

Rechtsmedizin 2025 · 35:73–79
<https://doi.org/10.1007/s00194-025-00742-y>
 Angenommen: 3. Februar 2025
 Online publiziert: 7. März 2025
 © The Author(s) 2025



„KI“ in der Rechtsmedizin – von der Forschung in die Praxis: Welche Herausforderungen ergeben sich?

M. Rütther¹ · S. B. Eickhoff^{1,2} · L. König³ · Stefanie Ritz³ · B. Schäffer⁴

¹ Forschungszentrum Jülich, Institut für Neurowissenschaften und Medizin, INM-7: Gehirn und Verhalten, Jülich, Deutschland

² Institut für Systemische Neurowissenschaften, Heinrich-Heine-Universität, Düsseldorf, Deutschland

³ Institut für Rechtsmedizin, Universitätsklinikum Düsseldorf, Düsseldorf, Deutschland

⁴ Institut für Rechtsmedizin, Universität München, München, Deutschland

Abstract

Hintergrund: Die zunehmende Einführung von künstlicher Intelligenz (KI) in die Rechtsmedizin, ähnlich wie in anderen Bereichen der Medizin, ist absehbar. Vor dem Hintergrund dieser Entwicklung stellt sich die Frage nach den Herausforderungen, die mit dem Einsatz von KI in der Rechtsmedizin verbunden sind – eine Thematik, die bisher kaum behandelt wurde.

Ziel der Arbeit: Der Text untersucht zentrale Fragestellungen sowie die daraus resultierenden rechtlichen und ethischen Herausforderungen bei der Implementierung von KI in die forensische Praxis. Zudem werden Lösungsansätze vorgestellt.

Material und Methoden: Die Publikation basiert auf einer systematischen Literaturanalyse zu KI-Anwendungen in der Rechtsmedizin, aber auch der verwandten Bereiche, wie der medizinischen Diagnostik und digitalen Forensik. Berücksichtigt werden auch rechtliche Rahmenbedingungen, insbesondere die europäische KI-Verordnung sowie Perspektiven der KI-Ethik.

Ergebnisse und Diskussion: Für KI-Anwendungen in der Rechtsmedizin werden 6 zentrale Herausforderungen identifiziert: Zuverlässigkeit, Transparenz und Nachvollziehbarkeit („Blackbox“-Problematik), Verantwortlichkeit für Fehler, Datenquantität und -qualität, Datenbias und Fairness sowie die Akzeptanz KI-generierter Diagnosen. Zur Bewältigung der identifizierten Herausforderungen werden verschiedene Ansätze vorgeschlagen. Dazu zählen der Einsatz von „explainable AI“ (XAI) in Verbindung mit KI-Sachverständigen, die Entwicklung von Standardprotokollen sowie multizentrisches Arbeiten zur Sicherung der Datenqualität und Systemzuverlässigkeit. Zudem wird eine verstärkte transdisziplinäre Zusammenarbeit zwischen Informatikern, Rechtsmedizinern, Juristen und Ethikern empfohlen, um die Akzeptanz und Integration von KI-Systemen in die Praxis zu fördern.

Schlüsselwörter

KI-Sachverständige · Blackbox-Problematik · Datenqualität · Ethik · Haftung

S. Ritz und B. Schäffer teilen sich die Letztautorenschaft.



QR-Code scannen & Beitrag online lesen

Der Einsatz künstlicher Intelligenz (KI) in der Rechtsmedizin ist absehbar, KI-Anwendungen könnten im Ermittlungsverfahren oder in foro bald eine zentrale Rolle in Entscheidungsprozessen einnehmen. Dann werden insbesondere Transparenz, Zuverlässigkeit, mögliche Datenbias sowie Haftungs- und Akzep-

tanzfragen zentrale Herausforderungen darstellen. Dieser Beitrag beleuchtet diese Aspekte und zeigt Lösungsansätze wie „explainable AI“ (kurz: XAI), multizentrische Standards und interdisziplinäre Zusammenarbeit für eine ethisch und rechtlich angemessene Integration auf.

Einleitung

„Künstliche Intelligenz“ (KI) oder „artificial intelligence“ (AI) kann man als die Fähigkeit eines Systems beschreiben, externe Daten korrekt zu interpretieren und aus ihnen zu lernen, um spezifische Ziele und Aufgaben zu erfüllen [1]. Bei der Umsetzung dieser Vorgänge spielen Algorithmen eine wichtige Rolle; diese enthalten klar definierte Anweisungen oder Schritte, denen ein System folgt. Für die Fähigkeit des Lernens werden etwa eigene Algorithmen, die auf statistischen Techniken wie Machine Learning (ML), Deep Learning (DL) oder Natural Language Processing (NLP) basieren, verwendet. In der Konsequenz werden in Zusammenschau aller Anweisungen Modelle etabliert, die der Klärung spezieller Fragestellungen dienen oder auch neue Inhalte, wie z. B. Texte, generieren können („generative KI“).

Wie in anderen Feldern der Medizin ist der praktische Einsatz von KI auch in der Rechtsmedizin nicht nur vorgezeichnet, sondern absehbar. In manchen medizinischen Bereichen, wie der Tumordiagnostik, wird sie bereits gegenwärtig eingesetzt (Kumar et al. [2]). Die Potenziale der Reduktion menschlicher Subjektivität in der Diagnostik, der Minimierung von Fehlern, der Entlastung bei Routineaufgaben oder der Reduzierung von monetären Kosten im Gesundheitssystem werden hervorgehoben oder erhofft [3–5].

Mittlerweile gibt es zahlreiche Publikationen, die KI-Applikationen in verschiedensten Feldern der Rechtsmedizin vorschlagen und diskutieren (Übersichten bei Piraianu et al. [3] und Tournois et al. [4], speziell zur Gesichtserkennung in der forensischen Begutachtung Ramsthaler et al. [6]). Allerdings wird auch festgestellt, dass es noch keine KI-Applikation gibt, die bereits in der forensischen Praxis eingesetzt wird [4, 5, 7]. Dafür gibt es sicher verschiedene Ursachen wie eine noch mangelnde Datenlage, die noch unzureichende Ergebnisgenauigkeit von Modellen oder auch die Herausforderungen eines möglichen „Datenbias“ (zusammenfassend Tournois et al. [4]).

Noch viel zu wenig diskutiert ist die Frage, welche Herausforderungen sich bei der Nutzung von KI in der forensischen Praxis ergeben werden. In der Zukunft

hypothetisch mögliche Fallkonstellationen machen sehr deutlich, dass eine Vielzahl komplexer Fragen und Unwägbarkeiten im Raum stehen wird, wenn KI-Anwendungen im Ermittlungsverfahren oder in foro eine zentrale Rolle in Entscheidungsprozessen einnehmen werden. Die vorliegende Arbeit untersucht zentrale Fragestellungen und identifiziert die daraus resultierenden rechtlichen und ethischen Herausforderungen bei der Implementierung von KI in die forensische Praxis.

Material und Methoden

Für die Analyse zentraler Fragestellungen und die Identifikation der Herausforderungen, die bei KI-Nutzung in der forensischen Praxis zu adressieren sind, wurde wie folgt vorgegangen: Zunächst erfolgte eine systematische Literaturanalyse zu KI-Anwendungen in der Rechtsmedizin unter Nutzung von Suchbegriffen wie z. B. „KI (Rechts-)Medizin“, „KI-Sachverständiger“, „KI-Verordnung“ und „Gutachterhaftung“, wobei die Analyse auch auf verwandte Bereiche, wie die medizinische Diagnostik und digitale Forensik, ausgedehnt wurde. Weiterhin wurden wesentliche medizinische, rechtliche und ethische Literaturdatenbanken identifiziert (z. B. PubMed, openJur, Beck-Online oder PhilPapers) und mit Blick auf die oben genannten Begriffe durchsucht. Berücksichtigt wurde zudem der Literatur- und Zeitschriftenbestand des Forschungszentrums Jülich und der Universitätsbibliothek der LMU München. Eine freie Internetrecherche wurde ebenfalls durchgeführt. Neben den genannten Quellen fanden auch die rechtlichen Rahmenbedingungen, insbesondere die europäische KI-Verordnung sowie weitere wesentliche Texte aus der KI-Ethik Eingang in die Analyse. Ausgehend von der Quellenanalyse wurde dann erarbeitet, welche wesentlichen Fragen der Einsatz von KI in der Rechtsmedizin aus ethischer, juristischer und rechtsmedizinischer Sicht aufwerfen wird. Im Anschluss an die Bewertung und Diskussion dieser Fragen wurden dann 6 zentrale Herausforderungen identifiziert, die mit der Nutzung von KI-Applikationen in der forensischen Praxis verbunden sind.

Ergebnisse und Diskussion

Mögliche Anwendungsfelder für die rechtsmedizinische Praxis zeichnen sich bereits ab [3, 4]. Dazu gehören Lebensaltersschätzungen über KI-Modelle auf Basis mehrerer Parameter, Eingrenzung des Todeszeitpunktes über ML-Modelle, Wundaltersschätzungen durch eine KI-gestützte Auswertung histologischer Bilder, Personenidentifikation mithilfe von KI-Modellen u. a.

In diesen und ähnlichen Fällen können KI-Anwendungen auf vielfältige Weise zum Einsatz kommen. In naher Zukunft ist es jedoch am wahrscheinlichsten, dass KI nicht sämtliche Arbeitsschritte des Rechtsmediziners übernimmt. Vielmehr dürften a) nur ausgewählte Prozesse an die KI ausgelagert werden und b) die KI in diesen Prozessen nicht vollständig autonom, sondern meist unterstützend als quasi „Co-Pilot“ des Sachverständigen eingesetzt werden.

Allerdings dürften sich auch bei einem nur selektiven und unterstützenden Einsatz von KI wichtige Fragen ergeben, insbesondere im Hinblick auf ihren Einsatz in der forensischen Praxis. Im Folgenden werden einige dieser zu erwartenden Fragen näher beleuchtet.

Wie hat eine KI-Applikation das Ergebnis ermittelt? Wie kann man das Problem der „Blackbox“ adressieren?

In oben genannten Beispielen des Einsatzes von KI würde (jedenfalls aus Sicht der Verteidigung) die Frage naheliegen, wie genau die KI zu diesem Ergebnis gekommen ist. Die Beantwortung dieser Frage aber ist eine Herausforderung, weil der „Lernprozess“ der KI bei manchen Modellen zwar beschrieben, aber nicht im Detail nachvollzogen werden kann.

Der Vorwurf, dass KI-Systeme mit Blick auf die in ihnen ablaufenden Prozesse als „Blackbox“ agieren, gehört zu den meistdiskutierten Kritikpunkten überhaupt (für einen Überblick: Vilone und Longo [8]). Dieser Umstand wird zusätzlich dadurch verschärft, dass mittlerweile eine europarechtliche Verordnung über künstliche Intelligenz (KI-Verordnung bzw. „AI Act“) existiert [25]. So werden in der KI-Verordnung explizit die „Transparenz“ und damit

Erklärbarkeit von KI-Systemen gefordert (Erwägungsgründe 27 und 59)¹.

Um diese Problematik zu entschärfen, fließen in der gegenwärtigen Forschung erhebliche Ressourcen in Projekte zur sog. „explainable AI“ (XAI) (für einen Überblick: Arrieta et al. [9]). Ziel von XAI ist es, vom Menschen interpretierbare Modelle zu erhalten, insbesondere für Anwendungen in sensiblen Bereichen wie Militär, Banking und im Gesundheitswesen. Denn KI soll nicht nur Probleme lösen, sondern die erzeugten Ergebnisse sollen für Anwender auch verständlich und vertrauenswürdig sein (Sajid et al. [10]). In diesem Zusammenhang werden verschiedene Methoden vorgeschlagen (z. B. die Nutzung von „saliency maps“, Grad-CAM oder die Darstellung von Shapley-Werten), die es ermöglichen sollen, die Prozesse von KI-Systemen transparent und nachvollziehbar zu machen. Der Einsatz von XAI ist in foro jedoch noch nicht erprobt. Mögliche Ansätze für den Einsatz könnten jedoch aus anderen Bereichen abgeleitet werden, in denen XAI bereits eine weitergehende Anwendung gefunden hat, etwa in der medizinischen Bildanalyse (z. B. bei der Tumorerkennung) oder der digitalen Forensik (z. B. der Erklärung von Cyberkriminalität). Das grundsätzliche Bestreben besteht auch in diesen Bereichen darin, die Transparenz der Ergebnisfindung durch XAI zu steigern.

Es sollte jedoch beachtet werden, dass XAI selbst mit Herausforderungen konfrontiert ist. Besonders im Hinblick auf die Einführung in forensischen Kontexten ist zu berücksichtigen, dass XAI notwendigerweise eine Abwägung zwischen Erklärbarkeit und Genauigkeit erfordert (Holzinger et al. [11], weiterführend: Raz et al. [12]). Eine erhöhte Erklärbarkeit kann zu einer geringeren Präzision der Ergebnisse füh-

ren, was bei Einsatz in der Rechtsmedizin zu erheblichen Unsicherheiten und potenziellen Fehleinschätzungen führen kann.

Müsste künftig stets ein Team aus „KI-Experten“ und Rechtsmedizinern geladen werden?

Des Weiteren könnte der Sachverständige wohl mit der Aufforderung überfordert sein, im Detail zu erklären, wie die eingesetzte KI-Applikation zu ihrer Diagnose kam. Die KI-Verordnung sieht für hochriskante KI-Systeme, wie sie z. B. in der Rechtsmedizin verwendet werden könnten, aber eine Verpflichtung zur Erklärbarkeit vor. Das bedeutet, dass die Entscheidungen der KI-Systeme verständlich und nachvollziehbar sein müssen.

In der Praxis könnte dies dazu führen, dass KI-Experten vor Gericht erforderlich sind, um die Funktionsweise und die Entscheidungsprozesse der KI zu erläutern. Es gibt in den USA bereits Fälle, in denen KI-Experten hinzugezogen wurden, um technologische Beweismittel zu bewerten oder technische Verfahren zu erläutern, beispielsweise im Bereich der forensischen Datenanalyse. Als stilbildend kann dafür das Gerichtsverfahren „Daubert vs. Merrell Dow 1993“ [13] angesehen werden, in dem der sog. Daubert-Standard etabliert wurde. Dieser legt fest, dass Beweismittel nicht allein deshalb zugelassen werden, weil sie in der wissenschaftlichen Gemeinschaft anerkannt sind. Stattdessen muss zusätzlich ein Experte deren Funktionsweise und Zuverlässigkeit nachweisen und dem Gericht nachvollziehbar erklären.

Wo liegen die Risiken des Einsatzes der KI-Modelle? Sind diese zuverlässig? Wie kann das überprüft werden?

Eine weitere naheliegende Frage in oben genannten Anwendungsbeispielen wäre diejenige nach der Zuverlässigkeit der KI-basierten Prognosen. Ein Faktor, der in diesem Zusammenhang eine besondere Rolle spielt, wird häufig unter dem Stichwort „Datenbias“ diskutiert (für einen Überblick Mehrabi et al. [14]). Damit sind Verzerrungen bei der Erhebung und Analyse von Daten gemeint, die in foro zu fehlerhaften Entscheidungen führen können. Beson-

ders kritisch wurden etwa KI-Anwendungen im Justizbereich (z. B. das bekannte und viel rezipierte COMPAS-System) analysiert, die sich nach eingehender Prüfung als auf unausgewogenen Datensätzen basierend herausstellten und diskriminierende Tendenzen bei der Risikobewertung von Straftätern aufwiesen (Dressel und Farid [15]). In der rechtsmedizinischen Praxis dürfte ein Datenbias ebenfalls eine Rolle spielen, insbesondere da die Datengrundlage oft relativ klein und somit möglicherweise nicht repräsentativ ist. Für die oben genannten Anwendungsbeispiele könnten diesbezüglich eine ungleichmäßige Repräsentation unterschiedlicher Altersgruppen und/oder Populationen problematisch sein.

Dass es notwendig ist, diskriminierende Tendenzen vor Gericht zu vermeiden, ergibt sich aus dem Grundgesetz (Art. 3 Abs. 1 GG). Auch die KI-Verordnung greift dieses allgemeine Anliegen auf (s. Erwägungsgrund 59 in [25])². Zwar wird dort die Prüfung grundsätzlich von der Risikoklasse abhängig gemacht, sodass nicht alle Anwendungen in der gleichen Weise geprüft werden müssen. Da jedoch gerade Applikationen in der Rechtsmedizin aller Wahrscheinlichkeit nach als Hochrisiko-KI-Systeme eingestuft werden dürften, ergeben sich ein erhöhter Prüfungsbedarf und -aufwand (s. bes. Kap. 3 zur Einstufung als Hochrisiko-KI in [25]). Um hierauf vonseiten der forensischen Praxis zu reagieren, bieten sich mehrere Ansätze an. Grundsätzlich sollte verstärkte Sorgfalt auf die Datenprüfung, -erhebung, -aufbereitung und -präsentation gelegt werden. Bei seltenen Fällen und geringen Mengen an Trainingsdaten könnte zudem geprüft werden, ob Verfahren wie Datenaugmentierung (d. h. die künstliche Erzeugung neuer Daten aus vorhandenen Daten) eingesetzt werden können, um eine ausgewogenere Datengrundlage zu schaffen (Shorten

¹ So wird im Erwägungsgrund 27 der Verordnung explizit auf die Wichtigkeit von Transparenz (und anderen Aspekten wie Sicherheit und Fairness) eingegangen. Hierbei meint Transparenz, dass „... KI-Systeme so entwickelt und verwendet werden, dass sie angemessen nachvollziehbar und erklärbar sind, wobei den Menschen bewusst gemacht werden muss, dass sie mit einem KI-System kommunizieren oder interagieren, und dass die Betreiber ordnungsgemäß über die Fähigkeiten und Grenzen des KI-Systems informieren und die betroffenen Personen über ihre Rechte in Kenntnis setzen müssen“.

² Mit Blick auf die Rechtsprechung wird in Erwägungsgrund 59 erläutert: „Darüber hinaus könnte die Ausübung wichtiger verfahrensrechtlicher Grundrechte wie etwa des Rechts auf einen wirksamen Rechtsbehelf und ein unparteiisches Gericht sowie das Verteidigungsrecht und die Unschuldsvermutung behindert werden, insbesondere, wenn solche KI-Systeme nicht hinreichend transparent, erklärbar und dokumentiert sind“.

und Khoshgoftaar [16]). Darüber hinaus könnten spezielle Software-Tools zur Erkennung eines Datenbias verwendet werden (für eine Auflistung und Erläuterung Bellamy et al. [17]).

Wer ist, rechtlich gesehen, verantwortlich im Fall von Fehleinschätzungen? Wer ist (oder sollte) rechtlich haftbar (sein)?

Die Frage nach der Zuschreibung von Verantwortung und, davon ausgehend, auch der rechtlichen Haftung ist sicherlich eine der zentralen Fragen, die sich beim Einsatz von KI stellen. An dieser Stelle ist zunächst darauf hinzuweisen, dass im Rahmen einer guten wissenschaftlichen Praxis der Sachverständige den Einsatz einer KI-Applikation zu dokumentieren hat. Darüber hinaus ist zudem § 839a BGB maßgeblich, der festlegt, dass ein von deutschen Gerichten oder (analog) von der Staatsanwaltschaft im Ermittlungsverfahren beauftragter Sachverständiger grundsätzlich zum Ersatz eines Schadens, der einem Verfahrensbeteiligten durch eine gerichtliche Entscheidung entsteht, die auf einem unrichtigen Gutachten beruht, verpflichtet werden kann (Dettmeyer [18]). Hierbei kann sich das fehlerhafte Gutachten zum einen daraus ergeben, dass von einem unzutreffenden Sachverhalt ausgegangen wird, und zum anderen, wenn aus dem festgestellten Sachverhalt unhaltbare Schlüsse gezogen werden. Maßgeblich ist zudem ein „qualifiziertes“ Verschulden, wobei § 839a BGB mindestens grobe Fahrlässigkeit voraussetzt, d. h., der Sachverständige muss seine erforderliche Sorgfalt in ungewöhnlich hohem Maße verletzt haben [19]. Im Zusammenhang mit dem Einsatz von KI ist sowohl denkbar, dass die KI den Sachverhalt unzutreffend erfasst (z. B. bei KI-gestützter Auswertung der staatsanwaltlichen Akten) als auch unzutreffende Schlussfolgerungen zieht (z. B. bei KI-gestützter Lebensaltersschätzung anhand radiologischer oder molekularer Befunde). Sobald sich der Fehler der KI im Gutachten realisiert (z. B. durch ungeprüfte Übernahme des von der KI falsch bestimmten bzw. geschätzten Alters eines Beschuldigten), ist dann grundsätzlich eine Haftung des Sachverständigen gemäß § 839a BGB

denkbar (Institut für Sachverständigenwesen e. V. [20]).

Darüber hinaus bleibt jedoch rechtlich noch offen, wie weitgehend die Prüfung des Sachverständigen ausfallen muss, um seiner Sorgfaltspflicht nachzukommen. Die KI-Verordnung stuft KI-Anwendungen im Bereich der Medizin jedenfalls als hochriskant ein, woraus sich eine erhöhte Verantwortung für den Einsatz ergibt. Denkbar wäre etwa, zukünftig den Einsatz von XAI als Grundlage der eigenen Prüfung und Dokumentation zu etablieren.

Neben der Frage nach der spezifischen Verantwortung des Sachverständigen stellt sich auch die Frage nach der allgemeinen Verantwortungsverteilung. Es besteht die Gefahr einer diffusen Verantwortungszuschreibung, da strittig sein könnte, ob die Haftung primär bei den Entwicklern der KI-Systeme oder bei den Anwendern (d. h. dem rechtsmedizinischen Sachverständigen) liegt (für die grundsätzliche Problematik de Sio und van den Hoven [21]). Man denke etwa an Fälle, in denen eine Fehleinschätzung, die durch einen Mangel an repräsentativen Daten oder durch algorithmische Verzerrungen verursacht wurde, vorliegt. Wie soll in diesen Fällen die Verantwortung aufgeteilt werden?

Die KI-Verordnung sanktioniert Verstöße grundsätzlich durch Geldbußen. Eine ergänzende KI-Haftungsrichtlinie (AI Liability Directive) befindet sich noch im Gesetzgebungsverfahren. Flankierend wurde die Produkthaftungsrichtlinie (Product Liability Directive) novelliert, die v. a. die Hersteller von KI-Anwendungen adressiert, in Deutschland aber erst noch in nationales Recht überführt werden muss (zum Stand der Dinge: Noerr [22, 23]). Bei Anwendungen an Lebenden könnte es erforderlich sein, auch die EU-Medizinprodukte-Verordnung (Medical Device Regulation) zu berücksichtigen.

Sachverständige sollten bei Unsicherheiten über den Umfang des Versicherungsschutzes in ihrer (insbesondere privat-)gutachterlichen Tätigkeit vor dem Einsatz von KI-Systemen Rücksprache mit ihrer Berufshaftpflichtversicherung halten. Zudem empfiehlt es sich, bei Haftungsfragen ggf. einen spezialisierten Anwalt zu konsultieren.

Was geschieht, wenn verschiedene KI-Modelle zu unterschiedlichen Ergebnissen gelangen (KI gegen KI)?

Für die verschiedenen Anwendungsfälle in der Rechtsmedizin werden derzeit mit unterschiedlichen Ansätzen zahlreiche KI-gestützte Modelle entwickelt. So ist es leicht vorstellbar, dass die Situation „KI gegen KI“ entsteht. Abweichende Ergebnisse unterschiedlicher KI-Modelle stellen jedoch die Glaubwürdigkeit und Verlässlichkeit von KI infrage und erschüttern das Vertrauen in die Objektivität dieser Modelle, insbesondere, wenn sie in sensiblen Bereichen wie in der Forensik genutzt werden.

Dass KI in der rechtsmedizinischen Praxis zu inkonsistenten Ergebnissen führen können, ist aufgrund der mangelnden Erprobung dieser Applikationen zwar nicht nachgewiesen. Zieht man jedoch die Anwendung solcher Systeme in anderen Bereichen heran, etwa bei Texten oder Bildern, so erscheinen konträre Ergebnisse durchaus wahrscheinlich. Beispiele aus der Praxis zeigen, dass sogar hinreichend vergleichbare generative Modelle, die auf denselben oder ähnlichen Datensätzen trainiert wurden, unterschiedliche und teils widersprüchliche Resultate liefern können. Diese Problematik könnte ebenso im rechtsmedizinischen Kontext auftreten, wenn KI-Modelle zur Analyse von ähnlichen Beweismitteln herangezogen werden.

Um die Unsicherheiten, die durch solche abweichenden KI-Ergebnisse entstehen, in einem Gerichtsverfahren zu bewältigen, gibt es verschiedene Herangehensweisen. Eine zentrale Methode ist wiederum der Einsatz von XAI, um Transparenz in den Entscheidungsprozess der KI zu bringen. Die Hoffnung besteht darin, mithilfe von XAI die Faktoren nachzuvollziehen, die zu unterschiedlichen Ergebnissen geführt haben. Dabei kann auch geprüft werden, ob die abweichenden Ergebnisse auf methodische Fehler oder auf unterschiedliche Trainingsdatensätze zurückzuführen sind. Darüber hinaus erscheint es auch hilfreich, auf die bestehende Praxis in der Rechtsmedizin zu rekurrieren, da sich dort bereits ein Umgang mit sich widersprechenden Gutachten etabliert hat, der auf die neue Situation „KI gegen KI“ angewandt werden könnte.

Kann eine KI eine Aussage „mit an Sicherheit grenzender Wahrscheinlichkeit“ treffen?

Die Frage der Leistungsfähigkeit und der Genauigkeit stellt sich im forensischen Kontext mit besonderer Dringlichkeit, insbesondere im strafrechtlichen Setting. Eine strafrechtliche Verurteilung setzt die „Überzeugung“ des Gerichts voraus (§ 261 StPO), dass die erhobenen Beweise ausreichen, um den Angeklagten im Sinne der Anklage schuldig zu sprechen. Im Rahmen der Beweiswürdigung durch das Gericht wird regelhaft das Maß einer „mit an Sicherheit grenzender Wahrscheinlichkeit“ gefordert [24].

Tournois et al. [4] schätzen den derzeitigen Stand der Entwicklung von KI-Systemen sehr kritisch ein. Sie betonen, dass KI-Systeme kaum evaluiert werden. Tatsächlich wird sogar noch weitergehend kritisiert, dass auch bei denjenigen Fällen, in denen ein Test durchgeführt wurde, diese größtenteils auf zweifelhaften Verfahren beruht haben, sodass einige grundsätzliche Schwierigkeiten von KI-Systemen nicht erkannt werden. Ein häufig unerkanntes Problem ist das sog. Overfitting. Dabei handelt es sich um ein Phänomen in der Modellentwicklung, bei dem ein KI-System die Trainingsdaten so präzise „lernt“, dass es nicht nur relevante Muster, sondern auch zufällige Details oder Störfaktoren in diesen spezifischen Daten erfasst. Dies führt dazu, dass das Modell auf neue, unbekannte Daten schlecht generalisiert und unzuverlässige Ergebnisse liefert.

Insgesamt lässt sich daher festhalten, dass eine Diskrepanz zwischen dem entsteht, was man sich von solchen Modellen erhofft, und dem, was sie derzeit zu leisten imstande sind. Es dürfte noch ein weiter Weg vor uns liegen, bis KI-Applikationen die Beantwortung rechtsmedizinischer Fragen mit der im Strafverfahren erforderlichen Sicherheit unterstützen können.

Welche Herausforderungen gilt es zu adressieren?

Die Vorstellung, dass KI-Applikationen auch die Rechtsmedizin revolutionieren werden, ist durchaus begründet. Ausgehend von der Diskussion der obigen

Fragen ist allerdings absehbar, dass relevante Herausforderungen vor uns liegen, bevor solche Applikationen in der rechtsmedizinischen Praxis etabliert werden können (Tynan [25]).

Zuverlässigkeit von KI-Applikationen

Applikationen, die auf KI basieren, müssen vor ihrem praktischen Einsatz ausführlich getestet und validiert werden. Dafür und für die Implementierung von KI-Applikationen in das forensische Setting sollten Standardprotokolle und Richtlinien entwickelt werden. Deren Inhalt wird natürlich von der Art der jeweiligen Applikation abhängen. Grundsätzlich ist ein solches Vorgehen in der Rechtsmedizin aber geübte Praxis (z.B. Richtlinien der GTFCh für toxikologische Analysen [26]).

Transparenz und Nachvollziehbarkeit

Der „Blackbox“-Charakter von KI-Anwendungen kann die Beweiswürdigung des Gerichts zu einer Herausforderung machen. Der Einsatz von XAI und die Zusammenarbeit mit KI-Sachverständigen könnten daher auch in der Begutachtungspraxis angeraten sein.

Verantwortlichkeit

KI-Anwendungen treffen ihre Aussagen aufgrund von Wahrscheinlichkeitsberechnungen. Sie sind daher prinzipiell fehleranfällig, was durch Datenbias (s. Herausforderung „Datenbias und Fairness“ unten) und mangelnde Datenqualität (s. Herausforderung „Datenquantität und -qualität“ unten) noch zu einem dringlicheren Problem werden kann. Daraus ergibt sich die rechtliche Frage, wer im Fall einer fehlerhaften Einschätzung verantwortlich ist – der Entwickler der KI-Anwendung oder der rechtsmedizinische Sachverständige, der sich in seinem Gutachten auf die KI-basierte Aussage stützt?

Datenquantität und -qualität

KI-Modelle setzen eine gute Datengrundlage in Qualität und Quantität voraus. Dass diese derzeit nur sehr schwer zu erreichen ist, ergibt sich, wie bereits erläutert, aus der

geringen Datenmenge und Vielfalt in der Rechtsmedizin. Diesem Umstand entgegenzuwirken, bedürfte es wohl multizentrischer Zusammenarbeit, um in Zukunft die Datengrundlage zu verbessern.

Datenbias und Fairness

Die Datenqualität ist nicht nur für die Verlässlichkeit der Einschätzungen und Prognosen einer KI-Anwendung (s. Herausforderung „Zuverlässigkeit von KI-Applikationen“ oben) relevant, sondern auch für die Frage, ob ggf. diskriminierende Aussagen entstehen, weil bestimmte Personengruppen in den Ausgangsdaten über- oder unterrepräsentiert werden (ausführlicher: Heinrichs [27]). Das spricht wiederum dafür, die relevanten Prozesse und Ausgangsdaten einer KI-Applikation transparent zu machen, z.B. durch die Involvierung von XAI.

Akzeptanz KI-generierter Diagnosen

Nicht zuletzt müssen KI-generierte Diagnosen von Gerichten akzeptiert werden. Dieser Herausforderung kann nur durch einen interdisziplinären Dialog und/oder interdisziplinäre Forschung begegnet werden. Völlig zu Recht hat Tynan [25] eine transdisziplinäre Kollaboration zwischen *Computer Scientists*, Rechtsmedizinern, Juristen und Ethikern eingefordert, um KI-Applikationen in einer Weise zu implementieren, die dem Rechtssystem dient; dazu seien wissenschaftlich fundierte Standards und Richtlinien gemeinsam zu erarbeiten. Auch die gefühlte Rechtssicherheit sollte adressiert werden, indem der Einsatz von KI-Applikationen transparent gemacht wird.

Fazit für die Praxis

In verallgemeinerter Form, z.B. mit Blick auf die medizinische Praxis und das Gesundheitssystem, sind die oben genannten Herausforderungen bereits seit längerem Gegenstand der ethischen und rechtlichen Diskussion. Die Einengung dieser Problemkomplexe auf die rechtsmedizinische Praxis ist bisher jedoch noch nicht geleistet worden – und ist eine wichtige, interdisziplinäre Aufgabe, die unbedingt angegangen werden sollte. Mit Einzug KI-gestützter Methoden in die gutachterliche Praxis wird es aus Sicht der Autoren empfehlenswert sein, den Einsatz

von KI-Systemen für die Gutachtenerstellung gegenüber den Auftraggebern anzuzeigen, KI-generierte Daten stets kritisch zu prüfen und bei Unklarheiten – nach Rücksprache mit dem Auftraggeber – einen KI-Sachverständigen zurate zu ziehen.

Korrespondenzadresse

Prof. Dr.med. Stefanie Ritz

Institut für Rechtsmedizin, Universitätsklinikum Düsseldorf
Moorenstraße 5, 40225 Düsseldorf,
Deutschland
Stefanie.Ritz@med.uni-duesseldorf.de

Funding. Open Access funding enabled and organized by Projekt DEAL.

Datenverfügbarkeit. In dieser Publikation wurden keine Daten generiert oder analysiert.

Einhaltung ethischer Richtlinien

Interessenkonflikt. M. Rütther, S.B. Eickhoff, L. König und S. Ritz geben an, dass kein Interessenkonflikt besteht. B. Schäffer hält Aktien der NVIDIA Corporation.

Für diesen Beitrag wurden von den Autor/-innen keine Studien an Menschen oder Tieren durchgeführt. Für die aufgeführten Studien gelten die jeweils dort angegebenen ethischen Richtlinien.

Open Access. Dieser Artikel wird unter der Creative Commons Namensnennung 4.0 International Lizenz veröffentlicht, welche die Nutzung, Vervielfältigung, Bearbeitung, Verbreitung und Wiedergabe in jeglichem Medium und Format erlaubt, sofern Sie den/die ursprünglichen Autor(en) und die Quelle ordnungsgemäß nennen, einen Link zur Creative Commons Lizenz beifügen und angeben, ob Änderungen vorgenommen wurden. Die in diesem Artikel enthaltenen Bilder und sonstiges Drittmaterial unterliegen ebenfalls der genannten Creative Commons Lizenz, sofern sich aus der Abbildungslegende nichts anderes ergibt. Sofern das betreffende Material nicht unter der genannten Creative Commons Lizenz steht und die betreffende Handlung nicht nach gesetzlichen Vorschriften erlaubt ist, ist für die oben aufgeführten Weiterverwendungen des Materials die Einwilligung des jeweiligen Rechteinhabers einzuholen. Weitere Details zur Lizenz entnehmen Sie bitte der Lizenzinformation auf <http://creativecommons.org/licenses/by/4.0/deed.de>.

Literatur

- Kaplan A, Haenlein M (2019) Siri, Siri, in my hand: who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. *Bus Horiz* 62:15–25
- Kumar Y, Koul A, Singla R, Ijaz MF (2023) Artificial intelligence in disease diagnosis: a systematic literature review, synthesizing framework and future research agenda. *J Ambient Intell Humaniz Comput* 14:8459–8486

“AI” in forensic medicine—From research into practice: what are the challenges?

Background: The increasing implementation of artificial intelligence (AI) in forensic medicine, similar to other medical fields, is foreseeable. Against this backdrop, the question arises regarding the challenges associated with the use of AI in forensic medicine, a topic that has not yet been addressed.

Objective: This article examines key issues and the resulting legal and ethical challenges in the implementation of AI into forensic practice. Additionally, potential solution approaches are presented.

Material and methods: The publication is based on a systematic literature analysis of AI applications in forensic medicine and related fields, such as medical diagnostics and digital forensics. It also considers legal frameworks, particularly the European AI Act, as well as perspectives from AI ethics.

Results and discussion: For AI applications in forensic medicine six key challenges were identified: reliability, transparency and explainability (the black box problem), accountability for errors, data quantity and quality, data bias and fairness, and the acceptance of AI-generated diagnoses. Various approaches are proposed to address the identified challenges. These include the use of “explainable AI” (in short: XAI) in conjunction with AI experts, the development of standard protocols and multicenter collaborations to ensure data quality and system reliability. Furthermore, enhanced interdisciplinary collaboration between computer scientists, forensic pathologists, legal experts and ethicists is advocated to promote the acceptance and integration of AI systems into practice.

Keywords

AI experts · Black box problem · Data quality · Ethics · Liability

- Piraiyanu A-I, Fulga A, Musat CL et al (2023) Enhancing the evidence with algorithms: how artificial intelligence is transforming forensic medicine. *Diagnostics* 13:2992
- Tournois L, Troussel V, Hatsch D, Delabarde T, Ludes B, Lefèvre T (2024) Artificial intelligence in the practice of forensic medicine: a scoping review. *Int J Legal Med* 138:1023–1037
- Ondruschka B, Seifert D, Rotermund N, Fehnker A (2024) A critical review of “Artificial intelligence in the practice of forensic medicine: a scoping review”. *Int J Legal Med* 138:1667–1668
- Ramsthaler F, Huckenbeck W, Kettner M, Lux C, Verhoff MA (2024) Kann Amazons Rekognition (AWS) die Identifizierungssicherheit erhöhen? *Arch Kriminol* 254:1–43
- Galante N, Cotroneo R, Furci D, Lodetti G, Casali MB (2023) Applications of artificial intelligence in forensic sciences: current potential benefits, limitations and perspectives. *Int J Legal Med* 137:445–458
- Vilone G, Longo L (2020) Explainable artificial intelligence: a systematic review. *arXiv preprint arXiv:2006.593*
- Arrieta AB, Díaz-Rodríguez N, Del Ser J et al (2020) Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Inf Fusion* 58:82–115
- Sajid Ali TA, Shaker ES, Khan M, Jose M, Alonso-Moral, Del Roberto C, Riccardo G, Javier DS, Natalia D-R, Francisco H (2023) Explainable Artificial Intelligence (XAI): What we know and what is left to attain Trustworthy Artificial Intelligence. *Inf Fusion* 99:101805
- Holzinger A, Saranti A, Molnar C, Biecek P, Samek W (2022) Explainable AI methods—a brief overview. In: Holzinger A, Goebel R, Fong R, Moon T, Müller KR, Wojciech S (Hrsg) *AI beyond deep models and classifiers*. Springer, Berlin, SS 13–S 38
- Raz A, Heinrichs B, Avnoon N, Eyal G, Inbar Y (2024) Prediction and explainability in AI: Striking a new balance? *J Big Data* 11:1
- Gerichtsurteil: Daubert v. Merrell Dow Pharmaceuticals I (1993), 509 U.S. 579
- Mehrabi N, Morstatter F, Saxena N, Lerman K, Galstyan A (2021) A survey on bias and fairness in machine learning. *ACM Comput Surv* 54:1–35
- Dressel J, Farid H (2018) The accuracy, fairness, and limits of predicting recidivism. *Sci Adv* 4:eaa05580
- Shorten C, Khoshgoftaar TM (2019) A survey on image data augmentation for deep learning. *J Big Data* 6:60
- Bellamy RK, Dey K, Hind M et al (2019) AI Fairness 360: an extensible toolkit for detecting and mitigating algorithmic bias. *IBM J Res Dev* 63(4):1–4–115
- Dettmeyer R (2016) Die Rolle des medizinischen Sachverständigen – gesetzliche Vorgaben und Entscheidungen der Rechtsprechung. *Z Rechtsmed* 26:316–324
- Kondusch O (2023) Haftung für fehlerhafte Gutachten – Grundlagen und Vermeidungsstrategien. *DS* 2023:238–248
- Institut für Sachverständigenwesen e.V. (2022) Merkblatt Haftung gerichtlichen Sachverständigen. www.ifsforum.de/fileadmin/user_upload/Downloads_Merkblaetter_etc/Merkblatt_Haftung_Gerichtssachverstaendiger.pdf. Zugegriffen: 17. Nov. 2024
- Santoni de Sio F, Van den Hoven J (2018) Meaningful human control over autonomous systems: a philosophical account. *Fronti Robot AI* 5:15

22. Haftungsrichtlinie KI www.noerr.com/de/insights/ki-haftungsrichtlinie-studie-des-eu-parlaments-zur-haftung-fuer-ki. Zugriffen: 21. Nov. 2024
23. Produkthaftungsrichtlinie. www.noerr.com/de/insights/neue-produkthaftungsrichtlinie-massive-pflichten-fuer-industrie. Zugriffen: 21. Nov. 2024
24. Frisch W (2016) Beweiswürdigung und richterliche Überzeugung. ZIS 2016:707
25. Tynan P (2024) The integration and implications of artificial intelligence in forensic science. Forensic Sci Med Pathol 20(3):1103–1105
26. Richtlinien der GTFCh für toxikologische Analysen. www.gtfch.org/cms/index.php/richtlinien. Zugriffen: 10. Jan. 2025
27. Heinrichs B (2022) Discrimination in the age of artificial intelligence. AISoc 37:143–154

Hinweis des Verlags. Der Verlag bleibt in Hinblick auf geografische Zuordnungen und Gebietsbezeichnungen in veröffentlichten Karten und Institutsadressen neutral.

Deutsche Gesellschaft für Humanes Sterben e. V. **Weißbuch Freitodbegleitung 2022**

Stuttgart: Kohlhammer 2024, 1. Auflage, 180 S., (ISBN: 978-3-17-044391-4), Softcover 29,00 EUR



Das Weißbuch Freitodbegleitung 2022 der Deutschen Gesellschaft für Humanes Sterben e. V. (DGHS) ist im Wesentlichen eine Dokumentation über die im Jahr 2022 durchgeführten 229 Freitodbegleitungen. Alle Fälle werden kurz in anonymisierter Form dargestellt. Zwei Fälle werden exemplarisch mit den (ebenfalls anonymisierten) vollständigen Dokumentationen wiedergegeben. Das Buch ist in zwei Teile gegliedert: „Hintergründe und Reflexionen“ und „Falldokumentationen“, flankiert von einem sehr langen Vorwort, das 10 Seiten umfasst, und einem kurzen Anhang, in welchem sich die DGHS vorstellt. Ein wenig irritierend ist das lange Vorwort, in dem ausführlich auf die juristische, wie auch auf die öffentliche Debatte eingegangen wird, was sich dann in großen Teilen inhaltlich im ersten Teil des Buches wiederholt.

Die 229 Falldarstellungen stellen den Lebensweg und die Motivlage der Suizidenten kurz dar, die Durchführung des Suizids selbst wird nicht im Detail erwähnt. Die Darstellungen bleiben aufgrund ihrer Kürze oberflächlich, vermitteln aber in ihrer Gesamtheit ein kaleidoskopartiges Bild des Personenkreises, der sich wegen einer Freitodbegleitung an die DGHS gewandt hat. Die beiden Fälle, bei denen die Dokumentation in ihrer Gesamtheit wiedergegeben wird, gleichen das etwas aus. Man vermisst jedoch eine Darstellung jener Fälle, deren Anträge auf Sterbebegleitung von der DGHS zurückgewiesen wurden; immerhin handelt es sich wohl um 34 Anträge, die abgelehnt wurden, meist wegen psychiatrischer Erkrankungen. Für ein weiteres Weißbuch würde man sich das wünschen. Das Buch ist in dieser Form vor allem für diejenigen interessant, die sich mit dem Thema „Suizid“ aus soziologischer und psychologischer Sicht beschäftigen.

Michael Bohnert, Würzburg