

Professor Dr. Axel Adrian

# Der Umgang mit natürlicher und formaler Sprache im Recht – von Aristoteles zum Richterautomaten?

Dieser Beitrag möchte einige Hinweise geben, wie sich der Umgang der Juristinnen und Juristen mit Sprache im Laufe der Zeit verändert haben könnte. Als „Stationen“ zur Verdeutlichung dieser Entwicklung werden Aristoteles und der Justizsyllogismus der klassischen juristischen Methodenlehre, der linguistic turn und schließlich die Frage, wie man juristische Argumente sprachlich formalisieren kann, um diese dann sogar maschinenverarbeiten zu können, herausgegriffen. So gesehen beginnt und endet der Beitrag mit Logik und spannt den Bogen von der aristotelischen Logik bis zur Informatik.

## Aristoteles als Ausgangspunkt

Unser allgemeines Verständnis von Logik wird im Wesentlichen auf Aristoteles (384–322 v. Chr.) zurückgeführt. So ist zunächst festzuhalten, dass das Wort Logik sich von griechisch „logos“ herleitet. Dies wiederum ist auf „legein“ zurückzuführen, was „sprechen“, „reden“ bedeutet. Man kann also sagen, das Wort sollte deutlich machen, es handele sich um eine „vernünftige Rede“. Dabei benutzte Aristoteles „logos“ im Sinn von „Definition“, Platon (428/427–348/347 v. Chr.) in den Bedeutungen „Darstellung“, „Erklärung“, „Aussage“. Bei Heraklit (535–475 v. Chr.) bedeutete es die „ewige Struktur der Welt“. Die Stoa sah im „logos“ das „Vernunftprinzip des Weltalls und das, woraus alle Tätigkeit entsteht“. Im Johannesevangelium bedeutet „logos“ schliesslich „Wort Gottes“.<sup>1</sup>

Aristoteles unterschied zunächst Meinung (Doxa) von Wissen (Episteme). Letzteres bezieht sich auf eine unveränderliche, ewige (objektive) Wahrheit (Aletheia). Diese wird u. a. in syllogistischen Schlüssen „entdeckt“. Meinungen dagegen können nicht wahr sein, sondern nur wahrscheinlich (Eikos). Es geht bei Meinungen also mehr um bloße Wahrscheinlichkeit (Endoxa) im Sinne von Wahrscheinlichkeit.

Weiter unterscheidet Aristoteles nun die Gegenstandsbereiche der Theorie („Augen des Geistes“) von denen der Praxis. Theorie bezieht sich auf bzw. „blickt“ durch das Veränderliche hindurch auf das Wesentliche, nämlich auf die unveränderliche Struktur als Gegenstand des Wissens, also z. B. auf Fra-

gen, die die heutige Naturwissenschaft beschäftigen.

Die Bereiche der Praxis beziehen sich dagegen auf das menschliche Handeln in sozialer und interaktiver Gestaltung der Lebensverhältnisse, insbesondere in der Ethik, Ökonomie und Politik. Diese Gegenstände sind stets veränderlich. Hier gibt es (eigentlich) keine Theorie, kein Wissen.<sup>2</sup>

Nach Aristoteles müsste man also z. B. Fragen der Rechtswissenschaften wohl dem Bereich der Praxis zuordnen und könnte nur mittels topischer und rhetorischer Methoden über Meinungen argumentieren, während man z. B. Fragen der Naturwissenschaften dem Bereich der Theorie zuordnen würde, in dem man mit syllogistischen – also logischen – Methoden ewige Wahrheiten entdecken könnte. Alles wurde dabei nur in natürlicher Sprache ausgedrückt, da moderne Logik, die in formalen Sprachen notiert wird, noch nicht erfunden war.

## Der Justizsyllogismus

Nach der klassischen juristischen Methodenlehre wendet man zum Teil noch heute den sog. Justizsyllogismus an.<sup>3</sup> Es handelt sich um ein formallogisches Schlussverfahren.

Dabei wird das abstrakt-generelle Gesetz, also der Obersatz auf den konkret-individuellen Rechtsfall, also den Sachverhalt, angewendet, d. h. Sachverhaltsmerkmale (SV) werden unter Tatbestandsmerkmale (TB) des Gesetzes subsumiert. Dann ergibt sich „logisch zwingend“ die konkrete Rechtsfolge (RF) für den Rechtsfall.

Die Struktur sieht beispielhaft so aus:

- „1. (Obersatz/TB) Wer zu schnell fährt, muss ein Bußgeld bezahlen.“
- „2. (Untersatz/SV) X ist zu schnell gefahren.“
- „3. (Schluß/RF) Also muss X ein Bußgeld bezahlen.“<sup>4</sup>

Diese „Methode“ wurde der Syllogistik des Aristoteles entlehnt, die eigentlich – wie gezeigt – für ewige, unveränderliche Gegenstände des Wissens in Bereichen der Theorie vorgesehen war, und nun doch auf Bereiche der Praxis, nämlich auf die „Entdeckung“ von Recht, angewandt wurde. Dies obwohl man leicht vermuten kann, dass das Recht wohl eher einem ständigen Wandel unterliegen dürfte, insbesondere, weil es von Menschen „gemacht“ wird.

## Die Begriffsjurisprudenz

Nach der historischen Rechtsschule<sup>5</sup> von Friedrich Carl von Savigny (1779–1861),

1) Rowohlt Philosophielexikon, 2013; Soentgen, Selbstdenken, 2020, S. 129 ff.

2) Wohlrappl, Der Begriff des Arguments, 2021, S. 22 ff., 24 ff., 438.

3) Adrian, Grundprobleme einer juristischen (gemeinschaftsrechtlichen) Methodenlehre, 2009, S. 778 ff. m. w. N.

4) Adrian, Wie wissenschaftlich ist die Rechtswissenschaft? – Gibt es eine bindende Methodenlehre?, in: Rechtstheorie Heft 4/2010, S. 521 ff.

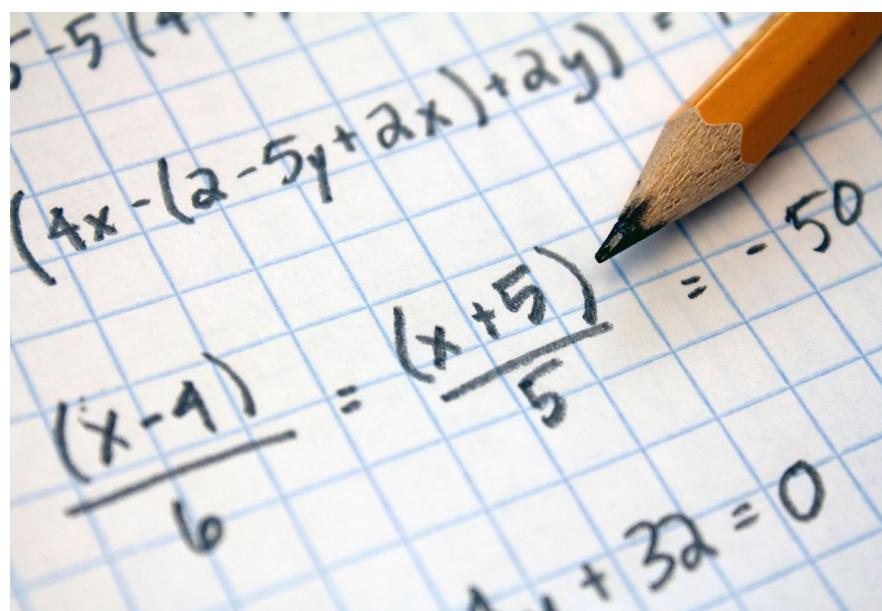
5) Adrian (Fn. 4), S. 564 m. w. N.: „Für die historische Schule ist Recht (...) an seine geschichtlichen Voraussetzungen gebunden; der ‚Stoff‘ des (...) Rechts also durch die gesamte Vergangenheit der Nation gegeben.“ (...) Savigny glaubte, daß das Recht kein Erzeugnis der Vernunft, sondern des in der Geschichte waltenden „Volksgeistes“ ist. Danach gibt es kein unwandelbares, für alle Völker gleiches Recht, da jedes Volk seine eigene Individualität und „Volksseele“ hat. So wird dem „Rationalismus des neuzeitlichen Naturrechts“ die „triebmäßige Unbewußtheit des Rechtsgefühls“ entgegengehalten.

dem Begründer der heute noch bekannten juristischen Methoden, stellte man sich das im Sinne der sich daraufhin entwickelnden<sup>6</sup> Begriffsjurisprudenz (Georg Friedrich Puchta, 1798–1846, Bernhard Windscheid, 1817–1892, und Rudolf v. Jhering, 1818–1892, in seiner ersten Phase ca. 1866) aber tatsächlich so vor, dass die Sprache des Rechts bzw. des Gesetzes aus sich selbst heraus, ohne Rekurs auf die Lebensverhältnisse, fruchtbar sein würde.

Das bedeutet, es wurde angenommen, dass es ein ideales logisches System in Form einer „Begriffspyramide“ geben würde, an deren Spitze ein allgemeiner Begriff steht, unter den sich alle übrigen Begriffe subsumieren lassen könnten. Die Begriffe selbst sollten also tatsächlich als Erkenntnisquelle dienen. Kaufmann mit einem Verweis auf Jhering, der ausführt: „Die Begriffe sind produktiv, sie paaren sich und zeugen neue“.<sup>7</sup>

Heute geht man allerdings in der zeitgenössischen Methodenlehre mehrheitlich davon aus, dass nicht alle juristischen Entscheidungen so aus einem gegebenen System von Rechtsätzen und Rechtsgrundsätzen mit logischen Mitteln und ohne eigene Wertung abgeleitet werden können. Ein dafür notwendiges ideales logisches System nach dem Vorbild der axiomatischen Mathematik könnte nur formale Aussagen treffen und sei daher inhaltsleer.

*Mathe und Jura passen gut zusammen: Wer rechnen kann, wird auch im Jurastudium punkten.*



Inhaltliche Regelungen des Zusammenlebens der Menschen werden durch Rechtsnormen als Ordnungsbegriffe nur (systematisch) bezeichnet, aber nicht geboren.<sup>8</sup> „Die wissenschaftliche Begriffsbildung sei nur Hebamme, nicht Wöchnerin.“<sup>9</sup>

Der späte Rudolf v. Jhering wechselte schließlich sogar „[...] radikal seine Position und wurde zum erbitterten Gegner seiner zunächst selbst vertretenen Ansichten, für die er selbst das negativ besetzte Wort „Begriffsjurisprudenz“ fand. Mit seinen Arbeiten über den „Zweck im Recht“ wurde Jhering zum Vorläufer der modernen Rechtstheorien.“<sup>10</sup>

### Die heutige juristische Methodenlehre bis zum linguistic turn

Den Diskussionsstand der überwiegenden Lehre zur heutigen juristischen Methodenlehre kann man prägnant bei Zippelius in dessen Buch „Juristische Methodenlehre“ nachlesen. Zippelius gesteht zu, dass das Recht in dem Maße an Exaktheit gewinnt, wie es sich in eine formalisierte und eindeutige Sprache bringen ließe.

Die formalen Regeln müssen aber mit einer inhaltlichen Bedeutung versehen werden. Erst wenn den logischen Aussagenvariablen eine Bedeutung beigelegt wird, werde aus der logischen Formel eine inhaltliche Aussage. Die

Verknüpfung der Bedeutung mit formalen Zeichen sei nach Zippelius eine Frage der Semantik. Eine Sprache, die damit einerseits formal exakt und andererseits inhaltlich ausdrucksfähig sein soll, setzte zweierlei voraus, nämlich einmal eine exakte Semantik und zum zweiten eine „formale Kalkülierbarkeit der rechtlichen Aussagen“. Nach Zippelius sind beide Voraussetzungen im Recht nicht erfüllt.

Zippelius zeigt auch, dass es bereits an einer exakten Semantik fehlt. Dies führt er darauf zurück, dass Rechtsnormen weitgehend Erfahrungsinhalte bezeichnen. Diese Wörter, die Erfahrungstatsachen bezeichnen, werden exemplarisch eingeführt – durch praktisches Einüben – und hätten daher einen inexakten Bedeutungsumfang, geradezu einen „Bedeutungsspielraum“. Die in der Juristerei zu verwendenden Wörter seien also regelmäßig mehrdeutig, sie haben nicht für alle Menschen die „genau gleiche Bedeutung“ und nach Zippelius oft auch für den einzelnen selbst „nicht einen ganz exakten Sinn“.<sup>11</sup>

Damit hat man aus der Menge der möglichen Wortbedeutungen des Gesetzestextes die auf den konkreten Fall zutreffende Wortbedeutung durch Auslegung nach dem Wortlaut, der Teleologie, dem System und der Historie und durch Vergleich des zu prüfenden Falles mit ansonsten etwa schon bekannten Präzedenzfällen der Rechtsprechung zu ermitteln.

### Künftige juristische Methodenlehre nach dem linguistic turn?

Analysiert man diese von der herrschenden Meinung vertretene juristische Methodenlehre mit Erkenntnissen der heutigen sprachpragmatischen Philosophie, stellt sich aber die ganz

6) Bydlinski, Juristische Methodenlehre, 2011, S. 109f. weist auf folgendes hin: „Savigny wollte bei der Bearbeitung des Rechtsstoffs (...) das ‚organische‘ und das ‚logische‘ (oder ‚systematische‘ bzw. ‚philosophische‘) Element unterscheiden und gleichberechtigt beachten (...), ohne über deren Verhältnis (...) zureichende Klarheit zu schaffen. Die Begriffsjurisprudenz hat einseitig das ‚logische‘ Element weitergepflegt.“.

7) Adrian (Fn. 4), S. 566 f. m.w.N.

8) Adrian (Fn. 4), S. 567 m.w.N.

9) Zippelius, Rechtsphilosophie, 2011, S. 207; nach Eugen Ehrlich zitiert.

10) Hähnchen, Rechtsgeschichte, 2016, S. 333.

11) Zippelius, Juristische Methodenlehre, 2021, S. 89f., 38.

prinzipielle Frage, ob man nach dem sog. „linguistic turn“ überhaupt noch verlässlich feststellen kann, welche Bedeutung ein bestimmtes Wort bzw. Zeichen einer „natürlichen Sprache“ hat.

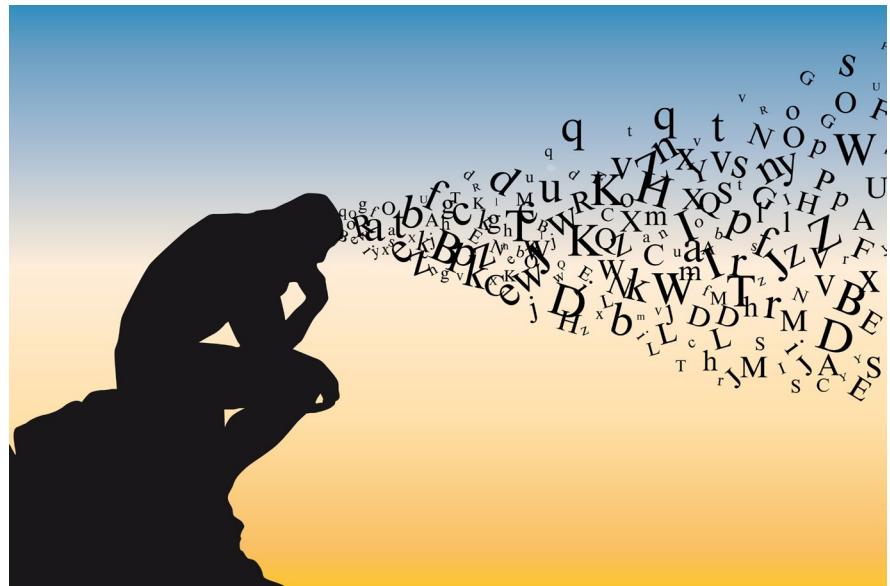
Als linguistic turn (linguistische Wende, sprachkritische Wende, sprachanalytische Wende, Wende zur Sprache) wird üblicherweise die Abwendung von der Idee verstanden, dass Sprache ontologisch fundiert ist, hin zur Frage, wie durch (bloße) Beziehungen Bedeutung von Sprache begriffen werden kann.<sup>12</sup> Der „frühe“ Wittgenstein glaubte zunächst, „der Satz ist ein Bild der Wirklichkeit. Der Satz ist ein Modell der Wirklichkeit, so wie wir sie uns denken.“<sup>13</sup>

So kann Wittgenstein (noch) sagen: „Die meisten Sätze und Fragen, welche über philosophische Dinge geschrieben worden sind, sind nicht falsch, sondern unsinnig.“<sup>14</sup>

Wittgenstein durchlebte dann aber sozusagen selbst einen „linguistic turn“. Der „späte“ Wittgenstein verwirft nämlich sein ursprüngliches ontologisches Fundament (d. h. die Idee, die Welt als Gesamtheit von Tatsachen zu sehen) sowie die Möglichkeit einer eindeutigen Beziehung zwischen Welt und deren Abbildung in Sprache. Es kommt danach nicht mehr auf eine exakte Analyse des Wortes zur Beschreibung der Welt an, sondern auf die Untersuchung seines jeweiligen Gebrauchs, welcher die Bedeutung freilegt. Alles ist (nur) „Sprachspiel“, wobei die Regeln nicht endgültig festlegbar sind.<sup>15</sup> So bleibt das berühmte Wort Wittgensteins: „Die Bedeutung eines Wortes ist sein Gebrauch in der Sprache.“<sup>16</sup>

### (Natürlicher) Sprachgebrauch, formale Sprachen und Legal-Tech

Nun dürfte die Idee nicht mehr fernliegen zu versuchen, den Gebrauch der Sprache zur Bestimmung der Bedeutung der Wörter auch in der Rechtswissenschaft mit Hilfe von Computern zu analysieren. Gerade weil viele Daten zu prüfen sind, um die Bedeutungen desselben Zeichens/Wortes in seinen verschiedenen Gebräuchen, d. h. Kontexten, zu erkennen, erscheint es sinnvoll, Computer einzusetzen, die mehr Daten verarbeiten können als der Mensch.<sup>17</sup> Dabei sind verschiedene computerwissenschaftliche Verfahren zu unterscheiden.



© pict rider – stock.adobe.com

*Wittgenstein: Die Bedeutung eines Wortes ist sein Gebrauch in der Sprache.*

Damit ein Computer z. B. mit statistischen Verfahren (sog. subsymbolische Künstliche Intelligenz), also mittels Machinelearning, z. B. unterscheiden kann, ob in einem Satz das Wort „Baum“ eine Pflanze oder einen Menschen, nämlich Herrn Baum, beschreibt, müssen Wörter in Abhängigkeit von ihrem jeweiligen Kontext durch Vektoren in hochdimensionalen mathematischen Räumen repräsentiert werden, indem z. B. einem neuronalen Netz viele Wörter mit ihren jeweiligen Kontexten als Trainingsdaten „gefüttert“ werden. Der Vektor, der dann die Pflanze bezeichnet, dürfte dann im hochdimensionalen mathematischen Raum näher bei anderen Vektoren/Wörtern „liegen“, die ebenfalls nur aufgrund ihres Kontextes auch Pflanzen repräsentieren, während der Vektor, der Herrn Baum repräsentiert, „näher“ bei Vektoren liegen dürfte, die andere Namen anderer Menschen repräsentieren. So kann natürliche Sprache formalisiert und durch Vektoren im mathematischen Raum repräsentiert werden.<sup>18</sup>

Sollte es am Ende technisch möglich werden, Texte in sehr großen Kontexten, d. h. z. B. BGH-Urteile im Kontext mit Urteilen aller Vorinstanzen samt allen zugehörigen Schriftsätze, zu verarbeiten und miteinander zu vergleichen, wären vielleicht eines Tages Entscheidungsasistenten für Gerichte denkbar.<sup>19</sup> Damit ein Computer dagegen z. B. mit regelbasierten Verfahren (sog. symboli-

sche Künstliche Intelligenz), also mittels Expertensystemen maschinell juristische Entscheidungen treffen kann, muss das Wissen über das Recht, die Welt, den zu entscheidenden Sachverhalt und das (handwerkliche) Know-How, wie Juristen

12) Der Begriff wurde von Richard Rorty eingeführt; Klaus F. Röhrl/Hans Christian Röhrl, Allgemeine Rechtslehre, 2008, S. 51 ff., verbinden diese Wende mit dem Neopragmatismus: „(...) Als Neopragmatismus bezeichnet man eine Richtung der Gegenwartsphilosophie, wiederum vor allem amerikanischer Autoren (Rorty, Putnam, Brandom), die zwar an den klassischen Pragmatismus anknüpfen, aber die Sprache inhaltlich und methodisch in den Mittelpunkt rücken. Die philosophische Untersuchung wird zur Sprachanalyse. Das ist die sog. Sprachphilosophische Wende des Pragmatismus. An die Stelle einer (pragmatischen) Theorie der Wahrheit tritt eine pragmatische Theorie der Bedeutung.“ Vgl. auch z. B. Carriero, in: Bartels/Stöckler, Wissenschaftstheorie. Ein Studienbuch, 2009, S. 43f., und Bartels, in: Bartels/Stöckler, ebd., S. 214.

13) Wittgenstein, Tractatus logico-philosophicus, Werkausgabe von Suhrkamp Taschenbuch, Bd. 1, 1984, 4.01, S. 30.

14) Wittgenstein (Fn. 13), S. 25.

15) Zum Ganzen siehe auch Adrian, Grundzüge einer allgemeinen Wissenschaftstheorie auch für Juristen. Konsequenzen aus den Zweifeln zeitgenössischer/postmoderner Philosophie für jene juristische Methodenlehre, 2014, S. 65 ff. m. w. N.

16) Wittgenstein (Fn. 13), S. 262 a. E.

17) S. z. B. die Videos zur Vorlesung KI und rechtliches Entscheiden aus dem SS 2020; Beginn mit Stunde 1 auf youtube [https://www.youtube.com/watch?v=OxQIDG-5T40&list=PLIAmFMfHr2U3Gs-VJxkV3jzvlj\\_lV2UGV](https://www.youtube.com/watch?v=OxQIDG-5T40&list=PLIAmFMfHr2U3Gs-VJxkV3jzvlj_lV2UGV).

18) <https://entwickler.de/python/grundsatze-moderner-textklassifizierung-fur-machine-learning-word-embeddings/>.

19) Adrian, Der Richterautomat ist möglich – Semantik ist nur eine Illusion, in: Rechtstheorie Heft 1/2017, S. 77 ff.

Fälle lösen, formalisiert und maschinenverarbeitbar dem Computer zur Verfügung gestellt werden.

Verschiedene Aspekte, die in der natürlichen Sprache problemlos ausgedrückt werden können, wie dass z. B. Tatsachenbehauptungen etwas anderes sind als Sollensnormen etc., müssen durch verschiedene logische Formalisierungen erst „mühsam“ repräsentiert werden, um mit formalen Sprachen die starke „Aussagekraft“ natürlichsprachlicher Ausdrücke zu erreichen.

So ist z. B. die Rechtsnorm der §§ 985 f. BGB, dass der Eigentümer (E) den Anspruch (A) hat, die Sache vom Besitzer (B) herauszuverlangen, wenn diesem kein Recht zum Besitz (R) zusteht, in der Aussagenlogik wie folgt formalisierbar:

$$(E \wedge B \wedge \neg R) \rightarrow A;$$

Wenn dies allgemein für alle ausgedrückt werden soll, die Eigentümer sind, so dass diese Ansprüche gegen alle haben, die Besitzer ihrer jeweiligen Sachen sind, dann könnte dieser Aspekt u. U. wie folgt durch Prädikatenlogik formalisiert werden:

$$\forall x (\exists y. \exists z. Eyx \wedge Bzx \wedge \neg Rzxy \rightarrow Ayxz).$$

Eyx: y ist Eigentümer der Sache x;

Bzx: z ist Besitzer der Sache x;

Rzxy: z hat Recht zum Besitz an der Sache x gegenüber y;

Ayxz: y hat Anspruch nach §§ 985 f. BGB auf eine Sache x gegen z;

*Algorithmen kommen schon heute in der Rechtsprechung zum Einsatz.*



also:

Für alle x (Sachen), wenn ein y (Ansprechsteller) und ein z (Anspruchsgegner) existieren und y Eigentümer und z Besitzer sind und z ist ohne Recht zum Besitz gegenüber y, dann ist der Anspruch nach §§ 985 f. BGB gegen z gegeben.

Wenn dagegen der Aspekt relevant sein soll, dass es sich hier nicht um die Feststellung der Tatsache handelt, dass es diese Ansprüche gibt, sondern dass ein Besitzer die Sache dem Eigentümer herausgeben *soll*, dann wäre dies mit Obligationen in der deontischen Logik auszudrücken (wobei O so viel bedeutet wie „dem Gesetz nach“), z. B. wie folgt:

$$O(E \wedge B \wedge \neg R) \rightarrow OA.$$

Gelesen als: „Wenn Ansprechsteller dem Gesetz nach Eigentümer einer Sache ist und Anspruchsgegner dem Gesetz nach Besitzer derselben ist und der Anspruchsgegner nicht dem Gesetz nach ein Recht zum Besitz der Sache hat, dann soll dem Gesetz nach der Anspruchsgegner dem Anspruchsinhaber die Sache nach §§ 985 f. BGB herausgeben.“

Gelingt es, juristische Argumente zu formalisieren, können diese mit Rechnern z. B. auf Widerspruchsfreiheit überprüft werden. Auch ist es dann denkbar, dass Computer uns dabei helfen festzustellen, ob wir überhaupt alle für eine Begründung erforderlichen Argumentations-

schritte angegeben haben oder ob unsere Argumentation lückenhaft ist, etc.<sup>20</sup>

Im Bereich des sog. Natural Language Processing (NLP) werden schließlich sowohl statistische als auch regelbasierte Verfahren angewandt, um natürliche Sprachen mit Maschinen zu verarbeiten. Will man in Texten z. B. Adressangaben finden, weil diese oft ein hohes Datenschutzrisiko darstellen, lässt sich ein Teil der Adressangaben regelbasiert erkennen, wenn es sich um Adressen der Form „Saarbrücker Straße 46, 83259 Schleching“ handelt, da sich diese mit dem abstrakten formalisierten Muster (einer sog. Corpus Query)

[Prep] [A] [N] N Zahl (, | in) Zahl N  
[Prep] [A] N

auf Basis einer automatischen Wortartenannotation beschreiben lassen. Dabei stehen eckige Klammern für optionale Elemente, Prep für eine Präposition, A für ein Adjektiv, und N für ein Nomen (Substantiv oder Eigenname).

## Résumé

Wie bisher sollten sich Studierende juristisches Wissen aneignen, aber sich auch „methodisch“ mit dem Gebrauch der natürlichen Sprache, insbesondere im Zusammenhang mit der juristischen Argumentation bei Falllösungen und bei Entscheidungen, befassen und diesen fachspezifischen Sprachgebrauch trainieren.

Künftig erscheint es dann weiter ebenso vielversprechend, sich „freiwillig“ auch noch mit der Formalisierung dieser so erlernten juristischen Argumentation zu befassen, um zu verstehen, wie diese sowohl mit statistischen als auch mit regelbasierten „KI“-Verfahren repräsentiert und verarbeitet werden können.

Es macht also Sinn, sich heute neben der natürlichen Sprache auch mit Techniken zur Formalisierung und mit den Möglichkeiten formaler Sprachen zu befassen, wenn wir Juristinnen und Juristen auch keine Informatikerinnen und Informatiker werden müssen. Das Ziel sollte aber sein, Freude an einer interdisziplinären Zusammenarbeit haben zu können.

20) S. z. B. die Videos zum Seminar KI und rechtliche Argumentation aus dem WS 2021/22; Beginn mit Stunde 1 auf youtube <https://www.youtube.com/watch?v=nEmVlYanD4g>.

**ZUM AUTOR**

**Dr. Axel Adrian** ist Notar in Nürnberg und Honorarprofessor für Rechtstheorie und Rechtsgestaltung an der Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU). Er hat über Fragen der Juristischen Methodenlehre promoviert und u.a. zu Rechtstheorie, Wissenschaftstheorie sowie zu Legal Tech veröffentlicht. Seine wissenschaftlichen Interessen betreffen insbesondere interdisziplinäre Fragen von Recht und Struktur- bzw. Computerwissenschaft (Logik, Mathematik, Informatik) sowie Verfassungsrecht, Wahlrecht, Demographie und Zivilrecht.



Prof. Dr. Axel Adrian,  
Notar, Nürnberg,  
Honorarprofessor, Universität Erlangen-Nürnberg  
axel.adrian@fau.de

Dr. Bettina Mielke / Professor Dr. Christian Wolff

## Möglichkeiten und Perspektiven der Korpuslinguistik für die Analyse von Rechtstexten

Recht und Sprache bilden einen untrennbar zusammenhang, wobei nicht die gesprochene Sprache, sondern die schriftliche Sprache und damit Texte im Vordergrund stehen. Da juristische Texte mittlerweile nahezu durchgehend in digitaler Form vorliegen, ist es naheliegend, korpus- bzw. computerlinguistische Verfahren zu ihrer Analyse heranzuziehen. Unterschiedliche Ziele, die vom automatisierten Erkennen inhaltlicher Zusammenhänge bis zur Anonymisierung von Urteilen reichen, können dadurch unterstützt werden.

### Einführung

Unabhängig von der Frage, ob die Rechtswissenschaft als Geisteswissenschaft anzusehen ist – dies wird nicht einheitlich gesehen – spielt die Auseinandersetzung mit (Rechts-)Texten bei der juristischen Tätigkeit eine ebenso zentrale Rolle wie bei den Geisteswissenschaften. In diesen textzentrierten Wissenschaften sind einerseits *inhaltsbezogene* Analysen von Interesse: Welche Konzepte in einem Dokument sind wichtig? Wie ähnlich sind die Inhalte verschiedener Dokumente? Welches Dokument passt am besten zu einer Suchanfrage? Andererseits gibt es zunehmend Untersuchungen, die *stilistische* oder auch *emotionale* Aspekte in Texten herauszuarbeiten versuchen (*sentiment analysis*).

Hinzu kommen Fragestellungen hinsichtlich der Zuordnung von Texten zu bestimmten *Genres* oder der Weiterentwicklung von einzelnen (Text-)Gattungen. Ein typischer Anwendungsfall für stilometrische Analysen ist etwa die Überprüfung oder Feststellung von Autorenschaft: Von welchem Autor wurde ein

Text geschrieben? Welche Autoren haben welche Teile eines Textes verfasst?<sup>1</sup> Derartige Analysen sind mittlerweile auch im juristischen Kontext erfolgt.

### Ausgangspunkt Digitalisierung

Seit einiger Zeit stehen zunehmend digitale Textressourcen auch im Rechtswesen bereit, z. B. digitale Ausgaben juristischer Fachliteratur, Urteils- und Normdatenbanken, Akten in elektronischer Form, so dass sich gute Voraussetzungen für den Aufbau von Dokumentkollektionen/Korpora als Grundlage texttechnologischer Anwendungen ergeben. Mit der Verfügbarkeit großer digitaler Textkorpora aus einigen Zehntausenden, Hunderttausenden oder gar Millionen von Dokumenten ist die Hoffnung verbunden, andere Erkenntnisse gewinnen zu können als bei der traditionellen, auf notwendigerweise wenige Dokumente konzentrierten intellektuellen Textinterpretation. Der amerikanische Literaturwissenschaftler Franco Moretti hat dafür den Begriff des „distant reading“<sup>2</sup> geprägt, sozusagen das Lesen einer

großen Menge von Texten durch den Computer im Unterschied zum „close reading“ des Menschen, der eine kleine Textmenge intensiv studiert. Die unten beschriebenen Projekte im deutsch- und englischsprachigen Raum, bei denen jeweils Textkorpora aus dem Rechtswesen untersucht wurden, geben eine Vorstellung von den Möglichkeiten.

### Korpuslinguistische Methoden

Die Korpuslinguistik ist ein Teilgebiet der (angewandten) Sprachwissenschaft, die zum Ziel hat, mit Hilfe von Computern (und damit als Teil der Computerlinguistik) Sprach- und Textkorpora standardisiert und repräsentativ aufzubauen und auszuwerten.<sup>3</sup> Derartige Corpora können von Verfahren und Anwendungen der Sprach- und Texttechnologie (engl.

1) Eder, Rolling Stylometry, Digital Scholarship in the Humanities, 31 (2016), S. 457–469.  
2) Moretti, Distant reading. Verso Books, 2013.  
3) Lemnitzer/Zinsmeister, Korpuslinguistik: Eine Einführung, 2015.