

Resümee der Beitragsreihe „Input Control“

Beitragsreihe „Input Control – Datenqualität und Datenvalidität als Grundlage rechtlicher Automatisierungsprozesse“,
Abschnitt „Schluss“

Prof. Dr. Thomas Riehm | Lehrstuhl für deutsches und europäisches Privatrecht,
Zivilverfahrensrecht und Rechtstheorie | Universität Passau

Prof. Dr. Kai von Lewinski | Lehrstuhl für Öffentliches Recht, Medien- und
Informationsrecht | Universität Passau

14. Mai 2020

LR 2020, Seiten 140 bis 144 (insgesamt 5 Seiten)

Aus rechtlicher Perspektive stellt sich bezüglich von Daten nicht nur die Frage der Zuordnung (Immaterialgüterrecht, Datenschutzrecht) und des Zugangs (Informationszugangsrecht, Informierungs- und Auskunftsansprüche und -pflichten), sondern auch die Frage nach deren inhaltlicher Qualität. Diese ist mit den Mitteln der überkommenen (Sachmängel-)Gewährleistung nicht bzw. nur unzureichend zu beantworten, weil es jeweils auf den Kontext ankommt, der zudem dynamisch sein kann, und weil sich Probleme mit der Datenqualität häufig außerhalb bipolarer vertraglicher Beziehungen auswirken.

1

I. Der analytische Wert des kybernetischen Systemmodells

Die Dreiteilung des kybernetischen Systemmodells, das auf der linearen Verarbeitungsreihenfolge „Input-Throughput-Output“ beruht, kann auch für die rechtliche Analyse herangezogen werden. Hier kann analytisch zwischen den drei Phasen unterschieden werden, in denen Daten jeweils in einem unterschiedlichen Kontext stehen.

2

- Daten stehen im Ausgangspunkt in einem **ursprünglichen Speicherkontext**. Wetterdaten werden von Meteorologen nach meteorologischen Kriterien zur Verfügung gestellt, Maschinendaten geben die Ingenieursperspektive auf einen Apparat wieder.
- Bei der (automatisierten, also algorithmenbasierten) Verarbeitung dieser Daten ändert sich deren Kontext. In einem Legal-Tech-Kontext bedeutet dies, dass er in den **Kontext der Beurteilung einer Rechtsfrage** gestellt wird. Die Schwierigkeit liegt hier, wie jeder praktisch tätige Jurist und schon jeder Referendar bestätigen kann, weniger in der Komplexität der Norm als in der Beurteilung der Tatsachen. Regelmäßig ist die Kontextänderung auch mit einer Umformatierung von Kategorien verbunden. Was landläufig als Nieselwetter gilt, ist in Hamburg

Sommer, was für die Deutsche Bahn noch pünktlich ist (und in der Bahn-App als Ankunftszeit in Grün dargestellt wird), reicht auf dem Würzburger Hauptbahnhof u.U. trotzdem nicht mehr zum Umsteigen.

- Wenn ein algorithmisches Legal-Tech-System dann zu einem Ergebnis kommt, steht es in einem **Entscheidungskontext**. Noch und absehbar sind Legal-Tech-Systeme nur entscheidungsunterstützend, schon weil sie keine Rechtspersönlichkeit haben, die Ausgangs- und Bezugspunkt einer Entscheidung sein könnte. Das Ergebnis wird also von einer Person übernommen, manchmal als ein Entscheidungsfaktor von mehreren, manchmal als alleinige Entscheidungsgrundlage, manchmal mit Plausibilitätsprüfung, oft auch ohne.

Da Daten jeweils (nur) im Kontext zu Informationen werden, und umgekehrt Informationen durch Speicherung als Daten ihren ursprünglichen Kontext verlieren, wird man zwischen den Mechanismen und Instrumenten zur Sicherung von Qualität und Validität von Daten diese drei Ebenen unterscheiden können. Dabei gibt es – hierauf hat *Matthias Rossi* eingangs seines Beitrags hingewiesen¹ – durchaus eine Wechselbeziehung zwischen der Qualität der Daten auf den unterschiedlichen Ebenen und deren Sicherung.

II. Input

Die rechtlichen Vorgaben für Daten, die Input in Legal-Tech-Systeme darstellen können, sind jeweils spezifisch und meist auch situativ²: Das Datenschutzrecht denkt bei der Richtigkeit an die Richtigkeit in Bezug auf eine Person (Art. 5 Abs. 1 lit. d DS-GVO), beim Betrug (§ 263 StGB) kommt es auf die Irrtumserheblichkeit der falschen Tatsache an, innerhalb des Vertragsrechts (bei Datenlieferungsverträgen bzw. Verträgen mit einer Datenlieferungskomponente) ist der Kontext Vertragsgegenstand oder jedenfalls Vertragsgrundlage³. Denkbar sind auch staatliche Standardsetzungen, die es bislang aber bestenfalls punktuell gibt⁴, etwa in der Geodäsie oder dem Bilanzrecht.

Insgesamt gibt es bislang nur wenige Vorgaben für ein *Input control*. So ist die Gewährleistung einer bestimmten Datenqualität im Informationszugangsrecht allgemein sogar ausgeschlossen⁵. In Sonderfällen kann sogar auch eine falsche Information rechtlich gewünscht oder jedenfalls akzeptiert sein, etwa bei dem arbeitsrechtlichen „Recht auf Lüge“ auf die Frage nach einer Schwangerschaft sowie beim *Salting* in der Kryptologie.

¹ *Rossi, Informationsgewährungspflichten und -möglichkeiten des Staates –Niveau und Sicherung der Qualität staatlicher Daten, LR 2020, 90.*

² *Hennemann, Datenrichtigkeit, LR 2020, 77 (82).*

³ *Fries, Trau, schau, wem? Qualitätssicherung durch Datenverwender, LR 2020, 87 (88).*

⁴ *Rossi, Informationsgewährungspflichten und -möglichkeiten des Staates –Niveau und Sicherung der Qualität staatlicher Daten, LR 2020, 90 (96).*

⁵ *Rossi, Informationsgewährungspflichten und -möglichkeiten des Staates –Niveau und Sicherung der Qualität staatlicher Daten, LR 2020, 90 (92).*

Auf die Herkunft und den situativen Kontext von Daten sowie die nur situativ und nur in bestimmten Kontexten geltenden Vorgaben zur Sicherstellung der Qualität und Validität von Daten muss beim Design von Legal-Tech-Systemen und deren Konstruktion geachtet werden. Angaben können datenschutzrechtlich richtig (und in ihrer Unschärfe sogar „datensparsam“) sein – etwa die Altersangabe: „zwischen 15 und 19 Jahren“ –, für die Altersverifikation beim Erwerb von Alkohol und im Kontext mit Erwachsenenunterhaltung hingegen ungeeignet und deshalb nicht von hinreichender Qualität.

6

Schwierig wird es, wenn Daten aus unterschiedlichen Quellen und (damit) unterschiedlichen Kontexten zusammengeführt werden. Bei solchen Aggregationen geht notwendigerweise Kontext verloren, beim Einsatz zum „Anlernen“ neuronaler Systeme erst recht⁶. Vor diesem Hintergrund sind rechtliche Anforderungen an die Qualität und Validität von Input-Daten für automatisierte oder gar autonome Entscheidungs(unterstützungs)systeme zu definieren.

7

III. Throughput

Bei der Verarbeitung von Daten aus verschiedenen Quellen und aus unterschiedlichen Kontexten müssen diese Umstände als Randbedingungen berücksichtigt werden. Denn eine passende Streitentscheidung setzt nicht nur das Vorhandensein der notwendigen Informationen voraus, sondern auch deren Passung, also den Kontext.

8

In „kontrollierten Umgebungen“ wie Streitbelegungsverfahren und Legal Tech, die auf eine bestimmte personale und/oder thematische Konfliktsituation zugeschnitten sind, wie sie etwa in Bezug auf Ansprüche wegen gestörten Reisen⁷ oder auf Online-Handelsplattformen⁸ bestehen, ist dies ohne großen Aufwand möglich. Nicht ganz zufällig sind dies auch die Bereiche, in denen sich Automatisierung von Streitbeilegung für die Praxis zuerst etabliert hat.

9

Damit dies über mehrere Verfahrensschritte hinweg und durchgehend dokumentiert geschieht, wird die Distributed-Ledger-Technologie (DLT) eingesetzt⁹. DLT liegt der Blockchain und insb. Kryptowährungen zugrunde. Während in der allgemeinen Wahrnehmung (und unterstelltermaßen unter IT-Beratern und -Vermarktern) die „Blockchain“ als Allheilmittel angepriesen wird, markiert sie für Legal Tech die Möglichkeit (und dann wohl auch die Pflicht), Herkunft und Kontext von Daten zu dokumentieren (zu können). Im Datenschutzrecht (Art. 5 Abs. 2 DS-GVO) heißt dies „Rechenschaftspflichtigkeit“ (*accountability*), in Verträgen über (große) Software-Projekte

⁶ Fries, [Trau, schau, wem? Qualitätssicherung durch Datenverwender](#), LR 2020, 87 (89).

⁷ Quarch, [Die Bedeutung von Daten für die Geltendmachung von Verbraucherrechten im LegalTech-Zeitalter](#), LR 2020, 111.

⁸ Deichsel, [Sachverhaltsfeststellung in der automatisierten Konfliktbeilegung](#), LR 2020, 98.

⁹ Urbach/Völter, [Grundlagen und Potenziale der Distributed Ledger Technology](#), LR 2020, 119.

ist dies eine Selbstverständlichkeit. Und in Daten(lieferungs)verträgen soll und wird es auch zum Standard gehören.

IV. Output

Das Ergebnis von Legal-Tech-Verfahren sind nicht verarbeitete Daten, sondern die Beurteilung eines Sachverhalts. Bislang handelt es sich – anders als bei herkömmlichen Gerichtsverfahren – nicht um beliebige Lebenssachverhalte, sondern typische Konstellationen, eben Konstellationen, für die diese Verfahren entwickelt und eingesetzt werden; im Vordergrund stehen hier immer noch (nur) Verspätungsentschädigungen¹⁰ und der Online-Handel¹¹ sowie die massenhafte Geltendmachung von gleichartigen Ansprüchen (etwa im „Abgasskandal“ oder aufgrund des „Mietendeckels“). Andere, einfach zu strukturierende Massengeschäfte werden folgen. 10

Das bedeutet aber nicht, dass Legal Tech für andere Aspekte des Rechtslebens und der Konfliktlösung nicht in Frage käme. Hier wird der Output solcher Systeme allerdings nicht in der umfänglichen Lösung des Falls liegen, sondern nur hinsichtlich derjenigen Sachverhaltselemente, die sich dank der gesicherten Qualität und Validität¹² und überschaubarer Entscheidungsarchitekturen¹³ eignen. Dass und wie sich verschiedene Elemente zusammenschalten lassen und welche Verschaltungsfragen dies für das dahinterstehende Rechts- und Aufsichtsregime bedeutet¹⁴, wird dann bald offenbar. 11

Womöglich kann bei der Diskussion von Legal Tech unter Rückgriff auf das eingangs erwähnte kybernetische Systemmodell noch stärker zwischen Output und Outcome unterschieden werden. Es ist nicht notwendigerweise der Anspruch von und an Legal Tech Systeme(n), einen Rechtsfall komplett zu entscheiden. Sondern sie nehmen dem menschlichen, richterlichen oder schiedsrichterlichen Entscheider dort Arbeit ab, wo Computer besser sind als Menschen, wenn es nämlich um die (Ein-)Ordnung von vergleichbaren Datenbeständen in einem einheitlichen und (dadurch) kontrollierbaren Kontext geht. 12

V. Outcome

Wenn die analytische Trennung zwischen Output und Outcome das Ergebnis der kleinen Blogbeitragreihe anlässlich der ausgefallenen Tagung „Input Control – Datenqualität und Datenvalidität als Grundlage rechtlicher Automatisierungsprozesse“ wäre, dann wäre das 13

¹⁰ [Quarch, Die Bedeutung von Daten für die Geltendmachung von Verbraucherrechten im LegalTech-Zeitalter, LR 2020, 111.](#)

¹¹ [Deichsel, Sachverhaltsfeststellung in der automatisierten Konfliktbeilegung, LR 2020, 98.](#)

¹² [Riehm/von Lewinski, Resümee der Beitragsreihe „Input Control“, LR 2020, 140 \(141\).](#)

¹³ [Riehm/von Lewinski, Resümee der Beitragsreihe „Input Control“, LR 2020, 140 \(142\).](#)

¹⁴ [Voß, Blockchain-Lösungen für die Modeindustrie und den Kapitalmarkt – die LUKSO-Entscheidung\(en\) der BaFin als Beitrag zur Herstellung von Rechtsinteroperabilität, LR 2020, 130.](#)

nicht nur – und auch hier ist diese Unterscheidung sinnvoll – fachlicher Output, sondern auch ein positiver Erkenntnis-Outcome. In diesem Sinne hoffen wir, Sie, liebe Leserinnen und Leser, durch die Lektüre nicht nur mit Daten, sondern auch mit Informationen inspiriert zu haben, die Sie in Ihren Kontext übertragen und dort weiterverwenden können.