Fachbereich WD 7

Künstliche Intelligenz: Einsatz in der öffentlichen Verwaltung und in der Justiz, Haftungssysteme und Risiken bei der Nutzung

Künstliche Intelligenz: Einsatz in der öffentlichen Verwaltung und in der Justiz, Haftungssysteme und Risiken bei der Nutzung

Aktenzeichen:	WD 7 - 3000 - 004/25
Abschluss der Arbeit:	04.03.2025 (zugleich letzter Abruf aller Internetquellen)
Fachbereich:	WD 7: Zivil-, Straf- und Verfahrensrecht, Medienrecht, Bau und Stadtentwicklung

Inhaltsverzeichnis

1. Einleitung

Die Wissenschaftlichen Dienste wurden um Auskunft darüber gebeten, für welche Aufgaben Künstliche Intelligenz (KI) von der öffentlichen Verwaltung und der Justiz eingesetzt wird. Darüber hinaus wurde gefragt, welche haftungsrechtlichen Probleme beim Einsatz von KI auftreten und welche Risiken damit verbunden sein können.

Der Begriff des KI-Systems wird in der europäischen KI-Verordnung in Art. 3 Nr. 1 definiert.[1] Danach handelt es sich um „ein maschinengestütztes System, das für einen in unterschiedlichem Grade autonomen Betrieb ausgelegt ist und das nach seiner Betriebsaufnahme anpassungsfähig sein kann und das aus den erhaltenen Eingaben für explizite oder implizite Ziele ableitet, wie Ausgaben wie etwa Vorhersagen, Inhalte, Empfehlungen oder Entscheidungen erstellt werden, die physische oder virtuelle Umgebungen beeinflussen können“. Kurz gesagt handelt es sich um Software, die mit Methoden des maschinellen Lernens entwickelt wurde.[2]

2. Künstliche Intelligenz in der öffentlichen Verwaltung

2.1. Beratungszentrum für Künstliche Intelligenz

Einen Überblick über den Einsatz von KI in der Bundesverwaltung bietet das Beratungszentrum für Künstliche Intelligenz (BeKI) des Bundesministeriums des Innern und für Heimat.[3] Die Übersicht des BeKI enthält derzeit (Stand Februar 2025) insgesamt 179 Einträge mit Informationen zu jedem KI-System in Form eines Steckbriefs. Diese Informationen umfassen Name und Status, zuständiges Ressort und Organisationseinheit, bisherige Kooperationspartner und weitere Kooperationswünsche, Auftragnehmer, Thema inkl. Kurzbeschreibung, Risikostrategie sowie das Datum des Projektbeginns.

Die thematischen Schwerpunkte der gemeldeten KI-Systeme liegen in den Bereichen Energie und Umwelt (63 KI-Systeme), Forschung (60), öffentliche Verwaltung (32), Arbeit und Soziales (20), Information und Kommunikation (18), Landwirtschaft (17) und Gesundheit (14).

Die meisten registrierten KI-Systeme stammen aus dem Umweltbundesamt (48), gefolgt von der Bundesagentur für Arbeit (15), dem Bundesamt für Verbraucherschutz und Lebensmittelsicherheit (14), dem Bundesamt für Naturschutz (12), dem Bundesinstitut für Risikobewertung (9) und dem Bundesamt für Wirtschaft und Ausfuhrkontrolle (7). Als verwendete KI-Methoden werden

---

1 Verordnung (EU) 2024/1689 des Europäischen Parlaments und des Rates vom 13. Juni 2024 zur Festlegung harmonisierter Vorschriften für künstliche Intelligenz und zur Änderung der Verordnungen (EG) Nr. 300/2008, (EU) Nr. 167/2013, (EU) Nr. 168/2013, (EU) 2018/858, (EU) 2018/1139 und (EU) 2019/2144 sowie der Richtlinien 2014/90/EU, (EU) 2016/797 und (EU) 2020/1828 (Verordnung über künstliche Intelligenz), ABl. EU 2024/1989.

2 Borges: Die Haftung für Software und KI-Systeme nach der neuen Produkthaftungsrichtlinie, CR 1/2025, 1, 5.

3 Bundesministerium des Innern und für Heimat: KI-Systeme der Bundesverwaltung, abgerufen unter: https://maki.beki.bund.de/a/bmi-makimo-app?kiosk.

vor allem Machine Learning, Natural Language Processing, Generative KI, Computer Vision und Deep Learning genannt.[4]

2.2. Studie zu den Potenzialen von KI in der öffentlichen Verwaltung

In einer Studie der IW Consult, einem Dienstleister für Auftragsforschung und Beratung des Instituts der Deutschen Wirtschaft, wurden im Auftrag des Google-Konzerns die Potenziale von KI für die öffentliche Verwaltung untersucht.[5] Dazu wurden Verwaltungsmitarbeiterinnen und -mitarbeiter zu den Anwendungsbereichen von KI befragt und die Auswirkungen des KI-Einsatzes in verschiedenen Arbeitsbereichen ermittelt.

58 Prozent der Befragten gaben an, KI zur Unterstützung bei der Internetrecherche zu nutzen. 57 Prozent der Befragten nutzen KI zur Übersetzung von Informationen oder Dokumenten. Bei der Datenanalyse und beim Verfassen langer Texte lassen sich jeweils 54 Prozent von KI unterstützen, beim Verfassen von Berichten und Dokumenten sind es 53 Prozent.

Die Studie geht davon aus, dass KI bei rund 70 Prozent der Arbeitsplätze in der Verwaltung unterstützend eingesetzt werden könnte, zum Beispiel bei IT-Fachkräften. Rund 12 Prozent der Arbeitsplätze könnten durch KI ganz oder teilweise automatisiert werden, beispielsweise bei Sekretariats- und Bürokräften. Für 18 Prozent der Arbeitsplätze werden keine oder nur geringe Auswirkungen durch den Einsatz von KI erwartet. Dies betrifft beispielsweise die Gebäudetechnik oder andere handwerkliche Arbeitsbereiche.[6]

2.3. Studie zu Anwendungsfeldern und Szenarien

Das Fraunhofer-Institut für Arbeitswirtschaft und Organisation hat gemeinsam mit „The Open Government Institute (TOGI)“ der Zeppelin Universität Friedrichshafen untersucht, wie KI die Organisation und Arbeitsweise der öffentlichen Verwaltung in den kommenden Jahren verändern wird.[7] Die Potenzialstudie gibt einen Überblick über Fähigkeiten und Einsatzmöglichkeiten von KI im öffentlichen Sektor. Als mögliche Anwendungsbeispiele stellt die Studie Chatbots und persönliche Sprachassistenten, Serviceroboter als digitale Assistenten, Identitätsmanagement als sichere Methode für den persönlichen Zugang zu verschiedenen Systemen oder Anwendungen für die Hintergrundverwaltung in der Sachbearbeitung vor. Neben den Anwendungsbeispielen für KI werden zudem Verfahren und Prozesse beschrieben, Stärken und Schwächen analysiert und konkrete Umsetzungsempfehlungen gegeben. Abschließend werden die drei Zukunftsszenarien „Von

---

4 Bundesministerium des Innern und für Heimat: KI-Systeme der Bundesverwaltung, abgerufen unter: https://maki.beki.bund.de/a/bmi-makimo-app?kiosk.

5 IW Consult: Der digitale Faktor – Wie Deutschland von intelligenten Technologien profitiert. Potenziale künstlicher Intelligenz in der öffentlichen Verwaltung, 2024, abgerufen unter: https://der-digitale-faktor.de/; https://der-digitale-faktor.de/download/IW_Google-Studie_DeepDive_PublicSector_DE.pdf.

6 IW Consult: Der digitale Faktor – Wie Deutschland von intelligenten Technologien profitiert. Potenziale künstlicher Intelligenz in der öffentlichen Verwaltung, 2024, abgerufen unter: https://der-digitale-faktor.de/.

7 Etscheid, von Lucke, Stroh, in: Fraunhofer-Institut für Arbeitswirtschaft und Organisation IAO (Hrsg.): Künstliche Intelligenz in der öffentlichen Verwaltung – Anwendungsfelder und Szenarien, 2020, abgerufen unter: https://publica-rest.fraunhofer.de/server/api/core/bitstreams/d3d9f520-1fd4-4516-98d6-a3370c134155/content.

KI-Systemen dominierte Verwaltung“, „KI-gestützter Überwachungsstaat“ und „Konstruktive Kombination von menschlicher und künstlicher Intelligenz“ beschrieben, um einen Eindruck zu vermitteln, wie die Arbeitswelt mit KI in Zukunft aussehen könnte.

3. Einsatz von KI in der Justiz

3.1. Whitepaper Künstliche Intelligenz und Recht

Die vom Bundesministerium für Bildung und Forschung geförderte Plattform Lernende Systeme hat ein Whitepaper veröffentlicht, das einen Überblick über mögliche Anwendungen von KI-Systemen im Umfeld gerichtlicher Entscheidungen gibt.[8] Ausgehend vom Autonomiegrad der eingesetzten KI-Systeme wird deren Potenzial beispielsweise zur Entlastung der Justiz untersucht.

Als Anwendungsbeispiele werden u.a. die automatisierte Auskunftserteilung durch Chatbots in der Rechtsberatung, KI-gestützte Tools zur Unterstützung von Rechtsanwälten bei der Suche nach Gesetztestexten, Urteilen und Dokumenten oder bei der Erstellung von Schriftsätzen vorgestellt.

3.2. Digitalisierungsinitiative für die Justiz

Im Rahmen der „Digitalisierungsinitiative für die Justiz“ unterstützt das Bundesministerium der Justiz (BMJ) fachlich und finanziell Digitalisierungsvorhaben der Länder, die der gesamten Justiz zugutekommen und für die eine Bundeszuständigkeit besteht.[9]

Dazu gehört unter anderem die Erarbeitung einer KI-Strategie, die dazu beitragen soll, ein einheitliches Vorgehen von Bund und Ländern bei der Beschaffung, Entwicklung und Nutzung von KI-Anwendungen abzustimmen. Darüber hinaus wird der Aufbau einer KI-Plattform angestrebt, um technische Standards für noch in der Entwicklung befindliche KI-Systeme zu schaffen.[10]

3.2.1. Generatives Sprachmodell

In einem gemeinsamen Projekt des Justizministeriums des Landes Nordrhein-Westfalen und des Bayerischen Staatsministeriums der Justiz wird ein KI-Sprachmodell entwickelt, das beispielsweise zur Beantwortung von Fragen zum Inhalt langer Dokumente oder zum Auffinden relevanter Textstellen eingesetzt werden soll.[11] Weitere Anwendungsmöglichkeiten sind die strukturierte

---

8 Rostalki, Janal et al., in: Plattform Lernende Systeme (Hrsg.): Künstliche Intelligenz und Recht, München, 2024, S. 3, abgerufen unter: https://cta4.plattform-lernende-systeme.de/publikationen-details/ki-im-rechtswesen-chancen-und-herausforderungen-fuer-die-demokratie.html.

9 Bundesministerium der Justiz (BMJ): Digitalisierungsinitiative für die Justiz, abgerufen unter: https://www.bmj.de/DE/themen/digitales/digitalisierung_justiz/digitalisierungsinitiative/digitalisierungsinitiative_node.html.

10 BMJ: KI-Strategie und KI-Plattform, abgerufen unter: https://www.bmj.de/DE/themen/digitales/digitalisierung_justiz/digitalisierungsinitiative/laendervorhaben/_doc/artikel_vorhaben_05_ki.html.

11 BMJ: Generatives Sprachmodell der Justiz (GSJ), abgerufen unter: https://www.bmj.de/DE/themen/digitales/digitalisierung_justiz/digitalisierungsinitiative/laendervorhaben/_doc/artikel_vorhaben_06_gsj.html.

Darstellung von Argumenten der Parteien in Gerichtsakten oder die Zusammenfassung wichtiger Rechtssätze aus Urteilen höherer Instanzen. Darüber hinaus sollen Richtern zum Urteil passende Leitsätze vorgeschlagen und als Teil eines komplexen Systems vergleichbare Fälle zum Zitieren empfohlen werden.

3.2.2. Strukturierung von Justizverfahrensakten

Das Ministerium der Justiz und für Migration Baden-Württemberg koordiniert das Projekt „Strukturierung mit KI“, dessen Ziel die Entwicklung eines universellen Strukturierungstools für gerichtliche Verfahrensakten auf Basis einer KI-Anwendung ist.[12] Damit soll unter anderem die Bearbeitung von Prozesskostenhilfeanträgen, die Erstellung von Kostenbeschlüssen in Massenverfahren oder die Anonymisierung von Dokumenten verbessert werden. KI-Anwendungen, die sich bereits in der Praxis bewährt haben, z.B. in Fluggastrechteverfahren, sollen weiterentwickelt werden.

3.2.3. Maschinelle Übersetzungsplattform

Das Land Baden-Württemberg leitet federführend das Projekt „Maschinelle Übersetzungsplattform der Justiz“, in dem eine KI-gestützte Übersetzungssoftware für rechtliche Dokumente, Urteile oder Zeugenaussagen entwickelt werden soll.[13] Das geplante System soll die bidirektionale Übersetzung von mindestens 40 Sprachen ermöglichen.

3.2.4. Assistenz für Massenverfahren

Das Niedersächsische Justizministerium wurde mit der Entwicklung eines KI-Assistenztools für Massenverfahren beauftragt, bei dem die Software durch ein einmaliges Anlernen eines Musterfalls (sog. „one shot anotation“) auf wiederkehrende Fallmuster trainiert wird.[14] Beispiele für mögliche Anwendungen sind wiederkehrende Parteivorträge in zivilrechtlichen Streitfällen, Abrechnungspositionen in Kostenverfahren oder Daten aus maschinell erstellten Messprotokollen in Ordnungswidrigkeitenverfahren. Ein Schwerpunkt der Entwicklung liegt auf der Bearbeitung von Asylverfahren.

3.2.5. Anonymisierungs- und Leitsatzerstellungs-Kit

Der Freistaat Bayern hat in einem Forschungsprojekt mit der Friedrich-Alexander-Universität Erlangen-Nürnberg die automatisierte Anonymisierung von Gerichtsentscheidungen mit Hilfe

---

12 BMJ: Strukturierung von Justizverfahrensakten mit Hilfe von KI und KI-Apps, abgerufen unter: https://www.bmj.de/DE/themen/digitales/digitalisierung_justiz/digitalisierungsinitiative/laendervorhaben/_doc/artikel_vorhaben_07_struki.html.

13 BMJ: Maschinelle Übersetzungsplattform der Justiz, abgerufen unter: https://www.bmj.de/DE/themen/digitales/digitalisierung_justiz/digitalisierungsinitiative/laendervorhaben/_doc/artikel_vorhaben_08_maschinell.html.

14 BMJ: Massenverfahren mit KI schneller und effizienter bearbeiten, abgerufen unter: https://www.bmj.de/DE/themen/digitales/digitalisierung_justiz/digitalisierungsinitiative/laendervorhaben/_doc/artikel_vorhaben_13_MAKI.html.

computerlinguistischer Verfahren untersucht und einen ersten Prototyp entwickelt.[15] Der gesamte Veröffentlichungsprozess soll KI gestützt automatisiert werden, indem ein justizspezifisches Anonymisierungstool personenbezogene Daten in den Entscheidungen erkennt und durch Kürzel ersetzt. Dadurch soll die manuelle Anonymisierung von Entscheidungen überflüssig werden.

Darüber hinaus sollen Large Language Models bei der Erstellung von Leitsätzen und der Verschlagwortung von Entscheidungen helfen. Die Erfassung der Metadaten und der Versand der Entscheidungen sollen automatisiert erfolgen.

4. Haftungsfragen der KI

Eine Haftung der KI selbst ist nach geltendem Recht mangels eigener Rechtspersönlichkeit nicht vorgesehen.[16] Stattdessen kommt eine Haftung des Herstellers, des Eigentümers, des Betreibers bzw. Anwenders in Betracht.

Als Haftungssubjekt bzw. Risikoträger kommt grundsätzlich jeder in Frage, der mit dem jeweiligen automatisierten System in Berührung kommt.[17] Dies können Hersteller und Zulieferer, IT-Dienstleister, Zulassungsstelle, Inhaber oder Nutzer sein. Verantwortlich für KI-generierte Inhalte ist im Zweifel der Anwender, der KI-generierte Inhalte im eigenen Namen verwendet oder sich die Ergebnisse einer KI zu eigen macht.

4.1. Haftung des Anwenders oder Systembetreibers

Für den Anwender bzw. Betreiber von KI-Technologie kann sich eine Haftung entweder aus vertragsrechtlichen Grundsätzen oder aus dem Deliktsrecht ergeben.[18] Die deliktische Haftung ergibt sich dabei aus § 823 BGB.

Für den Geschädigten dürfte es sich häufig als Problem erweisen, hinsichtlich des Verschuldens nachzuweisen, dass der Anwender bzw. Betreiber von KI die im Verkehr erforderliche Sorgfalt außer Acht gelassen hat.[19] Dies betrifft insbesondere die deliktische Haftung, bei der der Geschädigte die schuldhafte Verletzung nachweisen muss. Auf die vertragsrechtliche Haftung trifft dies ebenfalls zu, allerdings gibt es hier Ausnahmen in Form der Beweislastumkehr beim Nachweis des Fehlers und des Verschuldens. So besteht beispielsweise im Bereich des Arzthaftungsrechts

---

15 BMJ: Gerichtsentscheidungen durch KI leichter veröffentlichen, abgerufen unter: https://www.bmj.de/DE/themen/digitales/digitalisierung_justiz/digitalisierungsinitiative/laendervorhaben/_doc/artikel_vorhaben_14_ALeKS.html.

16 Hotz: Persönlichkeitsrechtliche Haftung beim Einsatz autonomer Systeme, ZUM 2025, 89, 92.

17 Essers: Haftungsfragen automatisierter Systeme, Berlin, 2024, S. 111.

18 Lampe in: Hoeren, Sieber, Holznagel: Handbuch Multimedia-Recht, 62. EL Juni 2024, Teil 29.2 Rn. 16 ff.

19 Wildhaber: Außervertragliche Haftung trotz Blackbox? – Verschiedene mögliche Ansätze für eine Betreiberhaftung, in: Gebauer, Huber (Hrsg.): Künstliche Intelligenz – Zurechnung, Vertrag, Verantwortung, Tübingen 2024, S. 128.

mit § 630h Abs. 1 BGB eine Beweislastumkehr, die auch auf KI-gestützte Untersuchungen und Diagnosesysteme anwendbar ist.

Als Beispiel für die Haftung des Systembetreibers kann die Entscheidung des Landgerichts Kiel im Zusammenhang mit der Veröffentlichung von KI-generierten Falschinformationen auf einer Plattform dienen, die sich als geschäftsschädigend herausstellten.[20] Das Gericht entschied, dass der Betreiber dieser Plattform nach den allgemeinen Grundsätzen als Verwender der Software haftet. Der Vorwurf lautete, dass der Betreiber willentlich eine eigene KI-basierte Software zur Beantwortung von Suchanfragen auf seiner Website nutzte. Diese bewusst eingesetzte KI habe nicht erkannt, dass die Information falsch war, da sie für solche Fälle unzulänglich trainiert gewesen sei. Insofern hafte der Betreiber auch für von Dritten eingestellte falsche Informationen, wenn er sich diese aus Sicht der Nutzer zu eigen mache und erkennbar die inhaltliche Verantwortung für die Informationen übernehme.[21]

4.2. Herstellerhaftung

Eine Haftung des Herstellers kommt regelmäßig in Betracht, wenn die KI z.B. nicht die vertraglich zugesicherte Beschaffenheit aufweist oder der Hersteller keine ausreichenden Sicherheitsvorkehrungen innerhalb der KI getroffen hat, um einen Schaden zu vermeiden.[22] In diesem Fall haftet der Hersteller vor allem im Rahmen der Produkthaftung und der deliktischen Produzentenhaftung.

4.2.1. Produkthaftung

Eine Haftung des Herstellers kann sich aus den Vorschriften über die spezialgesetzliche Produkthaftung nach dem Produkthaftungsgesetz (ProdHaftG)[23] ergeben. Voraussetzung hierfür ist nach § 1 Abs. 1 ProdHaftG die Verletzung eines geschützten Rechtsguts (Tötung einer Person, Verletzung von Körper oder Gesundheit, Beschädigung einer Sache) durch ein fehlerhaftes Produkt mit daraus resultierendem (Vermögens-) Schaden. Darüber hinaus darf kein Ausschlussgrund nach § 1 Abs. 2, 3 ProdHaftG vorliegen.

Produkte sind nach § 2 ProdHaftG bewegliche Sachen und Elektrizität. Als Teil eines Produkts wird auch KI angesehen, soweit sie in Hardware implementiert ist.[24] Fehlt es an einer Hardware-Implementierung, so war bislang umstritten, ob auch eine solche Software zu den Produkten zählt oder nicht. Nach dem Wortlaut des § 2 ProdHaftG erfolgt eine ausdrückliche Ausdehnung neben beweglichen Sachen nur auf Elektrizität als nicht verkörperte Einheit. Software wird

---

20 Landgericht Kiel, Urteil v. 29.2.2024, Az. 6 O 151/23, GRUR-RS 2024, 29599, Rn. 36.

21 Landgericht Kiel, Urteil v. 29.2.2024, Az. 6 O 151/23, GRUR-RS 2024, 29599, Rn. 36.

22 Essers: Haftungsfragen automatisierter Systeme, Berlin, 2024, S. 111, 149 ff.

23 Produkthaftungsgesetz vom 15. Dezember 1989 (BGBl. I S. 2198), das zuletzt durch Artikel 5 des Gesetzes vom 17. Juli 2017 (BGBl. I S. 2421) geändert worden ist.

24 Wilhelmi: Haftung beim Einsatz von Künstlicher Intelligenz, in: Aufderheide/Dabrowski (Hrsg.): Digitalisierung und Künstliche Intelligenz: wirtschaftsethische und moralökonomische Perspektiven, Berlin, 2020, S. 121 f.

dagegen nicht erwähnt, so dass bisher fraglich war, ob für KI überhaupt einer Haftung nach dem ProdHaftG unterliegt.[25]

Diese Frage wird durch die neue Produkthaftungsrichtlinie (ProdHaftRL)[26] geklärt, die am 10. Oktober 2024 vom Rat der Europäischen Union formell verabschiedet und am 18. November 2024 im Amtsblatt veröffentlicht wurde. Die Mitgliedstaaten haben bis zum 9. Dezember 2026 Zeit, um die Richtlinie in nationales Recht umzusetzen. In Art. 4 Nr. 1 S.2 ProdHaftRL wird Software ausdrücklich als Produkt eingestuft und in Erwägungsgrund 3 auf KI Bezug genommen, wodurch klargestellt wird, dass die Produkthaftung auch für KI gelten soll.[27]

Ein Produkt ist nach § 3 ProdHaftG fehlerhaft, wenn es nicht die Sicherheit bietet, die berechtigterweise erwartet werden kann. Dabei kommt es auf die berechtigten Sicherheitserwartungen zum Zeitpunkt des Inverkehrbringens an.[28] Ob das Inverkehrbringen als maßgeblicher Zeitpunkt auch für autonome Systeme geeignet ist, wird unterschiedlich beurteilt.[29] Ein Fehler kann bei autonomen Systemen auch dadurch entstehen, dass sich das System im Rahmen des Selbstlernens ein unerwünschtes Verhalten antrainiert.[30] Der Fehler müsste dann dem Zeitpunkt des Inverkehrbringens zugeordnet werden können. Dies hängt wiederum davon ab, inwieweit die Gefahr einer solchen Entwicklung bereits beim Inverkehrbringen vorhersehbar war.[31]

Probleme beim Nachweis eines Fehlers ergeben sich regelmäßig dann, wenn die Frage nach der Vermeidbarkeit eines Schadensszenarios nicht beantwortet werden kann, weil die im autonomen System ablaufenden Vorgänge und damit die für die Verhaltenssteuerung maßgeblichen Regeln nicht einsehbar sind.[32]

Nur bei den Haftungsausschlüssen nach § 1 Abs. 2 und 3 ProdHaftG trägt der Hersteller bisher die Beweislast, die ansonsten nach § 1 Abs. 4 S. 1 ProdHaftG dem Geschädigten für den Fehler, den Schaden und den ursächlichen Zusammenhang zwischen Fehler und Schaden obliegt.

---

25 Wilhelmi: Haftung beim Einsatz von Künstlicher Intelligenz, in: Aufderheide/Dabrowski (Hrsg.): Digitalisierung und Künstliche Intelligenz: wirtschaftsethische und moralökonomische Perspektiven, Berlin, 2020, S. 121 f.

26 Richtlinie (EU) 2024/2853 des Europäischen Parlaments und des Rates vom 23. Oktober 2024 über die Haftung für fehlerhafte Produkte und zur Aufhebung der Richtlinie 85/374/EWG des Rates, ABl. L 2024/2853 vom 18.11.2024.

27 Borges: Die Haftung für Software und KI-Systeme nach der neuen Produkthaftungsrichtlinie, CR 1/2025, 1, 4.

28 BGH, Urteil vom 09.05.1995 - VI ZR 158/94, NJW 1995, 2162, 2163.

29 Essers: Haftungsfragen automatisierter Systeme, Berlin, 2024, S. 154.

30 Wilhelmi: Haftung beim Einsatz von Künstlicher Intelligenz, in: Aufderheide/Dabrowski (Hrsg.): Digitalisierung und Künstliche Intelligenz: wirtschaftsethische und moralökonomische Perspektiven, Berlin, 2020, S. 122.

31 Wagner: Produkthaftung für autonome Systeme, AcP, 217 (2017), 707, 750.

32 Wilhelmi: Haftung beim Einsatz von Künstlicher Intelligenz, in: Aufderheide/Dabrowski (Hrsg.): Digitalisierung und Künstliche Intelligenz: wirtschaftsethische und moralökonomische Perspektiven, Berlin, 2020, S. 125.

Die neue ProdHaftRL sieht hierzu in Art. 10 Beweiserleichterungen vor. Danach gilt das Produkt als fehlerhaft, wenn der Hersteller wichtige Informationen nicht offenlegt, wenn das Produkt nicht den vorgeschriebenen Sicherheitsstandards entspricht oder wenn es offensichtlich bei normaler Verwendung versagt. Darüber hinaus wird nach Art. 10 Abs. 3 ProdHaftRL ein Kausalzusammenhang zwischen dem Produktfehler und dem Schaden vermutet, wenn das Produkt offensichtlich defekt und der Schaden für diesen Defekt typisch ist. Ist der Nachweis für den Kläger in besonders komplizierten Fällen sehr schwierig, so wird nach Art. 10 Abs. 4 ProdHaftRL auch vermutet, dass das Produkt fehlerhaft ist und der Schaden dadurch verursacht wurde, sofern der Kläger nachweist, dass dies wahrscheinlich ist.[33]

4.2.2. Produzentenhaftung

Die deliktische Produzentenhaftung ist in dem allgemeinen Schadensersatzanspruch des § 823 Abs. 1 BGB geregelt. Sie verfolgt wie die Produkthaftung das Ziel, Verwender und Dritte vor den Gefahren zu schützen, die sich aus sicherheitsrelevanten Defiziten von Produkten ergeben können.[34]

Im Gegensatz zur Gefährdungshaftung im ProdHaftG ist die Produzentenhaftung verschuldensabhängig.[35] Eine Besonderheit der Produzentenhaftung liegt jedoch in der Beweiserleichterung bis hin zur Beweislastumkehr, so dass der Geschädigte nicht in jedem Fall die Fehlerhaftigkeit des Produkts beweisen muss.[36] Vielmehr muss sich der Hersteller entlasten, indem er nachweist, dass er kein fehlerhaftes Produkt in Verkehr gebracht hat und seinen Verkehrssicherungspflichten nachgekommen ist.

Der Fehlerbegriff des § 3 ProdHaftG deckt sich weitgehend mit dem deliktischen Fehlerbegriff.[37] Allerdings umfasst die Produzentenhaftung des § 823 Abs. 1 BGB nur Konstruktions- und Instruktionsfehler, während die Produkthaftung auch Fabrikationsfehler einschießt.[38]

Zudem ist der Anwendungsbereich der Produzentenhaftung, der praktisch jede Verletzung von Herstellerpflichten erfasst, weiter als bei der Produkthaftung. Der Anwendungsbereich nach § 2 ProdHaftG ist bislang noch auf bewegliche Sachen und Elektrizität beschränkt.

---

33 Zu weiteren Änderungen durch die Produkthaftungsrichtlinie siehe auch die Arbeit des Fachbereichs Europa: Maßnahmen der EU zur Regulierung von KI, EU 6 – 3000 – 001/25.

34 Förster, in: Hau/Poseck (Hrsg.): BeckOK BGB, 73. Edition 01.02.2025, BGB § 823 Rn. 677.

35 Krimphove: Künstliche Intelligenz im Recht – eine Übersicht, JURA – Juristische Ausbildung, 2021 (7) 764, 768.

36 Essers: Haftungsfragen automatisierter Systeme, Berlin, 2024, S. 164 ff.

37 BGH, Urteil vom 16.06.2009 - VI ZR 107/08, NJW 2009, 2952, Rn. 12.

38 Wilhelmi: Haftung beim Einsatz von Künstlicher Intelligenz, in: Aufderheide/Dabrowski (Hrsg.): Digitalisierung und Künstliche Intelligenz: wirtschaftsethische und moralökonomische Perspektiven, Berlin, 2020, S. 121.

Geschützte Rechtsgüter nach § 823 Abs. 1 BGB sind Leben, Körper, Gesundheit, Freiheit, Eigentum und sonstige Rechte. § 1 Abs. 1 S. 1 ProdHaftG nennt statt des Eigentums die Sachbeschädigung und erwähnt Freiheit oder sonstige Rechte nicht.

Im Rahmen der Produzentenhaftung sind Sachschäden ausnahmsweise auch am fehlerhaften Produkt selbst zu ersetzen, wenn es sich um eine Eigentumsverletzung in Form eines sog. „Weiterfresserschadens“ handelt, der nicht „stoffgleich“ mit dem bereits anfänglichen, mangelbedingten Minderwert der Sache ist.[39] Dagegen verlangt § 1 Abs. 1 S. 2 ProdHaftG ausdrücklich, dass eine andere Sache als das fehlerhafte Produkt beschädigt wird.

In zeitlicher Hinsicht stellt die Produkthaftung z.B. in § 1 Abs. 2 Nr. 5, § 3 Abs. 1 lit. c oder § 3 Abs. 2 ProdHaftG auf den Augenblick des Inverkehrbringens des schadensverursachenden Gegenstandes ab.[40] Die deliktische Haftung legt dem Hersteller darüber hinaus eine nachsorgende Produktbeobachtungspflicht auf, aus der sich weitergehende Warn- und Rückrufpflichten ergeben können.[41]

Der Kreis der Ersatzpflichtigen beschränkt sich bei der deliktischen Haftung regelmäßig auf den tatsächlichen Hersteller und den Zulieferer. Der sog. „Quasi-Hersteller“ haftet nur in engen Grenzen,[42] Importeure, Vertriebshändler oder Lieferanten haften nicht.[43] Nach § 4 ProdHaftG haften dagegen grundsätzlich alle Genannten.

§ 823 Abs. 1 BGB sieht keine Beschränkungen hinsichtlich des Umfangs des Schadensersatzes vor. § 10 Abs. 1 ProdHaftG beschränkt die Haftung bei Personenschäden durch gleiche Produkte mit gleichem Fehler auf 85 Mio. Euro und § 11 ProdHaftG sieht bei Sachschäden einen Selbstbehalt von 500 Euro vor.

§ 14 ProdHaftG regelt die Unabdingbarkeit der Ersatzpflicht des Herstellers, während die deliktische Haftung des Herstellers dispositiv ist und von den Parteien vertraglich abbedungen werden kann.[44]

---

39 BGH, Urteil vom 14.05.1985 - VI ZR 168/83, NJW 1985, 2420, 2421.

40 Wagner: Produkthaftung für autonome Systeme, AcP, 217 (2017), 707, 749.

41 Wilhelmi: Haftung beim Einsatz von Künstlicher Intelligenz, in: Aufderheide/Dabrowski (Hrsg.): Digitalisierung und Künstliche Intelligenz: wirtschaftsethische und moralökonomische Perspektiven, Berlin, 2020, S. 128; Förster, in: Hau/Poseck (Hrsg.): BeckOK BGB, 73. Edition 01.02.2025, BGB § 823 Rn. 686.

42 BGH, Urteil vom 14.06.1977 – VI ZR 247/75, BeckRS 1977, 30397288.

43 Förster, in: Hau/Poseck (Hrsg.): BeckOK BGB, 73. Edition 01.02.2025, BGB § 823 Rn. 686.

44 Förster, in: Hau/Poseck (Hrsg.): BeckOK BGB, 73. Edition 01.02.2025, BGB § 823 Rn. 686.

5. Risiken durch KI-Nutzung

5.1. International AI Safety Report 2025

Der von der britischen Regierung veröffentlichte Abschlussbericht „International Scientific Report on the Safety of Advanced AI“ untersucht die Risiken, die mit der Nutzung von breit anwendbarer KI (General-Purpose-KI) verbunden sind.[45] In drei Hauptabschnitten fasst der Bericht wissenschaftliche Erkenntnisse zu den drei Kernfragen zusammen, was General-Purpose-KI leisten kann, welche Risiken mit ihr verbunden sind und welche Techniken zur Risikominimierung bestehen. Der Bericht unterscheidet zwischen Risiken durch böswillige Nutzung, Risiken durch Fehlfunktionen und systemischen Risiken.

Bei KI-Systemen besteht demnach die Gefahr des Missbrauchs, wenn mit ihrer Hilfe gefälschte Inhalte erzeugt werden, die einzelne Personen gezielt schädigen sollen.[46] Beispiele dafür sind Stimmenimitationen, manipulierte Bilder oder Videos. Darüber hinaus wird die Technologie für Betrug, Desinformation und Fake-News eingesetzt. Das Wissen über die gesellschaftlichen Auswirkungen von Falschinformationen ist laut der Studie jedoch begrenzt. Bei leistungsfähigeren Modellen besteht ein erhöhtes Risiko, dass sie für Cyberangriffe genutzt werden, etwa indem autonome Bots nach neuen Schwachstellen in Software suchen oder zur Programmierung neuer Schadsoftware eingesetzt werden.

In einem zweiten Komplex befasst sich die Studie mit Fehlfunktionen, z.B. durch Halluzinationen.[47] Wenn KI falsche Informationen liefert und diese dann ungeprüft übernommen werden, kann dies laut Studie unter anderem zu Reputationsschäden sowie finanziellen und rechtlichen Nachteilen für Einzelpersonen und Organisationen führen. Darüber hinaus könne die vielen Modellen in Bezug auf Aspekten wie Ethnie, Geschlecht, Alter und Behinderung innewohnende Voreingenommenheit zu diskriminierenden Ergebnissen führen und Stereotypen verstärken. Zudem könnten KI-Systeme außer Kontrolle geraten, indem Anwendungen mehr oder weniger autonom agieren, ganze Arbeitsschritte und Prozesse selbständig übernehmen und sich möglicherweise irgendwann verselbständigen. Schon heute seien Sprachmodelle in der Lage, zu lügen und eigene Ziele zu verfolgen.

Zu den systemischen Risiken zählt der Bericht den Einfluss der künstlichen Intelligenz auf den Arbeitsmarkt.[48] Dazu gehört die Gefahr, dass die KI Aufgaben übernimmt und Menschen dadurch ihre Arbeitsplätze verlieren. Die Automatisierung durch Software habe bereits Auswirkungen auf

---

45 Government United Kingdom: International AI Safety Report. The International Scientific Report on the Safety of Advanced AI, 2025, abgerufen unter: https://assets.publishing.service.gov.uk/media/679a0c48a77d250007d313ee/International_AI_Safety_Report_2025_accessible_f.pdf.

46 Government United Kingdom: International AI Safety Report. The International Scientific Report on the Safety of Advanced AI, 2025, S. 62 ff. abgerufen unter: https://assets.publishing.service.gov.uk/media/679a0c48a77d250007d313ee/International_AI_Safety_Report_2025_accessible_f.pdf.

47 Government United Kingdom: International AI Safety Report. The International Scientific Report on the Safety of Advanced AI, 2025, S. 88 ff.

48 Government United Kingdom: International AI Safety Report. The International Scientific Report on the Safety of Advanced AI, 2025, S. 110 ff.

Arbeitsplätze in der Industrie, aber auch im Personalwesen und in der Medizin. Sie könnte auch die globale Ungleichheit verstärken und die Welt in Regionen, die KI-Systeme nutzen, und andere Regionen, die keinen Zugang zu dieser Technologie haben, spalten. Abhängigkeiten von großen KI-Unternehmen könnten entstehen, so dass technische Probleme einzelner Unternehmen viele Menschen betreffen. Mögliche Risiken sieht der Bericht auch im hohen Verbrauch von Energie, Wasser und Rohstoffen.

5.2. Einschätzungen des Bundesamtes für die Sicherheit in der Informationstechnik

Das Bundesamt für Sicherheit in der Informationstechnik (BSI) hat ein Dokument zu den derzeit wichtigsten Bedrohungen und den daraus resultierenden Risiken während der Planungs- und Entwicklungsphase, des Betriebs und der Nutzung von generativen KI-Modellen veröffentlicht.[49] Die aufgeführten Risiken werden vom BSI in drei unterschiedliche Kategorien eingeteilt: Risiken im Rahmen der ordnungsgemäßen Nutzung, Risiken durch missbräuchliche Nutzung und Risiken durch Angriffen auf generative KI-Modelle.

Risiken im Zusammenhang mit der ordnungsgemäßen Nutzung könnten z.B. mangelnde Kontrolle seitens der Nutzerinnen und Nutzer oder unerwünschte Ausgaben sein.[50] Sie ergeben sich unter anderem aus der stochastischen Natur der Modelle, der Zusammenstellung und dem Inhalt der Trainingsdaten sowie der Bereitstellung der Modelle als Dienstleistung durch externe Unternehmen.

Als zweite Kategorie nennt die Studie Risiken, die sich aus dem Missbrauch von KI-Modellen ergeben. Dabei bleibt die ursprüngliche Funktionsweise der Inhaltserzeugung unverändert und das Modell arbeitet in seiner ursprünglichen Funktion, jedoch für unerwünschte, schädliche und illegale Zwecke.[51] Insofern handele es sich nicht um Angriffe auf KI im Sinne der IT-Sicherheit, sondern um eine Ausnutzung der Modelle an sich. Beispiele seien die Generierung gefälschter Inhalte, das Vortäuschen einer Identität, das Sammeln und Aufbereiten von Wissen im Kontext krimineller Aktivitäten, die Re-Identifizierung von Personen aus anonymisierten Daten oder die Generierung von Schadsoftware.

In einer dritten Kategorie werden Risiken durch Angriffe auf generative KI-Modelle beschrieben.[52] Dabei kann es sich um sogenannte Poisoning Attacks handeln, die durch Vergiftung des angegriffenen Modells (der Trainingsdaten, der hinterlegten Wissensdaten oder des Modells selbst) eine Fehlfunktion oder Leistungsverschlechterung herbeiführen sollen.

---

49 Bundesamt für Sicherheit in der Informationstechnik: Generative KI-Modelle, 2025, S. 13 - 35, abgerufen unter: https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/KI/Generative_KI-Modelle.html.

50 Bundesamt für Sicherheit in der Informationstechnik: Generative KI-Modelle, 2025, S. 13 ff.

51 Bundesamt für Sicherheit in der Informationstechnik: Generative KI-Modelle, 2025, S. 18 ff.

52 Bundesamt für Sicherheit in der Informationstechnik: Generative KI-Modelle, 2025, S. 24 ff.

Privacy Attacks zielen hingegen darauf ab, Informationen über Trainingsdaten, im Betrieb verarbeitete Daten oder das generative KI-Modell zu rekonstruieren.[53] Evasion Attacks verfolgen das Ziel, die Eingabe in ein generatives KI-Modell so zu verändern, dass das Ausgabeverhalten des Modells gezielt manipuliert oder bestehende Schutzmechanismen umgangen werden können.[54]

6. Zusammenfassung

Die zentrale Anlauf- und Koordinierungsstelle für KI-Projekte in der Bundesverwaltung ist das Beratungszentrum für Künstliche Intelligenz. Im Rahmen der „Digitalisierungsinitiative für die Justiz“ unterstützt das Bundesministerium der Justiz auch KI-gestützte Projekte der Länder. Für Hersteller von KI-Systemen besteht mit der deliktischen Produzentenhaftung und der Produkthaftung bereits ein umfassendes Haftungsregime, das durch die neue Produkthaftungsrichtlinie weiter ergänzt wird. Risiken durch den Einsatz von KI lassen sich unter anderem in Risiken durch böswillige Nutzung, Risiken durch Fehlfunktionen und systemische Risiken unterscheiden.

***

53 Bundesamt für Sicherheit in der Informationstechnik: Generative KI-Modelle, 2025, S. 27.

54 Bundesamt für Sicherheit in der Informationstechnik: Generative KI-Modelle, 2025, S. 30.