

Rezension der Monographie „Künstliche Intelligenz und Strafrecht“

Mit der im Verlag Österreich publizierten Dissertation ist der Autorin Kerstin Waxnegger die umfassende Darstellung eines Bereichs gelungen, der im übertragenen Sinn „freischwebend“, also weitestgehend unreguliert, aber für die Zukunft der Menschheit entscheidend ist. Mit dem Einsatz der KI stehen wir an der Schwelle einer Disruption, die mit bisherigen politischen Revolutionen nicht vergleichbar ist. Die Autorin erhebt den Anspruch einer umfassenden Klärung vieler Fragen: Wieweit ist das Strafrecht auf den Einsatz von KI vorbereitet? Welche Implikationen sind bei der Feststellung der strafrechtlichen Verantwortlichkeit zu beachten? Welche Rahmenbedingungen sind für einen erfolgreichen Einsatz von KI im Strafrecht erforderlich? Sie kann dabei allerdings nur scheitern, durchaus ehrenhaft, denn sie bewegt sich in einem Umfeld, in dem wichtige Eckpunkte und Prämissen nicht feststehen oder sich gerade erst herauszubilden beginnen. Umso verdienstvoller sind diese erste Orientierung und der Versuch, mögliche Antworten zu skizzieren und Szenarien zu entwerfen.

Die Problematik beginnt schon bei der letztlich nicht fassbaren Definition der KI. Historisch betrachtet beschäftigte das schon Descartes im 17. Jahrhundert und in der Folge setzte sich sogar Kant ansatzweise damit auseinander, wenn er schreibt, dass sich eine Person ihrer Identität nicht nur im Dasein bewusst sein, sondern diese auch immer wieder neu gewinnen muss. Locke führte den Begriff „Intelligenter Agent“ ein, der Vorstellungen über die Vergangenheit und Zukunft haben sollte.

Juristische Personen und eingetragene Personengesellschaften als Gesamtgesellschaften erfüllen diese Kriterien nicht, dennoch sind sie Rechtspersonen. Die Autorin setzt sich damit vor dem Hintergrund auseinander, dass eine Erweiterung auf das Rechtssubjekt „KI“ ein philosophisches und ethisches Problem darstellt. Wesensmerkmale von KI sind deren Lernfähigkeit und – nach derzeitigem Stand – das Arbeiten mit Wahrscheinlichkeitsabwägungen. Mit KI zu arbeiten bedeutet nach all dem, was wir heute wissen, erhebliche Risiken, wie Kontrollverlust durch den Menschen, einprogrammierte Voreingenommenheit oder die Unvermeidbarkeit von Fehlern. Jedenfalls kann KI nicht bestraft werden, so die Autorin, solange die dahinter stehenden Menschen für deren Programmierung und damit für das jeweilige Ergebnis unmittelbar verantwortlich sind.

Derzeit ist wohl jede Überlegung zur KI schon beim Druck überholt und mit diesem Handicap starte ich dennoch den Versuch, das Werk von Waxnegger zu reflektieren und zu beurteilen. Offen liegen vor uns die Problemfelder der Menschenrechte und der grundrechtlichen Absicherung, die vielfältig betroffen sind, im Datenschutz und im Privatleben, im Recht auf ein faires Verfahren, in der Verhältnismäßigkeit uvm.

In laienhafter Kenntnis grundsätzlicher technischer Vorgänge bin ich bislang davon ausgegangen, dass der Einsatz (auch) der KI unter dem anerkannten Stand der Forschung und Entwicklung folgt, dass also höchstwertige Technik verwendet und mit größtmöglicher Transparenz und unter penibler Supervision gearbeitet wird. Dass dafür noch kaum Rechtsgrundlagen und Rahmenbedingungen vorhanden sind, ist eine Tatsache, an der auch der kürzlich verabschiedete AI-Act der EU noch nicht viel ändern dürfte, wurde er doch „vor Chat-GPT“, also noch ohne Berücksichtigung der Änderungen konzipiert, die sich durch den Einsatz der sogenannten Large Language Models ergeben. Damit bahnt sich aber eine rasante und medial viel diskutierte Weiterentwicklung an, die dazu führt, dass unsere bisherigen IT-Systeme rasch abgelöst und früher als angenommen überholt werden.

In Bereichen, die die Zivilisationsgeschichte mit gutem Grund dem menschlichen Geist vorbehält, also auch im Strafrecht, ist zu konstatieren, dass der KI jedwede Empathie, Intuition und Sozialkompetenz fehlt, denn all dies ist einem Algorithmus und damit der Programmierung unzugänglich, das dürfte sich auch in Zukunft nicht so rasch ändern.

Die öffentliche Verwaltung in Österreich beginnt sich mit dem KI-Einsatz langsam auseinanderzusetzen. So hat z.B. das BMKÖS Leitlinien für einen ethischen KI-Einsatz herausgegeben, der viele der Risiken und Gefahren programmatisch aufzeigt. Für den Einsatz in der Justiz, der ja in Assistenzsystemen (wie der Anonymisierung von Entscheidung und der Verknüpfung mit Rechtstexten) schon erfolgt, wird aber eine umfassende strategische Planung und Abstimmung mit allen beteiligten Akteuren erforderlich sein.

Wie Mayerhofer im Gutachten zum 21. Österreichischen Juristentag nachzeichnet, begrenzt den KI-Einsatz in der Verwaltung nur der Gleichheitssatz, die Aufgaben der Gerichtsbarkeit behält aber das B-VG natürlichen Personen vor. In der Justiz stößt die „Maschinengerichtsbarkeit“ auf Grenzen, denn der Grundsatz der Menschenwürde schließt eine Herrschaft der KI über Menschen und damit unmittelbare Entscheidungen aus. Vielfältig und weitgehend noch nicht zu Ende gedacht sind in der Justiz die Möglichkeiten einer Unterstützung der Entscheidungsgorgane durch die KI, also nicht im Kernbereich der richterlichen Tätigkeit, dem Verhandeln und Entscheiden. Soweit nur Randbereiche betroffen sind, wie bei der Entscheidungsvorbereitung, ist der Einsatz auch nach dem Gutachten der deutschen Datenethikkommission denkbar.

Wohl erst nach Drucklegung der Arbeit ist mit dem Schufa-Urteil eine fundamentale Entscheidung des EuGH zu Art. 22 Abs. 1 DSGVO ergangen. Damit wurde der Anwendungsbereich dieser Bestimmung und damit auch die datenschutzrechtliche Regulierung der KI erheblich erweitert. Die DSGVO gewährt dabei nicht nur eine Produktregulierung, sondern sieht auch ein unmittelbares Beschwerderecht vor.

Für mich ist plausibel, dass der Algorithmus einer KI wesentlich mehr Daten speichern und verarbeiten kann als der Mensch. Zur „geistigen Kompetenz“ einer KI gehört es jedenfalls, abstrakt zu denken, komplexe Ideen zu verstehen und so Probleme zu lösen. Eine menschliche Interaktion erfordert dafür aber Emotionalität und Esprit. Auch wenn die EU-KI-VO nun Kategorien für KI-Anwendungen vorgibt, also risikolos und erlaubt, nur unter strengen Voraussetzungen zulässig oder sonst zu riskant und daher verboten, bleibt der Anbieter einer KI-Anwendung als Betreiber der nach der DSGVO Verantwortliche. Für den KI-Einsatz im Strafrecht sind Prognoseinstrumentarien naheliegend, bei denen aber erhebliche Diskriminierungstendenzen indiziert sind. Damit verbunden ist die Gefahr einer zu geringen Transparenz sowie einer kaum gegebenen Nachvollziehbarkeit (black box). All dies hätte erfordert, dass von Anfang an klare Regeln vorgesehen werden, das hat sich auf europäischer Ebene aber verzögert, meines Erachtens uneinholbar.

Die Autorin möchte durch ihre Monographie klären, ob das Strafrecht auf den Einsatz von KI vorbereitet ist, welche Implikationen bei der Zuweisung von Verantwortung bestehen und inwiefern vorhandene Regelungen angepasst werden sollten, um KI strafrechtlich adäquat erfassen zu können. Sie hält fest, dass eine abschließende Definition von KI auch mit ihrer Arbeit nicht gefunden werden kann. Zu den wesentlichen Eigenschaften von KI zählt sie deren Selbstveränderungsfähigkeit, ihre Lernfähigkeit sowie die Fähigkeit, ein Ziel ohne menschlichen Input zu erreichen.

KI wird dort sehr effizient sein, wo es dem Computer allein nicht mehr gelingt, mit einer Masse an Daten sinnvoll umzugehen, wenn Daten also schon auf Grund des Umfangs nicht direkt mit einem Programmcode erfasst werden können. KI wird so auch als Abwehrmittel verwendet werden können, indem sie exemplarisch zur Bewältigung von Cyber-Attacken eingesetzt wird. Damit stellt sich aber auch die Frage nach der Verantwortlichkeit der KI selbst. Dazu wird die aktuelle Definition von Vorsatz und Fahrlässigkeit neu formuliert werden müssen, denn KI selbst ist weder diskretions- noch dispositionsfähig. Die Autorin schlägt dazu vor, einen Sonderstatus zu schaffen. Das Kernproblem dabei sei, dass die KI Rechtsgüter verletzen kann, für die sie verantwortlich sein müsste, die Verletzung aber durch eine Ausführung geschieht, die nicht dem Willen des Programmierers oder Verwenders entspricht. Im Ergebnis sei aber eine „neue“ Strafbarkeit von KI abzulehnen. Die Verfasserin analysiert dazu auch die Verantwortlichkeit des Verwenders und schlägt als Differenzierung vor, ob das Verhalten in einem aktiven Tun oder einem Unterlassen liegt. Zudem sei im Einzelfall die Garantenstellung des Verwenders zu klären und die Unterscheidung zwischen vorsätzlichem und fahrlässigem Handeln zu beachten. Mittelpunkt jeder Fahrlässigkeit ist dabei immer die objektive Sorgfaltswidrigkeit; Verstöße gegen Normen und ein Abweichen vom Verhalten einer Maßfigur indizieren sodann sozial inadäquates Verhalten.

Sehr problematisch sei die Lösung der Frage, ob, ab wann und bis zu welchem Grad überhaupt ein Vertrauen auf eine KI gegeben sein kann, wobei die Autorin mit wohlüberlegten Argumenten eine Ausweitung des Vertrauensgrundsatzes auf die KI ablehnt. Ob beim Einsatz einer KI eine Sozialadäquanz gegeben ist, sich also ein erlaubtes Risiko verwirklicht, wird nach dem Inkrafttreten der KI-VO der EU möglicherweise neu zu beurteilen sein.

Ich gehe mit der Autorin konform, dass KI grundsätzlich immer vom Menschen zu kontrollieren sein wird, was insbesondere im medizinischen Bereich relevant ist, wenn zu beurteilen ist, was lege artis ist und ob Compliance-Systeme eingehalten wurden. Da ein Regressverbot in Österreich nicht existiert, ist mit Erwägungen über eine fehlende Kausalität nicht viel anzufangen. Die Verfasserin lässt offen, ob die Risikoverminderungstheorie bei der KI überhaupt zum Tragen kommt, und meint, dass diese Theorie grundsätzlich zur Diskussion stellen sein wird.

Zur Haftung des Herstellers merkt die Verfasserin an, dass sich die Bedeutung der Produktverantwortlichkeit in Zukunft erhöhen wird. Der Hersteller hat jedenfalls ein fahrlässiges Verhalten zu verantworten, wenn er die Gefahren einer KI nicht sorgfältig beurteilt und das Produkt zum Verkauf freigibt. Allerdings führt sie in der Folge aus, dass es nicht fahrlässig sein wird, wenn die Gesellschaft die Vermarktung als sozial adäquat beurteilt. Dazu werden auch gegen Ende des Werkes noch konzis die Verantwortlichkeiten weiterer Personen und die Behandlung von Algorithmen im Zielkonflikt behandelt.

Die Autorin stellt abschließend fest, dass die KI ein Produkt von Menschen für den Menschen ist, das black-box-Problem durch die Berücksichtigung möglichst zuverlässiger und passender Daten eingegrenzt werden kann und die mit dem Einsatz verbundenen erheblichen Risiken umfassende und genau definierte Kontrollen erfordern werden. Ihre Schlussüberlegungen, inwieweit das Schuldprinzip zu überdenken sein wird, scheinen mir angemessen, auch wenn hier wohl noch für längere Zeit auf die (strafrechtliche) Verantwortlichkeit des „menschlichen Täters“ hinter einer „unmittelbar agierenden“ KI abzustellen sein wird.

Schon seit einiger Zeit ist die KI im strafrechtlichen Bereich im Einsatz, u.a. bei den Datenstrukturierungstools und – im Ergebnis noch sehr fragwürdig – bei der Kriminalitätsprognoseerstellung. Im Ermittlungsverfahren ist ihr

Einsatz insbesondere zur Aufklärung im Bereich der Wirtschaftskriminalität und vor allem bei missbrauchsbezogenen Sexualdarstellungen Minderjähriger wertvoll.

Welche Strategie wird die Justiz bezüglich des KI-Einsatzes verfolgen? Wie wird der Grundsatz der Waffengleichheit gewahrt, wenn Verteidigung und Verfolgungsbehörde nicht über denselben Wissensstand verfügen? Nicht zuletzt, inwieweit ist der fundamentale Grundsatz der Unmittelbarkeit gewährleistet, wenn Algorithmen Entscheidungen zwar nicht treffen, aber dennoch weitestgehend vorbereiten? All dies sind Fragen, die in naher Zukunft geklärt werden müssen, wie die Autorin überzeugend darlegt. Ihr ist auch zuzustimmen, wenn sie große Problemfelder in der Beweisqualität, der Nachvollziehbarkeit von Prognosen und der Legalität von Ermittlungsmaßnahmen sieht.

Abschließend stelle ich mir, aber vielmehr den Verantwortlichen die Frage, in welcher Gesellschaft wir in Zukunft leben wollen und wieviel uns Sicherheit in Abwägung zur Freiheit wert sein soll. Für diesen Diskurs ist das Werk der Autorin zum jetzigen Zeitpunkt wertvoll; schon allein der rasche Fortschritt in diesem Bereich, aber auch eine in Teilen noch mögliche größere Schlüssigkeit ihrer Argumente werden aber mit großer Wahrscheinlichkeit eine Grundlage für eine alsbaldige Neuauflage ihres (dann überarbeiteten) Werkes bilden.

Nikolaus Lehner