg gebracht. Die

Bundesregierung treibt etwa mit dem Projekt „Elektronisches Gesetzgebungsverfahren des Bundes“ (E-Gesetzgebung) die vollständige Digitalisierung und Standardisierung des Rechtsetzungsverfahrens voran.¹⁰ Mit eLexa hat das Bundesministerium der Finanzen (BMF) gemeinsam mit der Freien Hansestadt Bremen einen KI-gestützten Assistenten entwickelt. Auf Basis eines Wissensgraphen¹¹ analysiert eLexa juristische Begriffe, stellt diese strukturiert dar und verknüpft sie mit Verwaltungsdatenbeständen. Auf diese Weise kann das System

- 1 BT-Drs. 20/11746, 2.
- 2 Vgl. Tosun/Bruch, Gesetze, die wirken: Ein Transformationsplan für die ministerielle Gesetzesvorbereitung, 2025, abrufbar unter: <https://agoradigital.de/projekte/gesetzgebungsreform>; Nationaler Normenkontrollrat, Deutschland muss einfacher werden, 19.2.2025, abrufbar unter: <https://www.normenkontrollrat.bund.de/Webs/NKR/SharedDocs/Downloads/DE/Positionspapiere/2025-02-nkr-strategiepapier-bundesregierung.html>.
- 3 Breidenbach, Was Gesetze sein könnten, 2025, S. 116 ff.; vgl. Digital Service, Harmonisierung des Begriffs „Einkommen“, 2022, abrufbar unter: <https://digitalservice.bund.de/fellowships/work4germany/projekte/harmonisierung-einkommensbegriff-bmf-2022>.
- 4 Zum Einsatz von Visualisierungen im Rahmen digitaltauglicher Gesetzgebung s. Nationaler Normenkontrollrat, Durch mehr Visualisierung zu digitaltauglichen Gesetzen, 2024, abrufbar unter: https://www.normenkontrollrat.bund.de/Webs/NKR/SharedDocs/Newsletter/DE/2023/2023-1/artikel-2-digitalcheck.html?mtm_campaign=linksFromNewsletter. Ein gelungenes Beispiel für eine Visualisierung findet sich im Gesetzesentwurf zur Änderung des Ehenamens- und Geburtsnamensrechts, s. Nationaler Normenkontrollrat, NKR-Stellungnahme Nr. 6652 zum Entwurf eines Gesetzes zur Änderung des Ehenamens- und Geburtsnamensrechts (BMJ), Stellungnahme v. 6.9.2023, abrufbar unter: <https://www.normenkontrollrat.bund.de/Webs/NKR/SharedDocs/Stellungnahmen/DE/2023/nkr-nr-6652.html>.
- 5 Mohun/Roberts, Cracking the code: Rulemaking for humans and machines, 2020, S. 16 ff.
- 6 Vgl. NEGZ, Digitalisierung der Gesetzgebung zur Steigerung der digitalen Souveränität des Staates, 2021, abrufbar unter: https://negz.org/wp-content/uploads/2022/12/19_Kurzstudie_Digitalisierung-der-Gesetzgebung-2021.pdf.
- 7 Hierfür kommen strukturierte Dokumentenformate wie „LegalDocML“ zum Einsatz, einer auf dem sog. Akoma-Netso-Standard beruhenden XML-Standard für Rechtsdokumente, vgl. BT-Drs. 20/13400, 6; BT-Drs. 19/25654; Paschke KIR 2024, 86 (88).
- 8 Der Anfang 2023 eingeführte Digitalcheck prüft Gesetzesvorhaben systematisch auf ihre Digitaltauglichkeit und soll mögliche Vollzugshürden frühzeitig identifizieren. Dazu wurden fünf Leitprinzipien entwickelt, s. näher Bundesministerium für Digitales und Staatsmodernisierung, Digitalcheck für Gesetze, 2023, abrufbar unter: <https://www.digitale-verwaltung.de/Webs/DV/DE/transformation/digitalcheck/digitalcheck-node.html>; Nationaler Normenkontrollrat: Deutschland muss einfacher werden, 2025, S. 5. Auch in einigen Bundesländern finden sich entsprechende Prüfungsinstanzen, so etwa in Bayern und Baden-Württemberg.
- 9 Das Zentrum für Legistik leistet einen wichtigen Beitrag zur methodischen Weiterentwicklung der Gesetzgebung, indem es strukturelle Verbesserungen vorschlägt und Qualifizierungsangebote für Legistinnen und Legisten entwickelt, vgl. Schmid ZParl, 2025, 193.
- 10 Zum Funktionsumfang geplanter Software iRd E-Gesetzgebung s. BT-Drs. 20/13400; auch auf EU-Ebene gibt es mit LEOS (Legislation Editing Open Software) vergleichbare Initiativen, näheres unter: <https://interoperable-europe.ec.europa.eu/collection/justice-law-and-security/solution/leos-open-source-software-editing-legislation#:~:text=follow%20the%20rules%20and%20avoid,mistakes>.
- 11 Wissensgraphen stellen Informationen als Knotenpunkte und ihre Beziehungen untereinander als Kantenverbindungen strukturiert dar. Zur semantischen Wissensrepräsentation für juristische KI-Systeme durch Wissensgraphen s. Paschke KIR 2024, 86 (88).

bei der Formulierung konsistenter und eindeutiger Rechtsbegriffe gezielt unterstützen.¹² Dass die eigentliche Formulierung der Gesetzestexte in der Praxis nach wie vor überwiegend „handcrafted“ erfolgt, zeigt, dass gleichwohl Bedarf an weiteren digitalen Werkzeugen besteht, die Legistinnen und Legisten umfassend bei der konkreten Ausformulierung von Gesetzesvorhaben begleiten.

II. Machbarkeitsstudie „Digitaltaugliches Steuerrecht“

Gemeinsam mit dem Zentrum für Digitalisierung des Steuerrechts der Ludwig-Maximilians-Universität München (LMU-DigiTax)¹³ und mit Unterstützung des BMF führt die Arbeitsgruppe Legal Tech der Technischen Universität München (TUM)¹⁴ seit Mitte 2024 die Machbarkeitsstudie „Digitaltaugliches Steuerrecht“ durch. Diese baut auf einer bestehenden Kooperation auf, idR bereits Hackathons im Bereich der Formalisierung von Steuerrechtsnormen mit Low/No-Code Tools¹⁵ und digital vollzugstauglicher Gesetzesformulierung¹⁶ veranstaltet wurden.

Das Forschungsprojekt untersucht systematisch, wie moderne Technologien, insbesondere KI, zur Unterstützung und Verbesserung gesetzgeberischer Prozesse beitragen können. Es sollen iRv drei Arbeitspaketen tragfähige Zukunftsszenarien entworfen und hinsichtlich ihrer praktischen Umsetzbarkeit bewertet werden. Zunächst wurde durch qualitative Interviews eine umfassende Zustandsbeschreibung der Funktionen und derzeitigen Arbeitsweise von in den Gesetzgebungsprozess eingebundenen Legistinnen und Legisten erarbeitet. Praktische Herausforderungen, strukturelle Hürden und bestehende Erwartungen an digitale Unterstützungsmethoden sollten unmittelbar aus der Perspektive der im Gesetzgebungsverfahren Beteiligten erfasst werden. Darauf aufbauend wurde untersucht, welche digitalen Werkzeuge und technischen Lösungsansätze bereits existieren und inwieweit sie sich für eine Unterstützung im Gesetzgebungsprozess eignen. Nun werden praxisnahe Anwendungsszenarien entwickelt, die zeigen, wie neue technische Möglichkeiten, speziell die Nutzung von KI, in die Arbeitsabläufe der Gesetzgebung integriert werden könnten.

Im Folgenden werden zentrale Zwischenergebnisse aus der Interviewreihe vorgestellt sowie ein praktisches Anwendungsszenario als „Proof of Concept“ präsentiert: LegisLLM, eine browserbasierte Anwendung, die durch iterative Kommunikation mit einem großem Sprachmodell (Large Language Model, LLM) Legistinnen und Legisten ermöglicht, auf Grundlage einer formulierten Aufgabenstellung schrittweise Änderungsgesetze zu entwickeln.

III. Interviewreihe des LMUDigiTax

IRd Machbarkeitsstudie führte das LMUDigiTax von Ende Oktober bis Anfang Dezember 2024 eine qualitative Interviewreihe durch.¹⁷ In halbstandardisierten Leitfadeninterviews wurden Legistinnen und Legisten auf Bundes- und Landesebene sowie weitere relevante Stakeholder zu ihren Erfahrungen in der praktischen Gesetzgebungsarbeit befragt.¹⁸ Die Fragen deckten den gesamten Gesetzgebungsprozess von der Anstoßphase über die rechtstechnische Ausarbeitung und Ressortabstimmung bis hin zur Umsetzung und Evaluation ab. Ziel der Inter-

viewreihe war es, praktische und organisatorische Hürden im Gesetzgebungsprozess zu identifizieren und spezifische Anforderungen an digitale Unterstützungsmethoden herauszuarbeiten. Weiterhin wurde eine Bewertung der bisherigen Umsetzung digitaltauglicher Gesetzgebung vorgenommen. Schließlich ging es auch um die Frage, welche Kriterien ein „gutes“ Gesetz in struktureller, normsystematischer und rechtsdogmatischer Hinsicht auszeichnet.

Die Auswertung der Gespräche macht deutlich, dass der Gesetzgebungsprozess derzeit von wiederkehrenden strukturellen und praktischen Hürden geprägt ist. Besonders häufig wurde auf den massiven Zeitdruck hingewiesen, unter dem Gesetzesvorhaben entstehen. Diese seien häufig getrieben durch politische Vorgaben und enge Fristen.¹⁹ Diese Dynamik führe nicht nur zu verkürzten Abstimmungsprozessen, sondern erhöhe auch die Fehleranfälligkeit und erschwere eine gründliche Prüfung der Inhalte. Auch für eine fundierte Auseinandersetzung der Legistinnen und Legisten mit Stellungnahmen bei der Verbandsbeteiligung bleibe daher oft kein Raum. Weiter wurde die stark manuell-händisch geprägte Arbeitsweise im legistischen Handwerk bemängelt. Entwürfe entstünden überwiegend im Wege von mittels E-Mail ausgetauschten Word-Dokumenten, während Änderungen, Konsolidierungen und Abstimmungen mühsam per Hand nachvollzogen werden würden. Bestehende Unterstützungen sowie digitale Werkzeuge wie Gesetzesvisualisierungen seien entweder nicht vorhanden oder werden im realen Alltag der Gesetzgebung kaum genutzt. Auch die digitale Umsetzbarkeit neuer Regelungen werde in der Praxis nur unzureichend berücksichtigt. Prinzipien wie „Once-Only“ oder Einsatz von Visualisierungen spielten, wenn überhaupt, oft erst in nachgelagerten Verfahrensphasen eine Rolle, wo der Einfluss auf die grundlegenden Züge der neuen gesetzlichen Regelungen nicht mehr möglich ist. Technische Expertise und Vollzugstellen würden häufig zu spät eingebunden, wodurch wichtige Umsetzungsperspektiven verloren gingen. Schließlich zeichne sich ein „gutes“ Gesetz dadurch aus, dass es richtig strukturiert und sprachlich möglichst verständlich²⁰ sei, vom Allgemeinen zum Besonderen aufgebaut werde und möglichst wenige Ausnahmen (und noch weniger Rückausnahmen) enthalte. Zudem müsse es sich normsystematisch in bestehende Regelungen einfügen und auch für juristische Laien möglichst nachvollziehbar bleiben.

¹² Vgl. BMF, Interoperable Rechtsbegriffe für effektive Datennachnutzung v. 29.1.2025, abrufbar unter: https://www.bundesfinanzministerium.de/Content/DE/Standardartikel/Themen/Steuern/Steuerliche_Themengebiete/Once-Only/interoperable-rechtsbegriffe.html.

¹³ Näheres zum LMUDigiTax unter: <https://www.lmudigitax.uni-muenchen.de/index.html>.

¹⁴ Näheres zur Arbeitsgruppe Legal Tech unter: <https://www.cs.cit.tum.de/lt/tum-legal-tech-working-group/>.

¹⁵ Endres/Grabmair/Heinemann/Hinrichs/Jennrich REthinking: Tax 2023, 36.

¹⁶ Endres/Grabmair/Hinrichs/Jennrich REthinking: Tax 2024, 68.

¹⁷ Unter Mitwirkung von Chiara Endres, Jil Marie Hinrichs, Adeline Canstein und Olive Lenk.

¹⁸ Vergleiche ähnlich die Ergebnisse der Interviewreihe in Fitsilis/Mikros, AI-based solutions for legislative drafting in the EU, 2024, S. 12 f.

¹⁹ Zur ressortübergreifenden Abstimmung von Referentenentwürfen sowie zur Anhörung von Ländern und Verbänden gibt die Gemeinsame Geschäftsordnung der Bundesministerien (GGO) hinreichende Fristen vor, vgl. §§ 44 ff. iVm § 50 S. 1 GGO. Diese würden jedoch faktisch meist nicht eingehalten werden.

²⁰ Vgl. § 42 Abs. 5 der Gemeinsamen Geschäftsordnung der Bundesministerien.

IV. Zielsetzung LegisLLM

Die Ergebnisse der Interviewreihe zeigen, dass Anforderungen digitaltauglicher Gesetzgebung bislang nur unzureichend in den legislativen Alltag integriert sind. Insbesondere Zeitdruck und manuelle Arbeitsprozesse drosseln eine effiziente und umsetzungsorientierte Rechtssetzung. Ziel der „Proof-of-Concept“-Anwendung LegisLLM ist es, mithilfe von KI Legistinnen und Legisten ausgehend von einer Aufgabe bis hin zu einem ersten Gesetzesentwurf iSe Arbeitsgrundlage zu führen. Das System soll nicht nur den Arbeitsaufwand reduzieren, sondern insbesondere auch dazu beitragen, Digital- und Vollzugstauglichkeit systematisch und frühzeitig mitzudenken.

Der Fokus der Anwendung liegt aktuell auf der Begleitung von Änderungsgesetzen, also auf gesetzgeberische Vorhaben, die bestehende Gesetze gezielt anpassen, ergänzen oder aufheben. Anders als beim Neuentwurf selbstständiger Stammgesetze steht hierbei nicht die Konzeption vollständig neuer, sondern die Weiterentwicklung bereits bestehender gesetzlicher Grundlagen im Vordergrund. Für die prototypi-

sche Umsetzung wurde das Steuerrecht als erstes Anwendungsfeld gewählt.²¹ LegisLLM ist in der Lage für die Aufgabe relevante Rechtsnormen automatisiert zu identifizieren, Vorschläge für unterschiedliche Regelungsalternativen zu generieren und diese unter Berücksichtigung juristischer Kriterien zu bewerten. Die Anwendung ist als Assistenzsystem für strukturierte Entscheidungsfindung zu verstehen. Politische Entscheidungen über Ziel und Zweck der gesetzlichen Regelung bleiben dem Gesetzgeber selbst vorbehalten.

V. Technische Grundlagen

Konzeptionell bildet LegisLLM einen iterativen fünfstufigen Prozess ab, der den Weg von einer legislativen Aufgabenstellung bis hin zu einem konkret formulierten Änderungsvorschlag in einzelne Arbeitsschritte aufteilt. Die Anwendung kommuniziert hierbei wiederholt mit einem Sprachmodell über eine Reihe gezielter Abfragen. Die Stärken großer Sprachmodelle können mittels qualifizierten Prompts und Informationsanreicherung gezielt genutzt werden.²²

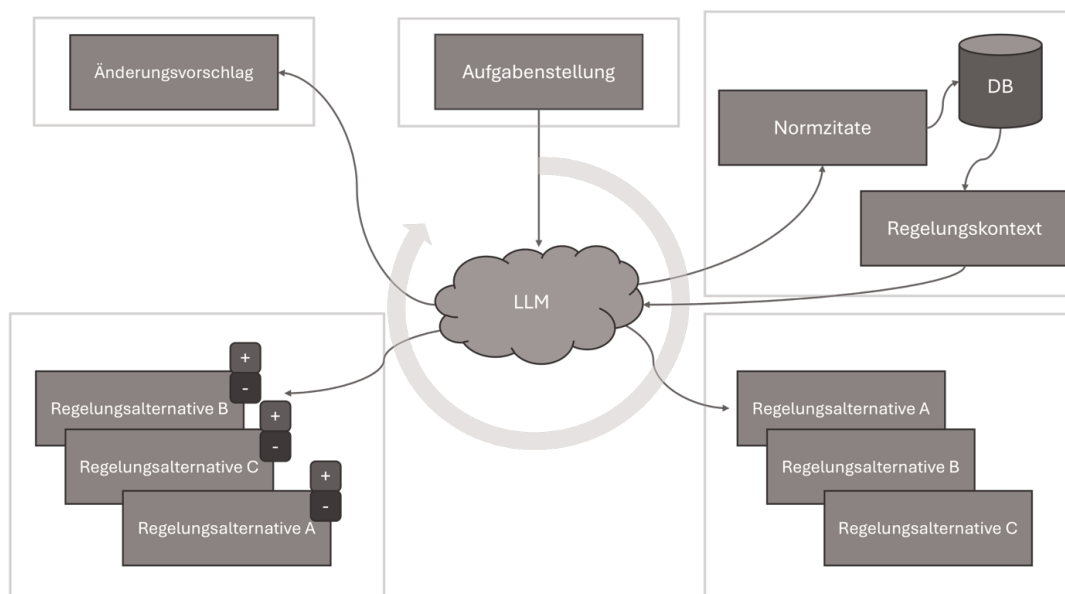


Abb. 1: Darstellung der Ablauflogik von LegisLLM

Im derzeitigen Entwicklungsstadium werden kommerzielle, proprietäre Sprachmodelle führender Anbieter über eine API-Anbindung genutzt.²³ In der prototypischen Umsetzung zeigten sich die besten Ergebnisse mit sog. „Reasoning-Modellen“. Diese sind speziell darauf ausgelegt, komplexe mehrstufige Aufgaben durch vorheriges „Durchdenken“ zu lösen, womit sie in diesem Bereich der vorherigen Generation „klassischer“ Sprachmodelle überlegen sind.²⁴

Der Anwendung vorgelagert ist die Ermittlung der legislativen Aufgabe in Textform. Diese erwächst idR aus einem politischen Impuls oder der Umsetzung eines Koalitionsvorhabens. Im ersten Arbeitsschritt erfolgt die Ermittlung des „Regelungskontexts“, also sämtlicher mit der jeweiligen Aufgabenstellung in Verbindung stehenden bzw. von dieser möglicherweise betroffenen Rechtsnormen.²⁵

Um sicherzustellen, dass der als Arbeitsgrundlage dienende Regelungskontext aktuell gehalten ist,²⁶ wird dieser in fünf Unterschritten ermittelt: (1) das LLM identifiziert betroffene

21 Vgl. NEGZ, Digitalisierung der Gesetzgebung zur Steigerung der digitalen Souveränität des Staates, 2021, abrufbar unter: https://negz.org/wp-content/uploads/2022/12/19_Kurzstudie_Digitalisierung-der-Gesetzgebung-2021.pdf, S. 7.

22 S. zum Einsatz von LLM-Anwendungen im Gesetzgebungsprozess auch: Fitsilis/Mikros, AI-based solutions for legislative drafting in the EU, 2024, S. 22 ff.

23 Kommerzielle Anbieter stellen idR besonders leistungsfähige Modelle zur Verfügung. Die API-basierte Nutzung ermöglicht zudem eine schnelle prototypische Umsetzung ohne hohen Infrastruktur- und Wartungsaufwand. Eine spätere Migration auf Open-Source-Modelle mit lokaler Bereitstellung bleibt offen.

24 Besta et al., Reasoning Language Models: A Blueprint, 2025, arXiv:2501.11223.

25 Nach Breidenbach setzt sich der für den Gesetzgeber relevante Regelungskontext aus bestehenden Rechtsregimen einerseits und einschlägiger Rechtsprechung andererseits zusammen, vgl. Breidenbach, Was Gesetze sein könnten, 2025, S. 46.

26 Sprachmodelle stellen Momentaufnahmen ihrer Trainingsdaten dar. Ohne externe Datenanbindung bestünde daher die Gefahr, auf der Basis von im Modell kodierter veralteter Gesetzesinformation zu arbeiten, wodurch das Risiko faktisch unzutreffenden generierten Textes erhöht würde (sog. Halluzinationen).

Gesetzeswerke, bevor (2) für jedes Gesetz das zugehörige Inhaltsverzeichnis aus einer tagesaktuellen XML-Datei extrahiert und dem LLM bereitgestellt wird. (3) Auf Grundlage der Inhaltsverzeichnisse identifiziert das LLM einschlägige Paragraphen aus, deren (4) vollständige Wortlaute wiederum aus den XML-Dateien extrahiert und bereitgestellt werden. (5) Das LLM analysiert den jeweiligen Wortlaut und bestimmt die spezifisch betroffenen Absätze.

Darauf aufbauend erfolgt die Entwicklung möglicher Regelungsalternativen. Das LLM generiert auf Basis des Regelungskontexts mehrere abstrakte Regelungsvorschläge. Diese spiegeln unterschiedliche legistische Strategien zur Umsetzung des Vorhabens wider und ermöglichen es, verschiedene rechtstechnische Ansätze vergleichend zu betrachten. Auf abstrakter Ebene zu nennen seien hier etwa variierende Eingriffstiefen (zB punktuelle Änderung eines Absatzes vs. Neufassung eines gesamten Paragraphen), unterschiedliche Regelungstechniken (Freigrenze vs. Freibetrag; Ausnahmeregel vs. Generalklausel) oder abweichende systematische Ansätze (Integration in bestehende Tatbestände vs. Schaffung eines eigenständigen Absatzes).

Im dritten Schritt werden die vorgeschlagenen Alternativen anhand zuvor definierter Kriterien einer strukturierten Abwägung unterzogen. Diese erfolgt wiederum durch gezielte Rückfragen an das LLM, das Texte generiert, in denen Stärken und Schwächen der jeweiligen Alternative dargestellt und gegeneinander abgewogen werden. Hierbei fließen rechtstechnische Erwägungen, das Einfügen in bestehende Normsystematik sowie eine vorläufige Folgenabschätzung ein.

Schließlich wird im vierten Schritt auf Grundlage der vom Legisten gewählten Alternative sowie des relevanten Wortlauts der aktuellen Gesetzesfassung ein konkreter Änderungsvorschlag generiert. Das Modell formuliert dabei den Entwurf

einer konsolidierten Gesetzesfassung. Im parlamentarischen Verfahren werden Änderungen allerdings nicht als konsolidierten Endfassung der ihr zugehörigen Stammgesetze wiedergegeben, sondern als sich anschließende Änderungsbefehle beschlossen.²⁷ Eine naheliegende weitere Funktion wäre daher, den konsolidierten Änderungsentwurf automatisch (per LLM oder regelbasiert) in entsprechende Änderungsbefehle zu überführen.

VI. Anwendungsfall: Freistellung von Überstundenzuschlägen

Im Folgenden wird die Funktionsweise von LegisLLM beispielhaft anhand einer legistischen Aufgabe veranschaulicht.

Dabei wird kein vergangenes oder abgeschlossenes Gesetzesvorhaben herangezogen. Zwar könnte dieses mit beschlossenen und als „Goldstandard“ dienenden Gesetzesentwürfen verglichen werden, jedoch besteht das Risiko, dass die Modelle bereits auf dem verabschiedeten (bzw. öffentlichen) Wortlaut eben jener Entwürfe trainiert wurden und diesen lediglich reproduzieren. Dies würde eine realistische Bewertung der tatsächlichen Leistungsfähigkeit des Modells erschweren. Aus diesem Grund wird an dieser Stelle als legistische Aufgabe ein Beispiel aus dem aller Wahrscheinlichkeit nach noch nicht in LLM-Trainingsdaten befindlichen aktuellen Koalitionsvertrag zur steuerlichen Begünstigung von Überstundenzuschlägen gewählt:

„Zuschläge für Mehrarbeit, die über die tariflich vereinbarte bzw. an Tarifverträgen orientierte Vollzeitarbeit hinausgehen, werden steuerfrei gestellt. Als Vollzeitarbeit gilt dabei für tarifliche Regelungen eine Wochenarbeitszeit von mindestens 34 Stunden, für nicht tariflich festgelegte oder vereinbarte Arbeitszeiten von 40 Stunden.“²⁸

>
Deploy ⋮

LegisLLM - KI-gestützte Legistik

1 Aufgabenstellung
2 Regelungskontext
3 Regelungsalternativen
4 Evaluierung
5 Finalisierung

1. Legistische Aufgabenstellung

Beschreiben Sie die legistische Aufgabenstellung ?

Zuschläge für Mehrarbeit, die über die tariflich vereinbarte bzw. an Tarifverträgen orientierte Vollzeitarbeit hinausgehen, werden steuerfrei gestellt. Als Vollzeitarbeit gilt dabei für tarifliche Regelungen eine Wochenarbeitszeit von mindestens 34 Stunden, für nicht tariflich festgelegte oder vereinbarte Arbeitszeiten von 40 Stunden.

Press **⌘**+Enter to apply

Vorschläge:

Einführung einer Freigrenze für Einnahmen aus Vermietung und Verpachtung

Die Pendlerpauschale wird zum 01.01.2026 auf 38 Cent ab dem ersten Kilometer dauerhaft erhöht

Der Gewerbesteuer-Mindesthebesatz wird von 200 auf 280 Prozent erhöht

Abb. 2: User Interface des LegisLLM Prototypen

²⁷ Bundesministerium der Justiz, Handbuch der Rechtsförmlichkeit, 3. Aufl. 2024, Rn. 552 ff., abrufbar unter: https://www.bmj.de/SharedDocs/Publikationen/DE/Fachpublikationen/Handbuch_der_Rechtsfoermlichkeit.html.

²⁸ Angelehnt an: Koalitionsvertrag zwischen CDU, CSU und SPD, 2025, S. 46, abrufbar unter: <https://www.koalitionsvertrag2025.de>; Ergebnisse der Sondierungen von CDU, CSU und SPD v. 8.3.2025, S. 7, abrufbar unter: https://www.spd.de/fileadmin/Dokumente/Sonstiges/20250308_Sondierungspapier_CDU_CSU_SPD.pdf.

LegisLLM identifiziert auf Grundlage dieser Aufgabe diejenigen Vorschriften, die für die beabsichtigte Steuerbefreiung relevant sein können. Im konkreten Beispiel wird für relevant befunden:

- § 3 EStG (Katalog der steuerfreien Einnahmen)
- § 3b Abs. 1 EStG (Steuerfreiheit von Zuschlägen für Sonntags-, Feiertags- oder Nachtarbeit)
- § 19 Abs. 1 EStG (Bestimmung von Einkünften aus nicht-selbstständiger Arbeit)

- § 39b Abs. 2 EStG (Bestimmungen zum Lohnsteuerabzugsverfahren)
- § 52 Abs. 1 EStG (Zeitliche Anwendungsvorschriften)

Die ausgegebenen Normzitate bilden zusammen mit dem dahinterliegenden, tagesaktuellen Normtext fortan den Regelungskontext. Im nächsten Schritt werden hierfür mehrere, bewusst noch abstrakt gehaltene Lösungsansätze generiert. Diese umfassen jeweils eine Kurzbezeichnung, eine Beschreibung der Änderung und die von der Änderung betroffenen Rechtsnorm(en).

>

Zurück

Deploy

:

Weiter

Vorgeschlagene Änderungen:

Vorschlag 1: 1. Einfügung einer neuen Nummer in § 3 EStG

Alternative:

1. Einfügung einer neuen Nummer in § 3 EStG

Beschreibung:

Die steuerfreie Behandlung der Zuschläge für Mehrarbeit oberhalb einer tarifvertraglich vereinbarten Vollzeit (mindestens 34 bzw. 40 Stunden) wird als eigener Tatbestand in § 3 EStG verankert. Damit würden Zuschläge, die ausschließlich auf diese zusätzliche Arbeitszeit entfallen und über die normale Vollzeitarbeit hinausgehen, steuerfrei gestellt. Diese neue Nummer könnte die Voraussetzungen konkretisieren (Mindestwochenstunden für tarifliche und außertarifliche Beschäftigungen) und definieren, in welcher Höhe und unter welchen Bedingungen die Zuschläge entstehen.

Betroffene Rechtsnormen:

EStG § 3 (Einfügen einer neuen Nummer), ggf. korrespondierende Änderung in EStG § 52 (Anwendungszeitpunkt)

Vorschlag 2: 2. Erweiterung von § 3b EStG um einen neuen Absatz

Vorschlag 3: 3. Ergänzung in § 19 EStG zur Definition des Arbeitslohns

Abb. 3: Darstellung einer abstrakten Regelungsalternative zur Steuerfreistellung von Überstundenzuschlägen

Die gefundenen Regelungsalternativen werden im nächsten Schritt einer vorläufigen Abwägung unterzogen. Im konkreten Beispiel schlägt die Anwendung folgende, bereits nach Eignung sortierte, Regelungsalternativen vor:

- Kombination aus neuer Nummer in § 3 EStG und Verordnungsermächtigung

- Einfügung einer neuen Nummer in § 3 EStG
- Erweiterung § 3b EStG um einen neuen Absatz
- Ergänzung § 19 EStG zur Definition des Arbeitslohns

Fordert man etwa für die Regelungsalternative „Erweiterung § 3b EStG um einen neuen Absatz“ eine vertiefte Abwägung an, erhält man folgende Überlegungen (hier zusammengefasst):

	Pro	Contra
Juristische Beurteilung	Keine EU-rechtlichen Bedenken vor dem Hintergrund der Arbeitszeitrichtlinie	Fragen hinsichtlich Einfügung in den bisherigen Normzweck des § 3b EStG mit Blick auf Art. 3 GG
		Abgrenzungsprobleme zwischen Mehrarbeits- und sonstigen Zuschlägen
Rechtstechnik	Strukturelle Überschaubarkeit und formale Integration in § 3b EStG	Vermischung von Zuschlagstatbeständen führt zu Unschärferisiken
Dogmatik	Systematisches Einfügen in § 3b EStG	Mögliches Abweichen vom bisherigen Normzweck des § 3b EStG
Folgenabschätzung		Erhöhter Aufwand für Verwaltung/Unternehmen für Prüfung/Dokumentation von Arbeitszeitgrenzen und Mehrarbeitszuschlägen
		Missbrauchsgefahr (Unternehmen zahlen mehr „Überstundenzuschläge“ als regulären Lohn)

Entscheidet sich die Legistin bzw. der Legist nunmehr für die Implementierung der den § 3b EStG erweiternden Regelungsalternative, generiert die Anwendung folgenden Änderungsvorschlag (Änderungen in Fettdruck):

§ 3b Steuerfreiheit von Zuschlägen für Sonntags-, Feiertags- oder Nachtarbeit

(1) Steuerfrei sind Zuschläge, die für tatsächlich geleistete Sonntags-, Feiertags- oder Nachtarbeit neben dem Grundlohn gezahlt werden, (...).

(2) Grundlohn ist der laufende Arbeitslohn, der dem Arbeitnehmer bei der für ihn maßgebenden regelmäßigen Arbeitszeit für den jeweiligen Lohnzahlungszeitraum zusteht; (...).

(3) (...)

(4) Steuerfrei sind Zuschläge, die für tatsächlich geleistete Mehrarbeit neben dem Grundlohn gezahlt werden, soweit diese Mehrarbeit über die für Vollzeitarbeit maßgebende Wochenarbeitszeit hinausgeht. Als Vollzeitarbeit gilt dabei

1. bei tariflich vereinbarten oder an Tarifverträgen orientierten Arbeitszeiten eine Wochenarbeitszeit von mindestens 34 Stunden,

2. bei nicht tariflich festgelegten oder vereinbarten Arbeitszeiten eine Wochenarbeitszeit von 40 Stunden.

Mehrarbeit im Sinne dieses Absatzes liegt nur insoweit vor, als die tatsächlich geleistete Wochenarbeitszeit die jeweilige Grenze des Satzes 1 Nummer 1 oder Nummer 2 überschreitet. Für den Grundlohn gilt Absatz 2 entsprechend.

Das erzeugte Ergebnis wirkt insgesamt stimmig: Es verankert die Steuerfreiheit für Überstundenzuschläge systematisch im bestehenden § 3b EStG, wahrt die im legistischen Auftrag enthaltenen Schwellenwerte von 34 bzw. 40 Wochenstunden und greift sauber auf den bereits durch § 3b Abs. 2 EStG legaldefinierten Grundlohn-Begriff zurück.

VII. Herausforderungen und Ausblick

Auch wenn LegisLLM als lauffähiger Prototyp bereits vielversprechende Ergebnisse liefert, stellen sich gleichwohl wesentliche Herausforderungen.

Obgleich die Anwendung dem Legisten das mühsame sog. Prompt Engineering²⁹ abnehmen soll, besteht das Risiko, dass aus zu abstrakt gehaltenen Aufgabenstellungen mangelhafte Ergebnisse entstehen. Rückfrageschleifen des Modells könnten hier zur Präzisierung und Korrektur der Aufgabe beitragen und so die Ergebnisqualität deutlich verbessern.

Eine weitere kritische Komponente ist die Augmentierung³⁰ des Sprachmodells mit aktuellen und qualifiziert kuratierten Inhalten zur Verbesserung der legistischen

Qualität. Der Regelungskontext wird aktuell allein unter Hinzunahme der aktuellen Gesetzesfassung generiert. In Zukunft soll dieser Ansatz um einen Knowledge-Graphen angereichert, der den Regelungskontext und Abhängigkeiten des zu ändernden Gesetzes repräsentiert.³¹ Auf diese Weise können potenzielle normsystematische Wechselwirkungen, etwa durch Querverweise oder Folgeänderungen, idealerweise frühzeitig erkannt und transparent gemacht werden. Die Idee, Gesetze strukturell durch Wissensgraphen abzubilden, wurde bereits vor dem Aufkommen großer Sprachmodelle untersucht.³² Diese können insbesondere dazu dienen, komplexe Hierarchien, (implizite) Verweise und zuletzt auch die zeitliche Dimension in sich ändernden Rechtsdokumenten sichtbar zu machen. Große Sprachmodelle können hierauf im Wege der Augmentierung zurückgreifen.³³

Auch in der inhaltlichen Anreicherung der zugrunde liegenden Prompting-Strategie liegt Weiterentwicklungspotenzial. Erfahrungen aus der Ressortabstimmung, das Handbuch der Rechtsförmlichkeit³⁴ oder andere Leitlinien in der Gesetzgebungspraxis können miteinbezogen werden. Auch die ergänzende Integration einer Gesetzesfolgenabschätzung ist denkbar, um die Abwägungen von Regelungsalternativen zu verbessern.

LegisLLM demonstriert als lauffähiger Prototyp, wie KI-gestützte Assistenzsysteme den legislativen Alltag schon heute spürbar entlasten können und die fachliche Expertise der Legistinnen und Legisten wieder in den Mittelpunkt rücken können. In seinem aktuellen Entwicklungsstand kann die Anwendung bereits das sensible Modell-Prompting und die aufwendige Datenaugmentierung übernehmen. Die Verbindung von Praxiserfahrungen aus dem Gesetzgebungsalltag mit spezifischem Fachwissen (hier: Steuerrecht) und technisch strukturierter Wissensverarbeitung schafft eine robuste

²⁹ Prompt Engineering ist die gezielte Optimierung von Eingaben (Prompts) im Hinblick auf eine gewünschte Ausgabe, vgl. Sahoo/Singh/Saha/Jain/Mondal/Chadha, A Systematic Survey of Prompt Engineering in Large Language Models: Techniques and Applications, 2025, arXiv:2402.07927.

³⁰ Der Begriff der Augmentierung wird typischerweise bei der Anreicherung des Eingabekontexts für das Sprachmodell mit Ergebnissen einer separaten Suchanfrage verwendet (sog. Retrieval-Augmented Generation), s. Patrick/Perez/Piktus/Petroni/Karpukhin/Goyal/Küttler/Lewis/Yih/Rocktäschel, Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks, 2021, arXiv:2005.11401.

³¹ Wissensgraphen stellen Informationen als Knotenpunkte und ihre Beziehungen untereinander als Kantenverbindungen strukturiert dar. Zur semantischen Wissensrepräsentation für juristische KI-Systeme durch Wissensgraphen s. Paschke KIR 2024, 86 (88).

³² So etwa von Coupette/Hartung/Katz, Legal hypergraphs, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 2024, Vol. 382, No. 2270.

³³ Zur Retrieval-Augmented Generation (RAG) im Allg. Patrick/Perez/Piktus/Petroni/Karpukhin/Goyal/Küttler/Lewis/Yih/Rocktäschel, Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks, 2021, arXiv:2005.11401.. Zu Graph RAG s. de Martim, Graph RAG for Legal Norms: A Hierarchical and Temporal Approach, 2025, arXiv:2505.00039.

³⁴ Das HdB der Rechtsförmlichkeiten enthält Empfehlungen zur Form und Gestaltung von Gesetzen und Rechtsverordnungen auf Bundesebene. Es richtet sich vorwiegend an Legistinnen und Legisten, die mit der Erstellung oder rechtlichen Mitprüfung von Gesetzes- und Verordnungsentwürfen nach § 46 Abs. 1 GGO betraut sind, vgl. Bundesministerium der Justiz, Handbuch der Rechtsförmlichkeit, 3. Aufl. 2024, abrufbar unter: https://www.bmj.de/SharedDocs/Publikationen/DE/Fachpublikationen/Handbuch_der_Rechtsfoermlichkeit.html.

Grundlage zur achtsamen und wertschöpfenden Nutzung von Sprachmodellen im Sinne der Ziele der Machbarkeitsstudie.

Schnell gelesen...

- ▶ Die Machbarkeitsstudie „Digitaltaugliches Steuerrecht“ zeigt auf, wie KI-gestützte Methoden ein hochkomplexes, zeitkritisches Gesetzgebungsverfahren unterstützen können.
- ▶ Interviews mit Legistinnen und Legisten verdeutlichen Zeitdruck, manuelle Dokumentenarbeit und mangelnde Digitaltauglichkeit als zentrale Schwachstellen der gegenwärtigen Gesetzgebungspraxis.
- ▶ Der Anwendungsprototyp LegisLLM setzt hier an, indem es den Weg vom politischen Auftrag bis zum ersten Änderungsgesetzentwurf in Teilschritte zerlegt und iterativ mit einem großen Sprachmodell kommuniziert.
- ▶ Das Beispiel der Steuerfreistellung von Überstundenzuschlägen belegt die Praxistauglichkeit, während offene Punkte wie präzise Augmentierung weiteren Forschungsbedarf markieren.



Andreas Schultz

ist wissenschaftlicher Mitarbeiter der Arbeitsgruppe Legal Tech der Technischen Universität München und Doktorand am Zentrum für Digitalisierung des Steuerrechts der Ludwig-Maximilians-Universität München (LMUDigiTax).



Melanie Pospisil

ist studentische Mitarbeiterin der Arbeitsgruppe Legal Tech der Technischen Universität München und Masterstudentin im Studiengang "Mathematics in Data Science".



Kathleen Jennrich

ist Digitalstrategin im Bundesministerium der Finanzen mit dem Schwerpunkt digitaltaugliches Recht, Interoperabilität und Once-Only.



Prof. Matthias Grabmair, Ph. D., LL. M.,

ist Tenure-Track Assistant Professor für Legal Tech im Department for Computer Science der School for Computation, Information and Technology der Technischen Universität München.