Schwerpunkt Digitalisierung und Recht

Für diese Ausgabe haben Niklas Braun und Johannes Wichert mit Prof. Dr. Brian Valerius, Inhaber des Lehrstuhls für Künstliche Intelligenz im Strafrecht an der Universität Passau, gesprochen.

Interview zum Thema KI im Strafrecht – Chance für Justiz oder Kriminelle?

Prof. Dr. Brian Valerius folgte nach seinem Studium, Promotion und Habilitation an der Universität Würzburg und Lehrstuhlvertretungen in Frankfurt/Oder, Frankfurt/Main und Köln einem Ruf an die Universität Bayreuth und hielt von 2011 bis 2022 den Lehrstuhl für Strafrecht, Strafprozessrecht und Medizinstrafrecht. Seit Oktober 2022 ist er Inhaber des Lehrstuhls für Künstliche Intelligenz im Strafrecht an der Universität Passau.

BayZR: Welche Einsatzmöglichkeiten bieten Künstliche Intelligenzen (kurz: KI) im Strafverfahren – aktuell und in der Zukunft?

Prof. Dr. Valerius: KI spielt nicht nur in unserem privaten Alltag eine immer wichtigere Rolle. Auch im juristischen Bereich mehren sich die Einsatzbereiche. In Strafverfahren werden KI-Systeme etwa genutzt, um große Datenmengen (Big Data) wie E-Mails oder Chatverläufe zu durchsuchen und zu analysieren. Des Weiteren können so automatisiert Bild- und Videoaufnahmen gesichtet werden, um z. B. kinder- und jugendpornographische Inhalte zu identifizieren. Dies ist schon deswegen eine wertvolle Unterstützung, weil es die Ermittlungsbeamtinnen und -beamten psychisch sehr belastet, solche Aufnahmen selbst auszuwerten.

Zu welchen Zwecken KI künftig eingesetzt werden kann, lässt sich noch nicht absehen. Schon seit längerem wird diskutiert, ob und gegebenenfalls wie KI bei der Entscheidungsfindung (im Strafverfahren z. B. bei der Strafzumessung) unterstützend herangezogen werden kann. Hier gibt es aber noch viele, nicht zuletzt (verfassungs-)rechtliche Hürden. Realistischer dürfte es sein, dass Analyseplattformen Datenbanken verknüpfen und nach Mustern durchsuchen, um Straftäter zu ermitteln oder Zusammenhänge zwischen Straftaten herzustellen. Im präventiven Bereich werden solche Systeme – auch in Deutschland – bereits eingesetzt, um schwere Straftaten wie etwa terroristische Anschläge oder Kindesmissbrauch zu verhindern.

Etwas weniger spektakulär wären Anwendungen, die – sollte es in Deutschland jemals zur audiovisuellen Dokumentation oder auch nur zu Tonaufzeichnung von strafgerichtlichen Hauptverhandlungen kommen – die mündlichen Äußerungen transkribieren. Auch hierfür könnte KI eingesetzt werden.

BayZR: Nicht immer begnügt sich die KI mit einer authentischen Darstellung der Realität. KI-generierte audiovisuelle Inhalte, sogenannte Deepfakes, stellen die Gerichte derzeit vor große Herausforderungen. Wie ist mit solchen KI-generierten, real wirkenden aber tatsächlich falschen Fotos, Videos und Audioaufnahmen umzugehen, wenn sie als be- oder entlastendes Beweismittel in den Gerichtsprozess eingebracht werden?

Prof. Dr. Valerius: Beweismittel im Strafverfahren müssen die Gerichte stets auf Authentizität und Glaubhaftigkeit überprüfen. Dies verhält sich bei Audio-, Bild- und Videoaufnahmen nicht anders. Natürlich droht die Gefahr, dass zunehmend manipulierte Dateien als Beweismittel vor Gericht auftauchen oder dass umgekehrt der Beweiswert realer Aufnahmen dadurch untergraben werden soll, dass sie als „Deep Fake“ deklariert werden.

Sollte eine Manipulation im Raum stehen, könnten Sachverständige (sogenannte digitale Forensiker) die Echtheit der Aufnahme überprüfen. Auch spezialisierte KI-Systeme könnten herangezogen werden und auf Unregelmäßigkeiten in Audio-, Bild- oder Videoaufnahmen hinweisen. Und zum Teil hilft auch der klassische Zeugenbeweis, um zu klären, ob etwa eine Videoaufnahme ein reales Geschehen zeigt oder nicht.

All dies setzt ein entsprechendes Gefahrenbewusstsein der Staatsanwaltschaften und Gerichte voraus. Hieran scheint es mitunter noch zu mangeln, wie generell nicht geringe Teile der Justiz nach wie vor für digitale Beweismittel sensibilisiert werden müssten.

BayZR: Welche Rolle spielt die Nutzung von KI bei der Strafverfolgung, etwa zur Gesichtserkennung oder bei einer Kriminalprognose, also bei der Einschätzung darüber, ob eine Person oder Personengruppe künftig gegen ein Strafgesetz verstoßen wird?

Prof. Dr. Valerius: Bei Systemen zur Gesichtserkennung, die etwa Aufnahmen von Überwachungskameras auswerten, gibt es in der Tat schon erste Erfolgsberichte. Für Aufsehen sorgte vor nicht allzu langer Zeit die Festnahme der RAF-Terroristin Daniela

Klette, die von den Behörden 30 Jahre lang gesucht wurde, bevor wohl Journalistinnen und Journalisten mit Hilfe von KI-Gesichtserkennungssoftware entscheidende Hinweise zu deren Aufenthaltsort fanden.

Gleichwohl sind solche Systeme rechtlich wie technisch nicht unbedenklich. So wird bei deren Einsatz nicht unwesentlich in die Privatsphäre der Betroffenen eingegriffen. Zudem weist Gesichtserkennungssoftware vor allem bei Angehörigen von Minderheiten eine nicht unbeachtliche Fehlerquote auf. Ebenso wurde über Prognosesoftware, die anhand verschiedener Kriterien die sogenannte Rückfallwahrscheinlichkeit berechnet, schon wiederholt berichtet, dass bestimmte Bevölkerungsgruppen benachteiligt werden.

Auch aus diesen Gründen hat die 2024 in Kraft getretene KI-Verordnung der Europäischen Union den Rückgriff auf diese Techniken zur Strafverfolgung begrenzt. Biometrische Echtzeit-Fernidentifizierungssysteme sind etwa in öffentlich zugänglichen Räumen nur unter bestimmten Voraussetzungen (z. B. zum Aufspüren von Tatverdächtigen von Terrorismusstrafaten, Mord, Geiselnahme oder Vergewaltigung) zulässig. Noch weiter eingeschränkt wurden Risikobewertungen von KI-Systemen, welche die Wahrscheinlichkeit bewerten oder vorhersagen sollen, dass ein Mensch eine Straftat begeht.

BayZR: Inwiefern kann eine KI selbst an die Stelle von Menschen treten? KI als Täterin? KI als Zeugin? Vielleicht sogar als Richterin?

Prof. Dr. Valerius: Als Täterin kann man sich KI bislang kaum vorstellen. Selbst wenn wir einmal an ein KI-System denken, das sich in einer humanoiden Außenhülle befindet und außerdem durch ein Sprachsystem kommuniziert: Strafrechtliche Verantwortlichkeit setzt jedenfalls in Deutschland Handlungs- und Schuldfähigkeit voraus. Das deutsche Strafrecht ist somit auf Menschen zugeschnitten, die – vereinfacht gesagt – wissen, was und warum sie es tun, und die sich bewusst für oder gegen das Recht entscheiden können. Selbst lernfähige KI ist aber von ihrer grundlegenden Programmierung abhängig und demzufolge weder handlungs- noch schuldfähig. Und ob es jemals eine sogenannte starke KI gibt, die ein eigenes Bewusstsein entwickelt, steht noch in den Sternen und ist auch unter Fachleuten aus der Informatik sehr umstritten. Selbst wenn also eine KI etwa einen Menschen verletzt oder sogar tötet, müssen wir nach menschlichen Akteurinnen und Akteuren Ausschau halten, die wir bestrafen können, sei es etwa den sogenannten Hersteller von KI oder auch die Nutzerin bzw. den Nutzer im konkreten Einzelfall.

Auch als Richterin dürfte KI nicht in Betracht kommen. Hiergegen sprechen schlicht zu viele rechtliche Einwände. Selbst beschränkt auf die angesprochene Strafzumessung ist etwa fraglich, wie ein sogenannter Robo-Judge die Entscheidung „aus dem Inbegriff der Verhandlung“ treffen oder wie der Anspruch der Angeklagten auf rechtliches Gehör gewährleistet werden soll. KI-Systeme dürften daher – jedenfalls im Strafrecht – allenfalls unterstützend bei der Entscheidung herangezogen werden, nicht aber diese Entscheidung selbst treffen.

Am ehesten können wir uns KI noch im Zeugenstand vorstellen. Allerdings ist die Strafprozessordnung auf menschliche Zeuginnen und Zeugen zugeschnitten; für eine solche Verfahrensbeteiligung von KI dürfte es auch an der erforderlichen persönlichen Wahrnehmung fehlen. Die Aufzeichnungen und Bewertungen von KI wären demzufolge als Urkunds- oder Augenscheinbeweis zu behandeln. Über solche Fälle wurde auch schon berichtet. So konnte bei einem Mordfall in Griechenland die Einlassung des Täters mit Aufzeichnungen seines Smartphones und von Daten der Smartwatch des Opfers widerlegt und er zu einem Geständnis bewegt werden. Und in Deutschland gab es ein Strafverfahren, in dem die Aufzeichnungen des Fitness-Trackers des Opfers die Verantwortlichkeit des Angeklagten ausschlossen.

BayZR: Die KI könnte also vor allem auf ihren Input analysiert werden. Fehlerhafter Input veranlasst Künstliche Intelligenzen allerdings auch zur Ausgabe fehlerhaften Outputs. Speist die Künstliche Intelligenz vorurteilsbehaftete Daten ein, bei denen Minderheiten überdurchschnittlich häufig delinquent geworden sind, würde sie in der Folge die Minderheit mit höherer Wahrscheinlichkeit als straffällig betrachten. Gibt es mittlerweile Lösungsmöglichkeiten, um mit diesem Problem umzugehen?

Prof. Dr. Valerius: Sogenannte Bias oder auch nur Stereotype in Trainingsdaten sind nach wie vor ein großes Problem. Trainingsdaten sind diejenigen Daten, aus denen KI-Systeme Muster extrahieren und Zusammenhänge erkennen, um mit deren Hilfe später eingegebene Daten zu beurteilen. Dass die erwähnte Prognosesoftware zu diskriminierenden Ergebnissen gelangte, lag wohl insbesondere an den biasbehafteten Trainingsdaten. Denn auch die beste KI kann stets nur so gut sein, wie dies die Trainingsdaten zulassen. In der Informatik wird insoweit auch von „garbage in, garbage out“ gesprochen.

Das Bewusstsein für dieses Problem wächst aber immerhin. Es werden zunehmend Methoden entwickelt, um z. B. die Trainingsdaten besser zu diversifizieren und auszubalancieren. Bei diesem Ziel können auch Kolleginnen und Kollegen aus Ethik und Recht behilflich sein, die bereits im Entwicklungsprozess für Diskriminierungen sensibilisieren. Nach dem Inverkehrbringen von KI-Systemen kann ein regelmäßiges Monitoring dazu beitragen, problematische Verzerrungen zu erkennen und zu korrigieren. Und ein Schlagwort, das im Zusammenhang mit KI häufig gebraucht wird, kann ebenso zu einer weniger vorurteilsbehafteten KI beitragen, nämlich „Transparenz“. Wenn etwa die Entscheidungsprozesse der KI transparenter im Sinne von nachvollziehbarer sind, werden die Resultate dadurch zwar nicht unmittelbar besser. Dies erleichtert aber ebenfalls die Überprüfung und das Erkennen diskriminierender Tendenzen.

BayZR: Wie setzen Straftäter Künstliche Intelligenzen gegenwärtig zur Begehung von Verbrechen ein? Welche Potentiale und Gefahren bestehen hier in der Zukunft?

Prof. Dr. Valerius: Kriminelle sind häufig unter den Ersten, die technische Neuerungen und Errungenschaften ge- bzw. missbrauchen. Dies ist bei KI leider auch nicht anders. So verwenden Straftäter bereits KI, um neue Schadsoftware zu entwickeln, sogenannte Phishing-Angriffe zu optimieren oder um einfacher Sicherheitslücken in Computersystemen zu identifizieren. KI kann ferner genutzt werden, um betrügerische Transaktionen zu automatisieren oder zu diesen oder anderen kriminellen Zwecken Fake-Profilen in sozialen Medien zu generieren. Und auch die angesprochenen Deep Fakes werden verwendet, um Betrügereien oder Erpressungen zu begehen oder Falschinformationen zu verbreiten.

Diese Beispiele werden sich in Zukunft sicherlich noch mehr. Zu denken ist an die Beeinflussung von Aktienkursen oder an sonstige Manipulationen von Börsen und Finanzmärkten mit Hilfe von KI. Zu befürchten ist aber nicht zuletzt, dass Cyberangriffe weiter zunehmen und zu noch größeren Schäden führen. Ein potentiell Ziel solcher Angriffe könnten insbesondere Einrichtungen der Kritischen Infrastruktur sein. Dazu gehören unter anderem Energie- und Wasserversorgung, Telekommunikationsunternehmen und Krankenhäuser. Alle können sich ausmalen, welche Folgen es haben könnte, wenn solche Einrichtungen etwa infolge eines Hackingangriffs lahmgelegt werden.

BayZR: Die Europäische Union nimmt mit der KI-Verordnung die Vorreiterstellung in der Regulierung künstlicher Intelligenzen ein. Wer ist bei der Regulierung und Pönalisierung gefragt? Sollte der deutsche Gesetzgeber tätig werden oder ist nur eine europäisch harmonisierte Regelung sinnvoll?

Prof. Dr. Valerius: Künstliche Intelligenz betrifft uns alle. Daher erscheint es durchaus sinnvoll, zumindest im Europäischen Binnenmarkt einheitliche Regeln zu schaffen. Ob der regulatorische Ansatz der KI-Verordnung überzeugt, ist eine andere Frage. Insbesondere in den USA wird bekanntlich ein abweichender Weg beschritten.

In der Tat war es aber ein Anliegen der Europäischen Union, hier eine Vorreiterstellung einzunehmen. Die Hoffnung besteht wohl darin, dass in einer globalisierten Welt mit weltweit agierenden Unternehmen das Konzept der KI-Verordnung auch außerhalb der Europäischen Union, wenn nicht sogar weltweit beachtet werden müsste. Beim Datenschutz hatte dies mit der Datenschutz-Grundverordnung ganz gut funktioniert; hier wird sogar von einem Brüssel-Effekt gesprochen. Ob ein solcher Effekt auch bei der Regulierung von KI zu bemerken sein wird und sich der Ansatz der KI-Verordnung international durchzusetzen vermag, bleibt abzuwarten.

Gleiches gilt für die künftigen Aufgaben der Mitgliedstaaten bei der Umsetzung der KI-Verordnung. Wichtig wäre jedenfalls, deren – häufig nur vage – Bestimmungen alsbald zu konkretisieren. Es herrscht derzeit eine große Unsicherheit, ob und gegebenenfalls wie die Anforderungen der Verordnung bei einzelnen KI-Systemen umgesetzt werden müssen. Bleiben hier nähere Vorgaben seitens der Europäischen Kommission aus, würde den nationalen und für die Durchsetzung der KI-Verordnung zuständigen Regulierungsbehörden eine größere Bedeutung zuteilwerden. Bei besonders sensiblen Bereichen (z. B. auch im Strafrecht, in dem die Gesetzgebungskompetenz der Europäischen Union ohnehin eingeschränkt ist) könnte der nationale Gesetzgeber die Vorgaben der KI-Verordnung auch noch ergänzen und so verschärfen.

BayZR: Welche Bedeutung hat die KI-Verordnung für das deutsche Strafrecht, bestehen – neben den verbleibenden Auslegungsfragen – auch noch Regelungslücken?

Prof. Dr. Valerius: Für das Strafrecht hat die KI-Verordnung schon wegen der fehlenden Kompetenz bislang nur eine eingeschränkte Relevanz. Es finden sich etwa weder neue Straftatbestände in der KI-Verordnung noch wird die Verantwortlichkeit beim Einsatz von KI geregelt. Mittelbare Auswirkungen sind allerdings nicht ausgeschlossen. Sämtliche Pflichten, die etwa die KI-Verordnung für den Betrieb von KI-Systemen normiert, könnten bei Missachtung eine Sorgfaltspflichtverletzung begründen und somit einen Anknüpfungspunkt für eine Fahrlässigkeitsstrafbarkeit bilden, z. B. wenn bei dem Betrieb eines KI-Systems ein Mensch verletzt oder sogar getötet wird.

Für das Strafverfahren schon jetzt von Bedeutung ist, dass die KI-Verordnung nicht nur bestimmte Praktiken – wie etwa den Einsatz von biometrischen Echtzeit-Fernidentifizierungssystemen und Risikobewertungen – beschränkt, sondern auch sogenannte Hochrisiko-Systeme auflistet, die nur unter strengen Voraussetzungen betrieben werden dürfen. Hierzu zählen unter anderem Polygraphen und Emotionserkennungssysteme. Wer auf den baldigen Einsatz eines KI-gestützten Lügendetektors in Strafverfahren gehofft hat, dürfte daher enttäuscht werden.

BayZR: Wir bedanken uns herzlich für Ihre Teilnahme am Interview!