

Der Einsatz von Neural Language Models für eine barrierefreie Verwaltungskommunikation

Anforderungen an die automatisierte Vereinfachung rechtlicher Informationstexte

Michael Gille¹, Thorben Schomacker², Jörg von der Hülls³, Marina Tropmann-Frick^{4 5 6}

Abstract: Machine-learning-based text simplification can be used to meet legal obligations to provide comprehensibility-enhanced public service texts. The article examines the use of artificial intelligence for public administration communication in rule-based easy language. The authors outline essential technical, legal and normative requirements for the development and use of automated text simplification through neural language generation. As part of the Open-LS research project, a contribution is made to clarifying the possibilities and limits of the use of artificial intelligence systems for text simplification in public services.

Keywords: Text simplification, Transformer, AI Regulation, Leichte Sprache, DIN SPEC 33429

1 Einleitung

Das Textverständnis kognitiv und lernbeeinträchtigter Menschen⁷ wird durch Übertragungen in verständlichkeitsoptimierte Sprache unterstützt, gerade auch um Inklusion und soziale Teilhabe zu fördern [Bund08]. Bisher überführen menschliche Übersetzer Texte in vereinfachte Versionen. Die Umwandlung von Informations- und

¹ michael.gille@haw-hamburg.de, <https://orcid.org/0000-0001-9978-0817>

² thorben.schomacker@haw-hamburg.de, <https://orcid.org/0009-0002-3577-341X>

³ joerg.vonderhuelles@haw-hamburg.de

⁴ marina.tropmann-frick@haw-hamburg.de, <https://orcid.org/0000-0003-1623-5309>

⁵ Alle Autoren sind tätig an der Hochschule für Angewandte Wissenschaften Hamburg (HAW Hamburg), Berliner Tor 5, 20099 Hamburg, Deutschland.

⁶ Diese Arbeit wird gefördert durch UpdateHamburg (Schomacker) und durch Hamburg Call for Transfer (von der Hülls). Das Projektteam dankt der Hamburgischen Investitions- und Förderbank für die Projektförderung des Projekts Open-LS im Rahmen des Programms PROFI Impuls. Wir danken zudem den anonymen Gutachtern für ihr wertvolles Feedback.

⁷ Während in der Gesetzgebung von „Menschen mit geistigen Behinderungen und Menschen mit seelischen Behinderungen“ (§ 11 (1) Satz 1 BGG) die Rede ist, wird im Rahmen dieses Beitrags von kognitiv und lernbeeinträchtigten Menschen oder, in Übereinstimmung mit DIN SPEC 33429-E [Dinn23], die Bezeichnung „Menschen mit Lernschwierigkeiten“ verwendet. Der gesetzliche Begriff wird teilweise als (zu) defizitorientiert und diskriminierend empfunden.

Kommunikationstexten in eine barrierefreie Sprachfassung insbesondere nach § 11 Behindertengleichstellungsgesetz (BGG) [KoVM23] sind mit Kosten für den öffentlichen Haushalt verbunden. Zwar sind seit der Einführung gesetzlicher Vorschriften zur Barrierefreiheit vor über zwanzig Jahren erhebliche Fortschritte zu verzeichnen, jedoch bleibt die Umsetzung lückenhaft [Thap21, AsHZ23, Bund21, Ahle23], obwohl bis zu 12 % der Bevölkerung Deutschlands Schwierigkeiten haben, Standardsprache zu verstehen und zu verwenden [GrBu20]. Automatisierte Ansätze, die auf Techniken des maschinellen Lernens basieren, können helfen, die Kosten der Textvereinfachung zu senken [GrBu20, Gugg19]. Anders als Chatbots und ähnliche Dialogsysteme, die Unterhaltungen simulieren, geht es bei der Textvereinfachung um Übersetzungen in barrierefreie Sprache, nicht unähnlich einer Übersetzungsleistung in eine andere Sprache. Der Einsatz von Künstlicher Intelligenz (KI)⁸ kann die Effizienz der Verwaltung steigern und sogar individuelle Kommunikationslösungen ermöglichen [Gugg19, Deut22]. Das vergleichsweise neue Forschungsfeld der automatisierten Textvereinfachung (Text Simplification, TS) greift auf KI in Gestalt neuronaler Sprachmodelle (Neural Language Models, NLM) zurück, um Texte in komplexitätsreduzierte Fassungen zu übertragen. Vortrainierte Sprachmodelle erlernen die Vereinfachungsaufgaben maschinell anhand großer Mengen spezieller Trainingsdaten insbesondere in Gestalt paralleler Korpora, wobei es bislang kaum geeignete deutschsprachige Datensätze gibt [Stod22].

Der Beitrag untersucht technische, regulatorische und normative Rahmenanforderungen des Einsatzes machine-learning-basierter Anwendungen zur Textvereinfachung für den verständlichkeitsoptimierten behördlichen Kontakt mit Personen, die auf barrierefreie Text- und Kommunikationsformen angewiesen sind. Hintergrund ist das interdisziplinäre Forschungsprojekt Open LS. Mit Methoden der Computerlinguistik wird ein textvereinfachendes NLM für Textformate entwickelt, das auf Personen mit Lernschwierigkeiten zugeschnitten und gezielt für die Übertragung von Rechtstexten trainiert wird, um deren Verständlichkeit für diesen Adressatenkreis zu verbessern. Nachfolgend werden spezifische Anforderungen an textvereinfachende KI-Systeme für die Erstellung barrierefreier rechtlicher Informationstexte der öffentlichen Verwaltung betrachtet und erste Ergebnisse aus dem Projekt (Open-LS) vorgestellt. Zunächst werden der Soll-Output sowie die technischen Modellmerkmale der Neural Language Generation (NLG) beim Einsatz für die Textvereinfachungsaufgabe umrissen und wesentliche Schritte des Trainings beschrieben (Abschnitt 2). Anschließend wird der rechtlich-normative Rahmen auf Anforderungen für Modell und Betrieb hin untersucht (Abschnitt 3) und Modellanforderungen abgeleitet (Abschnitt 4). Der Beitrag schließt mit einem Ausblick einschließlich Angaben zu Limitationen und Forschungslücken (Abschnitt 5).

⁸ „KI“ meint eine Maschine, welche dazu gebracht wird, sich so zu verhalten, wie man es bei einem Menschen als intelligent bezeichnen würde (gemäß der Definition in [MMRS55]). Dem Begriff „KI-System“ liegt im Folgenden die Definition aus Art. 3 Nr. 1 KI-VO (Entwurf der EU-Kommission v. 21.4.2021) zugrunde.

2 Automatische Vereinfachung rechtlicher Informationstexte durch Neural Language Generation

2.1 Leichte-Sprache-Text als Output der Vereinfachung

Ein Text gilt gemäß § 4 BITV 2.0 als barrierefrei, wenn er in "Leichter Sprache" formuliert ist. Unter Leichter Sprache (LS) wird eine verständlichkeitsoptimierte und regelbasierte Kunstform des Deutschen verstanden, nicht zu verwechseln mit "Einfacher Sprache", womit vereinfachte Varianten im Graubereich zwischen Standardsprache und LS bezeichnet werden [Maaß20]. Existierende Sprachmodelle wie ChatGPT wurden überwiegend auf Alltags- und Standardsprache trainiert. LS unterscheidet sich von Alltagssprache, die mit geringen regionalen Abweichungen von der Mehrheit der Bevölkerung deutschsprachiger Länder verwendet wird. Derzeit übersetzen spezialisierte menschliche Übersetzer standardsprachliche Texte in Leichte Sprache, wobei juristische Texte aufgrund ihres fachsprachlichen Charakters und ihres normativen Inhalts eine besondere Herausforderung darstellen [GaGe22].

Rechtliche Texte weisen in den verschiedenen Rechtsgebieten ähnliche sprachliche Merkmale auf, wie z.B. die Verwendung juristischer Fachsprache, Formalisierung, lange und komplexe Sätze, ein hohes Maß an Intertextualität, häufig gemischte Autorenschaft, ein breites Spektrum von Adressaten und ein fachspezifisches Spannungsverhältnis zwischen Genauigkeit der Formulierungen und begrifflicher Vagheit [Baum20]. Darüber hinaus sind viele Rechtstexte funktional auf Rechtsverbindlichkeit angelegt und begründen Rechte und Pflichten. In der Begründung, Organisation und Kommunikation rechtlicher Beziehungen unterscheiden sich diese Texte grundlegend von Texten, die Gegenstand der meisten zum NLM-Training genutzten intra- und monolingualen Korpora sind.

Die Vereinheitlichung von Leichte-Sprache-Regeln ist auch im Hinblick auf die barrierefreie Kommunikation durch die öffentliche Hand relevant. Die Empfehlungen des Standards des Deutschen Instituts für Normung, DIN SPEC 33429 (Entwurf), richten sich an den gesamten Personenkreis der an der Erstellung von Texten in LS Beteiligten einschließlich Auftraggebern, wobei diese Regeln auch bei Ausschreibungen sowie bei der Qualitätssicherung herangezogen werden können [Dinn23]. Dieses Regelwerk wird hier aufgrund der absehbaren Maßgeblichkeit für Übersetzungen in LS bei der Formulierung von Anforderungen an Output-Texte zugrunde gelegt.

2.2 Neural Language Models

Spezialisierte Modelle wie Neuronale Netze für die Sprachverarbeitung und -modellierung (= NLM) werden ständig weiterentwickelt. Die Grundlage für das Gelingen jedes Vorhabens mit neuronalen Netzen sind Daten. Die Verarbeitung von Sprachdaten

durch NLM beginnt zunächst mit Tokenisierung. Hier wird anfänglich der Text in linguistische Bausteine (wie beispielsweise Wörter) unterteilt. Diese Bausteine bilden dann semantische Einheiten, die sog. Token. Z.B. ist der Token „Ludwig XIV.“ aus zwei Wörtern zu einer semantischen Einheit zusammengefasst worden. Im nächsten Schritt, dem Embedding, werden die Token in eine verteilte mathematische Repräsentation (Vektor) transformiert. Dadurch können semantische Beziehungen durch die Verteilung im Vektorraum dargestellt werden [TuWW22].

Im Bereich des NLP gibt es zwar textuelle Daten in großer Zahl, gleichwohl mangelt es für viele Anwendungen an task-spezifischen, aufbereiteten Daten [ScTr21], um NLMs ausreichend zu trainieren. Diese Schwierigkeit kann durch Transferlernen bewältigt werden. Transferlernen ermöglicht es, aus Daten, die nicht speziell für den Task aufbereitet wurden, Wissen abzuleiten und dieses später für spezifische Tasks abzurufen. Die wohl weitverbreitetste Architektur für Transferlernen ist die Transformer Architektur [VSPU17]. Diese Architektur ermöglicht, das Anlernen des NLM in zwei Schritte aufzuteilen: 1) Vortrainieren auf nicht aufbereiteten Rohdaten, wodurch dem Modell ein generelles Welt- und Sprachverständnis beigebracht wird. 2) task-spezifisches Fine-Tuning, beispielsweise Textklassifikation oder Übersetzen. Auch wenige task-spezifische Datenpunkte können reichen, da das Modell bereits Sprachverständnis durch Vortrainieren aufweist. [TuWW22] Dadurch können task-spezifische Daten präziser für das Lernen von task-spezifischem Verständnis verwendet werden und müssen kein grundlegendes Sprachverständnis mehr schaffen [RSRL20].

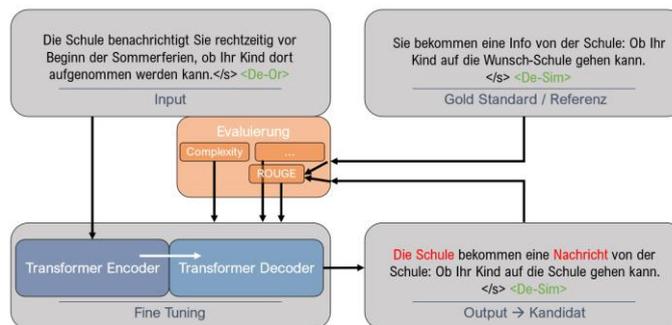
Ein klassischer Transformer besteht aus zwei Komponenten: 1) Der Encoder extrahiert Eigenschaften aus dem Input und speichert in einer verteilten Repräsentation (wie bspw. Vektoren) den Kontext. 2) Der Decoder verarbeitet sequenziell den Kontext zu einem task-spezifischen Ergebnis. Seit ihrer Einführung haben sich drei Stränge der Transformer Architektur herausgebildet [TuWW22, S.78–84]: 1) **Encoder-Only** Modelle eignen sich vor allem für NLU-tasks (Natural Language Understanding) wie zum Beispiel Klassifikation, Eigennamenerkennung oder das Beantworten von Fragen. 2) **Decoder-Only** sind vergleichsweise gut darin, das nächste Wort vorherzusagen. Daher werden sie häufig für Textgenerierung (= Natural Language Generation (NLG)) eingesetzt. Die bekanntesten Vertreter sind die Modelle aus der GPT-Familie. 3) **Encoder-Decoder Modelle** kombinieren die Eigenschaften dieser beiden Stränge. Sie funktionieren sowohl für NLU als auch für NLG-tasks. Für Textvereinfachung können also Encoder-Decoder und Encoder-Only Transformer eingesetzt werden, da es sich um einen NLG-task handelt.

2.3 Automatische Textvereinfachung

Bisherige Ansätze für deutschsprachige Textvereinfachung basieren auf parallelen Datensätzen (Standardsprache → Vereinfachung). Anfänglich wurde mit statistischen Regeln gearbeitet [SuEV16]. Ab 2020 wurden zunehmend NLMs zur Textvereinfachung eingesetzt [SäEV20, RSKK21, SpRE21, EBKP22, StMK23]. Diese Ansätze greifen auf

Encoder-Decoder Transformer Modelle zurück. Mittlerweile wurden Decoder-Only Transformer Ansätze vorgestellt [AOWJ23, DGLM23], welche auch nur mit monolingualen Daten (Einfache Sprache/LS) trainiert werden können. Einen ersten vollständigen Überblick über wissenschaftlich veröffentlichten deutschen Textvereinfachungsdatensätze sowie über KI-gestützte Ansätze zur automatischen Textvereinfachung liefern Stodden et. al. [StMK23] sowie Schomacker et. al. [SGMH23]. Da nach wie vor Encoder-Decoder Transformer Ansätze vielversprechendere Ergebnisse liefern [Alln23, AOWJ23, DGLM23] wird hier ein Encoder-Decoder Ansatz zugrunde gelegt.

Für die Evaluation automatisch-generierter Textvereinfachung werden in der Regel automatische Verfahren eingesetzt. Evaluation durch Menschen oder Expert*innen sind selten [GrSa22]. Häufig werden statistische Metriken, welche für die Auswertung interlingualer Übersetzungen entwickelt wurden, verwendet [GrSa22]. Prominente Vertreter sind BLEU [PRWZ02] und ROUGE [Lin04]. Beide basieren auf dem Vergleichen von einzelnen Worten oder Wortketten (N-Grams) der generierten Übersetzung und einer vorher ausgewählten richtigen Lösung. Es existieren auch Metriken, wie SARI [XNPC16], die speziell für Textvereinfachungen entwickelt wurden und neben der Ähnlichkeit mit einer richtigen Lösung u.a. auch die Komplexität des



Textes messen.

Abb. 1: Schematische Abbildung des Fine-Tuning eines vortrainierten Encoder-Decoder Transformers (z. B. mBART [LGGL20]) für Textvereinfachung. In Rot sind die fehlerhaften Stellen markiert. Die (html-artigen) grünen Anmerkungen markieren die Sprachvariante <De-Or> für Standarddeutsch und <De-Sim> für LS.

3 Rechtlicher und normativer Rahmen für automatisierte Textvereinfachung

Die Betrachtung des Rechtsrahmens für den verwaltungsseitigen Einsatz NLP-basierter Anwendungen zur automatisierten Sprachvereinfachung zielt auf die Identifikation von Modell- und Prozessanforderungen ab. Ihr liegt die Unterscheidung zwischen Anforderungen an das verwaltungsseitig bereitzustellende Textangebot in LS und betreiber- bzw. anbieterbezogenen Anforderungen.

3.1 Automatisierte behördliches Kommunikationsangebot in Leichter Sprache

Anforderungen an Texte juristischen Inhalts in LS zur behördlichen Kommunikation und Interaktion ergeben sich aus der UN Behindertenrechtskonvention, [Bund08] dem BGG, der BITV 2.0 sowie dem Onlinezugangsgesetz (OZG) [KoVM23] auf Grundlage der EU-Richtlinie zur Regelung des barrierefreien Zugangs zu Websites und mobilen Anwendungen öffentlicher Stellen [Dase16]. Bei den durch DIN SPEC 33429-E [Dinn23] noch einmal stärker kanonisierten Leichte-Sprache-Standards handelt es sich um *Soft Law*, das auf freiwillige Anwendung ausgerichtet ist. Durch Bezugnahme in Ausschreibungen und Verträgen wird dieser neue Standard bei Auftragsverhältnissen leistungs- und qualitätskonkretisierend für Dienstleistende verbindlich. Der Standard wurde vom Bundesministerium für Arbeit und Soziales initiiert und zielt als Standardisierungsvorhaben gerade auf die BITV 2.0 ab [Dind22]. Der neue Standard wird, ähnlich DIN EN 301 549, nach § 3 BITV 2.0 bei der öffentlichen Beschaffung zugrunde gelegt werden.

Die Anwendbarkeit automatisierter Textvereinfachung kann grundsätzlich nicht zu Verwaltungshandeln mit Regelungsgehalt führen, was die Bandbreite möglicher Textfunktionen des Ziel-Outputs einschränkt. Zwar können nach § 35a VwVfG (§ 155 (4) AO, § 31a SGB X) Verwaltungsakte vollautomatisiert erlassen werden, wenn kein Beurteilungsspielraum vorliegt und auch keine Ermessensentscheidung zu treffen ist. Nach § 24 (1) S 3 VwVfG muss die Behörde bei Einsatz „*automatische[r] Einrichtungen zum Erlass von Verwaltungsakten*“ gleichwohl eine individuelle Sachverhaltsaufklärung betreiben. In jedem Fall unterliegt der Einsatz selbstlernender Algorithmen bei verwaltungsrechtlichen Entscheidungen einem gesetzlichen Erlaubnisvorbehalt [Gugg19]. In Anbetracht der Risiken für die häufig besonders vulnerablen Adressaten von LS gilt dies erst recht. Mangels Rechtsgrundlage muss sich die automatisierte Textvereinfachung auf Grundlage des DIN Standards hinsichtlich ihrer Funktion auf Informations- und Kommunikationstexte beschränken, bei denen es sich um „*vollständig automatisiertes Realhandeln*“ handelt [ScSH22]. Für diese Textfunktionen greifen der risikoorientierte Haftungsrahmen (Abschn. 3.2) sowie die daraus ableitbaren Anforderungen (Abschn. 4).

3.2 Risiko- und Haftungsregeln für Betreiber NLG-basierter Anwendungen

Eine vertragliche oder außervertragliche Haftung des Betreibers oder Anbieters setzt eine Verletzung von Sorgfaltspflichten oder bei entsprechender gesetzlicher Haftungserweiterung die Erfüllung eines Gefährdungstatbestandes voraus. Die Grundzüge des risikobasierten Ansatzes der EU sind ausbuchstabiert und von den Mitgliedstaaten konsentiert [Cham23, Fino23, Hack23]. Erster Anknüpfungspunkt ist das „klassische“ deliktische Haftungsrecht, für das vor allem die Frage des Sorgfaltsmaßstabs sowie die Beweislastverteilung diskutiert werden. Hier ist die EU zudem mit dem Entwurf zweier KI-(Produkt)Haftungsrichtlinien [Bund22, BoSi22, Mayr23, Hack23] aktiv geworden. Der Ansatz beschränkt sich im Wesentlichen auf die EU-weite Harmonisierung von Beweiserleichterungen (Auskunftsansprüche sowie widerlegbare Vermutungen bzgl. Sorgfaltspflichtverletzungen und Ursächlichkeit für einen Schadenseintritt, Art. 3f. KI-Haftungsrichtlinie) im Rahmen des in den Mitgliedstaaten geltenden Haftungsrechts, allerdings bezogen auf sämtliche zivilrechtlichen Schadensersatzansprüche einschließlich deutscher Amtshaftungsansprüche [Stau23].

Sorgfaltspflichten des Betreibers bzw. Anbieters können verletzt sein, wenn das KI-System den technischen Standards nicht entspricht, wenn Programmier-, Trainings- oder Überwachungsfehler vorliegen. Die u.a. im Kontext des Produktsicherheitsrechts entwickelten Haftungsgrundsätze gelten gleichermaßen für digitale Produkte. [MüWa20] Obgleich DIN-Normen als *Soft Law* mangels Allgemeinverbindlichkeit keine umsetzungspflichtigen Rechtsvorschriften sind, besteht eine Pflicht zur Beachtung des Norminhalts bei der Beurteilung von Konformität und Risiken [Wilr23]. Haftungsrechtlich führt die Nichteinhaltung von DIN-Normen regelmäßig zu einer Sorgfaltspflichtverletzung, auch wenn eine Haftung vom Vorliegen weiterer Voraussetzungen abhängt [MüWa20]. Umgekehrt kommt der Hersteller bzw. Betreiber einer KI bei Einhaltung der Norm in den Genuss einer „*administrativen Konformitätsvermutung*“ [Wilr23], ohne automatisch von einer Haftung befreit zu sein. [MüWa20].

Das Risiko von schadensbegründenden fehlerhaften Outputs eines autonomen KI-Systems wird rechtlich vor allem als Autonomie- und Opazitätsrisiko im Rahmen der Betreiberhaftung diskutiert [Burc22]. Mögliche Szenarien in Bezug auf textvereinfachende Algorithmen wären bspw. unzutreffende, lückenhafte, missverständliche oder unverständliche Textübertragungen, die zu einer Nichtgeltendmachung von Ansprüchen führen könnten. Schwierigkeiten bereitet bei autonomen Systemen zudem die Abgrenzung zwischen der Verwirklichung von Opazitätsrisiken und Sorgfaltspflichtverletzungen des Geschädigten oder eines Dritten. Autonomierisiken sind jedenfalls dann nicht verwirklicht, wenn eine menschliche Sorgfaltspflichtverletzung schadensursächlich geworden ist, etwa bei falscher Nutzung, fehlerhafter Programmierung, unzureichendem Training oder fehlender Überwachung [Burc22]. Bei reinen Textübersetzungen wird dies i.d.R. klar zuordenbar sein. Bei

Verständnis- und Interpretationsfehlern kann die Risikoallokation oft nicht eindeutig erfolgen, v.a. dann nicht, wenn die Vulnerabilität der Adressaten von LS berücksichtigt werden muss [Fior21].

Weiterer rechtlicher Anknüpfungspunkt ist die Anwendung von Gefährdungshaftungstatbeständen, wie sie die EU Kommission für die Produkthaftung sowie in der (Entwurfassung der) sog. KI Verordnung vorgestellt hat [KGAK23, Wend22]. Bei der Gefährdungshaftung entfällt das Erfordernis einer Sorgfaltspflichtverletzung. Bereits der bloße Betrieb einer Gefahrenquelle (= KI-System) führt – unabhängig von Vorsatz oder Fahrlässigkeit – bei einem verursachten Schaden zu einer Haftung. Durch die Neuregelungen wird auch mit einflussreichen Vorgaben für den Einsatz von KI-Systemen durch die öffentliche Hand gerechnet. [ScSH22] Der risikobezogene Ansatz der KI-VO-E drückt sich in der Unterscheidung von verbotenen Praktiken, Hochrisiko-KI-Systemen, Systemen mit begrenztem Risiko sowie KI-Systeme mit geringem Risiko aus.

Für den Einsatz von Textvereinfachungssystemen zu Kommunikations- und Informationszwecken (Textfunktion) ist im Regelfall nicht davon auszugehen, dass es sich um ein Hochrisiko-KI-System handelt, obwohl sich der Adressatenkreis aus vulnerablen Personen zusammensetzt. Es fehlt, jedenfalls bei einer reinen Übersetzung zu Informationszwecken, in aller Regel an einem einschlägigen Hochrisiko-Einsatzbereich nach Anhang III zu Art. 6 (2) KI-VO-E. Somit wäre eine Textübersetzungs-KI als System mit begrenztem oder geringem Risiko einzustufen. Für Systeme mit begrenztem Risiko sollen künftig die Transparenzpflichten sowie weitere Anforderungen gelten (s. Abschn. 4). Diese Regeln greifen auf Überlegungen zur sogenannten *Responsible AI* [GöTB23] in Übereinstimmung mit den Ethikleitlinien der Hochrangigen Expertengruppe für KI zurück [Euro22]. Der umfassende risikozentrierte und differenzierende Ansatz der EU für KI-Systembetreiber und -anbieter resultiert in konkreten Anforderungen, die zudem dynamisch mit dem jeweiligen Stand der Technik anzupassen sind [Cham23]. Der geplante EU-Haftungsrahmen für KI-Systeme basiert dabei auf planvoller Risikobewertung und damit auf einem im Kern nicht-rechtlichen Konzept. Aus der Sicht der Institutionentheorie sind die Risiken transformierte Unsicherheiten, die ein aktuarisches Management in Gestalt von Versicherungen ermöglichen [Nort90]. Das Haftungsrecht schafft folglich negative Anreize durch die Definition von Haftungsrisiken, die Betreiber und Anbieter von KI-Systemen anstreben zu minimieren und zu versichern [LiFH22, Schu23]. Im folgenden Abschnitt werden die daraus ableitbaren Anforderungen weiter konkretisiert.

4 Ableitung von Anforderungen an eine rechts- und standardkonforme Textvereinfachung

Automatische Textvereinfachung, vor allem mittels eines Encoder-Decoder Ansatzes, bedarf paralleler hochqualitativer Daten. [HG20, DGLM23] bemängeln, dass die bisherige (Trainings-)Datengrundlage in zu geringem Maße die Regeln der LS umsetzt. Auf Basis von DIN SPEC 33429 kann ein LS-Framework sowohl für die Datenannotation als auch die Evaluation von KI-Modellen entwickelt werden. Die Analyse von DIN SPEC 33429-E hat einen vorläufigen Katalog mit 54 Anforderungen an eine standardkonforme Textvereinfachung ergeben. DIN SPEC 33429-E liefert Empfehlungen für die Umsetzung von LS. Da ein rein textbasierter Datensatz das Ziel ist, sind nur die sprachlichen Empfehlungen (Kap. 5) relevant. Es wurden alle Unterkapitel betrachtet und Anforderungen abgeleitet. Zu einigen Unterkapiteln wurde mehr als eine Anforderung definiert. Durch diesen Detailgrad wird die Zuweisung einer Überprüfungsmöglichkeit vereinfacht und transparent. Die Zuweisungen (pro Anforderungen eine Überprüfungsmöglichkeit) wurde zunächst durch einen in Computerlinguistik-erfahrenen Informatiker geprüft, für welche der Anforderungen eine automatisierte Form der Auswertung, durch beispielsweise LanguageTool-Regeln für LS, umsetzbar sind. Für die nicht-automatisierbaren Anforderungen, wurden zwei, in LS-geschulte Experten konsultiert, ob diese Anforderungen selbständig oder partizipativ⁹ überprüft werden können. Dadurch ergeben sich drei Kategorien für die Auswertungsform: 1) AUTOMATISCH, 2) EXPERTE, und 3) PARTIZIPATION. Wir geben eine konkrete Umsetzungsmethode an, wie beispielsweise eine Skala; zudem werden qualitative und partizipationsbezogene Evaluationskriterien abgeleitet. Eine aktuelle Fassung des Katalogs ist auf der Open-LS Projektseite zu finden.¹⁰

Eine Untersuchung des geplanten Rechtsrahmens ergibt weitere Modellanforderungen. Das KI-System muss gem. Art. 52 KI-VO-E so konzipiert sein, dass die Nutzer über die automatisierte Texterstellung informiert sind, soweit dies nicht aufgrund der Umstände erkennbar ist. Da der maschinell vereinfachte Text nicht durch eine Prüfgruppe i.S.v. DIN SPEC 33429 geprüft wurde, wäre wohl auch hierauf hinzuweisen. Des Weiteren sind gem. Art. 69 KI-VO-E i.V.m. Art. 8-15 KI-VO-E weitere Gesichtspunkte im Rahmen der Selbstregulierung zu beachten: Daten (Training, Validierung, Testung von Datensets) und Daten-Governance sind an detaillierten Designanforderungen das Training und die Evaluation betreffend zu orientieren (Art. 10 (2)-(4)). Besonders Augenmerk legt der Verordnungsentwurf auf Verzerrungen (*Bias*). Da in dem hier verfolgten Ansatz ungeachtet der Heterogenität der Zielgruppe ein grds. einheitlicher Regelkatalog betrachtet wird, kommt der Abmilderung dieses Homogenitätsbias im Rahmen der Data Governance eine Schlüsselrolle zu.

⁹ Partizipativ meint, gemäß DIN SPEC 33429-E, durch eine Prüfgruppe aus der Zielgruppe für LS

¹⁰ <https://open-ls.entavis.com/dinspec33429anforderungen/>

Vor Inbetriebnahme müssen die technische Dokumentation (Art. 11 i.V.m. Annex IV) geplant sowie Aufzeichnungen (Art. 12) ermöglicht sein. Diese Vorgaben sind auch im Hinblick, auf die in Abschnitt 3.2 genannten (widerleglichen) gesetzlichen Vermutungsregeln (Sorgfaltspflichten, Ursächlichkeit) und Beweislastverteilungen sowie zur Verringerung von Opazitätsrisiken auch im Zusammenhang mit Auskunftsansprüchen sinnvoll. Weiteren Anforderungen folgen aus Art. 13-15. Ein Risikomanagementsystem (Art. 9), wohl gerade auch mit Fokus auf die Vulnerabilität der Übersetzungsadressaten und der Möglichkeit eines *Unfair Bias*, ist einzurichten. Es liegt grds. in der Verantwortung des KI-Anbieters, auf welchem technischen Weg die jeweilige Anforderung umgesetzt wird. Der Stand der Technik wird dabei in jedem Fall zu berücksichtigen sein, wozu auch DIN EN 15038 (Übersetzungen), DIN EN 15838 (Kundenkontaktzentren), EN 301 549 (EU-Konformität digitaler Objekte zur Barrierefreiheit von IKT) gehören.

5 Zusammenfassung, Limitationen und Ausblick

LS Texte als Output automatisierter Textvereinfachung durch NLM müssen vielfältigen Anforderungen genügen. Neben den skizzierten Anforderungen für Encoder-Decoder-Modelle ist die Qualität der Trainingsdaten zentral, auch aufgrund gesetzlicher Vorgaben. Für die Vereinfachung rechtlicher Texte bedarf es paralleler, regelkonform annotierter Korpora. Regeln für den Vereinfachungoutput müssen bei behördlichen Informations- und Kommunikationstexten zwischen output-bezogenen und betreiberbezogenen Anforderungen unterscheiden. Erstere legen nahe, der Annotation und Evaluation die einschlägige DIN-Vorschrift zugrunde zu legen, was an einer entsprechenden Auswertung des Entwurfs von DIN SPEC 33429 illustriert wurde. Betreiber müssen sich an dem risikobasierten Haftungsrahmen orientieren, mit je nach Einsatzgebiet unterschiedlichen Risikokategorien und Vorgaben der KI Verordnung.

Dem Beitrag liegt der aktuelle Stand der Technik mit Bezug zur deutschsprachigen Textvereinfachung zugrunde. Es wurden NLM zur Vereinfachung deutscher Texte und ihre Eignung zur Übertragung juristischer Texte sowie die einschlägigen Anforderungen der Barrierefreiheit untersucht. Die allgemeine Methodik dieser Arbeit ist im Grundsatz für Textvereinfachungen für jede Domäne und jede Sprache anwendbar. Für andere Sprachen als Deutsch können keine allgemeingültigen Schlussfolgerungen getroffen werden. Diese Arbeit stützt sich auf die aktuelle Entwurfsfassung einer DIN-Norm, die endgültige Fassung und ihre Implikationen können davon abweichen. Gleiches gilt für den Rechtsrahmen der EU. Außerdem basiert die DIN-Norm auf Annahmen über ihre Adressaten, die eine Homogenitätsverzerrung hinsichtlich der demgegenüber heterogenen Zielgruppe mit Lernschwierigkeiten beinhaltet. Eine weitere Einschränkung bilden funktionale Verwendungsbeschränkungen auf reine Informationstexte sowie die Übertragung von Texten mit anderen Textfunktionen in reine Informationstexte, d.h., dass die Ziel-Textfunktion auf Informationsfunktionen beschränkt ist.

Im Rahmen des Open-LS Projekts werden weitere Forschungslücken adressiert:

1. Aufbau eines anforderungskonformen Trainingsdatensets für rechtliche Texte;
2. Anforderungskonforme Entwicklung, Training und Evaluation eines NLM;
3. Erforschung kosteneffizienter Möglichkeiten, um die Daten mit zielgruppenspezifischen KI-generierten Illustrationen zu ergänzen;
4. Untersuchung der Übertragbarkeit des Modells auf Domänen und Unterdomänen (z.B. juristische Unterdomänen), für die es nicht trainiert wurde;
5. Methodische Weiterentwicklung geeigneter normkonformer Evaluationsansätze.

Literaturverzeichnis

- [Ahle23] Ahlers, Rechtsanwalt Moritz: Der Referentenentwurf zum „Onlinezugangsgesetz 2.0“ aus vergaberechtlicher Sicht. In: Neue Zeitschrift für Baurecht und Vergaberecht. (2023), S. 147
- [Alln23] Alkaldi, Wejdan ; Inkpen, Diana: Text Simplification to Specific Readability Levels. In: Mathematics Bd. 11, Multidisciplinary Digital Publishing Institute (2023), Nr. 9, S. 2063
- [AOWJ23] Anschutz, Miriam ; Oehms, Joshua ; Wimmer, Thomas ; Jeziarski, Bartłomiej ; Groh, Georg: Language Models for German Text Simplification: Overcoming Parallel Data Scarcity through Style-specific Pre-training, arXiv (2023). — arXiv:2305.12908 [cs]
- [AsHZ23] Asghari, Hadi ; Hewett, Freya ; Züger, Theresa: On the Prevalence of Leichte Sprache on the German Web. In: Proceedings of the 15th ACM Web Science Conference 2023, WebSci '23. New York, NY, USA : Assoc for Computing Machinery, 2023, S. 147–152
- [Baum20] Baumann, Antje: Rechtstexte als Barrieren – Einige Merkmale der Textsorte „Gesetz“ und die Verständlichkeit. In: Maaß, C. ; Rink, I. (Hrsg.): Hdb Barrierefreie Kommunikation, Bd. 3, Berlin : Frank & Timme, 2020
- [BoSi22] Bomhard, David ; Siglmüller, Jonas: Europ. KI-Haftungsrichtlinie Der aktuelle Kommissionsentwurf und seine praktischen Auswirkungen (2022), S. 506ff.
- [Bund08] Bundesgesetzblatt Jahrgang 2008 Teil II Nr. 35: Rechte von Menschen mit Behinderungen — Gesetz zu dem Übereinkommen der Vereinten Nationen vom 13. Dezember 2006 über die Rechte von Menschen mit Behinderungen sowie zu dem Fakultativprotokoll vom 13. Dezember 2006 zum Übereinkommen der Vereinten Nationen über die Rechte von Menschen mit Behinderungen, 2008
- [Bund21] Bundesministerium für Arbeit und Soziales: Bericht der Bundesrepublik Deutschland an die Europäische Kommission über die periodische Überwachung der Einhaltung der Barrierefreiheitsanforderungen von Websites und mobilen Anwendungen öffentl. Stellen gemäß Artikel 8 der Richtlinie (EU) 2016/2102, 2021
- [Bund22] Bundesrat: Vorschlag für eine Richtlinie des Europäischen Parlaments und des Rates über die Haftung für fehlerhafte Produkte COM(2022) 495 final (Nr. BR Drucksache 515/22), 2022

- [Burc22] Burchardi, Dr Sophie: Risikotragung für KI-Systeme. In: , EuZW 2022., S. 685ff.
- [Cham23] Chamberlain, Johanna: The Risk-Based Approach of the European Union’s Proposed Artificial Intelligence Regulation. In: European Journal of Risk Regulation Bd. 14, Cambridge University Press (2023), Nr. 1, S. 1–13
- [Dase16] Das Europäische Parlament und der Rat der Europäischen Union: Richtlinie (EU) 2016/2102 des Europäischen Parlaments und des Rates über den barrierefreien Zugang zu den Websites und mobilen Anwendungen öffentlicher Stellen.
- [Deut22] Deutscher Bundestag: Technikfolgenabschätzung (TA): Künstliche Intelligenz und Distributed-Ledger-Technologie in der öffentlichen Verwaltung (Nr. BT Drucksache 20/3651), 2022
- [DGLM23] Deilen, Silvana ; Garrido, Sergio Hernández ; Lapshinova-Koltunski, Ekaterina ; Maaß, Christiane: Using ChatGPT as a CAT tool in Easy Language translation (2023). — arXiv:2308.11563 [cs]
- [Dind22] DIN e.V.: Geschäftsplan für ein DIN SPEC-Projekt nach dem PAS-Verfahren zum Thema „Empfehlungen für Deutsche Leichte Sprache“, 2022
- [Dinn23] DIN-Normenausschuss: Empfehlungen für Deutsche LS (DIN SPEC 33429).
- [EBKP22] Ebling, Sarah ; Battisti, Alessia ; Kostrzewa, Marek ; Pfütze, Dominik ; Rios, Annette ; Säuberli, Andreas ; Spring, Nicolas: Automatic Text Simplification for German. In: Frontiers in Communication Bd. 7, Frontiers Research Foundation (2022), S. 706718
- [Euro22] Europäische Kommission: Hochrangige Expertengruppe für künstliche Intelligenz | Gestaltung der digitalen Zukunft Europas. URL <https://digital-strategy.ec.europa.eu/de/policies/expert-group-ai>. - abgerufen am 2023-06-19
- [Fino23] Finocchiaro, Giusella: The regulation of AI. In: AI & SOCIETY (2023)
- [Fior21] Fioravanti, Chiara: Communicating the Law and Public Information to Vulnerable Audiences: Contexts and Strategies. In: Journal of Open Access to Law Bd. 9 (2021), Nr. 1, S. 8–8
- [GaGe22] Gallegos, Isabel ; George, Kaylee: The Right to Remain Plain: Summarization and Simplification of Legal Documents.
- [GöTB23] Göllner, Sabrina ; Tropmann-Frick, Marina ; Brumen, Boštjan: Aspects and Views on Responsible Artificial Intelligence. In: Machine Learning, Optimization, and Data Science, Lecture Notes in Computer Science. Cham, 2023
- [GrBu20] Grotlüschen, A. ; Buddeberg, K. (Hrsg.): LEO 2018: Leben mit geringer Literalität. Bielefeld : wbv, 2020 — ISBN 978-3-7639-6072-9
- [GrSa22] Grabar, Natalia ; Saggion, Horacio: Evaluation of Automatic Text Simplification: Where are we now, where should we go from here. In: Actes de la 29e Conférence sur le Traitement Automatique des Langues Naturelles. Volume 1 : conférence principale. Avignon, France : ATALA, 2022, S. 453–463
- [Gugg19] Guggenberger, L.: Einsatz künstlicher Intelligenz in der Verwaltung. In: , NVwZ. Bd. 12 (2019), S. 844

- [Hack23] Hacker, Philipp: The European AI Liability Directives -Critique of a Half-Hearted Approach and Lessons for the Future, arXiv (2023). — arXiv:2211.13960 [cs]
- [HGNM20] Hansen-Schirra, Silvia ; Gutermuth, Silke ; Nitzke, Jean ; Maaß, Christiane ; Rink, Isabel: Technologies for the Translation of Specialised Texts into Easy Language. In: , 2020 — ISBN 978-3-7329-0688-8, S. 99–127
- [KGAK23] Kazim, Emre ; Güçlütürk, Osman ; Almeida, Denise ; Kerrigan, Charles ; Lomas, Elizabeth ; Koshiyama, Adriano ; Hilliard, Airlie ; Trengove, Markus: Proposed EU AI Act. In: AI and Ethics Bd. 3 (2023), Nr. 2, S. 381–387
- [KoVM23] Kossens ; V. d. Heide ; Maaß: BGG § 11 Verständlichkeit und leichte Sprache.
- [LGGL20] Liu, Yinhan ; Gu, Jiatao ; Goyal, Naman ; Li, Xian ; Edunov, Sergey ; Ghazvininejad, Marjan ; Lewis, Mike ; Zettlemoyer, Luke: Multilingual Denoising Pre-training for Neural Machine Translation. In: Transactions of the Association for Computational Linguistics Cambridge, MA, MIT Press (2020)
- [LiFH22] Li, Shu ; Faure, Michael ; Havu, Katri: Liability Rules for AI-Related Harm: Law and Economics Lessons for a European Approach. In: European Journal of Risk Regulation Bd. 13, Cambridge University Press (2022), Nr. 4, S. 618–634
- [Lin04] Lin, Chin-Yew: ROUGE: A Package for Automatic Evaluation of Summaries. In: Text Summarization Branches Out. Barcelona, Spain : Association for Computational Linguistics, 2004, S. 74–81
- [Maaß20] Maaß, Christiane ; Hansen-Schirra, S. ; Maaß, C. (Hrsg.): Easy Language – Plain Language – Easy Language Plus: Balancing Comprehensibility and Acceptability, Easy–Plain–Accessible. Bd. 3., Berlin : Frank & Timme, 2020.
- [Mayr23] Mayrhofer, Ann-Kristin: Produkthaftungsrechtliche Verantwortlichkeit des „Trainer-Nutzers“ von KI-Systemen. In: , RD 2023. (2023), S. 20ff.
- [MMRS55] McCarthy, J. ; Minsky, M. L. ; Rochester, N. ; Shannon, C.E.: A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence (1955)
- [MüWa20] Münchener Kommentar zum BGB ; Wagner: BGB § 823, 2023.
- [Nort90] North, Douglass C.: Institutions, Institutional Change and Economic Performance, Political Economy of Institutions and Decisions. Cambridge : Cambridge University Press, 1990 — ISBN 978-0-521-39416-1
- [PRWZ02] Papineni, Kishore ; Roukos, Salim ; Ward, Todd ; Zhu, Wei-Jing: Bleu: a Method for Automatic Evaluation of Machine Translation. In: Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics. Philadelphia, Pennsylvania, USA : Ass. for Computational Linguistics, 2002, S. 311–318
- [RSKK21] Rios, Annette ; Spring, Nicolas ; Kew, Tannon ; Kostrzewa, Marek ; Säuberli, Andreas ; Müller, Mathias ; Ebling, Sarah: A New Dataset and Efficient Baselines for Document-level Text Simplification in German. In: Proceedings of the Third Workshop on New Frontiers in Summarization. Online and in Dominican Republic : Association for Computational Linguistics, 2021. — tex.ids= riosNewDatasetEfficient2021a, S. 152–161
- [RSRL20] Raffel, Colin ; Shazeer, Noam ; Roberts, Adam ; Lee, Katherine ; Narang, Sharan ;

- Matena, Michael ; Zhou, Yanqi ; Li, Wei ; u. a.: Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer. In: arXiv:1910.10683 [cs, stat] (2020). — arXiv: 1910.10683
- [SäEV20] Säuberli, Andreas ; Ebling, Sarah ; Volk, Martin: Benchmarking Data-driven Automatic Text Simplification for German. In: Proceedings of the 1st Workshop on Tools and Resources to Empower People with READING Difficulties (READI). Marseille, France : European Language Resources Association, 2020
- [Schu23] Schuett, Jonas: Risk Management in the Artificial Intelligence Act. In: European Journal of Risk Regulation, Cambridge University Press (2023), S. 1–19
- [ScSH22] Schoch ; Schneider ; Hornung: VwVfG § 35a Rn. 3-7.
- [ScTr21] Schomacker, Thorben ; Tropmann-Frick, Marina: Language Representation Models: An Overview. In: Entropy Bd. 23, Multidisciplinary Digital Publishing Institute (2021), Nr. 11, S. 1422
- [SGMH23] Schomacker, Thorben ; Gille, Michael ; Marina Tropmann-Frick ; von der Hülls, Jörg: Data and Approaches for German Text Simplification - Next Steps toward an Accessibility-enhanced Communication. In: , 2023
- [SpRE21] Spring, Nicolas ; Rios, Annette ; Ebling, Sarah: Exploring German Multi-Level Text Simplification. In: Proceedings of the International Conference on Recent Advances in Natural Language Processing (RANLP 2021). Held Online : INCOMA Ltd., 2021, S. 1339–1349
- [Stau23] Staudenmayer, Dr Dirk: Haftung für Künstliche Intelligenz. In: , NJW 2023. (2023), S. 894ff.
- [StMK23] Stodden, Regina ; Momen, Omar ; Kallmeyer, Laura: DEPLAIN: A German Parallel Corpus with Intralingual Translations into Plain Language for Sentence and Document Simplification, arXiv (2023). — arXiv:2305.18939 [cs]
- [Stod22] Stodden, Regina: Erstellung eines parallelen Vereinfachungskorpus für die deutsche Sprache – Unter Verwendung des HHU Annotationstools TS-anno (2022)
- [SuEV16] Suter, Julia ; Ebling, Sarah ; Volk, Martin: Rule-based Automatic Text Simplification for German. In: Proceedings of the 13th Conference on Natural Language Processing, 2016, S. 279–287
- [Thap21] Thapa, Basanta Für mehr Barrierefreiheit in der digitalen Verwaltung, 2021
- [TuWW22] Tunstall, Lewis ; von Werra, Leandro ; Wolf, Thomas: Natural Language Processing with Transformers. Revised Edition. Erscheinungsort nicht ermittelbar : O'Reilly Media, Inc., 2022
- [VSPU17] Vaswani, Ashish ; Shazeer, Noam ; Parmar, Niki ; Uszkoreit, Jakob ; Jones, Llion ; Gomez, Aidan N ; Kaiser, Lukasz ; Polosukhin, Illia: Attention is All you Need. In: Guyon, I. ; Luxburg, U. V. ; Bengio, S. ; Wallach, H. ; Fergus, R. ; Vishwanathan, S. ; Garnett, R. (Hrsg.): Advances in Neural Information Processing Systems 30, 2017, S. 5998–6008

- [Wend22] Wendehorst, Christiane: Liability for Artificial Intelligence: The Cambridge Handbook of Responsible Artificial Intelligence: Interdisciplinary Perspectives, Cambridge University Press, 2022
- [Wilr23] Wilrich, Thomas: Rechtliche Bedeutung von DIN-Normen und technischen Regelwerken - beck-online.pdf. In: , NJW 2023. (2023), S. 1400ff.
- [XNPC16] Xu, Wei ; Napoles, Courtney ; Pavlick, Ellie ; Chen, Quanze: Optimizing Statistical Machine Translation for TS. In: Transactions of the Association for Computational Linguistics Bd. 4. Cambridge, MA, MIT Press (2016), S. 401–415