

## Contract Analytics

Lässt Künstliche Intelligenz Raum für Menschliche Intelligenz? –  
 Auf den Use Case kommt es an!

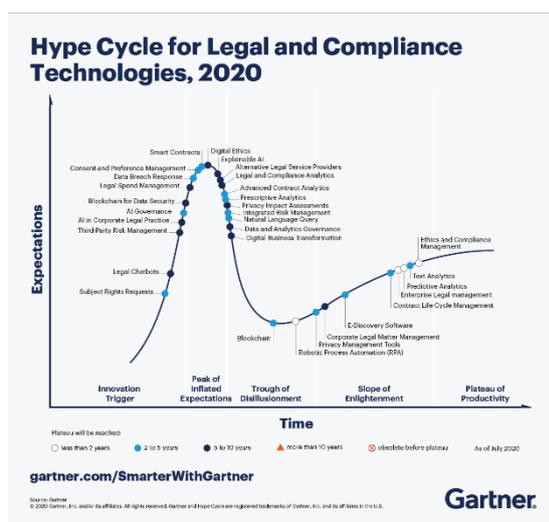
Michael Talevski | Legal Operations | TPR Legal

16. November 2020

LR 2020, Seiten 324 bis 328 (insgesamt 6 Seiten)

Betrachtet man den aktuellen Gartner Hype Cycle für „Legal and Compliance Technologies“, so zeigt sich, dass KI-basierte Contract Analytics (Text Analytics) allmählich die Akzeptanz im relevanten Markt finden.<sup>1</sup> Zugleich zeigt sich

1



aber bei genauerer Betrachtung, dass beim Einsatz der von Contract Analytics Anbietern verwendeten Technologien eine klare Definition des Use-Cases notwendig ist. Diese Definition muss Grundlage der Entscheidung sein, wann und insbesondere welche Art von Contract Analytics den versprochenen Mehrwert bei der Lösung eines bestehenden Problems bringen soll. Die Entscheidendes Kriterium ist hierbei, ob die vom der KI generierten Extraktionsergebnisse zur bloßen Analyse von Tendenzen oder gar zur inhaltlichen Bewertung eines Vertrages verwendet werden. Es zeigt sich, dass insbesondere juristische Auslegungen und

Diskussionen die Qualität der Extraktionsergebnisse beeinträchtigen können. Zudem kann bereits zu Beginn die Anzahl der vorliegenden Trainingsbeispiele, die zum Anlernen eines solchen Systems benötigt werden, über Scheitern und Erfolg eines Projekts bestimmen.

Dieser Artikel gibt einen Überblick über die Einsatzmöglichkeiten von Contract Analytics und sensibilisiert Anwender durch die Abgrenzung von Machine Learning und Semantic Analytics für die besonderen Herausforderungen im Umgang mit KI – ganz im Sinne eines Erwartungsmanagements.<sup>2</sup>

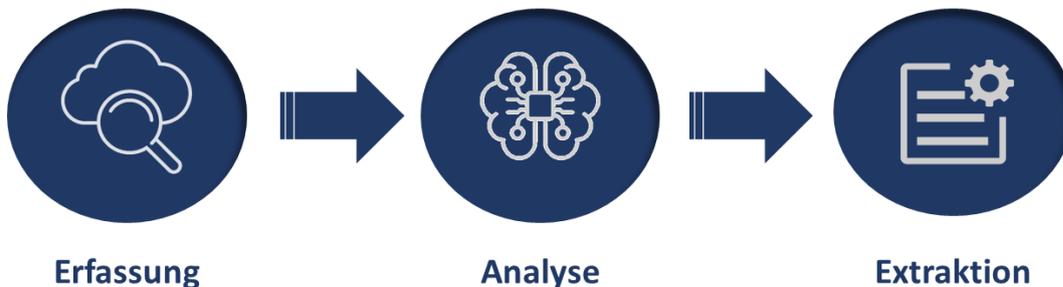
<sup>1</sup> <https://www.gartner.com/smarterwithgartner/4-key-trends-in-the-gartner-hype-cycle-for-legal-and-compliance-technologies-2020/> (abgerufen am 26.11.2020)

<sup>2</sup> Beispiele: [DocuSign Insight](#), [Kira](#), [Eigen](#), um exemplarisch drei Technologieanbieter zu nennen.

## I. Contract Analytics

Contract Analytics Anwendungen sollen bei der Analyse und Bewertung von Verträgen unterstützen, das Identifizieren von relevanten Klauseln beschleunigen oder als Grundlage für den Aufbau eines Contract Management-Systems dienen.<sup>3</sup> Effizienz, Qualitätsoptimierung und daraus resultierendere Zeitgewinne sind zumeist die zentralen Bewertungskriterien bei der Überlegung ein Contract Analytics Tool einzusetzen. Die Methoden dieser Anwendungen reichen dabei von der semantischen Analyse des Textes (Semantic Analytics) bis hin zu auf Machine Learning (ML) gestützten Contract Analytics. Aber auch eine Kombination aus beiden Methoden ist denkbar, um Verträge zu analysieren. Die zuvor genannten technologischen Betrachtungen werden meist unter dem Begriff „Künstliche Intelligenz“ zusammengefasst. Vereinfacht wird unter Künstlicher Intelligenz ein System verstanden, das Probleme selbstständig lösen kann.<sup>4</sup>

Methodisch kann man den Analyseprozess dieser Anwendungen grob in drei Schritte unterteilen:



- Schritt 1: Die Umwandlung von unstrukturierten Daten und verschiedenen Dateiformaten in ein maschinenlesbares Format (OCR) bildet dabei die Grundlage - **Erfassung**
- Schritt 2: Die Identifizierung von Textinhalten oder vertraglichen Klauseln - **Analyse**
- Schritt 3: Die Zuordnung der Klausel zu dem jeweiligen, inhaltlich korrespondierenden Themenbereich - **Extraktion**

## II. Machine Learning vs. Semantic Analytics

Um eine effiziente Vertragsanalyse durchführen zu können, setzen Anbieter vermehrt auf Machine Learning oder die semantische Analyse von Klauselinhalten.

<sup>3</sup> Stikeman Elliot LLP, Seven Legal Technologies Every Lawyer Should Know About, 2018, S. 3

<sup>4</sup> Thomas Söbbing, Fundamentale Rechtsfragen zur künstlichen Intelligenz, RethinkingLaw, Februar 2019, S. 33 ff.

## 1. Machine Learning

Voraussetzung einer zuverlässigen Vertragsanalyse auf Basis des Machine Learnings bildet dabei ein anwendergestützter und damit überwachter Trainingsprozess auf Grundlage vorhandener Datensätze.<sup>5</sup> Aufgrund von Scoping- bzw. Klassifikationsproblemen, also Schwierigkeiten bei der Zuordnung eines bestimmten Trainingsbeispiels in eine jeweilige thematische Kategorie, empfiehlt sich bei der Analyse von rechtlichen Texten das überwachte, auch als *Supervised* oder *User Defined* bezeichnete, Machine Learning.

4

Unter Machine Learning kann vereinfacht eine statistische Methodik verstanden werden, mit der neue Informationen auf Grundlage von bestehenden (Trainings-)Daten generiert werden.<sup>6</sup> Bei dieser Form des Machine Learnings trainiert eine versierte Person<sup>7</sup> - **ein:e Analyt:in** - das System auf Grundlage von vorgegebenen Beispielen und ordnet diese einem thematischen Überbegriff, einer Regel oder **Policy**, zu. Diese Beispiele werden in der Regel von sogenannten **Subject Matter Experts (SME's)**, also Fachkundige in einem bestimmten Tätigkeitsbereich, vordefiniert. Mit Hilfe dieser Trainingsbeispiele soll es dem System möglich sein, die erlernten Muster zu erkennen und diese auf neue, unbekannte Daten anzuwenden. Vereinfacht bedeutet dies; man „füttert“ die Maschine mit gleichgelagerten Klauselbeispielen, bis thematisch korrespondierende Klauseln selbstständig, in unbekanntem Dokumenten, erkannt werden (Analyse) und als Ergebnis in der Policy erscheint (Extraktion). Anwendungsfall dieses Instruments ist insbesondere die schnelle Analyse großer Vertragsmengen, um Tendenzen und thematische Schwerpunkte aus dem jeweiligen Vertragswerk erkennen zu können, um so direkt an die richtige Stelle im Vertrag geführt zu werden.

Es zeigt sich in der Praxis, dass insbesondere bei der Betrachtung von inhaltlich komplexen Klauseln die ML-basierten Contract Analytics Tools an ihre Grenzen stoßen können, denn Machine Learning benötigt im Idealfall eindeutige und vollständige Trainingsergebnisse. Im Trainingsprozess wird von einem oder mehreren SME's dem System vorgegeben, welche Klauselbeispiele zu einem bestimmten Themenkomplex gehören.<sup>8</sup> Dieses sog. Scoping wird bei eindeutigen Klauseln, wie etwa bei der Beantwortung der Frage: „Ist diese Klausel eine Haftungsklausel?“ noch zu einem wohl sehr eindeutigen Ergebnis und damit inhaltlich gleichgelagerten Extraktionen führen. So sieht es schon anders aus, soll die Frage beantwortet werden, ob eine Klausel in den Anwendungsbereich einer bestimmten Norm fällt, oder eben nicht. Bestehen für die Beantwortung dieser Frage zudem bspw. unterschiedliche Rechtsansichten (oder ansonsten zu ihrer inhaltlichen Einordnung oder Bewertung), führt dies zur geringeren Eindeutigkeit der Trainingsdaten und damit aus Sicht der Anwender zu einem verfälschten bzw. weniger genauen Extraktionsergebnis. Relevant für

5

---

<sup>5</sup> <https://www.bigdata-insider.de/was-ist-machine-learning-a-592092/> (abgerufen am 26.11.2020)

<sup>6</sup> Human-Centred Machine Learning, CHI EA '16: Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems, 2016

<sup>7</sup> Amr, Hands-On Machine Learning with scikit-learn and Scientific Python Toolkits, 2020, S. 11

<sup>8</sup> Aggregating Crowdsourced Labels in Subjective Domains. Proc. ACM Hum.-Comput. Interact.X, April 2018, S. 17

die Qualität des Extraktionsergebnisses ist folglich die Vollständigkeit und Eindeutigkeit der Trainingsdaten.<sup>9</sup> Da Diskussionen und abweichende Meinungen zum (juristischen) Arbeitsalltag gehören, sind Ungenauigkeiten und Divergenzen im Scoping und damit auch den bereitgestellten Trainingsdaten, insbesondere bei komplexen Themen, kaum zu vermeiden. Möchte der Anwender<sup>10</sup> nun nachvollziehen, wieso das eingesetzte Tool zu einem Ergebnis gekommen ist, ist dies bei rein ML-gestützter Contract Analytics nahezu unmöglich<sup>11</sup>. Dies betrifft auch die Nachvollziehbarkeit von unzutreffenden Extraktionsergebnissen (sog. *false positive*) aber auch von vorhandenen, aber durch das System nicht gefundenen Ergebnissen (sog. *false negative*).

Diese Methode des Contract Analytics eignet sich daher im Kern vorwiegend zum Erkennen von Tendenzen sowie dem schnellen und strukturierten Auffinden von eindeutigen Klauseln. Zudem benötigen Machine Learning Komponenten in der Regel eine sehr hohe Anzahl an Trainingsbeispielen (Verträgen), weshalb diese Methode bereits in der Vorbereitungsphase zum Scheitern verurteilt sein könnte, wenn diese nicht zur Verfügung stehen.

### **Exkurs: Explainable Artificial Intelligence**

Aufgrund des zuvor angesprochenen Problems, der nahezu unmöglichen Nachvollziehbarkeit von Extraktionsergebnissen beim Einsatz ausschließlich ML-gestützter Systeme, geht der aktuelle Trend in die Richtung der Explainable Artificial Intelligence (XAI). Ziel dieser Methode ist dem Anwender die Möglichkeit zu geben, das Ergebnis wirklich nachvollziehen und eine für den Menschen verständliche Lösung liefern zu können.

6

Hierbei werden aktuell verschiedenste Ansätze diskutiert. Insbesondere im Bereich des Text bzw. Contract Analytics stößt man auf Modelle, die baumartige Taxonomien (Klassifikationen) visualisieren, um nachverfolgen zu können, welche Trainingsdaten zum endgültigen Ergebnis geführt haben.<sup>12</sup> Andere Methoden zeigen dem Anwender die Merkmale eines Extraktionsergebnisses auf, die zur Klassifizierung (Zuordnung einer Vertragsklausel zu einer Policy) beigetragen haben. Hierbei werden den Merkmalen Gewichtungen zugeordnet, welche dabei helfen sollen eine bessere Nachvollziehbarkeit des Ergebnisses zu schaffen bzw. zu verbessern.<sup>13</sup>

---

<sup>9</sup> <https://www.hhi.fraunhofer.de/presse-medien/nachrichten/2019/neues-denken-ausserhalb-der-black-box-erklaerbarkeit-und-vertrauenswuerdigkeit-von-kuenstlicher-intelligenz.html> (abgerufen am 26.11.2020)

<sup>10</sup> Aus Gründen der Lesbarkeit wird bei Personenbezeichnungen die männliche Form gewählt, es ist jedoch immer die weibliche Form mitgemeint.

<sup>11</sup> Dieses Problem der schwierigen Nachvollziehbarkeit von KI-generierten Extraktionsergebnissen wird als die „KI-Blackbox“ beschrieben [https://www.kdnuggets.com/2019/03/ai-black-box-explanation-problem.html]. Der aktuelle Trend geht in die Richtung Explainable Artificial Intelligence (XAI), bei der es dem Anwender möglich sein soll, Extraktionsergebnisse der Software nachvollziehen zu können.

<sup>12</sup> Oduor, Qian, Li, Popa, XAIT: An Interactive Website for Explainable AI for Text, 2020, S.120 ff.

<sup>13</sup> Chhatwal, Gronvall, Huber-Fliflet, Explainable Text Classification in Legal Document Review, 2019

## 2. Semantic Analytics

Nähere Betrachtung verdient auch die semantische Analyse von Klauseln, mittels derer eindeutige und nachvollziehbare Extraktionsergebnisse erzielt werden können und sich auch dazu eignet die Anzahl der notwendigen Trainingsbeispiele zu reduzieren. Dabei kommt es, anders als bei Machine Learning, noch stärker auf die menschliche Komponente im Erstellungsprozess der Policy an.

7

Zum Erstellen einer Policy ist hier eine sprachwissenschaftliche Betrachtung der jeweils zu identifizierenden Klausel notwendig. Auf Grundlage dieser Betrachtung werden Wörter, Schriftzeichen oder ganzen Sätze identifiziert und deren Bedeutungszusammenhänge erarbeitet.<sup>14</sup> Ergebnis dieser Übung ist ebenfalls eine Policy, die genau das auffindet, was bei der Suche vorgegeben wurde. So wäre etwa eine Policy denkbar, die sämtliche Formen des Wortes „Haftung“ („haften“, „haftet“ etc.) enthält, um eine Haftungsklausel im Vertragstext zu identifizieren. Vorteil dieser Methode ist die recht eindeutige Nachvollziehbarkeit der Extraktionsergebnisse. Die Nutzung dieser Ergebnisse, beispielsweise zur inhaltlichen Beurteilung eines neuen Vertrages, kann weit über die reine Erkennung von Tendenzen hinaus gehen. Insbesondere bei komplexen Klauseln, bei welchen ein eindeutiges Extraktionsergebnis erzielt werden soll, aber im Scoping bereits unterschiedliche Meinungen bestehen, bietet sich diese Herangehensweise an. Zudem können mit dieser semantischen Analyse die Machine Learning Policies angereichert und verfeinert werden. Zu beachten ist in diesem Kontext der zeitliche Aufwand zur Erstellung einer solchen Policy. Dieser ist, im Unterschied zu einer rein auf Machine Learning basierenden Policy, bei gleicher Anzahl von Trainingsdaten, höher – wenngleich sich ein Einsatz dieser Methode bereits bei weitaus weniger Trainingsdaten anbietet und daher ein Vergleich schwierig ist. Insbesondere der händische Aufwand beim Erstellen einer auf semantischen Analyse basierenden Policy ist höher, da hier, neben sämtlichen Wortformen, auch Verschreiber und die schlechte (Scan-) Qualität der Trainingsdaten im Erstellungsprozess der Policy beachtet werden muss. Zudem steigt der zeitliche Aufwand in der Regel mit der Größe des Vertragssets, wohingegen eine auf Machine Learning basierende Policy unabhängig von der Größe des Vertragssets eine gleichbleibende Anzahl an Trainingsdaten benötigt.

## III. Fazit

Beim Einsatz von Contract Analytics kommt es stets auf den Use-Case an. Ein eindeutiges Ergebnis, welche Methode sich für den konkreten Anwendungsfall eignet, kann daher nicht pauschal gegeben werden. Es zeigt sich aber, dass bei der Identifizierung von inhaltlich komplexen Klauseln eine auf semantischer Analyse basierende Contract Analytics Komponente häufig das genauere Extraktionsergebnis liefert – wenngleich diese Methode einen höheren Arbeitsaufwand bedeuten kann. Sollen hingegen große und eindeutige Themen analysiert werden, erscheint ein Machine Learning Ansatz als besser geeignet. Bei beiden Ansätzen ist allerdings die menschliche Intelligenz Ausgangspunkt der Entwicklung, sodass die händische Komponente auch weiterhin nicht ersetzt werden wird.

8

---

<sup>14</sup> Fuchsenlueger, Semantische Analyse unstrukturierter Daten, 2016, S.68–77