

Juristische Fakultät

Professor Dr. Christian Wolf
Geschäftsführender Direktor

Königsworther Platz 1
30167 Hannover
Tel. +49 511 762 8268/69
Fax +49 511 762 19840
www.jura.uni-hannover.de/ipa/
www.jura.uni-hannover.de/wolf/
E-Mail:
wolf@jura.uni-hannover.de
lg.zpr@jura.uni-hannover.de

2. April 2021

Gegendarstellung

~~Legal Tech Inkubator gelingt der Durchbruch –
Examensbewertung künftig auf Knopfdruck möglich~~

**April, April –
oder warum wir nicht Pythia befragen sollten**

Ist es theoretisch denkbar, dass eine KI anhand formaler Kriterien Examensklausuren bewertet? Grundsätzlich müsste dies wohl möglich sein. Etwas verallgemeinert unterscheidet man zwischen zwei unterschiedlichen Methoden, mathematische Lösungsansätze auf juristische Sachverhalte anzuwenden. Der weitaus ältere Ansatz versucht, juristische Texte in die formale Sprache der Logik zu übersetzen und so juristische Fragestellungen rechenbar zu machen. Schon *Gottfried Wilhelm Leibniz* schwebte vor, mit einer *Jurisprudentia Rationalis*, aufbauend auf einer *Characteristica Universalis*, juristische Fragestellungen berechnen zu können. Bislang haben sich alle Versuche, rechtliche Probleme mit Hilfe der formalen Logik zu entscheiden, als nicht zielführend erwiesen.¹ Law ist eben nicht Code. Bereits die sprachliche Vagheit sowie die Notwendigkeit der Urteilkraft,

¹ *Branding*, *Artif Intell Law* 25 (2017) S. 5 f.

um das Allgemeine im Besonderen (Fall) zu erkennen, lässt sich in einem rein regelbasierten Modell nicht abbilden.

Anders verhält es sich mit dem zweiten Ansatz, mathematische Erkenntnis auf juristische Fragestellungen zu übertragen, dem datenzentrierten Ansatz. Zwischenzeitlich gibt es eine ganze Reihe von Computerlinguisten, Informatikern und Juristen, die sich insbesondere mit der Vorhersage von Gerichtsentscheidungen befassen. Dabei werden die Gerichtsentscheidungen nach formalen nicht-juristischen Kriterien analysiert. Hierzu werden die Gerichtsentscheidungen dekonstruiert. Der dogmatische Begründungszusammenhang wird aufgelöst und bestimmte Urteilspassagen als N-Gramm erfasst. Unter einem Gramm versteht man entweder einzelne Buchstaben oder Wörter. N drückt die Anzahl der Buchstaben oder Wörter aus, welche zu einem N-Gramm für die Analyse zusammengefasst werden. Ein Bi-Gramm besteht folglich entweder aus zwei Buchstaben oder, wenn die Einheit Gramm sich auf ein Wort bezieht, auf zwei Wörter.²

Die so gewonnenen Rohdaten müssen weiter aufbereitet werden. In einem ersten Schritt werden hierzu diejenigen N-Gramme gefiltert, welche in ganz unterschiedlichen Texten häufig verwendet werden, ohne charakteristisch für den Text zu sein, z.B. „das“ oder „wie“. Dies erfolgt nach der *Term Frequency-Inverse Document Frequency* (TF-IDF) Methode. Maßstab ist dabei nicht irgendeine Wortbedeutung, sondern nur die Häufigkeit der Verwendung der N-Gramme. Der Text lässt sich weiter nach bestimmten Kriterien untersuchen, die alle eins gemein haben: Sie haben nichts mit einem juristisch-dogmatischen Argument zu tun, sie lassen sich aber mit Vektoren darstellen und damit mit Referenztexten vergleichen. In unserem Aprilscherz also mit unterschiedlich bewerteten Original-Examensklausuren.

Die Programmierungen solcher Case Predictive Analytics sind sicherlich hervorragende Programmierungsleistungen. Sie sind von erheblichem intellektuellem

² Medvedeva/Vols/Wieling, *Artif Intell Law* 28 (2020) S. 237, 249 ff.

Reiz und üben deshalb ganz beachtliche Anziehungskraft aus. Sicherlich gibt es auch Bereiche, wo jenseits des intellektuellen Spiels Erkenntnisse daraus gewonnen werden können. Aber:

Die Entscheidungsfindung, warum die Arbeit mit z.B. 7 Punkten bewertet wurde, ist nicht mehr nachvollziehbar. Sie beruht ja, wie beschrieben, nicht auf einem inhaltlichen Fehler (z.B. es wurde ein Werkvertrag angenommen statt eines Werklieferungsvertrags). Was die Studierenden zum Examen lernen und wie ihre Leistung bewertet wurde, korrespondiert nicht mehr.

Obwohl ein solches Bewertungssystem wohl kaum auf Akzeptanz stoßen wird, wird es im Legal Tech-Bereich ernsthaft diskutiert. *Richard Susskind*, *Online Courts and the future of Justice*, 2019, empfiehlt den KI-Einsatz, um den Richter durch einen Computerrichter zu ersetzen. An die Stelle einer juristisch-argumentativ abgesicherten Entscheidung tritt die Wahrscheinlichkeitsprognose der KI, wie ein Richter entschieden hätte. Es geht dabei um eine für die Parteien nicht mehr erkennbare und damit adressierbare Mustererkennung. *Susskind* beschreibt daher auch die Arbeitsweise von KI-Systemen wie folgt: „(...) *many contemporary AI Systems operate not by copying human beings; instead, they function in quite different and unhuman ways.*“ (S. 274)

Wie immer man „unhuman“ übersetzen will, als unmenschlich, unhuman, grausam oder übermenschlich, mit unserem Denken, Handeln und Schreiben ist die Bewertung des Einzelfalls nicht mehr erreichbar.

Die Ironie dieses tief in der Kalifornischen Ideologie verhafteten Denkens, welche eine Welt ohne Menschenwürde, Freiheit und Demokratie für fortschrittlich und effizient hält (*Nemitz/Pfeffer*, *Prinzip Mensch* 2020, 91), ist, dass man sich im Machine Learning-Umfeld zumeist der Programmiersprache Python bedient. Der Name Python leitet sich von der Komikergruppe des surrealen britischen Humors, Monty Python, ab (<https://docs.python.org/3/tutorial/appetite.html>). Zugleich ist Python aber auch der Name der Drachin gewesen, die das Orakel von Delphi

bewachte. Von dort machte die Priesterin Pythia, durch Dämpfe in einen rauschhaften Zustand versetzt, ihre Weissagungen.

Die auf die KI gestützte Mustererkennung und die damit einhergehende Entscheidungsvorhersage – ob nun Examensbewertung oder Gerichtsurteil – macht uns zu Empfängern eines Orakelspruchs. Wir sind nicht mehr Subjekt eines solchen Verfahrens der Rechtsfindung, sondern Objekt. Die Entscheidung hätten wir hinzunehmen wie einen Orakelspruch. War Aufklärung die Befreiung aus der selbstverschuldeten Unmündigkeit, wäre eine derartige Examensbewertung bzw. gerichtliche Entscheidungsfindung die selbstverschuldete Begründung der Unmündigkeit.