

Predictive Analytics: Die Zukunft der juristischen Recherche?

von Christian Hartz



Christian Hartz ist Legal Engineer bei Wolters Kluwer und als Product Owner für verschiedene KI-Projekte verantwortlich.



Wolters Kluwer

In Deutschland ist Wolters Kluwer ein führender Anbieter von Fachinformationen, Software und Services im Bereich Recht, Wirtschaft und Steuern. Unsere Expertenlösungen verbinden profunde Expertise in klar definierten Fachgebieten mit Technologie und Services. So erzielen unsere Kunden bessere Analysen, Ergebnisse und höhere Produktivität.

A. Einleitung

Entscheidet in der Zukunft ein Roboter über die Strafbarkeit einer Person? Werden gerichtliche Standardentscheidungen innerhalb kürzester Zeit automatisch getroffen, sobald alle Parteien ihre Sachvorträge elektronisch eingereicht haben?

Auch wenn das alles noch ganz leise Zukunftsmusik ist: Wir sind bereits auf dem Weg in diese Zukunft. Ob sie dann tatsächlich Realität wird, ist eine ganz andere Frage.

Sicher ist allerdings, dass diese sogenannten Robo-Lawyers und -Judges eine Analyse-Komponente benötigen. Diese muss in der Lage sein, Sachverhaltsschilderungen entgegenzunehmen und Entscheidungen zu treffen, wie das Recht auf diesen Sachverhalt anzuwenden ist.

Das klingt so, als läge bis dahin noch ein weiter Weg vor uns. Aber ist das wirklich so? Schauen wir uns einmal an, wo wir stehen.

B. Was ist Predictive Analytics

Zur Erklärung, was Predictive Analytics ist, gehen wir zunächst einmal eine Stufe zurück und befassen uns mit der Frage was Analytics im juristischen Bereich bedeutet. Um es von Predictive Analytics abzugrenzen bezeichnen wir diese Form als Legal Analytics.

I. Legal Analytics

Legal Analytics verwendet verschiedene Hilfsmittel, um bereits vorhandene juristische Texte (i.d.R. Gerichtsentscheidungen) so für den Computer verständlich zu machen, dass er die Informationen verwenden kann. Diese Analyse basiert vor allem auf Natural Language Processing (NLP).

Auch Informationen wie Verfahrensdauer, Gericht, Gerichtsort, Richter oder die verwendeten Normen und Fundstellen können statistisch ausgewertet werden.

Schließlich können Muster in Entscheidungen sichtbar gemacht werden und bestimmte Entscheidungen als zentral für juristische Problemfälle identifiziert werden, indem man sich anschaut, wie sich die Entscheidungen in dem Gefüge anderer Entscheidungen bewegen. In den folgenden Bildern sieht man, dass es ein paar wenige Entscheidungen gibt, die besonders viele Verknüpfungen haben; diese bezeichnet man als Hubs:

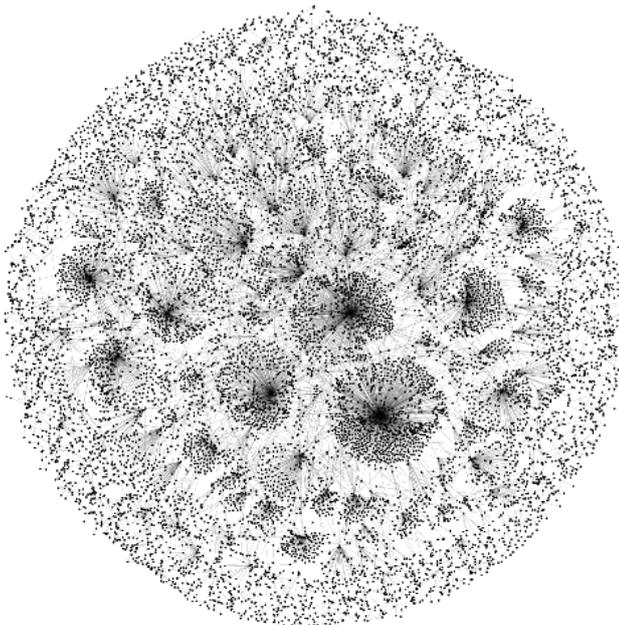


Abbildung 1: Citation-Network Curia of Hungary

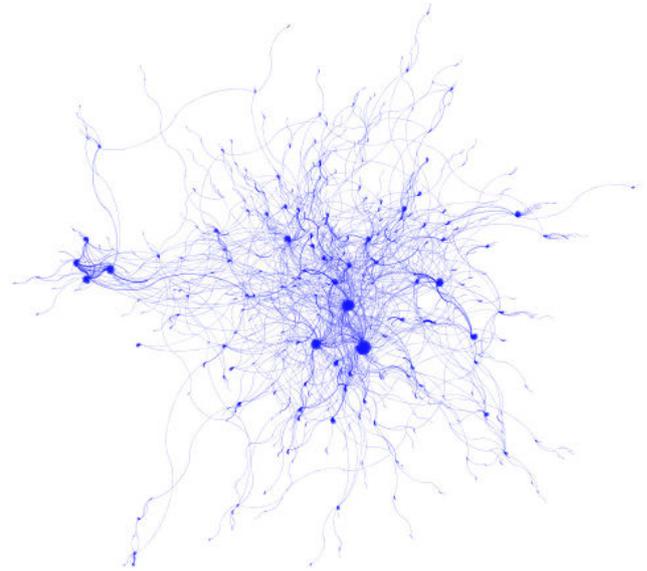


Abbildung 2: Citation-Network Curia of Hungary – Nur Hubs und starke Links

II. Predictive Analytics

Predictive Analytics ist der nächste Schritt: Das generierte Wissen wird genutzt, um für eine Sachverhaltseingabe eine Vorhersage darüber zu treffen, wie dieser Sachverhalt entschieden werden wird. Es werden also die durch Legal Analytics geschaffenen Informationen aus bekannten Entscheidungen verwendet, um Vorhersagen (Predictions) über bisher unbekannte Sachverhalte zu treffen.

Das ist kein grundsätzlich neues Thema. „Predictive Analytics“ befindet sich bereits seit 2010 auf dem „Gartner Hype Cycle for Emerging Technologies“ im Bereich des „Slope of Enlightenment“ mit einer allgemeinen Rezeption innerhalb von weniger als zwei Jahren. Allerdings gilt es zu beachten, dass hier mehrheitlich die Rezeption im Bereich des Maschinenbaus oder anderen Bereichen der Fertigung gemeint ist.

Für den juristischen Bereich geht es weniger darum, Sensordaten einzusammeln, um Vorhersagen über mögliche Ausfälle oder Abnutzungen von Maschinen zu treffen. Der Fokus liegt auf der Verarbeitung von Texten.

Vorhersagen im juristischen Bereich können Informationen über mögliche Anspruchsgrundlagen, mögliche Argumente und Gegenargumente oder Gewinnwahrscheinlichkeiten sein.

III. Prescriptive Analytics

Prescriptive Analytics schließlich soll eine Hilfestellung bieten, wie das gewünschte Ergebnis – beispielsweise den Fall vor Gericht zu gewinnen – erreicht werden kann.

Basis dafür sind sowohl Legal Analytics als auch Predictive Analytics. In einer stark vereinfachten Form könnte man sagen, dass alle Anspruchsgrundlagen und alle Argumente, die durch Legal und Predictive Analytics gefunden wurden, solange durchgespielt werden, bis die Gewinnwahrscheinlichkeit der Kombination am höchsten ist. Praktisch gesprochen könnte Prescriptive Analytics so aussehen, dass diejenige Argumentation empfohlen wird, die verglichen mit allen anderen möglichen Argumentationen den größten Erfolg verspricht. Es könnte aber auch angezeigt werden, welche Ansprüche geltend gemacht werden sollten, weil diese in ähnlich gelagerten Fällen zu einem erfolgreichen Ergebnis führten.

Natürlich könnten auch spezifische Informationen zu Gerichten mit einfließen und beispielsweise das Entscheidungsverhalten eines bestimmten Richters berücksichtigt werden. Das Ergebnis wäre dann eine auf den Richter abgestimmte Argumentation oder ein auf ihn abgestimmtes Verfahren.

Gerade letzteres entspricht oftmals dem Wissensvorsprung, den eine lokal ansässige Kanzlei gegenüber auswärtigen Vertretern haben kann.

C. Die Technik hinter Legal/Predictive/Prescriptive Analytics

Nachdem wir uns nun kurz die möglichen Nutzungsszenarien angesehen haben, werden wir uns in diesem Abschnitt mit der Technik hinter den Analytics-Lösungen beschäftigen. Die Ausführungen können aus Platzgründen allerdings nur ganz oberflächlich ausfallen.

Grundsätzlich können sowohl auf der Nutzereingabe als auch auf der Ausgabe der Lösung verschiedene Techniken eingesetzt werden. Auf der einen Seite gilt es den Korpus der bereits vorhandenen Entscheidungen zu analysieren bzw. aufzubereiten. Auf der anderen Seite gilt es die Nutzereingabe selbst zu verarbeiten.

Folgende Komponenten sind oftmals Teil solcher Lösungen:

- Natural Language Processing
- Machine Learning
- Deep Learning Ansätze
- Knowledge Graphs

Dies ist keinesfalls eine abschließende Liste, sondern soll eine relevante Auswahl darstellen, die wir uns im Folgenden etwas genauer anschauen.

I. Natural Language Processing

Unter Natural Language Processing (NLP) versteht man Techniken, mit denen der Computer in die Lage versetzt werden soll, Text besser verarbeiten zu können.

Insbesondere seit dem Aufkommen von Transformer-Modellen wie BERT, GPT-2 (mittlerweile gibt es GPT-3) oder XLNet ist man nicht mehr darauf angewiesen, einzelne Regeln zu formulieren, um ein Wissensmodell für den Computer zu erstellen.¹ Diese Modelle wurden auf sehr großen Datenmengen trainiert (BERT bspw. auf 800M Wörter aus dem BooksCorpus und 2,500M Wörter aus Wikipedia)², bei denen wie bspw. bei BERT einzelne Wörter maskiert werden und der Algorithmus Vorhersagen treffen muss, welches Wort ausgelassen wurde. Hierdurch lernen diese Algorithmen die Sprache ohne weiteres Zutun der Entwickler („unsupervised learning“).

Verfahren des Natural Language Processing sind beispielsweise Wort-Segmentierung (Tokenization), Lemmatisierung, Part-of-Speech-Tagging, Named Entity Recognition (NER) oder Sentiment Analysis. Aber auch Topic-Recognition, Text-Summarization oder Machine-Translation sind Elemente von Natural Language Processing-Verfahren.

Schauen wir uns diese Verfahren im Einzelnen an:

Tokenization zerlegt Sätze (oder auch Wörter) in einzelne Wörter, um ihnen dann in weiteren Verfahren bestimmte Attribute zuzuordnen.

Die Lemmatisierung versucht, die gefundenen Wörter auf ihre Grundform zurückzuführen (Bäume -> Baum).

Mittels Part-of-Speech-Tagging können Elemente von Sätzen Wortarten zugeordnet werden. Zusammen mit dem Dependency Parsing können Beziehungen zwischen den Wörtern in den Sätzen hergestellt werden:

¹ Transformer wurden eingeführt durch Google zusammen mit der Universität von Toronto, Vaswani/Shazeer u.a., „Attention is All You Need“, arXiv:1706.03762 [cs.CL], [hier](#) abrufbar (Stand: 28.05.2021).

² Devlin/Chang u.a., BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding, arXiv:1810.04805, [hier](#) abrufbar (Stand: 28.05.2021).

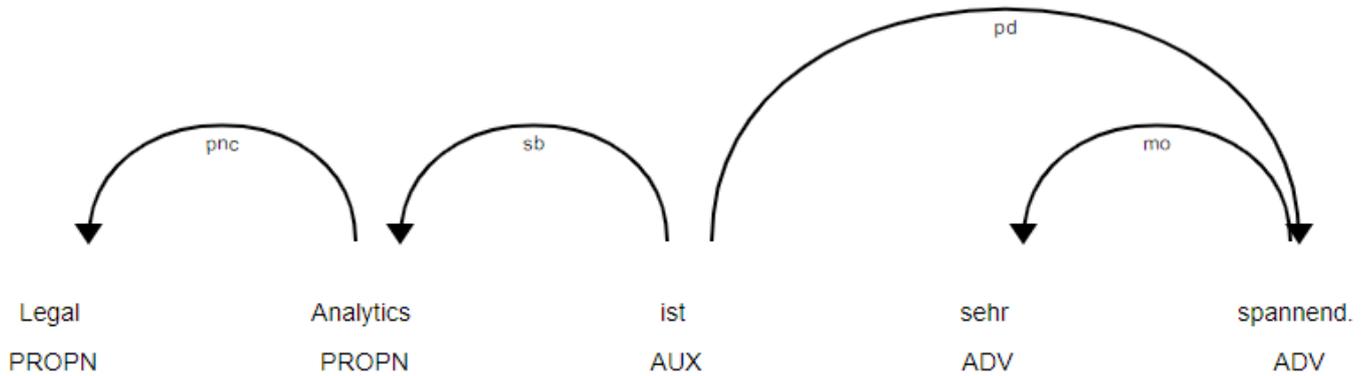


Abbildung 3: Dependency Parsing mit spaCy und displaCy

Named Entity Recognition (NER) ist ein Verfahren, mit dem Wörter bestimmten Klassen zugeordnet werden. Das können beispielsweise Namen, Organisationen, Orte aber auch Datumswerte oder Geldbeträge sein. In Abbildung 4 findet sich ein Beispiel aus einem Wikipedia-Text über die Universität Köln:

The University of Cologne **ORG** was established in 1388 **DATE** as the fourth **ORDINAL** university in the Holy Roman Empire, after the Charles University of Prague **ORG** (1348 **DATE**), the University of Vienna **ORG** (1365 **DATE**) and the Ruprecht Karl University **ORG** of Heidelberg (1386 **DATE**).

Abbildung 4: NER basierend auf dem Wikipedia-Text der Universität von Köln ([Hier](#) abrufbar (Stand: 28.05.2021)).

Die Sentiment-Analyse wird oftmals dafür verwendet, um positive Produkt-Rezensionen von negativen zu unterscheiden. Allerdings kann sie auch in juristischen Texten, beispielsweise zur Vorhersage in einem bereits laufenden Verfahren verwendet werden.³

Topic-Recognition ist der Versuch, entweder durch gänzlich unüberwachtes Lernen oder durch semiüberwachtes Lernen Themengebiete in Dokumenten zu erkennen und diese zu separieren. Dies kann sowohl auf der Seite der Nutzereingabe als auch in den Rechtsprechungsdokumenten selbst sinnvoll sein. Ein auf statistischer Auswertung beruhender Ansatz für das Topic-Modelling ist „Latent Dirichlet Allocation“.⁴

³ Liu/Chen, A two-phase sentiment analysis approach for judgement prediction, *Journal of Information Science*, 44(5), S. 594 ff.

⁴ Blei/Ng u.a., *Journal of Machine Learning Research* 2003, 993 ff., [hier](#) abrufbar (Stand: 28.05.2021).

Neuere Ansätze verwenden hierfür Vektormodelle⁵ oder den von Google entwickelten BERT.⁶

Text-Summarization-Methoden können verwendet werden, um Zusammenfassungen dessen zu erstellen, was für den Juristen wirklich relevant ist; beispielsweise können alle wichtigen Fakten zusammengetragen werden, um anschließend eine Zusammenfassung zu erstellen. Die Text-Summarization gelingt noch nicht immer; insbesondere die Extractive Summarization wird aber immer besser.⁷

Schließlich bietet Machine-Translation, wenn auf den ersten Blick auch nicht ganz einfach ersichtlich, große Vorteile. Dies liegt daran, dass die deutsche Sprache im Vergleich zur englischen Sprache von weniger Modellen und Algorithmen unterstützt wird. Durch ein automatisiertes Übersetzen können gute Ergebnisse erzielt werden und englischsprachige Algorithmen können zur Anwendung kommen.⁸ So könnte hierüber beispielsweise auch LEGAL-BERT⁹, der auf eng-

⁵ Angelov, Top2Vec, arXiv:2008.09470, [hier](#) abrufbar (Stand: 28.05.2021).

⁶ Vgl. statt vieler nur Grootendorst, Topic Modeling with BERT, [hier](#) abrufbar (Stand: 28.05.2021).

⁷ Vgl. zum Thema übersichtsartig Kanapala/Pal, Text summarization from legal documents: a survey, *Artificial Intelligence Review*, March 2019, DOI: 10.1007/s10462-017-9566-2.

⁸ Zum grundsätzlichen Thema: Schuster/Gupta u.a., Cross-lingual Transfer Learning for Multilingual Task Oriented Dialog, arXiv:1810.13327 [cs.CL], [hier](#) abrufbar (Stand: 28.05.2021).

⁹ Chalkidis/Fergadiotis u.a., LEGAL-BERT: The Muppets straight out of Law School, arXiv:2010.02559, [hier](#) abrufbar (Stand: 28.05.2021).

lischsprachigen Texten trainiert wurde, zum Einsatz kommen.

II. Clustering Algorithmen

Daneben können aber auch Clustering-Algorithmen verwendet werden. Ein Clustering-Algorithmus nutzt beispielsweise Wort¹⁰, Satz¹¹, Absatz¹² oder Dokumentembeddings¹³, um Ähnlichkeiten zwischen Texten festzustellen oder sie bestimmten Datentöpfen zuzuordnen. Vereinfacht dargestellt wird jedem Wort eine Vektor-Repräsentation zugeordnet. Dies führt dazu, dass auch einfache mathematische Kalkulationen auf diesen Vektoren vorgenommen werden können. Im Rahmen von Predictive Analytics kann dies beispielsweise dazu verwendet werden, um in einer Entscheidung des Gerichtes die Sätze, in denen das Gericht abstrakte Rechtsausführungen macht, von denen zu unterscheiden, in denen ganz konkret über den vorliegenden Fall gesprochen wird.

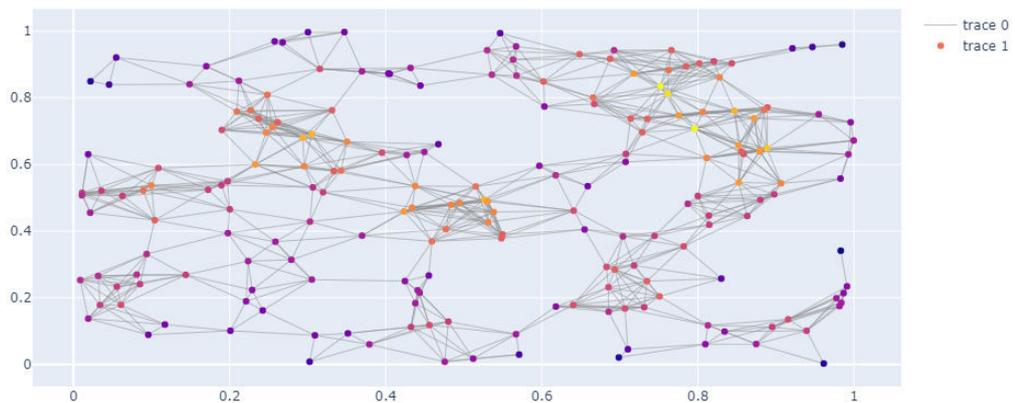


Abbildung 5: Beispielgraph erstellt mit Plotly

Was ist der Vorteil?

Abstrakte Rechtsausführungen können auf einen anderen Fall angewendet werden, der ähnliche juristische Probleme aufweist. Konkrete Ausführungen hingegen können auf einen sehr ähnlichen Fall ggf. sogar direkt übertragen werden.

¹⁰ Mikolov/Chen u.a., Efficient Estimation of Word Representations in Vector Space, arXiv:1301.3781 [cs.CL], [hier](#) abrufbar (Stand: 28.05.2021).

¹¹ Pagliardini/Gupta u.a., Unsupervised Learning of Sentence Embeddings Using Compositional n-Gram Features, NAACL 2018, 528 ff., [hier](#) abrufbar (Stand: 28.05.2021).

¹² Dai/Olah u.a., Document Embedding with Paragraph Vectors, arXiv:1507.07998 [cs.CL], [hier](#) abrufbar (Stand: 28.05.2021).

¹³ Le/Mikolov, Distributed Representations of Sentences and Documents, arXiv:1405.4053 [cs.CL], [hier](#) abrufbar (Stand: 28.05.2021).

III. Sonstige Techniken

Es können aber auch andere Techniken zum Einsatz kommen, wie klassische Entscheidungsbäume oder ein Knowledge-Graph. Entscheidungsbäume können genutzt werden, um Entscheidungswege abzubilden, die erneut abgerufen werden könnten. So wäre es eine Option, Aufbauschemata zu hinterlegen. Die Nutzereingabe könnte automatisch damit abgeglichen werden, um dann dem Nutzer eine Information darüber anzubieten, ob die Anspruchsvoraussetzungen erfüllt sind. So setzt der Anspruch aus § 433 I 1 BGB das Vorliegen eines Kaufvertrages voraus. Ob dieser vorliegt, könnte aus dem Sachverhalt automatisch entnommen, geprüft und dann als Voraussetzung bejaht oder verneint werden.

Ein Knowledge-Graph kann verwendet werden, um alle Informationen so zu speichern, dass die Beziehungen über das hinterlegte Wissen abgebildet werden können. So können dem Nutzer die passenden Ergebnisse geliefert werden.¹⁴ Der Knowledge-Graph beinhaltet Wissen über die Relation zwischen zwei Elementen. Dies kann ein Über-Unterordnungsverhältnis, eine Eltern-Kind-Beziehung oder auch eine ganz andere Relation sein. Die Informationen im Knowledge-Graph werden i.d.R. in einem Triple abgebildet, das aus Subjekt, Prädikat und Objekt besteht.

Ein Beispiel: (Strafrecht) -> (ist Teil des) -> (Öffentlichen Rechts). Die Abbildung 5 stellt einen einfachen Graphen dar.

¹⁴ Zum Einsatz eines Knowledge-Graphen bei Wolters Kluwer Deutschland, [hier](#) abrufbar (Stand: 28.05.2021); bzw. Crotti Junior/Orlandi u.a., Knowledge Graph-Based Legal Search over German Court Cases, ESWC 2020, 293 ff.

Bei all diesen verschiedenen Techniken ist es wichtig, sie sinnvoll zu orchestrieren. Nur dann kann Predictive Analytics funktionieren.

D. Der Workflow

Der Workflow, der mit Predictive Analytics abgedeckt werden kann, ist sehr vielfältig. Bereits die Eingabemöglichkeit in eine Predictive Analytics-Solution ist mannigfaltig:

- Texteingabe in ein Suchfeld
- Übernahme von E-Mails oder Word-Inhalten direkt in eine Analytics-Funktion
- Übernahme aus Dokumentenmanagementsystemen oder Kanzleisoftware
- Spracheingabe
- ...

Auch die Ausgabe kann unterschiedlich aussehen; eine einfache Trefferliste mit Ergebnissen von Dokumenten, die auf die Sucheingabe passen oder die Top 10 Argumente, die genutzt werden sollten. Aufgrund dieser umfangreichen Einsatzmöglichkeiten sollte gerade der Nutzen für den Kunden nicht aus dem Fokus geraten. Es muss zunächst das implementiert werden, was den höchsten Kundennutzen bringt. Hierzu kann beispielsweise ein Lean Product Cycle wie der in Abbildung 6 verwendet werden.

E. Regulatorische Maßnahmen und Hürden

Wenden wir unseren Blick nun den regulatorischen Anforderungen zu.

Je tiefer die Applikation in den Workflow des Nutzers eingebunden ist, desto höhere Anforderungen an Security und Privacy sind zu stellen; auch § 203 StGB wird relevanter.

In neuerer Zeit und besonders für die Zukunft wichtig ist der Verordnungsentwurf der Europäischen Kommission für KI (sog. Artificial Intelligence Act).¹⁵

Würde eine solche Lösung unter anderem von Richtern genutzt werden, so würde sie gemäß Rz. 40 des Vorschlages als High-Risk-AI-System klassifiziert werden:

“In particular, to address the risks of potential biases, errors and opacity, it is appropriate to qualify as high-risk AI systems intended to assist judicial authorities in researching and interpreting facts and the law and in applying the law to a concrete set of facts.”

Dies führt allerdings nicht dazu, dass ein solches System verboten wäre.

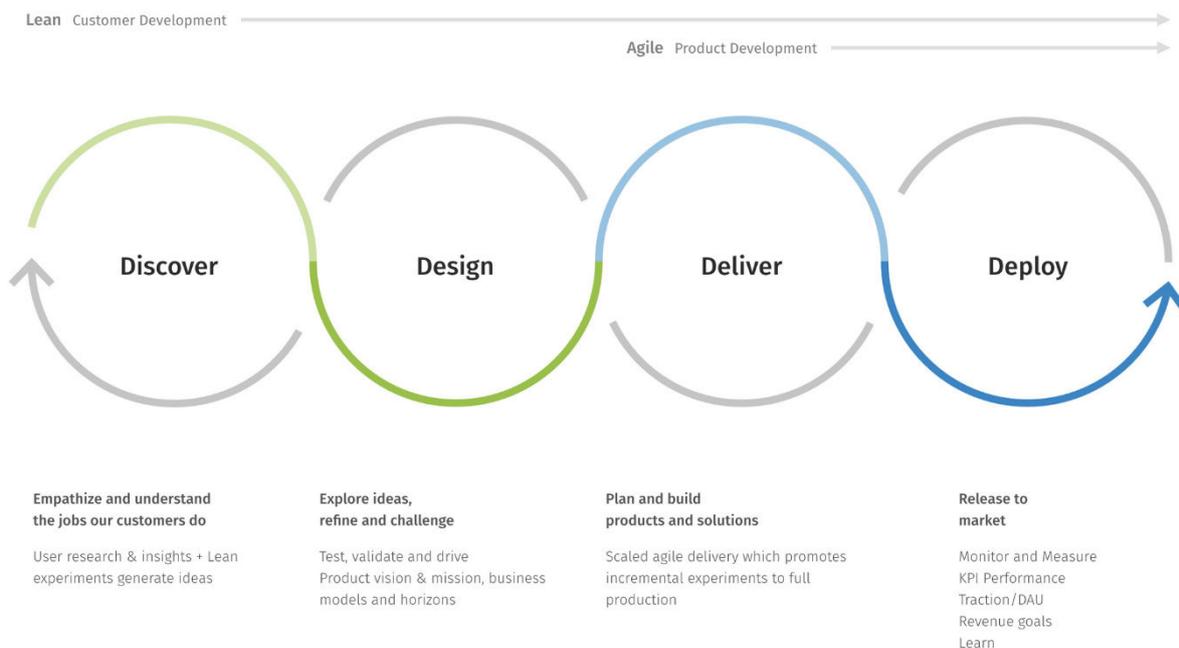


Abbildung 6: Lean Product Cycle von Wolters Kluwer

¹⁵ Proposal for a Regulation laying down harmonised rules on artificial intelligence (Artificial Intelligence Act), hier abrufbar (Stand: 28.05.2021).

Es führt lediglich dazu, dass besondere Anforderungen an die Entwicklung zu stellen sind, wie sie in Rz. 43 aufgezählt werden:

“Requirements should apply to high-risk AI systems as regards

- the quality of data sets used,
- technical documentation and record-keeping,
- transparency and the provision of information to users,
- human oversight,
- and robustness,
- accuracy and
- cybersecurity”

Auch wenn der AI Act noch nicht in Kraft ist und das Inkrafttreten noch dauern wird, sollte er bereits jetzt berücksichtigt werden. Wird etwa nicht eingehalten, dass der Datensatz eine gewisse Qualität hat, so kann dies im schlimmsten Fall dazu führen, dass alle Algorithmen gänzlich neu trainiert werden müssen.

F. Fazit

Wie aufgezeigt, ist die Technik zur Umsetzung einer Predictive- oder Prescriptive-Analytics-Lösung keine Magie und auch keine Zukunftsmusik; sie ist bereits vorhanden. Es ist nur noch eine Frage der Zeit, bis auf dem deutschen Markt erste Predictive-Analytics-Anwendungen zu finden sein werden. Auch Wolters Kluwer arbeitet an solchen Lösungen. Gerade zur rechten Zeit kommen daher die regulatorischen Anforderungen, um sicherzustellen, dass Predictive Analytics nicht zu einer Gefahr für den Rechtsmarkt, sondern zu dem wird, was es sein kann: Eine Hilfe, um juristische Fälle einfacher und besser zu lösen.

CTRL

Cologne Technology **R**&Law
review

+
Hier geht es zur
ganzen Ausgabe.

Dort findest du auf über
100 Seiten in 15 Aufsätzen
alles von NFTs über Legal
Tech im Strafprozess bis
hin zum Stand des
E-Examens in NRW.

