

FINANZEN UND TECHNIK – GASTBEITRAG

Künstliche Intelligenz bei der Dokumenten- und Datenanalyse immer wichtiger

Börsen-Zeitung, 10.10.2020

Die Menge speicherbarer Unternehmensdaten und -dokumente nimmt täglich rasant zu: Verträge, Finanzdaten, Rechnungen, Formulare, Kunden- und Inventarlisten, interne Anweisungen und Richtlinien sowie Berichte, Schriftsätze und nicht zuletzt eine Flut von E-Mails und Chatprotokollen können heute dank nahezu unbegrenzter Speicherkapazitäten elektronisch archiviert werden. Da diese Daten Aufschluss über den Zustand eines Unternehmens geben, werden sie regelmäßig Objekt der Begierde bei der Vorbereitung von Unternehmenstransaktionen (etwa im Rahmen der Due-Diligence-Prüfung) oder bei Compliance-Untersuchungen.

Die Analyse von Unternehmensdaten kann zudem im Rahmen von straf- oder zivilrechtlichen Verfahren mit US-Bezug eine Rolle spielen, wenn anlässlich eines sog. Discovery-Verfahrens (förmliches Beweisverfahren im US-amerikanischen Prozessrecht) auf gerichtliche Anordnung hin bestimmte Daten als mögliche Beweismittel herausgegeben werden müssen. Bei der Sicherstellung elektronischer Daten spricht man dabei von der sog. E-Discovery.

Der erforderliche Personal- und Zeitaufwand solcher Analysen ist a priori immens. Bei einem E-Discovery-Verfahren beispielsweise werden zunächst Datensätze aller Art anhand von Stichwortsuchen, zeitlichen und thematischen Eingrenzungen oder gefiltert nach Autoren erhoben. Diese im Detail noch „ungefilterten“ Datensätze gilt es dann zur Aufklärung eines Sachverhalts oder Tatverdachts näher zu analysieren. Eine solche Auswertung „von Hand“ (bzw. vor dem Bildschirm) ist nicht nur mit erheblichem Personaleinsatz und finanziellem Aufwand verbunden, sondern birgt – jedenfalls bei großen Unternehmen – wegen der schieren Masse von Informationen auch stets die Gefahr von Inkonsistenzen, Abstimmungsdefiziten und menschlichen Fehlern. Oftmals ist es ohne den Ein-

satz spezieller Technik kaum möglich, derartiger Datenmengen Herr zu werden.

Vor diesem Hintergrund nimmt der Einsatz von künstlicher Intelligenz (KI) im Bereich der Datenanalyse schon seit einiger Zeit stetig zu. KI wird allgemein als automatisiertes Verfahren verstanden, das Software und Maschinen durch komplexe Algorithmen die Fähigkeit verleiht, menschliche Entscheidungsfindungsprozesse zu imitieren. Die hierfür eingesetzten Algorithmen können Datensammlungen nicht nur nach bestimmten Schlagwörtern durchsuchen, sondern auch anhand statistischer Wahrscheinlichkeiten relevante Zusammenhänge aufzeigen sowie Bewertungen und Kategorisierungen vornehmen. Richtig eingesetzt, arbeitet der Algorithmus dabei mit einer Genauigkeit, die einer manuellen Auswertung regelmäßig überlegen ist, bei deutlicher Zeit- und Kostensparnis.

Zunehmend wird dabei auch sog. Machine Learning oder Predictive Coding eingesetzt, bei dem die Software zunächst durch den menschlichen Anwender „trainiert“ wird und in der Folge selbstständig Regeln für die weitere Suche entwickelt. Der Algorithmus wird dafür zunächst mit bestimmten Suchbefehlen programmiert. Danach bearbeitet der Anwender einen zufällig ausgewählten Beispiel-Datensatz anhand derselben Suchparameter. Der eingesetzte Algorithmus entwickelt in der Folge basierend auf dieser Auswertung eigene Regeln, die er zusätzlich zu den bereits programmierten Suchanweisungen bei der weiteren Analyse anwendet und im Verlaufe der Auswertung weiterentwickelt.

KI ist kein „Roboter-Anwalt“

Der „Meister“ tut jedoch gut daran, seinen „Zauberlehrling“ laufend zu überwachen und zu korrigieren. Der Algorithmus „lernt“ zwar aus den Korrekturen, so dass seine Treffsicher-

heit kontinuierlich zunimmt. KI kann aber nur so gut sein wie ihr menschlicher Anwender: Fehler bei der Anleitung bzw. Überwachung werden nicht automatisch korrigiert, sondern fortwährend repliziert – eine KI ist eben kein „Roboter-Anwalt“. Auch den komplexesten Algorithmen fehlt ein tatsächliches, originäres (juristisches) Verständnis des Inhalts der analysierten Daten. Die Software kann nicht zwischen den sprichwörtlichen Zeilen lesen. Soll etwa im Rahmen einer internen Untersuchung ein Bestechungsverdacht verfolgt werden, ist z. B. bei der Analyse von E-Mails eine menschliche Mitarbeit immer noch unerlässlich: etwa verwendete Codewörter oder Umschreibungen müssen interpretiert und juristisch bewertet werden, bevor Algorithmen mit dieser Bewertung gespeist werden können. Zudem stößt KI dort an ihre Grenzen, wo Dokumente nicht oder nur ungenügend digitalisierbar sind. Insbesondere handschriftliche Notizen oder alte, schlecht gescannte Dokumente widerersetzen sich hartnäckig elektronischen Suchläufen. Auch die Auswertung von Sprachaufzeichnungen bleibt vorerst Sache der Experten.

Zukünftig werden umfassende Daten- und Dokumentenanalysen – insbesondere in Anbetracht der durch fortschreitende Digitalisierung stetig wachsenden Datenmengen – wohl nicht mehr ohne den Einsatz von KI zu leisten sein. Dabei kann jedoch (nach dem gegenwärtigen Stand der Technik) auf eine menschliche Anleitung und Überwachung keinesfalls verzichtet werden. In der Praxis birgt diese Entwicklung schon jetzt deutliche Kostenvorteile bei gleichzeitiger Steigerung der Analysequalität.

Christina Heil, Rechtsanwältin bei Debevoise & Plimpton LLP in Frankfurt und Jan Schoberwalter, Rechtsanwalt bei Debevoise & Plimpton LLP in Frankfurt