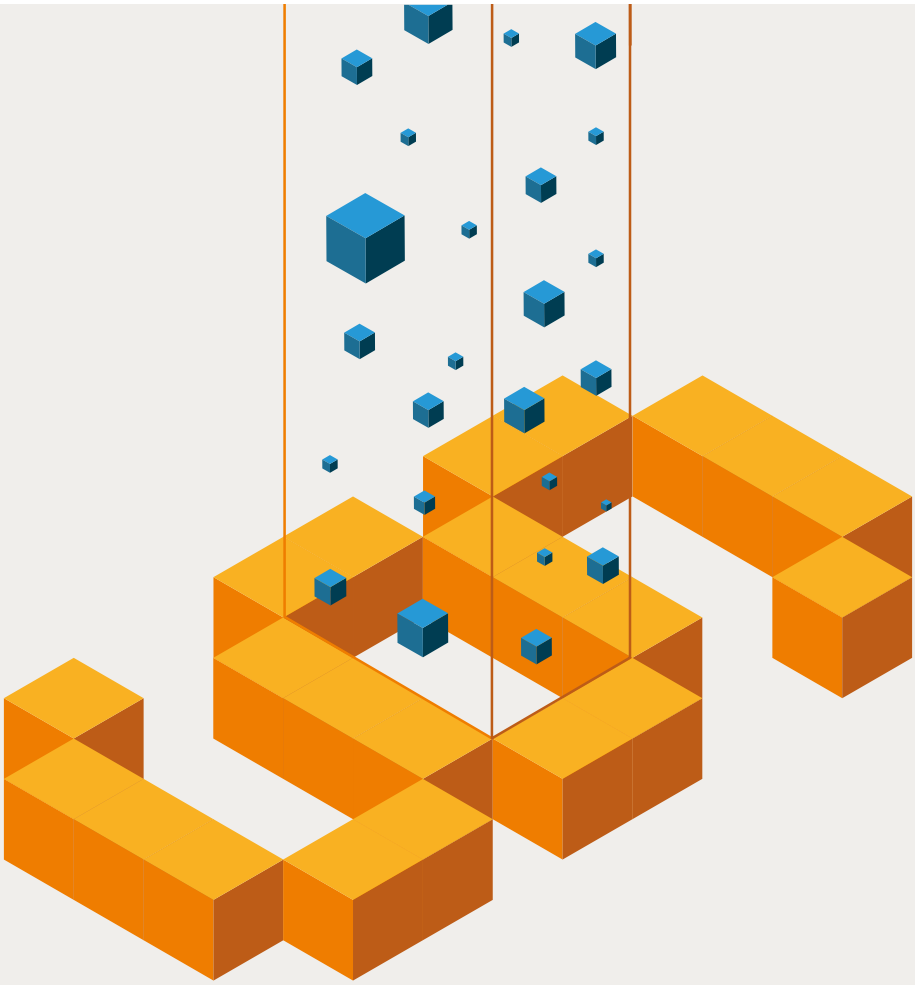
d-fine

Abschlussbericht (Kurzfassung)

Machbarkeitsstudie zu einer KI-gestützten Anwendung zur Identifikation des Vereinfachungspotentials von Regelungen

A&O SHEARMAN · Fraunhofer IAIS · INSTITUT FÜR DAS RECHT DER DIGITALISIERUNG · LEXEMO



analytisch. technologisch. quantitativ.

Juni 2025

Impressum

Studie im Auftrag des Bundesministeriums der Finanzen zum Forschungsvorhaben 5/23.

Erstellt durch
d-fine GmbH
Kurfürstendamm 21
10719 Berlin
Deutschland

www.d-fine.com

Unterauftragnehmer
A&O Shearman, Fraunhofer-Institut für Intelligente Analyse- und Informationssysteme IAIS (Fraunhofer IAIS), Lexemo GmbH, Prof. Dr. Florian Möslein (Institut für das Recht der Digitalisierung der Universität Marburg)

Hinweis: Das Fraunhofer IAIS ist weder Herausgeber des Mediums noch anderweitig haftungsbegründend eingeschaltet.

Studienleitung
Dr. Ferdinand Graf (d-fine)
Marc Girshgorn (d-fine)

Bearbeitet von
Dr. Petra Gutjahr (d-fine), Prof. Dr. Florian Möslein (Universität Marburg), Pascal Di Prima (Lexemo), Katharina Beckh (Fraunhofer IAIS), Dr. Jens Wagner (A&O Shearman)

Gestaltung
www.visionaere.de

17. Juni 2025

Executive Summary

Der Abbau unnötiger bürokratischer Hürden ist ein zentrales politisches Ziel in Deutschland und der Europäischen Union (EU), um Wachstum und Beschäftigung zu fördern. Die hohe Regulierungsdichte auf EU-, Bundes- und Landesebene verursacht oft Doppelungen, Überschneidungen oder gar Widersprüche, die den Aufwand für die Wirtschaft erhöhen und die Rechtssicherheit beeinträchtigen. Vor diesem Hintergrund beauftragte das Bundesministerium der Finanzen (BMF) diese Machbarkeitsstudie zur Entwicklung einer KI-Anwendung, die laufend Potentiale zur Vereinfachung bestehender sowie neuer Regelungen identifiziert und bei der Rechtsbereinigung unterstützt.

Die Studie zeigt, dass eine solche KI-gestützte Anwendung realisierbar ist, sofern die erforderlichen technischen und organisatorischen Rahmenbedingungen geschaffen werden:

01. **Erfassen geltender Regelungen:** Für eine belastbare KI-Analyse ist eine umfassende, aktuelle und qualitätsgesicherte Datenbasis der geltenden Regelungen und ihrer Auslegungen essenziell. Die Vielzahl und Heterogenität relevanter Datenquellen sowie Zugriffsbeschränkungen erschweren dabei eine vollständige Abdeckung. Zur praktischen Umsetzung empfiehlt die Studie daher, zunächst auf frei zugängliche staatliche Rechtsnormen zu fokussieren.

02. **Geeigneter methodischer Ansatz:** Die Studie ermittelt zentrale Voraussetzungen für eine erfolgreiche KI-gestützte Identifikation und Analyse der Vereinfachungspotentiale und schlägt geeignete methodische Ansätze vor. Zu den Voraussetzungen gehören die Fähigkeit, Rechtstexte kontextsensitiv zu analysieren, darauf aufbauend Geltungsbereiche und Regelungen rechtlich sachgerecht zu vergleichen sowie die Ergebnisse nachvollziehbar und überprüfbar zu begründen. Die möglichen methodischen Ansätze kombinieren dazu mehrere KI-Verfahren. Dabei werden kurze, mit Kontext angereicherte Textpassagen von Rechtsnormen automatisiert vorausgewählt, mithilfe von Sprachmodellen miteinander verglichen und die Ergebnisse für die Qualitätssicherung mit weiteren Informationen angereichert.

03. **Verlässlichkeit der Ergebnisse:** Eine hohe Verlässlichkeit der vorgeschlagenen Ansätze lässt sich durch geeignete automatisierte Vorbereitungs- und Validierungsschritte erzielen. Ergänzend dazu bleibt eine manuelle Überprüfung der KI-Ergebnisse durch die zuständige Fachebene unerlässlich.

Für die mögliche Entwicklung sollte ein iteratives und modulares Vorgehen gewählt werden, bei dem zuerst ein oder mehrere Prototypen für Kernfunktionalitäten in kurzer Zeit umgesetzt und vom BMF direkt erprobt werden. Bei Funktionalitäten, die nicht primär die Regelungsvereinfachung, sondern übergreifend die Rechtsetzung betreffen, sollten Synergien mit bestehenden nationalen und europäischen Initiativen wie E-Gesetzgebung und Legislation Editing Open Software (LEOS) gesucht werden.

Langfristig hat die KI-basierte Analyse das Potential, sowohl einen substanziellen Beitrag zum Bürokratieabbau zu leisten als auch den manuellen Aufwand bei der Erarbeitung von Gesetzen zu reduzieren.

Inhaltsverzeichnis

Executive Summary 3

01. Einleitung 6

02. Juristische Wissensbasis einer KI-Anwendung bereitstellen 8
- 02.01 Klassifizierung von Normen und Auslegungen 8
- 02.02 Analyse und Bewertung möglicher Datenquellen 9
- 02.03 Fazit 11

03. Vereinfachungspotentiale mit KI identifizieren 12
- 03.01 Ziele und methodischer Rahmen 12
- 03.02 Voraussetzungen für eine KI-gestützte Analyse 13
- 03.03 Machbarkeitsanalyse KI-gestützter Ansätze 14
- 03.04 Fazit 18

04. Verlässlichkeit KI-gestützter Ergebnisse optimieren 20
- 04.01 Theoretische Bewertung der Verlässlichkeit 20
- 04.02 Prototypische Erprobung 22
- 04.03 Fazit 23

05. Rechtsbereinigung mit KI unterstützen 24
- 05.01 Rechtsbereinigung und Potentiale für KI-Einsatz 24
- 05.02 Machbarkeitsanalyse KI-gestützter Ansätze 24
- 05.03 Fazit 25

06. KI-Anwendung umsetzen: Eine mögliche Roadmap 26

07. Gesamtfazit und Ausblick 27

08. Literaturverzeichnis 28

09. Ansprechpersonen 31

10. Glossar 32

01. Einleitung

Der Abbau unnötiger bürokratischer Hürden ist ein zentrales politisches Ziel in Deutschland, um Wachstum und Beschäftigung nachhaltig zu stärken (Die Bundesregierung, 2025; Nationaler Normenkontrollrat, 2024a, S. 82ff.). Auch die Europäische Kommission (Europäische Kommission, 2025a; von der Leyen, 2024) verweist auf den Bedarf an Vereinfachung, Konsolidierung und Kodifizierung der Rechtsvorschriften, um Überschneidungen und Widersprüche zu beseitigen und gleichzeitig hohe Standards zu wahren und hat zu diesem Zweck bereits erste vereinfachende Rechtsakte vorgeschlagen (Europäische Kommission, 2025a). Die rasanten Fortschritte im Bereich Künstlicher Intelligenz (KI) und deren zunehmender Einsatz in der öffentlichen Verwaltung und im rechtlichen Kontext legen nahe, dass KI-Methoden dabei maßgeblich unterstützen können (vgl. Tabelle 1).

Vor diesem Hintergrund beauftragte das Bundesministerium der Finanzen (BMF) die vorliegende Machbarkeitsstudie. Sie untersucht, inwieweit eine KI-Anwendung verlässlich und laufend Vereinfachungspotentiale bestehender und neu vorgeschlagener Regelungen ermitteln und bei der Bereinigung des Rechts unterstützen kann. Die etwaige KI-Anwendung ist als unterstützendes Werkzeug für die Ressorts der Bundesregierung angedacht, wobei die Verantwortung und Entscheidungsbefugnis uneingeschränkt bei der zuständigen Fachebene verbleiben. Alle KI-generierten Ergebnisse, wie identifizierte Verdachtsfälle auf Vereinfachungspotentiale, müssen daher von den Beschäftigten eines Ressorts überprüft und gegebenenfalls angepasst werden. Zudem gelten hohe Anforderungen an Verlässlichkeit, Erklärbarkeit und Verifizierbarkeit der KI-Ergebnisse.

Die Machbarkeit wird anhand ausgewählter KI-basierter Kernfunktionalitäten untersucht, die für die mögliche Anwendung als essenziell eingestuft werden. Sie umfassen die Ermittlung von Regelungen mit überlappendem Geltungsbereich und anschließende Identifikation der Vereinfachungspotentiale in diesen Regelungen. Der methodische Ansatz der Studie kombiniert dabei interdisziplinäre Forschungs- und Anwendungsperspektiven aus den Bereichen Rechtsetzung und Künstlicher Intelligenz. Grundlage der Ergebnisse sind eine Auswertung aktueller wissenschaftlicher Publikationen, Einschätzungen relevanter Fachexpertinnen und Fachexperten sowie punktuelle prototypische Tests vielversprechender Ansätze.

Die Studie wurde unter Federführung von d-fine in Kooperation mit A&O Shearman, dem Fraunhofer-Institut für Intelligente Analyse- und Informationssysteme (Fraunhofer IAIS), Lexemo sowie Prof. Dr. Florian Möslein (Institut für das Recht der Digitalisierung) durchgeführt und erstreckte sich über ca. sechs Monate (Mitte August 2024 bis Ende Januar 2025).

Zur besseren Übersicht ist die Studie in sechs Kapitel gegliedert. Kapitel 02 analysiert die Bereitstellung von Regelungen und Auslegungshilfen als Wissensbasis für die etwaige KI-Anwendung, zeigt zentrale Herausforderungen auf und gibt Handlungsempfehlungen. Kapitel 03 widmet sich der Machbarkeit der Identifikation der Vereinfachungspotentiale wie Doppelungen, Widersprüchen und Inkonsistenzen mit KI, legt die Anforderungen an KI-basierte Methoden dar und schlägt konkrete methodische Ansätze vor. Kapitel 04 bewertet die Verlässlichkeit der erzielbaren KI-Ergebnisse anhand theoretischer Analysen und prototypischer Tests. Kapitel 05 analysiert die Möglichkeiten KI-gestützter Vorschläge zur Rechtsbereinigung. Darauf aufbauend skizziert Kapitel 06 eine Roadmap zur möglichen Umsetzung der Anwendung.

Ausgewählte Initiativen mit Bezug zu dieser Studie	
Deutschland	Zum Thema **„Modernisierung des Rechtsetzungskreislaufs"** führt die Bundesregierung zahlreiche Projekte, wie die E-Gesetzgebung (Verwaltung Innovativ, 2023), E-Verkündung (Der Beauftragte der Bundesregierung für Informationstechnik, 2025) und das Neue Rechtsinformationssystem des Bundes (kurz NeuRIS (Bundesministerium für Justiz, 2024)) durch, um den Gesetzgebungsprozess von der Vorphase bis zur Verkündung und dem Vollzug zu digitalisieren.[1] Der **NKR sieht großes Potential für den Einsatz von KI-Methoden in der Gesetzgebung**, etwa bei der Erfüllung von Routineaufgaben sowie in der frühen Phase der Erarbeitung von Rechtstexten, um die Vollzugs- und Digitaltauglichkeit zu verbessern und unnötige Bürokratie zu vermeiden (Nationaler Normenkontrollrat, 2024b). Exemplarisch wurde mit Hilfe von KI-Methoden das Kostenbarometer Regulatorik (Kostbar) erstellt (Deloitte, 2021).
EU	**LEOS** (**L**egislation **E**diting **O**pen **S**oftware) (Europäische Kommission, 2025b) ist ein Web-Tool zur Unterstützung der Ausarbeitung von Rechtsvorschriften in der EU. In zwei Studien (Fitsilis, Mikros, & Leventis, 2024; Palmirani, Vitali, Van Puymbroeck, & Nubla Durango, 2022) wurde untersucht, wie LEOS durch KI-basierte „smarte Funktionalitäten" erweitert werden könnte und skizziert eine Roadmap für deren Umsetzung. **Publio** (Europäische Union, 2025a), der virtuelle Assistent des Amts für Veröffentlichungen, unterstützt Nutzende in Form eines mehrsprachigen Chatbots u. a. bei der Suche nach Inhalten in EUR-Lex. Er nutzt Microsoft Language Understanding (LUIS) (Microsoft, 2025) in Kombination mit einem großen generativen Sprachmodell (Europäische Union, 2023).

Tabelle 1: Ausgewählte Initiativen mit Bezug zu dieser Studie

1 Für einen detaillierten Überblick siehe z. B. Nationaler Normenkontrollrat (2023, S. 29ff.).

02. Juristische Wissensbasis einer KI-Anwendung bereitstellen

Um mögliche Vereinfachungen in bestehenden Normen zu erkennen, benötigt eine KI-Anwendung aktuelles und umfassendes Wissen über die geltenden Regelungen. Zugleich darf sie nur solche Vorschriften berücksichtigen, die von regelungsbefugten Normgebern wie Parlamenten, Behörden oder nicht-staatlichen Organisationen, soweit ihnen Normsetzungsbefugnis verliehen worden ist, erlassen wurden. Für eine fundierte Analyse und Qualitätssicherung sind zudem Auslegungshilfen wie Rechtsprechung und Gesetzesbegründungen wichtig. Dieses Kapitel untersucht daher die verfügbaren juristischen Datenquellen[2] hinsichtlich ihrer Eignung und Möglichkeit der Einbindung in die etwaige KI-Anwendung.

02.01 Klassifizierung von Normen und Auslegungen

Das Recht wird von einer Vielzahl von Rechtsnormen auf EU-, Bundes- und Länderebene geprägt. Zusätzlich existieren teilweise bindende Vorschriften der Behörden wie des Bundesamts für Wirtschaft und Ausfuhrkontrolle (BAFA), der Berufsgenossenschaften sowie weiterer Organisationen, wie der Internationalen Organisation für Wertpapieraufsicht (IOSCO). Für die Zwecke dieser Machbarkeitsstudie schlagen wir folgende Klassifizierung der relevanten Normen und Auslegungshilfen vor:

Ebene	Erläuterung
1. (Stamm-)Gesetze und Basisrechtsakte	Auf der EU-Ebene zählen hierzu Verordnungen, Richtlinien und Beschlüsse. Auf nationaler Ebene umfassen sie Bundes- und Landesgesetze (einschließlich der Umsetzungsgesetze europäischer Richtlinien). Beispiele sind: • Datenschutz-Grundverordnung (DSGVO)[3] • Europäische Lieferkettenrichtlinie (CSDDD)[4] • Lieferkettensorgfaltspflichtengesetz (LkSG)[5]
2. Nachgelagerte Regelungen sowie Normen und Standards	Hierzu zählen alle Regelungen, die auf Grundlage von Gesetzen und Basisrechtsakten erlassen werden und die dort enthaltenen Regelungen im Detail ausgestalten oder ihre Durchführung regeln, wie delegierte Rechtsakte, Durchführungsrechtsakte und -verordnungen, Rechtsverordnungen, Verwaltungsvorschriften und -richtlinien (z. B. allgemeine Verwaltungsvorschriften, Rundschreiben, Erlasse und Leitlinien von Behörden) sowie technische Normen und Standards. Beispiele sind: • Delegierte Rechtsakte und Durchführungsrechtsakte zur DSGVO • EDSA-Leitlinien und Empfehlungen • BSI-Standard 200-2 „IT-Grundschutz“ • DIN- oder ISO-Normen
3. Rechtsprechung (Judikative Auslegung)	Hierzu zählen Urteile und Beschlüsse europäischer und deutscher Gerichte, die zur Auslegung und Konkretisierung der Rechtsnormen beitragen.

2 Die bisherigen Erkenntnisse zur Nutzung Künstlicher Intelligenz im Rechtskontext zeigen, dass KI-Modelle allein auf Basis ihres inhärenten Trainingswissens nicht in der Lage sind, verlässliche Antworten auf komplexe rechtliche Fragestellungen zu liefern (siehe z. B. die Studie von Dahl et al. (2024)). Um das Vereinfachungspotential von Regelungen identifizieren zu können, muss die etwaige KI-Anwendung daher einen möglichst umfassenden Zugriff auf die juristische Datenbasis haben.
3 Verordnung (EU) 2016/679, ABl. (EU) L 119 vom 4.5.2016, S. 1–88 (aktuell geltende Fassung).
4 Richtlinie (EU) 2024/1760, ABl. (EU) L, 2024/1760, 5.7.2024 (aktuell geltende Fassung).
5 Gesetz über die unternehmerischen Sorgfaltspflichten zur Vermeidung von Menschenrechtsverletzungen in Lieferketten, BGBl. 2021 I S. 2959 (aktuell geltende Fassung).

4. Juristische Literatur, Kommentare und Auslegungshilfen von Normgebern	Diese Quellen dienen als Auslegungshilfen und umfassen insbesondere Interpretationshilfen, die direkt vom Normgeber veröffentlicht werden, juristische Kommentare zu den Regelungen und sonstige relevante juristische Fachliteratur. Zu diesen Quellen zählen auch Gesetzesbegründungen.

Tabelle 2: Klassifizierung von Regelungen und Auslegungshilfen im Kontext der Studie

Die Machbarkeitsanalyse fokussiert sich in den folgenden Kapiteln primär auf Rechtsnormen der Ebene 1, da diese die Rechtsgrundlage für nachgelagerte Regelungen darstellen und damit den größten Einfluss auf den bürokratischen Aufwand haben. Ergänzend werden nachgelagerte Regelungen der Ebene 2 berücksichtigt. Für eine präzise Interpretation und Anwendung der Gesetze werden Gesetzesbegründungen und behördliche Auslegungen sowie Rechtsprechung und ggf. juristische Literatur (Ebenen 3 und 4) als Auslegungshilfen berücksichtigt, um etwaige Unklarheiten aufzudecken und zu klären.

02.02 Analyse und Bewertung möglicher Datenquellen

Datenquelle	Staatl. RIS			Legislative			Amtsblätter			Behörden			Gerichte			NSO	juris/beck-o.
	EU	B	L	EU	B	L	EU	B	L	EU	B	L	EU	B	L		
Ebene																	
1. (Stamm-) Gesetze und Basisrechtsakte	●	●	●	1	1	1	●	●	●								●
2. Nachgelagerte Regelungen	●	○	●				●	●	●	●	●	○					●
2. Regelungen Nicht-Staatlicher Normgeber																●	○
3. Gerichtsurteile	●	●	●										●	●	○		●
4. Auslegungshilfen Normgeber	●			●	●	●				●	●	●				●	●
4. Auslegungshilfen Nicht-Normgeber																	●

Vollständigkeit: Stark lückenhaft ○ ● ● ● ● Weitestgehend vollständig

Hinweise:

1 In den Parlamentsdatenbanken sind in der Regel keine verkündeten Fassungen der Gesetze und Basisrechtsakte enthalten.

Abkürzungen:

Staatl. RIS	Staatliche Rechtsinformationssysteme[6]	B	Bundesebene
NSO	Offizielle Webseiten nicht-staatlicher Organisationen (z. B. Berufsgenossenschaften, Standardsetzer)	L	Landesebene
EU	Europäische Ebene	juris/beck-o.	Die Angebote der juris GmbH und von beck-online

Tabelle 3: Vollständigkeit der Datenquellen

6 Zu den betrachteten staatlichen Rechtsinformationssystemen gehören EUR-Lex, Gesetze-im-Internet, Verwaltungsvorschriften-im-Internet, Rechtsprechung-im-Internet sowie die Rechtsportale der Länder (z. B. RECHT.NRW.DE oder BAYERN.RECHT).

Die Identifikation und Integration geeigneter Datenquellen für Regelungen und Auslegungshilfen ist eine zentrale Voraussetzung für die Machbarkeit der etwaigen KI-Anwendung. Um verlässliche Ergebnisse zu erzielen, benötigt die Anwendung eine Datenbasis aus qualitativ hochwertigen, aktuellen und relevanten Texten.

Damit ausschließlich „echte“ Normen berücksichtigt werden, sollten Regelungstexte dabei nur aus vertrauenswürdigen Informationsquellen wie staatlichen Rechtsinformationssystemen (RIS) oder offiziellen Webseiten der Normgeber stammen. Darüber hinaus spielen die Vollständigkeit und die Einfachheit der Integration der Datenquelle sowie mögliche Lizenz- und Nutzungsbeschränkungen bei der Auswahl eine wichtige Rolle. So hängt der mit der Integration und Aktualisierung der Datenbasis verbundene manuelle Aufwand beispielsweise davon ab, ob die Datenquelle konsolidierte Fassungen geltender Regelungen bereitstellt, wie leicht diese auffindbar sind und wie zeitnah sie aktualisiert werden.

In Tabelle 3 werden potenzielle Datenquellen hinsichtlich ihrer Vollständigkeit bewertet, während Tabelle 4 die Integrationsfähigkeit dieser Quellen in die etwaige KI-Anwendung untersucht.

Datenquelle	Staatl. RIS			Legislative			Amtsblätter			Behörden			Gerichte			NSO	juris/beck-o.
	EU	B	L	EU	B	L	EU	B	L	EU	B	L	EU	B	L		
Kategorie																	
Auffindbarkeit z. B. über zentrales Portal	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
													1		3		1
Konsolidierte Fassungen	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
															3		
Maschinenlesbarkeit und API-Zugriff	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
															3		
Keine Lizenz-beschränkungen	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	
		2	2	2		2		2	2		2	2	2	2	2;3	2	
Verzicht auf Paywalls	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	
															3		

Erfüllung der Kategorie: Kaum erfüllt ● ● ● ● ● Weitestgehend erfüllt

Hinweise:

1 Teilweise findet lediglich eine Verlinkung auf EUR-Lex statt.

2 Teilweise fehlen für Datenquellen (explizite) Angaben zu Lizenzbeschränkungen bzw. es ist unklar auf welche Quellen sie sich beziehen. Dies hat in der Regel zu einer Abwertung geführt.

3 Bezieht sich ausschließlich auf die Verfassungsgerichtsbarkeit der Länder, da Entscheidungen der ordentlichen Gerichtsbarkeit in der Regel nur in den staatlichen RIS zu finden sind.

Abkürzungen:

Staatl. RIS	Staatliche Rechtsinformationssysteme[7]
NSO	Offizielle Webseiten nicht-staatlicher Organisationen (z. B. Berufsgenossenschaften, Standardsetzer)
EU	Europäische Ebene
B	Bundesebene
L	Landesebene
juris/beck-o.	Die Angebote der juris GmbH und von beck-online

Tabelle 4: Einfachheit der Integration und Verzicht auf Zugriffsbeschränkungen bei Datenquellen

7 Siehe Fußnote 6 auf Seite 9.

Das überwiegend im Bundeseigentum befindliche Rechtsinformationssystem juris sowie das private System beck-online bieten einen umfassenden Zugriff auf staatliche Normen und relevante Auslegungshilfen. Zudem bieten sie zusätzliche Vorteile durch vorhandene Klassifizierung und Verlinkung der Rechtstexte und Auslegungen, was die Analyse durch die KI signifikant vereinfacht. Allerdings sind beide kostenpflichtig und unterliegen Nutzungsbeschränkungen, die eine Auswertung durch KI derzeit untersagen (C.H.Beck oHG, 2025; juris GmbH, 2025).

In Bezug auf die weiteren Datenquellen zeigt die Analyse Folgendes:

- **Ebene 1:** Eine weitgehend vollständige Einbindung der Stammgesetze und Basisrechtsakte ist ebenfalls über staatliche Rechtsinformationssysteme wie *EUR-Lex, Gesetze-im-Internet* und *Rechtsportale der Länder* möglich.
- **Ebene 2:** Staatliche RIS wie *EUR-Lex, Verwaltungsvorschriften-im-Internet* sowie *Rechtsportale der Länder* enthalten (wie juris und beck-online) nur einen Teil der von nachgeordneten Behörden in der EU und Deutschland erlassenen Vorschriften. Eine umfassende Erfassung weiterer Regelungen wäre mit sehr hohem Aufwand verbunden, da diese meist nur über die Amtsblätter sowie Webseiten der zuständigen Behörden oder Organisationen in uneinheitlichen Formaten verfügbar sind. Normen nicht-staatlicher Organisationen sind bis auf wenige Ausnahmen nicht über zentrale Portale verfügbar und sind zudem häufig lizenz- und kostenpflichtig, was deren Integration zusätzlich erschwert.[8]
- **Ebene 3:** Aktuelle veröffentlichte Rechtsprechung, insbesondere höherer Instanzen, ist weitgehend über staatliche RIS *EUR-Lex* und *Rechtsprechung-im-Internet* sowie *Rechtsportale der Länder* verfügbar. Eine zentrale Herausforderung besteht allerdings darin, dass in Deutschland keine umfängliche Veröffentlichungspflicht für Urteile besteht und somit geschätzt nur ein Bruchteil der Urteile tatsächlich öffentlich zugänglich ist (Coupette & Fleckner, 2018; Hamann, 2021).
- **Ebene 4:** Gesetzesbegründungen können über die Webseiten der Parlamente abgerufen werden. Juristische Kommentare und sonstige juristische Fachliteratur werden neben juris bzw. beck-online auch von anderen Verlagen bereitgestellt, sind jedoch dort auch lizenzierten Zugriffsbeschränkungen unterworfen und zahlungspflichtig.

02.03 Fazit

Eine qualitätsgesicherte, aktuelle und umfassende Datenbasis von Regelungen ist essenziell, um eine KI-gestützte Analyse der Vereinfachungspotentiale zu ermöglichen. Zudem sind relevante Auslegungen (insb. Urteile und Gesetzesbegründungen) für die Analyse und Qualitätssicherung wichtig. Die Analyse juristischer Datenquellen zeigt, dass eine vollständige Integration aller Regelungsebenen in die etwaige KI-Anwendung, insbesondere nachgelagerter Regelungen von Behörden, Normen nicht-staatlicher Organisationen und Gerichtsentscheidungen, derzeit kaum realisierbar ist. Dies liegt an der fragmentierten Datenlandschaft, unvollständigen Veröffentlichungen sowie Lizenz- und Nutzungsbeschränkungen.

Um die Machbarkeit der KI-Anwendung zu gewährleisten, sollte der Fokus zunächst auf staatliche Rechtsnormen und Auslegungen gelegt werden, die über offizielle Rechtsinformationssysteme frei zugänglich und rechtlich unbedenklich integriert werden können. Das mehrheitlich im Bundeseigentum stehende Rechtsinformationssystem juris könnte ebenfalls eine umfassende Datenbasis bereitstellen, inklusive relevanter Auslegungshilfen, sofern Lizenzfragen geklärt werden können.

8 Siehe für lizenzrechtliche Einschränkungen zum Beispiel Bank of International Settlements (2025).

03. Vereinfachungspotentiale mit KI identifizieren

Im vorherigen Kapitel wurde untersucht, wie eine umfassende Datenbasis von geltenden Regelungen und relevanten Auslegungshilfen für eine KI-Anwendung bereitgestellt werden kann. Darauf aufbauend untersucht dieses Kapitel, ob und wie sich mithilfe einer KI-gestützten Analyse Vereinfachungspotentiale in diesen Regelungen identifizieren lassen.

Zunächst werden, ausgehend von praktischen Beispielen, die wichtigsten Arten solcher Vereinfachungspotentiale beschrieben. Anschließend wird aufgezeigt, unter welchen Voraussetzungen ihre KI-basierte Identifikation möglich ist. Daraufhin werden zwei mögliche methodische Ansätze für die etwaige KI-Anwendung vorgestellt, die auf Erkenntnissen aus aktueller Forschung und jüngsten technischen Entwicklungen basieren.

03.01 Ziele und methodischer Rahmen

Infobox 1 beschreibt die wichtigsten Kategorien der Vereinfachungspotentiale im Kontext dieser Studie. Dabei liegt der Fokus auf Doppelungen, Widersprüchen und Inkonsistenzen, die in der Praxis zu bürokratischem Mehraufwand führen.[9]

Im Mittelpunkt der Studie stehen Vereinfachungspotentiale, die durch das Zusammenwirken mehrerer Regelungen entstehen. Sie lassen sich im Gesetzgebungsprozess besonders schwer vorhersehen und zeigen sich ggf. erst in der praktischen Anwendung. Das Auftreten dieser Vereinfachungspotentiale setzt voraus, dass die betroffenen Regelungen gleichzeitig zur Anwendung kommen können, also denselben oder zumindest überlappenden Geltungsbereich aufweisen. In diesen Fällen unterscheidet diese Studie zwischen den folgenden Kategorien:

- **Doppelung:** Eine Doppelung liegt vor, wenn mehrere Regelungen identische oder sehr ähnliche Rechtsfolgen – insbesondere Pflichten – vorsehen, die zu erhöhtem bürokratischem Aufwand für die Betroffenen führen. Mitunter werden die Pflichten ausdrücklich wiederholt, in anderen Fällen ergeben sie sich implizit.
  Beispiel: Ein Unternehmen muss mehrfach identische oder weitgehend vergleichbare Dokumentation bei unterschiedlichen Behörden einreichen oder dieselben Verfahren mehrfach durchlaufen.

- **Widerspruch:** Ein Widerspruch liegt vor, wenn Regelungen Rechtsfolgen festlegen, die sich in der Anwendung gegenseitig ausschließen können. Die Betroffenen können dann sämtliche Pflichten nur schwer oder gar nicht gleichzeitig erfüllen und es ist nicht klar, welche Regelung vorrangig gelten soll.
  Beispiel: Die Erfüllung einer Regelung erfordert in der Praxis die Erfassung personenbezogener Daten, die nach Maßgabe der DSGVO untersagt ist.

- **Inkonsistenz:** Eine Inkonsistenz liegt vor, wenn das Verhältnis zweier Regelungen unklar ist, weil deren Tatbestand und Rechtsfolge Pflichten vorsehen, die zwar nicht notwendig widersprüchlich sind, aber zu Verwirrung, unterschiedlichen rechtlichen Interpretationen oder ineffizienter Umsetzung führen können.
  Beispiel: Abweichende Definitionen zentraler Begriffe (z. B. „Einkommen“, „abhängige Beschäftigung“) führen bei den Betroffenen zu Verwirrung oder Unklarheit darüber, ob die Regelungen die gleichen Pflichten implizieren, was deren Rechtssicherheit beeinträchtigt.

Infobox 1: Zentrale Kategorien von Vereinfachungspotentialen im Kontext der Studie

9 Dies wird beispielsweise von Hacker (2024) anhand der Wechselwirkungen zwischen der KI-Verordnung und anderen Rechtsakten erläutert.

Wie in Infobox 1 dargestellt, setzt das Auftreten dieser Vereinfachungspotentiale voraus, dass betroffene Regelungen zumindest teilweise denselben Geltungsbereich aufweisen. **Daher empfiehlt sich ein zweistufiger methodischer Ansatz für die KI-Anwendung, der folgende Kernfunktionalitäten umfasst:**

- **Geltungsbereichsanalyse:** Zunächst werden Regelungen identifiziert, deren persönlicher, sachlicher, räumlicher und zeitlicher Geltungsbereich sich ganz oder teilweise überschneiden.
- **Identifikation der Vereinfachungspotentiale:** Anschließend werden die so eingegrenzten Regelungen auf mögliche Vereinfachungspotentiale wie Doppelungen, Widersprüche oder Inkonsistenzen hin untersucht.

Diese beiden Kernfunktionalitäten bilden die Basis für eine Reihe möglicher Anwendungsfälle.

03.02 Voraussetzungen für eine KI-gestützte Analyse

Ausgehend von den Praxisbeispielen und einer umfassenden Literaturrecherche im Rahmen der Studie wurden die folgenden zentralen Voraussetzungen für den methodischen Ansatz einer KI-Anwendung ermittelt, um die genannten Kernfunktionalitäten erfolgreich umzusetzen:

01. **Verständnis der Struktur und Terminologie von Rechtstexten:** Rechtstexte zeichnen sich durch eine komplexe, hierarchische Gliederung und die Verwendung juristischer Fachterminologie aus. Eine KI-Anwendung muss daher in der Lage sein, die relevanten Gliederungselemente und ihre Funktion (z. B. Begriffsdefinitionen, Regelungen in Hauptteil, Voraussetzungen, Ausnahmeregelungen) korrekt zu erkennen. Ebenso sollen gängige Rechtsbegriffe wie „Adressat“ oder „Verpflichtung“ korrekt interpretiert werden.

02. **Fähigkeit zur kontextsensitiven Textanalyse:** Die Bedeutung einzelner Bestimmungen in Rechtstexten erschließt sich in der Regel erst aus dem rechtlichen Kontext. So stehen sie in Wechselwirkung nicht nur mit anderen Textpassagen desselben Rechtstextes, sondern auch mit weiteren Gesetzen, Richtlinien oder Urteilen, auf die teils explizit verwiesen wird. Zudem können identische Begriffe in verschiedenen Rechtstexten abweichende Bedeutung haben,[10] während gleiche Sachverhalte sehr unterschiedlich benannt werden.[11] Die KI-gestützte Analyse muss in der Lage sein, solche kontextuellen Nuancen zu erfassen.

03. **Grundlegende Fähigkeit zu Rechtsanalysen:** Aufbauend auf den ersten beiden Anforderungen muss die etwaige KI-Anwendung eine grundlegende Fähigkeit zur rechtlichen Analyse aufweisen. Dies umfasst in Hinblick auf die Kernfunktionalitäten:

    a. Geltungsbereichsanalyse: Die KI muss anhand einschlägiger Regelungstexte und weiterer Kontextinformationen (z. B. Verweise, Auslegungshilfen) erkennen können, welche Regelungen sich in ihrem Geltungsbereich überschneiden und daher gleichzeitig zur Anwendung kommen können. Die zentrale Herausforderung für eine KI-gestützte Erkennung besteht darin, dass Geltungsbereichsbestimmungen nicht einheitlich anhand fester Merkmale (z. B. Tätigkeitsbeschreibungen nach NACE-Klassifikation, Unternehmensstandort, Rechtsform) definiert werden

10 Beispielsweise existieren im Wirtschaftsrecht rund 220 unterschiedliche Begriffsdefinitionen von „Einkommen“ (Bundesministerium des Innern und für Heimat, 2025).

11 Beispielsweise verwendet die europäischen Verordnung über künstliche Intelligenz (KI-Verordnung, Verordnung (EU) 2024/1689, ABl. (EU) L, 2024/1689, 12.7.2024 (aktuell geltende Fassung)) Begriffe wie „KI-System“ oder „KI-Algorithmus“, während die EU-Richtlinie über Märkte für Finanzinstrumente (MiFID II, Richtlinie 2014/65/EU, ABl. (EU) L 173 vom 12. Juni 2014, S. 349–496 (aktuell geltende Fassung)) den Begriff „algorithmischer Handel“ nutzt. Obwohl algorithmischer Handel KI-Algorithmen umfassen kann, wird dies in der MiFID II nicht explizit erwähnt. Eine KI-Anwendung sollte jedoch in der Lage sein, solche konzeptuellen Zusammenhänge zu erkennen, um mögliche Vereinfachungspotentiale in Regelungen zu identifizieren.

und verschiedene Dimensionen (sachlich, persönlich, räumlich und zeitlich) oft ineinandergreifen. Zudem können spezifische Bestimmungen (z. B. spezifische Stichtagsregelungen, Ausnahmeregelungen oder das Proportionalitätsprinzip) den Geltungsbereich einzelner Regelungen weiter präzisieren oder einschränken.

b. Identifikation der Vereinfachungspotentiale: Wurden überlappende Geltungsbereiche festgestellt, muss die KI potenzielle Doppelungen, Widersprüche und Inkonsistenzen in den betroffenen Regelungen identifizieren. Die Analyse kann dabei in der Regel nur **Verdachtsfälle** auf mögliche Vereinfachungspotentiale liefern, weil zum Beispiel mehrfache ähnliche Pflichten für Unternehmen auf tatsächliche Doppelungen hinweisen können oder in manchen Fällen auch bewusst in Kauf genommen werden. Widersprüche oder Inkonsistenzen werden zudem oft nur in bestimmten Anwendungssituationen relevant und werden daher erst bei der Anwendung der Rechtsakte ersichtlich.[12] Die abschließende Klärung solcher Verdachtsfälle bleibt dem Normgeber oder ggf. Gerichten (durch „klärende Urteile“) vorbehalten.

04. **Erklärbarkeit und Verifizierbarkeit:** Damit KI-basierte Ergebnisse, etwa die Verdachtsfälle auf Vereinfachungspotentiale, im Gesetzgebungsprozess berücksichtigt werden können, müssen sie juristisch nachvollziehbar begründet, überprüfbar und ggf. korrigierbar sein. Auf diese Weise verbleibt die Verantwortung bei den Nutzenden, während die KI lediglich als Assistenzsystem fungiert. Zudem wird dadurch die Qualitätssicherung nachgelagerter Prozesse, wie der Gesetzesfolgenabschätzung, erleichtert.

05. **Hohe Verlässlichkeit der Ergebnisse:** Angesichts der erforderlichen manuellen Überprüfung der KI-Ergebnisse und ihren potenziellen Auswirkungen auf den Gesetzgebungsprozess ist zudem eine hohe Verlässlichkeit unabdingbar. Generative KI-Modelle können beispielsweise dazu neigen, in bestimmten Situationen überzeugend klingende, aber inhaltlich unzutreffende Aussagen zu produzieren (sogenannte „Halluzinationen“), was im Gesetzgebungskontext gravierende Folgen haben kann, und geeignete Maßnahmen zur deren Minimierung erfordert (Magesh, et al., 2024).

Im nächsten Abschnitt werden methodische Ansätze für die Kernfunktionalitäten vorgestellt, die die genannten Machbarkeitsvoraussetzungen erfüllen und eine KI-gestützte Identifikation der Vereinfachungspotentiale in Rechtstexten ermöglichen. Eine vertiefte Betrachtung zur Verlässlichkeit der Ansätze folgt im Kapitel 04.

03.03 Machbarkeitsanalyse KI-gestützter Ansätze

Der vorangegangene Abschnitt hat die zentralen methodischen Voraussetzungen für eine KI-gestützte Analyse von Vereinfachungspotentialen formuliert. Die variable, komplexe Struktur von Rechtstexten sowie die Notwendigkeit ihrer kontextbezogenen Auslegung (vgl. Voraussetzungen 01 und 02) stellen eine Hürde für den Einsatz regelbasierter Verfahren und „herkömmlicher“ Methoden des maschinellen Lernens (z. B. Entscheidungsbäume, Support Vector Machine) dar.[13]

Dank jüngster Fortschritte im Bereich KI, insbesondere durch die Weiterentwicklung der auf der Transformer-Architektur basierenden großen Sprachmodelle (im Folgenden „Sprachmodelle“ genannt), ist es inzwischen möglich, semantische und kontextuelle Zusammenhänge in Rechtstexten weitaus präziser zu erfassen. Zu diesen Sprachmodellen zählen sowohl große generative Sprachmodelle (Large Language Models, kurz: LLMs) wie GPT oder Mistral als auch große nicht-generative Sprachmodelle wie BERT, die zwar

12 Um derartige Probleme bereits in der Phase der Rechtsetzung zu entschärfen, regt der Normenkontrollrat (NKR) beispielsweise die Einführung sogenannter Ex-ante-Praxischecks an (Nationaler Normenkontrollrat, 2024a, S. 37ff.).

13 Zwar lassen sich mit diesen Methoden Aufgaben wie Entitätenerkennung oder Themenmodellierung umsetzen, mit denen Begriffe und thematische Schwerpunkte in Texten identifiziert werden (vgl. z. B. Blümmel et al. (2024)), dennoch ist aufgrund der Komplexität und Variabilität der Rechtstexte keine ausreichende Aussagekraft der Ergebnisse zu erwarten. Insbesondere die begrenzte Berücksichtigung kontextueller Informationen übersteigen die Fähigkeiten dieser Modelle.

ebenfalls auf einer Transformer-Architektur basieren, jedoch für andere Aufgaben, etwa Textklassifikation oder Entitätenerkennung, geeignet sind. Sprachmodelle, die gezielt auf juristischen Inhalten trainiert wurden, etwa Saul-7B (basierend auf Mistral) oder LEGAL-BERT (basierend auf BERT), erreichen dabei bessere Ergebnisse als ihre jeweiligen Basismodelle (Chalkidis, Fergadiotis, Malakasiotis, Aletras, & Androutsopoulos, 2020; Colombo, et al., 2024). Gleichwohl verdeutlichen aktuelle Studien, dass die Fähigkeit zur rechtlichen Subsumtion und zur konsistenten, fehlerarmen Regelungsanalyse (vgl. Voraussetzungen 03 und 05) bei diesen Sprachmodellen bislang nur eingeschränkt gegeben ist. So können Modelle wie GPT-4 und Mistral-7B und auf LLMs basierte juristische KI-Lösungen wie Lexis+ AI zwar grundsätzlich juristische Fragestellungen beantworten, doch treten dabei häufig Unstimmigkeiten, Fehlinterpretationen und sogenannte „Halluzinationen“ auf (Katz, Bommarito, Gao, & Arredondo, 2024; Magesh, et al., 2024).

Entsprechend existiert derzeit keine einzelne KI-Methode und keine am Markt verfügbare KI-Lösung, die alle in Abschnitt 03.02 definierten Voraussetzungen (01-05) vollständig erfüllt und damit deren Machbarkeit demonstrieren würde.

Vor diesem Hintergrund wird nachfolgend ein neuartiger methodischer Ansatz vorgestellt, der mehrere KI-Methoden im Rahmen einer Anwendung kombiniert. Auf diese Weise sollen alle definierten Voraussetzungen gemeinsam erfüllt und eine verlässliche Identifikation der Vereinfachungspotentiale in Regelungen ermöglicht werden.

03.03.01 Übergreifende Aspekte vorgeschlagener methodischer Ansätze

Der nachfolgend beschriebene methodische Ansatz nutzt ein oder mehrere Sprachmodelle (z. B. GPT, Mistral oder LEGAL-BERT) und bettet deren Analysen in eine Pipeline ein, die aus vorbereitenden, aufbereitenden und validierenden Teilschritten besteht. Dadurch wird die verlässliche Identifikation der Vereinfachungspotentiale ermöglicht und zugleich die Ergebnisse für eine manuelle Prüfung verifizierbar gemacht.

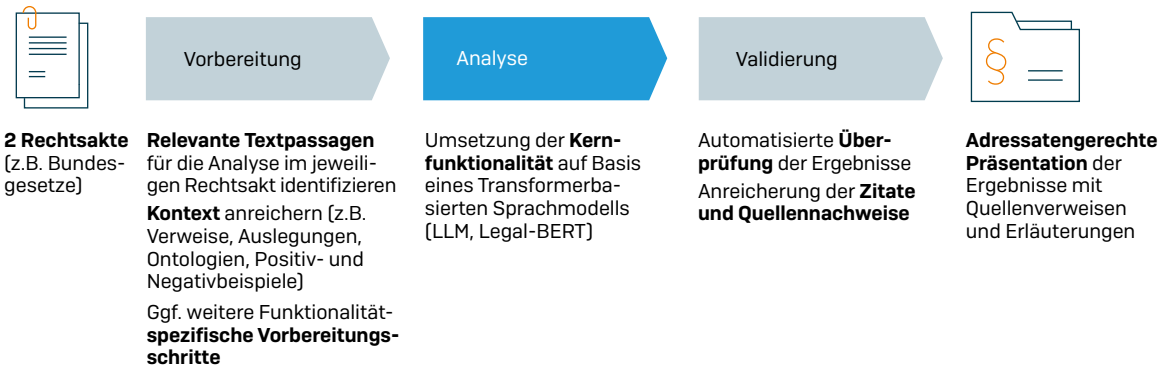


Abbildung 1: Schematische Darstellung eines methodischen Ansatzes zur Umsetzung der Kernfunktionalitäten

Zur vereinfachten Illustration wird angenommen, dass die etwaige KI-Anwendung jeweils zwei Rechtsakte (z. B. Gesetze oder Verordnungen) erhält[14] und je nach in Abschnitt 03.01 beschriebener Kernfunktionalität:

01. Regelungen mit überlappendem Geltungsbereich ermittelt, oder

02. in Regelungen mit überlappendem Geltungsbereich mögliche Doppelungen, Widersprüche oder Inkonsistenzen identifiziert.

14 Beispiele sind das LkSG und die CSDDD oder die DSGVO und die Europäische Medizinprodukteverordnung (MDR, Verordnung (EU) 2017/745, ABl. (EU) L 117 vom 05/05/2017, S. 1–175 (aktuell geltende Fassung)). Die Einschränkung auf zwei Rechtsakte dient lediglich der Veranschaulichung der methodischen Ansätze. In der Praxis soll die etwaige KI-Anwendung die zugrundeliegende juristische Wissensdatenbank automatisch durchsuchen können, um (beispielsweise ausgehend von einem Rechtsakt) weitere Regelungen mit überlappendem Geltungsbereich zu identifizieren.

Ergänzend wird angenommen, dass die etwaige KI-Anwendung auf eine umfassende juristische Wissensdatenbank mit weiteren relevanten Rechtstexten und Auslegungshilfen zugreifen kann (vgl. Kapitel 02).

Obgleich sich insbesondere der Analyseschritt je nach Kernfunktionalität unterscheidet, umfasst die Pipeline stets die folgenden Bausteine:

- **Vorbereitungsschritt:** Um eine präzisierte Analyse durch ein Sprachmodell zu ermöglichen und das Risiko von Fehlinterpretationen und Halluzinationen zu minimieren, werden zunächst kurze, besonders relevante Textpassagen für die jeweilige Analyse ausgewählt,[15] anstatt komplette Rechtsakte miteinander zu vergleichen. Für die Identifikation möglicher Doppelungen, Widersprüche oder Inkonsistenzen sollen beispielsweise Artikelpaare vorausgewählt werden, die potenzielle Pflichten für dieselben Adressaten festlegen. Zur Eingrenzung dieser Textpassagen empfiehlt sich ein hybrider Ansatz aus regelbasierten Verfahren (z. B. das Erkennen einschlägiger Schlüsselbegriffe wie „Verpflichtung", „Anwendungsbereich" oder spezifische Strukturmerkmale im Gesetzestext) und einer Sprachmodell-basierten Voranalyse (z. B. mithilfe von LEGAL-BERT oder einem feingetunten LLM).[16]

  Diese Textpassagen werden für die anschließende Analyse um zusätzliche Kontextinformationen wie Verweise auf andere Rechtsakte, nachgelagerte Regelungen, Auslegungshilfen oder juristische Ontologien und kontrollierte Vokabulare[17] angereichert. Darüber hinaus können domänenspezifische Wissensgraphen eingebunden werden, um implizite Beziehungen zwischen Normen, Begriffen und Rechtsprechung sichtbar zu machen und dadurch die Kontexttiefe für nachgelagerte KI-Analysen weiter zu erhöhen.

- **Analyseschritt:** Auf Basis der im Vorbereitungsschritt ausgewählten und um Kontextinformationen angereicherten Textpassagen erfolgt die eigentliche Untersuchung mit Hilfe eines Sprachmodells. Die Fragestellung (sog. Prompts) und der zugehörige Kontext werden dabei möglichst präzise formuliert, um eine verlässliche Analyse zu ermöglichen. Außerdem kann die Bereitstellung einiger aufgabenspezifischer Beispiele für die Analyse (Few-Shot In-Context-Learning) eine signifikante Leistungssteigerung der Sprachmodell-basierten Analyse bewirken. Die Ergebnisse dieses Schritts umfassen je nach Kernfunktionalität erste Vorschläge für Regelungen mit überlappendem Geltungsbereich oder initiale Verdachtsfälle auf Doppelungen, Widersprüche oder Inkonsistenzen in diesen Regelungen.

- **Validierung:** Um eine hohe Verlässlichkeit der Analyseergebnisse zu gewährleisten, ist eine automatisierte Validierung der KI-Ergebnisse unerlässlich. Mögliche Maßnahmen sind in Kapitel 04 detailliert erläutert. Ergänzend werden die Resultate mit relevanten Quellen verknüpft, um eine manuelle Prüfung durch die Fachebene zu erleichtern.

- **Adressatengerechte Präsentation der Ergebnisse:** Anschließend werden die Ergebnisse in einer für die Nutzenden verständlichen Form aufbereitet. Dies kann durch Hervorhebung relevanter Textpassagen in den betrachteten Texten, grafische Darstellungen der Zusammenhänge oder andere Visualisierungs- und Aggregationsverfahren unterstützt werden, um eine schnelle Bewertung und manuelle Qualitätssicherung der Ergebnisse zu ermöglichen.

---

15 Gesetze, Verordnungen und weitere relevante Texte werden dazu zunächst automatisiert in kleinere, thematisch abgegrenzte Segmente (sogenannte „semantische Chunks") unterteilt, etwa thematische Blöcke (wie „Begriffsbestimmungen", „Pflichten und Sanktionen") oder Paragrafen.

16 Die geplante Umstellung auf das Akoma Ntoso-Format im Rahmen der E-Gesetzgebung (Deutscher Bundestag, 2020; Verwaltung Innovativ, 2023) und des Projekts AK4EU (Europäische Union, 2025b) wird die automatisierte Verarbeitung von Rechtsnormen und Gerichtsentscheidungen erleichtern und dadurch die Qualität der KI-Ergebnisse weiter steigern. Das Format ermöglicht entsprechende Vorstrukturierung von Rechtstexten für die digitale Verarbeitung.

17 Die Nutzung von Ontologien und kontrollierten Vokabularen, wie beispielsweise dem European Legislation Identifier (ELI) oder im Handbuch der Rechtsförmlichkeit, kann die etwaige KI-Anwendung unterstützen. Diese Standards ermöglichen eine einheitliche Identifikation und Kategorisierung von Rechtsbegriffen und fördern die Interoperabilität zwischen verschiedenen Rechtssystemen und Datenquellen.

03.03.02 Kernfunktionalität: Geltungsbereichsanalyse

Ziel der Kernfunktionalität ist zu ermitteln, ob und in welchem Ausmaß sich die Geltungsbereiche der betrachteten Rechtsakte überschneiden. Die Ergebnisse bilden die Grundlage für die nachgelagerte Identifikation von Doppelungen, Widersprüchen oder Inkonsistenzen (vgl. Abschnitt 03.03.03).

Hierfür werden die im Vorbereitungsschritt vorausgewählten und mit Kontext angereicherten Textpassagen (d.h. jene, die mögliche Angaben zum persönlichen, sachlichen, räumlichen oder zeitlichen Geltungsbereich enthalten) mithilfe eines Sprachmodells verglichen. Für den Analyseschritt wurden zwei Hauptansätze identifiziert:

- **Expliziter Ansatz:** Die Geltungsbereichsbeschreibungen des jeweiligen Rechtsaktes werden zuerst aus den im Vorbereitungsschritt vorausgewählten Textpassagen extrahiert und anschließend mit Hilfe semantischer Ähnlichkeitsmaße[18] (z. B. durch ein spezialisiertes Sprachmodell wie LEGAL-BERT) paarweise verglichen.
- **Impliziter Ansatz:** Die vorausgewählten Textpassagen werden direkt miteinander verglichen, ohne vorherige explizite Extraktion der Beschreibungen der Geltungsbereichsdimensionen. Hierfür können Sprachmodelle wie LLMs und LEGAL-BERT kombiniert werden.

Der explizite Ansatz bietet ein hohes Maß an Transparenz, da die extrahierte Geltungsbereichsbeschreibung separat validiert und für die Wiederverwendung in der juristischen Datenbank abgelegt werden kann. Demgegenüber wird erwartet, dass der implizite Ansatz weniger fehleranfällig ist, wenn Gesetzestexte keine klar abgegrenzten „Geltungsbereichsabschnitte“ enthalten oder diese auf mehrere Textpassagen verteilt sind.

Auf Basis dieser Erkenntnisse ist zudem zu erwarten, dass eine KI branchenspezifische Regelungsaspekte[19] erkennen kann, insbesondere wenn diese explizit im Text vorhanden sind oder anhand charakteristischer sprachlicher Muster ableitbar sind.

Ergänzend stellt sich die Frage, unter welchen Voraussetzungen die KI-Anwendung automatisiert ermitteln kann, welche Regelungen für ein spezifisches Unternehmen gelten. Dies erfordert logisches Schlussfolgern und die Prüfung von Tatbestandsvoraussetzungen. Da Sprachmodelle zwar Texte verarbeiten, aber keine rechtsmethodischen Vorgehensweisen anwenden können (insbesondere keine inhärente Fähigkeit zur rechtlichen Subsumtion besitzen), gelten dafür zusätzliche Voraussetzungen:

01. **Detailliertes Unternehmensprofil:** Das Profil muss alle wesentlichen Merkmale enthalten, die für eine Subsumtion relevant sein können (z. B. Tätigkeiten, Unternehmensgröße, Rechtsform, Niederlassung).
02. **Strukturierte Normrepräsentationen:** Um die Anwendung der Geltungsbereiche auf konkrete Sachverhalte durch Sprachmodelle zu ermöglichen, sind entsprechende Repräsentationen hilfreich (etwa als Wenn-Dann-Regeln oder Entscheidungsbäume). Ein Beispiel dafür ist die Rulemapping-Methode (Breidenbach, 2025), bei der sich Regelungstexte in widerspruchsfreie, maschinenlesbare Entscheidungsbäume übersetzen lassen. Diese Visualisierung und Formalisierung erleichtert die automatisierte Prüfung, ob und welche Normen auf das spezifische Unternehmensprofil zutreffen.

18 Das Ähnlichkeitsmaß wird dabei als Maß für die inhaltliche Überschneidung der Geltungsbereiche interpretiert. Eine hohe inhaltliche Ähnlichkeit kann in diesem Fall als hohe Wahrscheinlichkeit für eine Überschneidung der jeweiligen Dimensionen der Geltungsbereiche interpretiert werden. Eine geringe Ähnlichkeit signalisiert, dass die Wahrscheinlichkeit für eine Überschneidung gering ist.

19 z. B. ob eine Regelung einen Wirtschaftszweig betrifft.

03.03.03 Kernfunktionalität: Identifikation der Vereinfachungspotentiale

Ziel dieser Kernfunktionalität ist festzustellen, ob und ggf. welche Vereinfachungspotentiale zwischen Rechtsakten mit überlappendem Geltungsbereich bestehen. Das gewünschte Ergebnis ist entweder die Feststellung, dass keine Vereinfachungspotentiale identifiziert wurden, oder die Beschreibung der Vereinfachungspotentiale (Doppelung, Widerspruch und/oder Inkonsistenz) jeweils unter Angabe der relevanten Textpassagen in den zwei Rechtsakten und erläuternder Hinweise.

Es wurden zwei Hauptansätze für den Analyseschritt identifiziert, die analog zur Geltungsbereichsanalyse Sprachmodelle verwenden, vorzugsweise solche, die speziell auf juristische Inhalte trainiert wurden (z. B. LEGAL-BERT). Der Vorbereitungsschritt umfasst dabei zusätzlich folgende für diese Kernfunktionalität spezifische Teilschritte:

- **Bildung von Paaren:** Um Vereinfachungspotentiale zu identifizieren, werden zunächst Paare mit relevanten Textpassagen aus jedem der beiden zu vergleichenden Rechtsakte gebildet (z. B. Artikelpaare). Dabei werden nur Textpassagen berücksichtigt, die Pflichten an Unternehmen formulieren.[20]
- **Filterung:** Anschließend werden die Paare entfernt, die direkt als irrelevant einzuschätzen sind. Das Filtern dient dazu, die Anzahl der Paare und damit den Rechenaufwand des folgenden Schritts weiter zu reduzieren. Um relevante Paare zu identifizieren sind drei Ansätze möglich: Semantische Ähnlichkeit, „Join"-Kriterium und Klassifikation der Relevanz.

Die darauf aufbauende Analyse nutzt die nach der Filterung verbliebenen Paare und die dazu angereicherten Kontextinformationen für die Identifikation der Vereinfachungspotentiale und basiert auf:

- **Klassifikation von Paaren:** Jedes verbliebene Paar wird in Bezug auf Vereinfachungspotentiale analysiert. Das Ergebnis ist eine Feststellung, ob ein Verdachtsfall auf Vereinfachungspotentiale vorliegt (z. B. „ja, sehr wahrscheinlich", „ja, weniger wahrscheinlich", „nein") sowie eine Begründung und weitere Hinweise zur Überprüfung.

Die Verlässlichkeit dieser Ansätze und die prototypische Erprobung werden im nächsten Kapitel erläutert.

03.04 Fazit

Dieses Kapitel untersucht die Machbarkeit einer KI-gestützten Identifikation der Vereinfachungspotentiale in Regelungen. Ausgehend von praktischen Beispielen werden zunächst die zentralen Arten der Vereinfachungspotentiale – Doppelungen, Widersprüche und Inkonsistenzen – definiert. Darauf aufbauend werden die wesentlichen Voraussetzungen für deren verlässliche Erkennung durch KI erarbeitet. Sie umfassen die Fähigkeit einer KI-Anwendung, strukturelle und kontextuelle Nuancen deutscher und englischer Rechtstexte korrekt zu erfassen, basierend darauf Geltungsbereiche und einzelne Regelungen rechtlich sachgerecht miteinander zu vergleichen sowie die Analyseergebnisse nachvollziehbar und überprüfbar zu begründen.

Anschließend werden basierend auf Erkenntnissen aus der Literaturrecherche und aktuellen Entwicklungen der natürlichen Sprachverarbeitung methodische Ansätze für die etwaige KI-Anwendung vorgestellt. Sie nutzen ein oder mehrere Transformer-basierte Sprachmodelle (z. B. GPT-4, Mistral Large und/oder LEGAL-BERT) und betten deren Analysen in eine Pipeline ein, die aus vorbereitenden, validierenden und aufbereitenden Teilschritten besteht. **Durch die vorgeschlagenen methodischen Ansätze**

20 Eine entsprechende Kategorisierung von Rechtsvorschriften mit Hilfe von KI-Methoden wurde z. B. in der Kostbar-Studie (Deloitte, 2021) durchgeführt.

lassen sich sämtliche identifizierte Voraussetzungen erfüllen und die technische Machbarkeit der KI-gestützten Identifikation der Vereinfachungspotentiale demonstrieren.

Im nächsten Kapitel wird die Verlässlichkeit dieser Ansätze weiter untersucht und anhand einer prototypischen Analyse zweier Rechtsakte illustriert.

04. Verlässlichkeit KI-gestützter Ergebnisse optimieren

Die im vorherigen Kapitel vorgeschlagenen Ansätze ermöglichen es, Vereinfachungspotentiale in den Regelungen zu identifizieren und zu analysieren. Sie bieten damit eine solide Grundlage für den angestrebten Bürokratieabbau und eine Entlastung der Verwaltung. Eine wesentliche Voraussetzung für deren Einsatz ist jedoch eine hohe Verlässlichkeit der mit KI erzielten Ergebnisse. Vor diesem Hintergrund stellt dieses Kapitel fünf zentrale Dimensionen der Verlässlichkeit vor – Plausibilität, Vollständigkeit, Konsistenz, Robustheit und Erklärbarkeit/Verifizierbarkeit – die als Konzeptstützen für die Bewertung der Verlässlichkeit dienen.[21] Es werden Maßnahmen abgeleitet, um eine hohe Verlässlichkeit entlang jeder dieser Dimensionen zu gewährleisten. Neben der theoretischen Bewertung werden erste Erkenntnisse aus einer prototypischen Erprobung ausgewählter Komponenten des zuvor beschriebenen Ansatzes zur Identifikation der Vereinfachungspotentiale präsentiert.

04.01 Theoretische Bewertung der Verlässlichkeit

Zentrale Dimensionen der Verlässlichkeit im Kontext einer KI-basierten Analyse:

Plausibilität (Plausibility): Beschreibt, inwieweit die KI-Ergebnisse juristisch stichhaltig und fachlich nachvollziehbar sind. Dies umfasst auch Genauigkeit (Accuracy) bei überprüfbaren Angaben (z. B. Fristen, Verweise).[22]

Vollständigkeit (Recall): Beschreibt den Anteil aller relevanten Verdachtsfälle auf Vereinfachungspotential, die von der KI-Anwendung erkannt werden.

Konsistenz (Consistency): Beschreibt die Fähigkeit der Anwendung, bei wiederholten Eingaben mit gleichen Daten hinreichend ähnliche Ergebnisse zu liefern.[23]

Robustheit (Robustness): Beschreibt die Stabilität der Anwendung bei Variationen oder Unsicherheiten in den Eingabedaten.

Erklärbarkeit/ Verifizierbarkeit (Explainability): Beschreibt, inwieweit die Ergebnisse der Anwendung nachvollziehbar begründet und durch relevante Rechtsquellen oder Dokumente einfach überprüft werden können.

Infobox 2: Zentrale Dimensionen der Verlässlichkeit im Kontext einer KI-basierten Analyse

Infobox 2 schlägt die fünf zentralen Dimensionen der Verlässlichkeit einer etwaigen KI-Anwendung vor. Hohe Plausibilität und Vollständigkeit stellen sicher, dass die Ergebnisse juristisch belastbar sind, sich nachvollziehen lassen und relevante Informationen möglichst umfassend einbeziehen. Hohe Konsistenz gewährleistet, dass gleiche Sachverhalte zu ähnlichen Ergebnissen führen, während hohe Robustheit für stabile Leistung bei der Anwendung auf verschiedene Rechtsgebiete oder unvollständige Sachverhalte bei Anwendung auf unterschiedliche Rechtsgebiete oder bei unvollständigen Informationen sorgt. Schließlich bietet Erklärbarkeit/Verifizierbarkeit die notwendige Grundlage, um die Analyseergebnisse manuell zu prüfen und zu bewerten.

21 Die Dimensionen ergänzen sich teilweise, z. B. kann Robustheit als weitergefasste Konsistenz betrachtet werden.

22 Anders als bei vielen traditionellen KI-Anwendungen (z. B. Bildklassifikation) existiert im juristischen Bereich oft nicht die eine „richtige“ rechtliche Interpretation. Entscheidend ist daher, dass die Ergebnisse der KI-Anwendung juristisch vertretbar und kohärent sind, anerkannten rechtlichen Methoden entsprechen und insbesondere überprüfbare Sachverhalte korrekt abbilden.

23 KI-Modelle wie die großen Sprachmodelle (LLMs) basieren auf probabilistischen Ansätzen mit stochastischen Prozessen und liefern daher selbst bei identischen Eingaben nicht immer identische Ergebnisse. Diese Ergebnisse sollten jedoch hinreichend konsistent sein.

Technologische Komponente	Konsistenz	Robustheit	Erklärbarkeit
Große Sprachmodelle (LLM)	●	●	●
BERT	●	●	●
ML-Modelle[24]	●	●	variiert
Regelbasierte Verfahren	●	●	●
Semantische Ähnlichkeitstools[25]	variiert	variiert	variiert

Verlässlichkeit: Sehr gering ● ● ● ● ● sehr hoch

Tabelle 5: Verlässlichkeit der technologischen Komponenten der im Kapitel 03 beschriebenen Ansätze hinsichtlich Konsistenz, Robustheit und Erklärbarkeit/Verifizierbarkeit

Tabelle 5 verdeutlicht die Stärken und Schwächen einzelner für die Lösungsansätze relevanter Technologien in Bezug auf die Verlässlichkeitsdimensionen. Plausibilität und Vollständigkeit hängen dabei hingegen vom konkreten Anwendungsszenario ab und werden näher im folgenden Abschnitt betrachtet. Ein zentrales Element der vorgeschlagenen Ansätze zur Geltungsbereichs- und Regelungsanalyse ist der Einsatz von Sprachmodellen wie LLMs und BERT bzw. deren Modellvarianten, die speziell auf juristische Inhalte trainiert wurden (z. B. LEGAL-BERT). Die Wahl dieser KI-Methoden ist dadurch bedingt, dass sie die identifizierten Machbarkeitsvoraussetzungen erfüllen und eine hohe bis sehr hohe Flexibilität in der Anwendung zeigen. Allerdings weisen insbesondere die LLMs in punkto Verlässlichkeit Schwächen auf, zum Beispiel durch das Risiko von „Halluzinationen“ und fehlenden etablierten Bewertungsstandards (Deloitte, 2021; Gehrmann, Clark, & Sellam, 2023). Zudem kann eine scheinbar kohärente und genaue Ausgabe eines LLMs zu fehlerhaften Entscheidungen verleiten, wenn menschliche Kontrolle vernachlässigt wird (Lee, et al., 2024).

Um eine hohe Verlässlichkeit der Gesamtanwendung zu gewährleisten, müssen diese Methoden daher gezielt kombiniert und zusätzliche Maßnahmen zur Steigerung der Verlässlichkeit ergriffen werden. Dies spiegelt sich in dem Vorbereitungs- und Validierungsschritt der vorgeschlagenen Ansätze. Mehrere Studien zeigen, dass sich die Verlässlichkeit von LLMs in Bezug auf einzelne Dimensionen verbessern lässt (Kalra, et al., 2024; Arabzadeh, et al., 2024; Spatharioti, Rothschild, Goldstein, & Hofman, 2023). Mögliche Maßnahmen umfassen beispielsweise:

- wiederholte Abfragen an das Modell oder einheitliche Modellparameter,
- Reduktion möglicher Fehler mittels Prompt-Engineering oder Modell-Ensembling,
- gezielte Anpassung von Konfidenzwerten an Korrektheit (etwa mithilfe eines „Richter-Modells“),
- Einbindung von Rechtsquellen aus der zugrunde liegenden Datenbank und Hervorhebung von Schlüsselstellen, um potenzielle Irrtümer schneller erkennbar zu machen.

24 Modelle des maschinellen Lernens (keine neuronalen Netze).
25 Die Ansätze, die für die semantische Ähnlichkeit verwendet werden, basieren in der Regel auf LLMs, BERT-Modellen oder auf Regeln. Daher ist die Zuverlässigkeit der Komponente basierend auf der zugrundeliegenden Technologiekomponente zu bewerten.

Insgesamt legt die theoretische Bewertung nahe, dass sich mit den vorgeschlagenen Ansätzen, ggf. ergänzt durch zusätzliche Maßnahmen, eine hohe Verlässlichkeit der KI-Ergebnisse erreichen lässt. Dabei ist allerdings immer der erforderliche Aufwand gegen die gewünschten Optimierungen abzuwägen. Ein iteratives Vorgehen, bei dem die Maßnahmen schrittweise implementiert, kontinuierlich evaluiert und bei Bedarf angepasst werden, empfiehlt sich, um dieses Spannungsfeld zu lösen (mehr im Kapitel 06). Angesichts des dynamischen technologischen Fortschritts bei den Sprachmodellen ist zudem damit zu rechnen, dass sich die Verlässlichkeit KI-gestützter Analysen kontinuierlich weiter steigern lässt (OpenAI, 2023; OpenAI, 2024).

04.02 Prototypische Erprobung

Um zusätzliche praktische Erkenntnisse über die Verlässlichkeit KI-generierter Analysen zu gewinnen, wurden ausgewählte Elemente des in Kapitel 03 vorgestellten Ansatzes zur Identifikation der Vereinfachungspotentiale mit Hilfe von LLMs prototypisch getestet. Der Fokus dieser Erprobung lag auf der Identifikation der Vereinfachungspotentiale in zuvor ausgewählten Paaren von Passagen aus zwei Rechtstexten, da dies ein zentraler Schritt der vorgeschlagenen Ansätze ist. Exemplarisch wurden dazu die Paragrafen des LkSG in der Fassung vom 16. Juli 2021 mit Artikeln der CSDDD in der Fassung vom 5. Juli 2024 verglichen, da beide Regelwerke starke inhaltliche Ähnlichkeiten aufweisen und potenziell ähnliche Pflichten für Wirtschaftsakteure begründen.[26]

Für jedes Paar gab die für die prototypische Erprobung entwickelte Applikation eine Einschätzung ab, ob Vereinfachungspotentiale vorliegen, und lieferte dazu eine Begründung mit einer Beschreibung der jeweiligen Potentiale.[27] Die Ergebnisse wurden anschließend von Juristinnen und Juristen qualitativ bewertet, um die Plausibilität der Antworten sowie deren Vollständigkeit und Erklärbarkeit/Verifizierbarkeit zu prüfen. Zur Ermittlung der Konsistenz (Reproduzierbarkeit) wurde der Test mit identischen Paaren und Prompts insgesamt viermal wiederholt.

Artikelpaar	Ausgabe der Applikation (auf Basis von GPT-4o)	Ausgabe der Applikation (auf Basis von OpenAI o1-preview, gekürzt)
Vergleich von **§ 23 LkSG** (Zwangsgelder) und **Art. 27 CSDDD** (Sanktionen)	**Vereinfachungspotential: sehr wahrscheinlich** **Begründung:** Die unterschiedlichen Regelungen zur Höhe der Zwangsgelder in den beiden Texten führen zu einer Inkonsistenz, die hohe Vereinfachungspotentiale bieten. Die absolute Obergrenze in Text 1 und die relative Berechnung in Text 2 können zu Verwirrung und ineffizienter Umsetzung führen, insbesondere wenn beide Regelungen gleichzeitig anwendbar sind.	**1. Höhe des Zwangsgeldes** **...** **Bewertung: Widerspruch, sehr wahrscheinlich** **Begründung:** Es besteht ein klarer Widerspruch ... hinsichtlich der Bemessung und Höhe der Zwangsgelder. Die Richtlinie fordert deutlich höhere Zwangsgelder basierend auf dem Umsatz, während das LkSG eine fixe Obergrenze von 50.000 Euro vorsieht ...

26 Die Regelungen der CSDDD sind zudem im Wesentlichen durch die Anpassung des LkSG in deutsches Recht umzusetzen.

27 Da es sich bei den Vereinfachungspotentialen nur um Verdachtsfälle handelt, wurden als Kategorien für die Einschätzung definiert: „ja, sehr wahrscheinlich“, „ja, weniger wahrscheinlich“ oder „nicht relevant“.

		2. Definition und Anwendung von Zwangsgeldern ... **Bewertung: Doppelung, weniger wahrscheinlich** **Begründung:** Beide Regelungen sehen die Verhängung von Zwangsgeldern vor, was zu einer Doppelung führen kann. Jedoch lässt die Richtlinie den Mitgliedstaaten Gestaltungsspielraum bei der Umsetzung. Die spezifische Ausgestaltung und Anwendung der Zwangsgelder kann daher unterschiedlich sein, wodurch das Vereinfachungspotential weniger wahrscheinlich ist.

Tabelle 6: Beispielhafte Gegenüberstellung KI-gestützter Analysen der Vereinfachungspotentiale (auf Basis von GPT-4o und OpenAI o1-preview) beim Vergleich von § 23 LkSG und Art. 27 CSDDD

Die Tabelle 6 veranschaulicht die Ausgabe der KI-Testapplikation am Beispiel eines Paars.[28] Aus juristischer Sicht liefern die Ergebnisse der prototypischen Tests grundsätzlich hilfreiche Hinweise auf mögliche Vereinfachungspotentiale. Für den praktischen Einsatz sind jedoch zusätzliche Optimierungen (z. B. automatisierte Validierung, optimierte Prompt-Gestaltung) und menschliche Kontrolle unerlässlich. Während bei der Plausibilität kein Modell eindeutig dominiert, erzielte die o1-preview-basierte Test-Applikation bezogen auf Vollständigkeit und Erklärbarkeit bessere Ergebnisse. Zudem deuten die Ergebnisse insgesamt auf eine solide Konsistenz hin.

Schließlich legt sowohl die theoretische Analyse als auch der vergleichende Test von GPT-4o und o1-preview nahe, dass weitere Modellverbesserungen künftig die Verlässlichkeit KI-generierter Analysen weiter erhöhen können.

04.03 Fazit

In diesem Kapitel wurde die Verlässlichkeit der in Kapitel 03 vorgestellten Ansätze zur Geltungsbereichsanalyse und Identifikation der Vereinfachungspotentiale anhand fünf zentraler Dimensionen bewertet: Plausibilität, Vollständigkeit, Konsistenz, Robustheit und Erklärbarkeit bzw. Verifizierbarkeit. **Die theoretische Betrachtung sowie die prototypische Erprobung legen nahe, dass sich eine hohe Verlässlichkeit der Ansätze entlang aller fünf Dimensionen grundsätzlich erreichen lässt, sofern diese durch geeignete Zusatzmaßnahmen ergänzt werden.** Bei allen Verbesserungsmaßnahmen ist allerdings abzuwägen, welcher zusätzliche Aufwand dem angestrebten Nutzen gerecht wird. Ein iteratives Vorgehen – bei dem Maßnahmen schrittweise eingeführt, fortlaufend evaluiert und bei Bedarf nachjustiert werden – bietet einen möglichen Ansatz, um dieses Spannungsfeld zu meistern (siehe Kapitel 06).

28 Dieses Beispiel wurde zu Testzwecken ausgewählt, die Inkonsistenz wäre im Rahmen der Umsetzung der CSDDD in nationales Recht aufzulösen.

05. Rechtsbereinigung mit KI unterstützen

Nachdem in den vorangegangenen Abschnitten die Machbarkeit und Verlässlichkeit der Identifikation der Vereinfachungspotentiale in Rechtsnormen erörtert wurde, stellt sich nun die Frage, inwieweit mit Hilfe von KI entsprechende Vorschläge zur Rechtsbereinigung ausgearbeitet werden können. Diese Vorschläge müssen im Einklang mit der bestehenden Rechtsordnung stehen und dürfen weder neue Regelungslücken noch neue Vereinfachungspotentiale erzeugen.

05.01 Rechtsbereinigung und Potentiale für KI-Einsatz

Die Möglichkeiten einer effektiven Rechtsbereinigung sind vielfältig: Sie reichen von klarstellenden, nicht-legislativen Maßnahmen (z. B. Leitlinien oder Praxisleitfäden) über eine verbesserte Verzahnung betroffener Rechtsakte (z. B. durch gezielte Verweise in den betroffenen Rechtstexten oder nachgelagerten Regelungen) bis hin zu einer übergreifenden Harmonisierung und Überarbeitung bestehender Regelungen.[29]

Eine KI-Anwendung kann den Rechtsbereinigungsprozess durch vielfältige Funktionalitäten (im Folgenden „intelligente Rechtsassistenten“ genannt) unterstützen und den damit verbundenen manuellen Aufwand signifikant reduzieren.[30] Solche Assistenten bieten den zuständigen Stellen (z. B. Fachabteilungen, Gesetzgebenden) zugeschnittene Funktionalitäten, die sich auf eine klar abgrenzbare Informationsbasis stützen und deren Ergebnisse leicht überprüfbar sind, so dass eine gezielte Qualitätssicherung ermöglicht wird.

Gleichwohl können die KI-Systeme nach dem derzeitigen Stand der Technik nicht eigenständig praxistaugliche, wirksame, und politisch ausgewogene Konkurrenzregelungen formulieren und diese transparent begründen. Maßgeblich dafür sind insbesondere die begrenzten Fähigkeiten aktueller KI-Systeme, die praktischen Auswirkungen einer Regelungsänderung realistisch einzuschätzen, sowie die Notwendigkeit einer variablen, kontextspezifischen Informationsbasis, die etwa aktuelle politische Zielsetzungen und relevante wirtschaftliche, gesellschaftliche und administrative Rahmenbedingungen berücksichtigt. Auch die erforderliche Interessensabwägung zwischen den beteiligten Akteuren kann eine KI nicht eigenständig leisten.

05.02 Machbarkeitsanalyse KI-gestützter Ansätze

Im Folgenden wird eine Auswahl möglicher intelligenter Rechtsassistenten vorgestellt. Derartige Rechtsassistenten können den Prozess der Vorbereitung, Formulierung und Validierung von Rechtsbereinigungen unterstützen.

05.02.01 Bereitstellung einer fundierten Informationsgrundlage

Eine KI-Anwendung kann dabei unterstützen, zahlreiche relevante Informationen zusammenzustellen, die als Grundlage der Entscheidungsfindung dienen, z. B.:

- **Relevanten Rechtskontext bereitstellen:** Mithilfe semantischer Suchverfahren und LLMs ähnlich zu Publio (Europäische Union, 2025a) lassen sich die im Kontext der Rechtsbereinigung relevanten Inhalte aus der juristischen Datenbank der etwaigen KI-Anwendung (vgl. Kapitel 02) weitgehend automatisiert ermitteln und vorschlagen. Dadurch können insbesondere Bezüge zu höherrangigem Recht, aber auch zu nachgelagerten Regelungen oder Auslegungshilfen hergestellt werden. Daneben kann die im Kapitel 03 beschriebene Kernfunktionalität genutzt werden, um weitere Vereinfachungspotentiale zu identifizieren, welche thematisch verwandt zur bereinigenden Stelle sind.

29 Diese und ähnliche Maßnahmen werden z. B. für die Wechselwirkungen zwischen der KI-Verordnung mit anderen Rechtsakten von Hacker (2024) diskutiert.
30 Vgl. auch Nationaler Normenkontrollrat (2024b).

- **Weitere relevante Anforderungen aufbereiten:** Relevante Bezüge zu politischen Rahmenbedingungen (z. B. aus dem Koalitionsvertrag) oder Anforderungen aus dem Bereich der Rechtsetzung (z. B. Arbeitshilfen und Praxischecks) lassen sich für die konkrete Rechtsbereinigung ebenfalls mit intelligenten Suchverfahren weitgehend automatisiert identifizieren. Dies erfordert den Zugriff auf die entsprechenden Dokumente. Daher kann es erforderlich sein, zusätzliche Datenquellen einzubinden, beispielsweise über das Portal verwaltung-innovativ.de bzw. die E-Gesetzgebung.

05.02.02 Validierung von Vorschlägen

Mit den in Kapitel 03 vorgestellten Verfahren lassen sich Vorschläge zur Vereinfachung des Rechts daraufhin überprüfen, ob sie neue Vereinfachungspotentiale erzeugen. Dabei wird auch deren Vereinbarkeit mit höherrangigem Recht, z. B. den EU-Vorgaben, und die Möglichkeit weiterer Normkollisionen berücksichtigt.

Zusätzlich können KI-Methoden helfen einzuschätzen, ob durch vorgeschlagene Regelungsvereinfachung neue Regelungslücken entstehen könnten. Eine abschließende automatisierte Beurteilung solcher Lücken ist jedoch nicht realisierbar, da dies die Fähigkeit zur rechtlichen Subsumtion und die Prüfung sämtlicher abgedeckter Sachverhalte und Adressaten konkreter Anwendungsfälle voraussetzt (vgl. Abschnitt 03.03.02).

05.02.03 Weitere mögliche Rechtsassistenten

Verschiedene Projekte und Initiativen auf europäischer und nationaler Ebene befassen sich bereits mit weiterführenden KI-basierten Rechtsassistenten, um Gesetzgebungsprozesse zu unterstützen oder zu validieren. So beschreibt die „Augmented LEOS“-Studie (Fitsilis, Mikros, & Leventis, 2024) diverse „Smart Functionalities“, die den Prozess der Gesetzeserstellung und -prüfung verbessern sollen. Dazu gehören kontextbezogene Rechtsprüfungen, beispielsweise die korrekte Verwendung von Verweisen und Fachterminologie, sowie automatisierte Hinweise zu juristisch mehrdeutigen Formulierungen oder widersprüchlichen Zeitbezügen. Die SETA-Studie (Hradec, et al., 2019) untersucht KI-gestützte Verfahren zu Policy-Analyse und Impact-Assessment. Auch im Bereich der E-Gesetzgebung sollen digitale Werkzeuge entwickelt werden, um den Prozess der Gesetzeserstellung zu erleichtern (Bundesministerium des Innern und für Heimat, 2024). Diese intelligenten Rechtsassistenten sind nicht nur für die Bereinigung bestehender Rechtsnormen relevant, sondern können potenziell in verschiedenen Stadien der Gesetzgebung eingesetzt werden. Entsprechend empfiehlt beispielsweise der Nationale Normenkontrollrat eine direkte Integration solcher Funktionalitäten in den Editor der E-Gesetzgebung (Nationaler Normenkontrollrat, 2024b).

05.03 Fazit

Dieser Abschnitt befasst sich mit der Vorbereitung von Vorschlägen zur Rechtsvereinfachung, die auf den zuvor identifizierten Vereinfachungspotentialen beruhen. **Während die Auswahl und Ausgestaltung konkreter Maßnahmen legislativer und politischer Verantwortung unterliegen, kann eine KI-Anwendung den dafür erforderlichen manuellen Aufwand erheblich reduzieren.**

Insbesondere lassen sich durch intelligente Rechtsassistenten relevante Informationen, sofern sie in der zugrundeliegenden juristischen Datenbank vorliegen, automatisiert identifizieren und strukturiert aufbereiten sowie Vorschläge zur Rechtsbereinigung validieren, beispielsweise im Hinblick auf Normkollisionen oder weitergehende Vereinfachungsmöglichkeiten. Zahlreiche weitere Rechtsassistenten werden bereits im Rahmen verschiedener Initiativen untersucht und erprobt. Da diese jedoch nicht speziell auf die Rechtsbereinigung, sondern allgemein auf die Rechtsetzung ausgerichtet sind, bietet sich eine Integration entsprechender Funktionalitäten in die E-Gesetzgebung an.

06. KI-Anwendung umsetzen: Eine mögliche Roadmap

Angesichts des innovativen Charakters der beschriebenen KI-Anwendung und des rasanten technologischen Fortschritts im Bereich Künstlicher Intelligenz empfehlen wir für eine mögliche Realisierung ein iteratives, modulares und nutzerzentriertes Vorgehen. Dieses Vorgehen trägt dazu bei, Risiken – etwa hohe Kosten, Verzögerungen oder ausbleibende Mehrwerte – zu reduzieren und den Nutzen im Hinblick auf Bürokratieabbau zu maximieren. Zudem ermöglicht es eine schnelle Umsetzung und flexible Anpassungen an neue technologische Entwicklungen im Kontext von KI.

Um diese Vision zu realisieren, bietet es sich an, zuerst einen lokalen Prototyp umzusetzen, der ausgewählte Rechtsquellen einbindet (z. B. EU-Verordnungen und Richtlinien sowie Bundesgesetze) sowie die Kernfunktionalitäten prototypisch umsetzt. Darauf aufbauend kann die Anwendung schrittweise zu einer voll funktionsfähigen Lösung ausgebaut werden. Tabelle 7 zeigt die mögliche Phasenaufteilung.

Prototyp	Minimum Viable Product	Produktive Anwendung
Datenbasis: Bundesgesetze sowie EU-Verordnungen und Richtlinien (z. B. aus Gesetze-im-Internet und EUR-Lex). **Funktionalitäten:** Prototypische Umsetzung · der Suche zur Identifikation relevanter Regelungen mit überlappendem Geltungsbereich · des Vergleichs ausgewählter Regelungen, um Vereinfachungspotentiale zu erkennen **Konzeption:** · Exploration der Möglichkeit zur Einbindung juristischer Datenbanken (z. B. juris) · Erste Prüfung rechtlicher und technischer Vorgaben für einen späteren Produktivbetrieb	**Datenbasis:** Juristische Datenbank. **Funktionalitäten:** · Weiterentwicklung des Prototyps (Analyse der Geltungsbereiche und der Vereinfachungspotentiale, optimierte Benutzeroberfläche) · Einbau Validierung und Qualitätssicherung **Konzeption:** · Aufbereitung und Abstimmung rechtlicher, technischer und organisatorischer Anforderungen für den künftigen Produktivbetrieb · Auswahl geeigneter Lizenzmodelle · Entwurf der Zielarchitektur und Integration in bestehende Verwaltungsprozesse	Die **Produktivnahme** umfasst: · **Verfeinerung und Umsetzung der Produktionsanforderungen** · **Inbetriebnahme** · **Ggf. Migration** des MVP auf eine leistungsfähige Plattform (z. B. KIPITZ) Anschließend können (iterative) **Weiterentwicklungen** erfolgen, z. B.: · **Qualitätsverbesserungen**, z. B. Performance und Verlässlichkeit · **Erweiterung der Funktionalitäten**, z. B. um Vorschläge für Rechtsbereinigung, Aufnahme weiterer Regelungsebenen, etc.

Tabelle 7: Mögliche Gliederung der Phasen bei der Entwicklung der beschriebenen KI-Anwendung (illustrativ)

Für die Entwicklung und den Pilotbetrieb des Prototyps sind auf Basis der im Rahmen der Studie durchgeführten Kostenschätzung rund 345 Personentage externer Dienstleister sowie 75 Personentage interner Ressourcen einzuplanen. Zusätzlich entstehen IT-Infrastruktur- und Lizenzkosten von etwa 11 500 € (brutto). Beim anschließenden Minimum Viable Product (MVP) ist abhängig von den Ergebnissen der prototypischen Umsetzung und den finalen Anforderungen mit dem Drei- bis Siebenfachen dieser Aufwände zu rechnen, was einem Budgetbedarf von ca. 2,0 – 4,6 Mio. Euro entspricht. Eine spätere produktive Skalierung verursacht weitere, nutzungsabhängige Kosten, ermöglicht jedoch gleichzeitig substanzielle Effizienz- und Automatisierungsgewinne. Ein iteratives Vorgehen sowie ein konsequenter Markt-Check vor jeder Ausbaustufe verhindern dabei Doppelentwicklungen und sichern budgetäre Flexibilität.

07. Gesamtfazit und Ausblick

Die vorliegende Machbarkeitsstudie zeigt, dass eine KI-gestützte Anwendung zur Identifikation der Vereinfachungspotentiale von Regelungen grundsätzlich realisierbar und vielversprechend ist. Mit Blick auf die zentralen Aspekte ergibt sich folgendes Gesamtbild:

01. **Juristische Wissensbasis einer KI-Anwendung bereitstellen** (Kapitel 02): Eine umfassende und aktuelle juristische Wissensbasis ist wesentlich für eine erfolgreiche KI-Anwendung. Aufgrund der fragmentierten Datenlandschaft und Lizenzbeschränkungen empfiehlt es sich, zunächst auf frei zugängliche staatliche Rechtsnormen zu fokussieren und gegebenenfalls Synergien mit existierenden oder geplanten Rechtsinformationssystemen (z. B. juris und NeuRIS) zu nutzen.

02. **Vereinfachungspotentiale mit KI identifizieren** (Kapitel 03): Die Studie definiert Doppelungen, Widersprüche und Inkonsistenzen als wesentliche Kategorien der Vereinfachungspotentiale. Die zentralen Voraussetzungen für eine KI-gestützte Identifikation dieser Potentiale sind die Fähigkeit einer KI-Anwendung, strukturelle und kontextuelle Nuancen von Rechtstexten korrekt zu erfassen, darauf aufbauend Geltungsbereiche und Vorschriften miteinander zu vergleichen sowie Analyseergebnisse nachvollziehbar und überprüfbar zu begründen. Da nur wenige vergleichbare Studien vorliegen, wurden im Rahmen dieser Machbarkeitsstudie auf Basis einer Literaturrecherche und aktueller technologischer Entwicklungen passende Lösungsansätze dafür entwickelt.

03. **Verlässlichkeit KI-gestützter Ergebnisse optimieren** (Kapitel 04): Die theoretische Analyse und die prototypische Erprobung legen nahe, dass eine hohe Verlässlichkeit einer KI-Anwendung in Bezug auf Plausibilität, Vollständigkeit, Konsistenz, Robustheit und Erklärbarkeit/Verifizierbarkeit erreicht werden kann, sofern die im Kapitel 03 beschriebenen Methoden durch zusätzliche Validierungs- und Qualitätssicherungsmechanismen ergänzt werden.

04. **Rechtsbereinigung mit KI unterstützen** (Kapitel 05): Eine KI-Anwendung kann das Verfahren der Rechtsbereinigung auf Basis der identifizierten Vereinfachungspotentiale deutlich beschleunigen, beispielsweise indem sie relevante Informationen strukturiert aufbereitet oder Vorschläge in Hinblick auf neue Vereinfachungspotentiale validiert. Gleichwohl obliegt die Entscheidung über die Auswahl und Ausgestaltung konkreter Maßnahmen den legislativen und politischen Instanzen.

05. **KI-Anwendung umsetzen: Eine mögliche Roadmap** (Kapitel 06): Angesichts des raschen technologischen Fortschritts im Bereich KI und der Komplexität einer solchen Anwendung empfiehlt sich ein iteratives, modulares und nutzerzentriertes Vorgehen bei der Umsetzung. Im ersten Schritt kann ein lokaler Prototyp die Kernfunktionalitäten, nämlich der Geltungsbereichsanalyse und der Identifikation der Vereinfachungspotentiale, demonstrieren und eine rasche Lernkurve ermöglichen. Darauf aufbauend kann der Prototyp schrittweise zu einer voll funktionsfähigen KI-Anwendung weiterentwickelt und in bestehende IT-Strukturen eingebettet werden.

08. Literaturverzeichnis

Arabzadeh, N., Huo, S., Mehta, N., Wu, Q., Wang, C., Awadallah, A. H., ... Kiseleva, J. (2024). Assessing and Verifying Task Utility in LLM-Powered Applications. *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing* (S. 21868-21888). Miami, Florida, USA: Association for Computational Linguistics. doi:10.18653/v1/2024.emnlp-main.1219

Bank for International Settlements. (2025). *Terms and conditions of use*. Von https://www.bis.org/terms_conditions.htm abgerufen

Blümmel, T., Rehn, J., Mereu, C., Graf, F., Bazing, F., Kneuer, C., ...Dobrikov, T. (2024). Exploring the use of Artificial Intelligence (AI) for extracting and integrating data obtained through New Approach Methodologies (NAMs) for chemical risk assessment. *EFSA Supporting Publications, 21*(1). doi:10.2903/sp.efsa.2024.EN-8567

Breidenbach, S. (2025). *Was Gesetze sein könnten*. München: C.H. Beck.

Bundesministerium des Innern und für Heimat. (2024). *E-Gesetzgebung. Documentation*. Von https://gitlab.opencode.de/bmi/e-gesetzgebung/documentation abgerufen

Bundesministerium des Innern und für Heimat. (2025). *"Dieselbe Sprache sprechen" – semantische Interoperabilität als Schlüssel für effiziente Verwaltung und Once-Only-Prinzip*. Von https://www.digitale-verwaltung.de/Webs/DV/DE/onlinezugangsgesetz/rahmenarchitektur/standardisierung/sprachliche-interoperabilitaet/sprachliche-interoperabilitaet-node.html abgerufen

Bundesministerium für Justiz. (2024). *Das neue Rechtsinformationsportal – für einen einheitlichen und modernen Zugang zu Rechtsinformationen des Bundes*. Von https://www.bmj.de/DE/themen/digitales/digitalisierung_justiz/digitalisierungsinitiative/_articles/rechtsinformationsportal_artikel.html abgerufen

C.H.Beck oHG. (2025). *Allgemeine Geschäftsbedingungen*. Von https://rsw.beck.de/beck-online-service/agb-beck-online abgerufen

Chalkidis, I., Fergadiotis, M., Malakasiotis, P., Aletras, N., & Androutsopoulos, I. (2020). LEGAL-BERT: The Muppets straight out of Law School. *Findings of the Association for Computational Linguistics: EMNLP 2020* (S. 2898–2904). Association for Computational Linguistics. doi:10.18653/v1/2020.findings-emnlp.261

Colombo, P., Pires, T. P., Boudiaf, M., Culver, D., Melo, R., Corro, C., . . . Desa, M. (2024). SaulLM-7B: A pioneering Large Language Model for Law. arXiv. doi:10.48550/arXiv.2403.03883

Coupette, C., & Fleckner, A. M. (2018). Quantitative Rechtswissenschaft: Sammlung, Analyse und Kommunikation juristischer Daten. *JuristenZeitung*, 73, 379–389. doi:10.1628/jz-2018-0020

Dahl, M., Magesh, V., Suzgun, M., & Ho, D. E. (2024). Large Legal Fictions: Profiling Legal Hallucinations in Large Language Models. *Journal of Legal Analysis, 16*, 64–93. doi:10.1093/jla/laae003

Deloitte. (2021). *Kostenbarometer Regulatorik (Kostbar): Regulatorische Aufwände für Unternehmen der Versicherungsbranche und des Maschinenbaus*. Von https://www.deloitte.com/content/dam/assets-zone2/de/de/docs/services/risk-advisory/2024/Deloitte-Kostbar-Abschlussbericht.pdf abgerufen

Der Beauftragte der Bundesregierung für Informationstechnik. (2025). *E-Verkündung*. Von https://www.cio.bund.de/Webs/CIO/DE/digitale-loesungen/it-konsolidierung/dienstekonsolidierung/it-massnahmen/e-verkuendung/e-verkuendung-node.html abgerufen

| Deutscher Bundestag. (2020). *Antwort der Bundesregierung auf die Kleine Anfrage der Abgeordneten Victor Perli, Niema Movassat, Dr. Gesine Lötzsch, weiterer Abgeordneter und der Fraktion DIE LINKE*. Bundestags-Drucksache 19/25654.

| Die Bundesregierung. (2025). *Bürokratie abbauen – Wirtschaft entlasten*. Von https://www.bundesregierung.de/breg-de/aktuelles/buerobratie-abbauen-2264628 abgerufen

| Europäische Kommission. (2025a). *Kommission will Bürokratie und Hürden für Unternehmen abbauen*. Von https://commission.europa.eu/news/commission-proposes-cut-red-tape-and-simplify-business-environment-2025-02-26_de abgerufen

| Europäische Kommission. (2025b). LEOS - *Open Source software for editing legislation*. Von https://interoperable-europe.ec.europa.eu/collection/justice-law-and-security/solution/leos-open-source-software-editing-legislation abgerufen

Europäische Union. (2023). V*irtual assistant : Explainability notice. Publio – Intelligent assistant of the Portal of the Publications Office of the European Union (OP Portal)*. Von https://op.europa.eu/en/web/about-us/explainability-notice abgerufen

Europäische Union. (2025a). *Publio, the Publications Office virtual assistant*. Von https://op.europa.eu/en/web/webtools/publio-the-publications-office-virtual-assistant abgerufen

| Europäische Union. (2025b). AKN4EU. *A Common Structured Format for EU Legislative Documents*. Von https://op.europa.eu/de/web/eu-vocabularies/akn4eu abgerufen

| Fitsilis, F., Mikros, G., & Leventis, S. (2024). *Overview of smart functionalities in drafting legislation in LEOS*. Brüssel, Belgien: Europäische Kommission. Von https://interoperable-europe.ec.europa.eu/sites/default/files/document/2024-06/Augmented LEOS with smart functionalities final report.pdf abgerufen

| Gehrmann, S., Clark, E., & Sellam, T. (2023). Repairing the Cracked Foundation: A Survey of Obstacles in Evaluation Practices for Generated Text. *Journal of Artificial Intelligence Research*, 77, 103–166. doi:10.1613/jair.1.13715

| Hacker, P. (2024). *Der AI Act im Spannungsfeld von digitaler und sektoraler Regulierung*. Gütersloh: Bertelsmann Stiftung. doi:10.11586/2024189

| Hamann, H. (2021). Der blinde Fleck der deutschen Rechtswissenschaft – Zur digitalen Verfügbarkeit instanzgerichtlicher Rechtsprechung. *JuristenZeitung*, 76, 656–665. doi:10.1628/jz-2021-0225

| Hradec, J., Ostlaender, N., Macmillan, C., Acs, S., Listorti, G., Tomas, R., & Arnes Novau, X. (2019). *Semantic Text Analysis tool: SeTA. Supporting analysts by applying advanced text mining techniques to large document collections*. Luxemburg: Publications Office of the European Union. doi:10.2760/577814

| juris GmbH. (2025). *Allgemeine Geschäftsbedingungen der juris GmbH*. Von https://www.juris.de/jportal/nav/services/agb/index.jsp abgerufen

| Kalra, R., Wu, Z., Gulley, A., Hilliard, A., Guan, X., Koshiyama, A., & Treleaven, P. C. (2024). HyPA-RAG: A Hybrid Parameter Adaptive Retrieval-Augmented Generation System for AI Legal and Policy Applications. *Proceedings of the 1st Workshop on Customizable NLP: Progress and Challenges in Customizing NLP for a Domain, Application, Group, or Individual (CustomNLP4U)* (S. 237-256). Miami, Florida, USA: Association for Computational Linguistics.

| Katz, D. M., Bommarito, M. J., Gao, S., & Arredondo, P. (2024). GPT-4 passes the bar exam. *Philosophical Transactions of the Royal Society A*, 382(2270), 20230254. doi:10.1098/rsta.2023.0254

| Lee, Y., Son, K., Kim, T. S., Kim, J., Chung, J. J., Adar, E., & Kim, J. (2024). One vs. Many: *Comprehending Accurate Information from Multiple Erroneous and Inconsistent AI Generations. Proceedings of the 2024 ACM Conference on Fairness, Accountability, and*

Transparency (S. 2518–2531). New York, NY, USA: Association for Computing Machinery. doi:10.1145/3630106.3662681

Magesh, V., Surani, F., Dahl, M., Suzgun, M., Manning, C. D., & Ho, D. E. (2024). Hallucination-Free? Assessing the Reliability of Leading AI Legal Research Tools. arXiv. doi:10.48550/ARXIV.2405.20362

Microsoft. (2025). *Language Understanding (LUIS)*. Von https://www.luis.ai/ abgerufen

Nationaler Normenkontrollrat. (2023). *Jahresbericht 2023. Weniger, einfacher, digitaler. Bürokratie abbauen. Deutschland zukunftsfähig machen*. Berlin: Nationaler Normenkontrollrat. Von https://www.normenkontrollrat.bund.de/Webs/NKR/SharedDocs/Downloads/DE/Jahresberichte/2023-jahresbericht.pdf?__blob=publicationFile&v=5 abgerufen

Nationaler Normenkontrollrat. (2024a). *Jahresbericht 2024. Gute Gesetze. Digitale Verwaltung. Weniger Bürokratie. Momentum nutzen, Wirkung steigern. Berlin: Nationaler Normenkontrollrat*. Von https://www.normenkontrollrat.bund.de/Webs/NKR/SharedDocs/Downloads/DE/Jahresberichte/2024-jahresbericht.pdf?__blob=publicationFile&v=3 abgerufen

Nationaler Normenkontrollrat. (2024b). *Potenziale und Grenzen von Künstlicher Intelligenz in der Gesetzgebung*. Von https://www.normenkontrollrat.bund.de/Webs/NKR/SharedDocs/Newsletter/DE/2024/2024-6/04_artikel_ki_gesetzgebung.html abgerufen

OpenAI. (2023). *GPT-4 Technical Report*. arXiv. doi:10.48550/ARXIV.2303.08774

OpenAI. (2024). *Learning to reason with LLMs*. Von https://openai.com/index/learning-to-reason-with-llms/ abgerufen

Palmirani, M., Vitali, F., Van Puymbroeck, W., & Nubla Durango, F. (2022). *Legal Drafting in the Era of Artificial Intelligence and Digitisation*. Brüssel, Belgien: Europäische Kommission. Von https://interoperable-europe.ec.europa.eu/sites/default/files/document/2022-06/Drafting%20legislation%20in%20the%20era%20of%20AI%20and%20digitisation%20%E2%80%93%20study.pdf abgerufen

Spatharioti, S. E., Rothschild, D. M., Goldstein, D. G., & Hofman, J. M. (2023). Comparing Traditional and LLM-based Search for Consumer Choice: A Randomized Experiment. arXiv. doi:10.48550/arXiv.2307.03744

Verwaltung Innovativ. (2023). *Projekt E-Gesetzgebung*. Von https://www.verwaltung-innovativ.de/DE/Gesetzgebung/Projekt_eGesetzgebung/projekt_E_gesetzgebung_node.html abgerufen

Viegas, C. F., Costa, B. C., & Ishii, R. P. (2023). JurisBERT: A New Approach that Converts a Classification Corpus into an STS One. *In Computational Science and Its Applications – ICCSA 2023* (S. 349–365). Springer Nature Switzerland. doi:10.1007/978-3-031-36805-9_24

von der Leyen, U. (2024). *Europa hat die Wahl. Politische Leitlinien für die nächste Europäische Kommission 2024-2029*. Von https://commission.europa.eu/document/download/e6cd4328-673c-4e7a-8683-f63ffb2cf648_de?filename=Political Guidelines 2024-2029_DE.pdf abgerufen

09. Ansprechpersonen

Dr. Petra Gutjahr
Leiterin Grundsatzabteilung,
d-fine

Petra.Gutjahr@d-fine.com

Marc Girshgorn
Senior Manager, d-fine

Marc.Girshgorn@d-fine.com

Dr. Ferdinand Graf
Partner, d-fine

Ferdinand.Graf@d-fine.com

10. Glossar

Application Programming Interface (kurz: API): Eine Schnittstelle, die es ermöglicht, dass Softwareanwendungen miteinander interagieren und bestimmte Funktionen oder Daten eines Betriebssystems, einer Anwendung oder eines Dienstes nutzen.

Akoma Ntoso-Format: Ein XML-Datenformat, das speziell für die strukturierte Darstellung von juristischen Dokumenten und deren Metadaten entwickelt wurde. Es fördert die Interoperabilität zwischen verschiedenen Systemen und erleichtert die automatisierte Verarbeitung und das Auffinden von rechtlichen Informationen.

BERT (Bidirectional Encoder Representations from Transformers): BERT ist ein nicht-generatives *Transformer-basiertes Sprachmodell*, das auf umfangreichen Textkorpora vortrainiert wird und für spezifische Aufgaben feinabgestimmt werden kann. Es existieren Varianten wie *LEGAL-BERT* (Chalkidis, Fergadiotis, Malakasiotis, Aletras, & Androutsopoulos, 2020) und *JurisBERT* (Viegas, Costa, & Ishii, 2023), die speziell für juristische Texte entwickelt wurden, um die spezifische Terminologie und Struktur juristischer Sprache besser zu verstehen und zu verarbeiten.

Entitätenerkennung (Named Entity Recognition, kurz: NER): Eine KI-Methode erkennt automatisch Eigennamen und Entitäten sowie ihre Relevanz im Text, wie bspw. „BaFin“ oder „Bundesrepublik Deutschland“, ebenso Daten oder Versionsnummern, wenn diese in immer wieder ähnlichen Kontexten auftauchen.

Große generative Sprachmodelle (Large Language Models, kurz: LLMs): Besonders leistungsstarke Klasse von Sprachmodellen, mit denen ein breites Spektrum domänenspezifischer Texte analysiert und generiert werden kann. Für Nutzerinnen und Nutzer ist die Bedienung der Modelle über einen *Prompt* sehr intuitiv, da explizite Aufgabenstellungen in natürlicher Sprache direkt an die KI formuliert werden können. Zu dieser Kategorie gehören sowohl proprietäre Modelle wie GPT-4o von *OpenAI* als auch Open-Source-Modelle von *Mistral AI, Meta, OpenGPT-X* oder *Aleph Alpha*.

Halluzination: Im Kontext von *großen generativen Sprachmodellen* beziehen sich Halluzinationen auf die Erzeugung von falschen oder erfundenen Informationen, die vom Modell als wahr dargestellt werden, was zu Missverständnissen oder Fehlinformationen führen kann.[31]

Künstliche Intelligenz (kurz: KI): Ein KI-System ist ein maschinengestütztes System, das für einen in unterschiedlichem Grade autonomen Betrieb ausgelegt ist und das nach seiner Betriebsaufnahme anpassungsfähig sein kann und das aus den erhaltenen Eingaben für explizite oder implizite Ziele ableitet, wie Ausgaben wie etwa Vorhersagen, Inhalte, Empfehlungen oder Entscheidungen erstellt werden, die physische oder virtuelle Umgebungen beeinflussen können (Definition gem. KI-Verordnung).[32]

Large Language Model (kurz: LLM): Englische Bezeichnung für *großes generatives Sprachmodell*.

Maschinelles Lernen (ohne neuronale Netze): Diese Methoden verwenden Algorithmen, die aus annotierten Daten lernen, um Textklassifizierung und -analyse durchzuführen. Beispiele hierfür sind Support Vector Machines oder regressive Modelle wie XGboost. Besonders in großen Datenmengen mit hoher Dimensionalität können diese Modelle Strukturen erkennen und Vorhersagen treffen, die aufgrund der Datenmenge manuell schwer zugänglich sind.

Pipeline: Festgelegte Abfolge von Prozessschritten, z. B. zur Vorverarbeitung und Validierung.

Prompt: Anfragentext, der einem *LLM* übergeben wird.

31 Siehe die Betrachtung von Halluzinationen im Bereich Legal AI durch Magesh et al. (2024).
32 Kapitel I, Art. 3, Punkt 1.

Prompt Engineering: Gezielte Abwandlung und Verbesserung des *Prompts*, der einem *LLM* übergeben wird, um die Qualität der Antwort zu verbessern.

Regelbasierte Verfahren: Diese Methoden nutzen vordefinierte Regeln und Muster, um spezifische Informationen aus Texten zu extrahieren. Während sie in klar strukturierten Domänen effektiv sein können, stoßen sie bei der Komplexität und Variabilität juristischer Sprache jedoch schnell an ihre Grenzen.

Semantische Ähnlichkeitsbestimmung: Algorithmen zur Messung der Ähnlichkeit und Stärke semantischer Beziehungen zwischen Begriffen, Phrasen oder Texten analysieren Wortfrequenzen, Koexistenz von Wörtern sowie Wortbeziehungen oder Kontextinformationen.

Transformer-basierte Sprachmodelle (kurz: Sprachmodelle): Diese Modelle basieren auf der Transformer-Architektur, die sich durch ihre Effizienz und Fähigkeit zur Verarbeitung von kontextueller Information auszeichnet. Sie können die Bedeutung und Struktur von Texten tief erfassen und kontextuelle Ähnlichkeiten erkennen, selbst wenn unterschiedliche Formulierungen oder Begriffe verwendet werden. Beispiele für solche Modelle sind Sprachmodelle wie *BERT* und auch *große generative Sprachmodelle* (LLMs) wie GPT oder Mistral.

d-fine

analytisch. technologisch. quantitativ.